Evaluating the Relative Predictive Validity of Measures of Self-Referential Processing for Depressive Symptom Severity

Ethel Siew Ee Tan[1], Hong Ming Tan[2], Kah Vui Fong[3], Chong Wei Ho[3], Chloe Teo[4], Zhao Yuan Tan[4], Rachel Jing Min Ong[5], Jerall Yu[3], Maxine Lee[3], An Rae Teo[3], Sin Kee Ong[6], Xin Ying Lim[5], Jin Lin Kee[7], Nikita Rane[1], Sheryl Yu Xuan Tey[1], Jussi Keppo[2], Geoffrey Chern-Yee Tan[1]*

[1]Institute of Mental Health, Singapore, Singapore
[2]Institute of Operations Research and Analytics, National University of Singapore, Singapore, Singapore
[3]College of Humanities and Sciences, National University of Singapore, Singapore, Singapore
[4]Lee Kong Chian School of Medicine
[5]Faculty of Social Sciences, National University of Singapore, Singapore, Singapore
[6]NHG Polyclinics, Singapore, Singapore
[7]Ministry of Education, Singapore, Singapore

***Corresponding author**

Keywords: self-schema$_1$, self-concept$_2$, self-referential processing$_3$, personality$_4$, depression$_5$

ABSTRACT

The self-referential encoding task (SRET) has a number of implicit measures which are associated with various facets of depression, including depressive symptoms. While some measures have proven robust in predicting depressive symptoms, their effectiveness can vary depending on the methodology used. Hence, understanding the relative contributions of population differences, word lists and calculation methods to these associations with depression, is crucial for translating the SRET into a clinical screening tool. This study systematically investigated the predictive accuracy of various SRET measures across different samples, including one hospital population matched with healthy controls and two university student populations, exposed to differing word lists. Participants completed the standard SRET and its variations, including Likert scales and matrix formats. Both standard and novel SRET measures were calculated and compared for their relative and incremental contribution to their associations with depression, with mean squared error (MSE) used as the primary metric for measuring predictive accuracy. Results showed that most SRET measures significantly predicted depressive symptoms in clinical populations but not in healthy populations. Notably, models with task modifications, such as Matrix Endorsement Bias and Likert Endorsement Bias, achieved the lowest mean squared error (MSE), indicating better predictive accuracy compared to standard Endorsement Bias measures. These findings imply that task modifications and the use of longer word lists may enhance the effectiveness of screening methods in both clinical and research settings, potentially improving early detection and intervention for depression.

INTRODUCTION

Screening and monitoring of Depression have traditionally relied on self-reported measures of depressive symptoms, capturing the individual's current mood but potentially overlooking underlying vulnerability or future risk (El-Den et al., 2018).[1] Self-referential processing (SRP), the processing of information related to one's self, provides a measure of the individual's perception of themselves, which is a core feature of depression (Northoff et al., 2006, Rogers et al., 1977; Symons and Johnson, 1997; Bentley et al., 2017).[2–5] The self-referential encoding task (SRET) has been widely used to measure individuals' self-schemas. In the SRET, an adjective is presented and participants are required to make binary decisions about whether the word describes themselves. Various measures including number of word endorsements, reaction time to decide, and number of words recalled are then collected to derive information regarding the participant's emotional self-biases (Derry & Kuiper, 1981; Bradley & Matthews, 1983).[6,7]

Numerous studies have explored the associations between SRET measures and depression, including symptoms, longitudinal course, treatment response, relapse and remission (Le Moult et al., 2017; Disner et al., 2017; Allison et al., 2021; van Kleef et al., 2023).[8–11] For instance, Dainer-Best et al. (2017) found that depressed individuals tended to endorse more negative words and nondepressed individuals endorsed more positive words.[12] Negative biases in SRP have also been linked to increased risk of recurrent depressive episodes (Le Moult et al., 2017), while deficits in SRP, such as slower drift rates in rejecting negative stimuli, have been shown to persist into remission among individuals with depression (Allison et al., 2021).[8,10]

Nonetheless, several SRET variables have been inconsistent in their associations with depressive symptoms across different studies. For example, Kiang et al. (2017) found no significant differences in the endorsement of positive and negative words among depressed individuals.[13] Nevertheless, those endorsing severe levels of depressive symptoms tended to endorse fewer positive words, suggesting a potential link with symptom severity.

This ambiguity in results extends to other SRET measures. Reaction time (RT), while initially shown to have no significant difference in clinically healthy and clinically depressed populations (Gotlib et al., 2004),[14] was found to be slower in depressed individuals when processing self-referential adjectives (Fritzche et al., 2013; Collins and Winer, 2023).[15,16] In addition, several studies enhanced the interpretability of the RT data by incorporating drift rates in their analysis, revealing a significant correlation between mean drift rates for positive and negative words and baseline depression levels. Similarly, recall bias, initially shown to be significant in individuals with recurrent depressive episodes (LeMoult et al., 2017),[8] showed no significant association with depressive symptom severity in a study by Dainer-Best et al. (2018).[12] These findings complicate the use of endorsement bias, RT and recall bias as reliable markers for depressive symptoms and underscores the complexity of assessing cognitive processes in MDD.

The inconsistencies between studies may be accounted for by population differences, procedural differences such as variations in the word lists used in the different studies and the method of calculating each variable across studies. For example, in Joorman et al.'s study (2006) paper[17], endorsement was calculated by taking the number of words endorsed in a valence category divided by the total number of words endorsed while in LeMoult et al.'s study (2017), endorsement was operationalised by taking the number of endorsed words in each valence category, thus introducing methodological inconsistencies.[8]

There is a need to better understand the relative contributions of population differences, word lists and calculation methods to these associations with depression if the SRET is to be clinically translated. In this study, we aim to systematically investigate the predictive validity of the various SRET measures across different samples exposed to differing word lists. The primary aim is to evaluate the relative strength of association with depressive symptoms between SRET-based measures. To this end, we draw together data collected from clinical, community and healthy populations where participants completed a common self-report measure for depressive symptoms and a self-referential encoding task encompassing a range of measures. The self-referential encoding task includes a wide range of words encompassing common word lists from other studies such as LeMoult et al. (2017)[8], and Frewen & Lundberg (2012)[18]. Each measure is then compared for its relative and incremental contribution to their associations with depression.

METHOD

Datasets and Paradigms

In this study, three primary datasets were utilized for the comparative analysis:

The first dataset (Dataset A) comprised a Self-Referential Encoding Task (SRET) consisting of 60 words, incorporating endorsement data, latencies, recall data, recognition data as well as Likert data for a separate set of 200 words. This task was administered to a sample of 191 participants, including 147 patients with past or current depressive or anxiety symptoms from the Institute of Mental Health (IMH) and 44 healthy controls. Participants were drawn from three studies conducted at IMH, including 86 IMH patients from the "Understanding the person, exploring change across psychotherapies" (Xchange) study, which also included 52 participants from the "Understanding the Person, Improving Psychotherapy: Preventing Relapse by targeting Emotional bias Modulation in PsychoTherapy" (PRE-EMPT) and 34 patients and 18 healthy controls from "The role of cholinergic dysfunction in the progression of depression" (CholDep) study. In the Choldep study, healthy controls were also recruited by word of mouth.

The second dataset (Dataset B) comprised another iteration of the SRET that consisted of 185 words, and has endorsement data, latencies and also components of an Other-Referential Processing (ORP) task. It was administered to 61 participants, who were recruited from the National University of Singapore (NUS) as part of an undergraduate thesis project.

The third dataset (Dataset C) also employed a SRET that consisted of 147 words and was administered to a separate sample of 97 participants, recruited at NUS for another undergraduate thesis project. This dataset included endorsement data, latencies and recall data.

Measures

Self-referential encoding task (SRET)

The SRET is a task designed to access an individual's self-relevant schemas (Derry and Kuiper, 1981),[6] that typically comprises three sequential segments: endorsement, distractor task, and incidental recall.

In the endorsement task, participants judged whether an adjective described them. Participants responded by pressing keys representing "not me" or "me" on a computer keyboard. Reaction times, measured in milliseconds were also recorded.

Following the endorsement task, participants engaged in a five-minute distractor task to minimise interference and memory consolidation of endorsed words. Subsequently, only participants in sample A and C completed an incidental recall task lasting one minute, during which they were prompted to recall the words that have been displayed to them in the endorsement task. Afterward, participants in sample A were also presented with a list of 120 words and asked to indicate whether each word had been presented during the endorsement task. The list comprised 60 words from the endorsement task and 60 distractor words. Finally, sample A participants were presented with a matrix containing 200 words and asked to select which words described them. For each selected word, participants from sample A rated how accurate the word described them on a scale of 1-4 (1 = Not at all, 4 = Completely accurate).

The following measures derived from the SRET were calculated to evaluate the observed responses.

Endorsement

Multiple endorsement variables were derived using varying calculation methods found in existing literature.

The Proportion of Negative Words Endorsed and Proportion of Positive Words Endorsed were computed as the number of positive/negative words endorsed divided by the number of positive/ negative words presented, respectively.

Endorsement bias was operationalised as the number of positive/negative words endorsed divided by the total number of words endorsed. Additionally, a variable representing the difference between negative and positive biases was calculated by subtracting the positive endorsement bias from the negative endorsement bias.

Latencies

Two reaction time (RT) variables were analyzed. Positive RT bias and Negative RT bias were calculated to assess the differences in reaction times between endorsing and rejecting negative or positive words.

Positive RT bias was calculated by using the formula: Positive RT Bias = (Mean RT of Endorsement of Positive Words - Mean RT of Non-Endorsement of Positive Words) / Average RT Across all Trial Types. Similarly, Negative RT bias was calculated with the formula: Negative RT Bias = (Mean RT of Endorsement of Negative Words - Mean RT of Non-Endorsement of Negative Words) / Average RT Across all Trial Types.

This method of determining RT bias corresponds with the approach employed in previous SRET studies (Connolly et al., 2016).[19]

Recall bias

Only words that were correctly recalled were considered. Positive Recall Bias was obtained by taking the number of positive recalled and endorsed words divided by the total number of recalled words. Similarly, Negative Recall Bias was calculated by dividing the number of negative recalled and endorsed words by the total recalled words. Additionally, the difference between the Negative Recall Bias and the Positive Recall Bias was calculated. This is to measure the relative strength of memory biases for positive and negative self-referential information.

Recognition bias

Negative Recognition Bias was operationalised as the number of correctly recognised and endorsed negative words divided by total number of words recognised correctly while Positive Recognition Bias was defined as the number of correctly recognised and endorsed positive words divided by total number of words recognised correctly.

Signal Detection Theory (SDT) examines how individuals distinguish meaningful information from "noise" (Lynn & Barrett, 2014).[20] The response bias metric, *c*, measures an individual's tendency to respond affirmatively or negatively. In this study, *c* reflects the tendency to endorse or reject a given adjective. A value of $c < 0$ represents a strong bias towards "Yes", while $c > 0$ represents a strong bias towards "No" (Stanislaw & Todorov, 1999).[21] As such, we calculated the measures $c+$, $c-$ and $c+$ minus $c-$ using the following formula:

Reference formula (Macmillan, 1993):

$$c = -\frac{\phi^{-1}(H) + \phi^{-1}(F)}{2}$$

*H denotes hit rate ($\frac{number\ of\ correctly\ recognised\ words\ that\ were\ endorsed}{Total\ number\ of\ words\ in\ categorisation\ task}$)

*F denotes false alarm rate ($\frac{number\ of\ incorrectly\ recognised\ words\ that\ were\ endorsed}{Total\ number\ of\ words\ in\ categorisation\ task}$)

Drift rate

Drift rates (*v*) for sample A were estimated using the drift diffusion model and it represents the speed and direction of information accumulation, with positive values suggesting a preference for the upper threshold ("Yes"). In the context of SRET, drift rate reflects the efficiency of processing whether negative or positive self-descriptive adjectives describe oneself.

Endorsement data and RT data underwent analysis using FAST-dm software, with parameters determined by the trial size of 60 words. Given the small sample size and expected presence of contaminated data, the Kolmogorov-Smirnov estimation test was chosen to run the program. Previous studies have also explored the association between SRET and depressive symptoms, incorporating drift rates as a measure (Dainer-Best et al., 2018; Castagna et al., 2023).[12,22] Two drift rate measures were derived: 1) Drift rates towards negative words (v-), indicating the speed and direction of information accumulation when processing negative stimuli, and 2) Drift rates towards positive words (v+), indicating the speed and direction of information accumulation when processing positive stimuli.

Matrix and Likert data

Similar to the endorsement measures, the Proportion of Matrix Negative words were calculated by taking the number of negative words endorsed in the matrix divided by the number of negative words in the matrix, and the same calculation was done for Proportion of Matrix Positive words. Negative Matrix Endorsement Bias was also calculated by taking the number of negative words endorsed in the matrix divided by total number of words endorsed in the matrix, and vice versa for Positive Matrix Endorsement Bias.

For the Likert measures, the data underwent recoding: responses of 1 were adjusted to 0, while responses of 2 to 4 were recoded as 1. Following recoding, the sum of negative and positive words was calculated. The same calculation was done for the Proportion of Likert Negative words and Proportion of Likert Positive words. Negative Likert Endorsement bias was computed by dividing the sum of negative words by the total number of non-zero responses, while Positive Likert Endorsement bias was calculated in a similar manner.

Overlapping words

As each of the three datasets utilized a unique set of word lists, with overlaps in words between them, we conducted regressions using only the overlapping words to investigate the influence of participant characteristics versus word lists on predictive differences. Across all three datasets, we identified a total of 23 overlapping words.

Depressive symptoms

The Quick Inventory of Depressive Symptoms (QIDS-16-SR) is a 16-item self-report measure that assesses the 9 criterion domains used to diagnose a major depressive disorder (Rush et al., 2003) and was a shortened version derived from the Inventory of Depressive Symptomatology (IDS).[23] Participants were asked to rate the severity of each of the 30 symptoms in the preceding 7 days on a scale of 0–3, with higher scores indicating greater symptom severity. The total score is calculated by summing 9 of the 30 items. The total score ranges from 0 to 27.

QIDS-16-SR was shown to have satisfactory psychometric properties: Cronbach's alpha was .86 in a sample of 596 adult outpatients diagnosed with chronic nonpsychotic major depressive disorder, displaying high internal consistency. QIDS-SR-16 total scores were also highly correlated with IDS-30-SR (r = .96) and HAM-D (r = .86) total scores. Overall, the QIDS-16-SR exhibited excellent psychometric properties, suggesting its utility as a brief assessment tool for depressive symptom severity in both clinical and research settings (Rush et al., 2003).

Multiple linear regression analyses

We conducted multiple linear regression analyses to examine the associations between the different SRET variables such as negative/ positive endorsement bias, negative/ positive latency bias with depressive symptoms, while controlling for demographic variables. Every SRET variable was entered as independent variables separately in regression models. Demographic variables, including age, gender, and education level, were included as covariates in the models to account for potential confounding effects.

Regression analyses were performed separately for each dataset (Dataset A, Dataset B, and Dataset C) to explore potential variations in the associations across different samples. R^2 values of the regression models will be utilised as a measure to evaluate the relative strength of associations between the SRET variables and depressive symptoms.

Mean Squared Error as Primary Metric

Mean square error (MSE) is measured as the average of the squared differences between each of the actual values and the predicted values by a model. The formula for calculating MSE is as follows:

$MSE = \frac{\sum_{i=1}^{n} (y_i - p_i)^2}{n}$, where y_i is the i^{th} instance of the actual value and p_i is the predicted value, and n is the total number of values (Tyagi et al., 2022).[24]

For regression tasks, MSE evaluates how well the actual data is represented by a model's predictions, with lower MSE values indicating better model performance. Hence, MSE values were used to select the models as they measure the prediction error of SRET measures (Li J, 2017)[25]. In this context, the model with a lower MSE value would indicate a better fit in determining the severity of depressive symptoms.

R^2 as Metric

R^2 quantifies the proportion of variance in depressive symptoms that is explained by the different SRET measures. Higher R^2 values indicate that a larger proportion of the variance in depressive symptoms can be accounted for by the SRET measures, suggesting stronger associations and provision of better fit, hence allowing for a comparison of the relative predictive power of the SRET variables in explaining variations in depressive symptoms[26].

All analyses were conducted using R (R Core Team, 2022) on RStudio version 12.0.353.

RESULTS

Comparison of SRET variables between datasets

In this study, we compared the SRET variables between three datasets: a hospital population with matched healthy controls (Dataset A) and two university populations (Datasets B and C) using R^2 values of endorsement, RT and recall. The demographic characteristics of participants of each dataset are presented in Supplementary Table 1. Regression analyses showed that endorsement bias variables and RT bias variables were significantly associated with depressive symptoms only for Dataset A, while no significant associations were found in datasets B and C, representing the non-depressed population. Please refer to Supplementary Tables 2, 3 for the regression models of the effect of endorsement bias and RT bias on depressive symptoms, across all three datasets and Supplementary Table 4 for the comparison of effects of recall bias variables for Datasets A and C. This also held true when only calculating measures from the 23 overlapping words between the datasets. Please refer to Supplementary Tables 5, 6 and 7 for the regression models of the effect of endorsement measures and RT measures on depressive symptoms for the 23 overlapping words, across all populations.

To investigate the consistent lack of significance observed in Datasets B and C, we conducted an F-test comparing the variance in depressive symptoms between Dataset A and the combined Datasets B and C at the alpha = 0.20 level of significance and there appeared to be a nonsignificant trend towards greater variance in depressive symptoms in Dataset A ($F(184, 157) =1.2372$, $p = .1695$). Subsequent analyses were conducted on Dataset A alone as it demonstrated more robust associations with depressive symptoms.

Comparison between measures

Regression analysis conducted on the 23 overlapping words between datasets revealed that the Proportion of Negative Words endorsed measure explained the most variance in depressive symptoms, followed by Negative Endorsement Bias and Proportion of Positive Words Endorsed. Additionally, significant contributions were observed from both Negative RT bias and Positive RT bias, albeit these measures explained lesser variance than endorsement measures. Please refer to Supplementary Tables 5 and 6 for comprehensive regression models of the 23 words across all datasets.

In Dataset A alone, Negative Endorsement Bias had a MSE value of 16.61 ($R2 = 0.444$, $F(8, 175) = 19.96$, $p = 1.54E-19$) and difference in endorsement bias variable with a MSE value of 16.65 ($R2 = 0.443$, $F(8, 175) = 19.86$, $p = 1.86E-19$). These two models had the lowest and comparable MSE values. This was followed by the Positive Endorsement Bias with a MSE value of 16.69 ($R2 = 0.441$, $F(8, 175) = 11.56$, $p = 5.81E-12$), Proportion of Negative Words endorsed with a MSE value of 16.72 (R2 = 0.440, F(8, 175) = 19.67, p = 2.67E-19) and lastly, proportion of positive words endorsed with a MSE value of 20.56 (R2 = 0.312, F(8, 175) = 11.31, p = 8.48E-12). Comparatively, the RT bias variables performed worse than endorsement bias variables. Negative RT bias had a MSE value of 26.96 (R2 = 0.097, F(8, 175) = 2.7, p =.01) and positive RT bias had the highest MSE value of 27.7 (R2 = 0.072, F(8, 175) = 1.95, p =.06). These findings suggest that endorsement bias variables displayed stronger predictive power compared to RT bias variables in Dataset A. Please refer to Supplementary Table 7 for the regression models of the 23 words in Dataset A.

Endorsement

The model with the Negative Endorsement Bias as a predictor had a MSE value of 15.47 and R^2 value of 0.482 ($F(8, 175) = 23.27$, $p = 3.73E^{-22}$) in Dataset A. Since Positive Endorsement Bias and the Difference in Endorsement Bias were statistically equivalent to Negative Endorsement Bias ($p = 3.73E^{-22}$), only the results for Negative Endorsement Bias are reported. The relatively low MSE values for Endorsement Bias measures indicate good predictive performance in dataset A. Overall, Endorsement Bias and its difference ranked sixth in its MSE values among the other SRET predictors.

The model with Negative RT Bias had a MSE value of 26.57 and a significant R^2 of 0.110 ($F(8, 175) = 3.10$, $p = .004$) in Dataset A. Positive RT bias had a slightly higher MSE value of 27.13 and a significant R^2 of 0.090 ($F(8, 175) = 2.52$, $p = .017$) in Dataset A. However, the MSE values for RT biases are outperformed by other predictors such as endorsement bias and matrix/likert endorsement bias. This suggests that the endorsement bias predictors provide a better fit in predicting depressive symptoms. Overall, Positive RT Bias ranked 21th and Negative RT Bias ranked 17th in its MSE values among the other SRET predictors.

Conversely, the Positive Recall Bias model had a MSE value of 27.37 and a R^2 of 0.080 ($F(8, 174) = 2.31$, $p = .028$) in Dataset A. Negative Recall Bias had a lower MSE value of 26.74 and a higher R^2 value of 0.106 ($F(8, 174) = 2.95$, $p = .006$). Similarly to RT Bias, the MSE values of Recall Bias are outperformed by the Endorsement Bias predictors. This once again suggests that the endorsement bias predictors provide a better fit in predicting depressive symptoms. For the Difference between Recall Biases model, it had a MSE value of 27.64. However, this model was not statistically significant ($R^2 = 0.070$, $F(8, 174) = 2.05$, $p = .051$). Overall, Negative Recall Bias ranked 19th, Positive Recall Bias ranked 22th, and the Difference in Recall Bias ranked 23th in their MSE values respectively. The Difference in Recall Bias ranked the lowest out of every predictor.

Recognition Bias and Drift Rates

Response biases for recognition of endorsed negative words (*c*-) and recognition of endorsed positive words (*c*+) yielded a MSE value of 22.42 (F (8, 171) = 10.62, $p = 4.51E^{-11}$, R^2

= 0.303) and 20.32 (F (8,170) = 7.48, p = 7.68E^{-8}, R^2 = 0.235) respectively. Notably, the Difference between the two response Biases (c+ *minus* c-) model yielded the lowest MSE value of 17.89 (F (8, 170) = 15.53, p = 1.18E^{-15}, R^2 = 0.39). These findings highlight that the model focusing on the difference between response biases may provide the most accurate prediction of depressive symptoms, as evidenced by its superior MSE performance. Overall, the model with difference between response biases ranked ninth among the other SRET predictors.

Furthermore, drift rates towards negative words and positive words were analysed in Dataset A to understand their relationship with depressive symptoms. Drift rate towards negative words (v-) yielded a MSE value of 20.88 (F (8, 171) = 9.69, p = 3.91E^{-10}, R^2 = 0.284). Drift Rate towards positive words (v+) had a lower MSE value of 20.35 (F (8, 171) = 10.59, p = 4.92E^{-11}, R^2 = 0.302). The MSE value suggests that Drift Rate towards positive words (v+) provides a slightly better fit for predicting depressive symptoms, compared to Drift rate towards negative words (v-). However, overall, the model with Drift Rate towards positive words (v+) ranked fourteenth among the other SRET predictors.

Endorsement bias using Matrix and Likert Scale

Negative Matrix Endorsement Bias, Positive Matrix Endorsement Bias and Difference in Matrix Endorsement Bias models in Dataset A all reported an MSE value of 12.87 and a R^2 value of 0.558 (F(8, 172) = 30.99, p = 1.74E^{-27}). The Matrix Endorsement Bias models have the second lowest MSE value among all the models tested on Dataset A. This suggests that the Matrix Endorsement Bias models are a good fit and are one of the best models for predicting depressive symptoms.

Following which, Negative Likert Endorsement Bias model reported a MSE value of 13.31 (F(8, 172) = 29.18, p = 2.80E^{-26}, R^2 = 0.542). Positive Likert Endorsement Bias had a MSE value of 14.31 (F(8, 172) = 25.43, p = 1.20E^{-23}, R^2 = 0.509). Both Positive and Negative Likert Endorsement Bias have low MSE values, suggesting that both models are good fits for predicting depressive symptoms in SRET measures. Notably, the difference between negative Likert endorsement bias demonstrated the lowest MSE value of all the models tested in Dataset A, with an MSE of 12.83 (F(8, 172) = 31.18, p = 1.30E^{-27}, R^2 = 0.559). This suggests that Difference in Negative Likert endorsement bias has the highest predictive accuracy of depressive symptoms.

Overall, the MSE values reported from all the models tested in Dataset A are consistent with the R^2 values, indicating which models provide the best fit for predicting depressive symptoms. The model with the lowest MSE value was the Difference in Positive and Negative Likert Endorsement Bias, followed by the Positive and Negative Matrix Endorsement Bias and their differences.

The Negative and Positive Likert endorsement bias ranked third and fourth, respectively. The Proportion of Matrix Negative Words Endorsed model reported the fifth lowest MSE value. Finally, the Positive and Negative Endorsement Bias and its differences ranked 6th place.

Conversely, the models with the highest MSE values included the Positive and Negative RT bias, Positive and Negative Recall bias with its differences, and finally, the Positive and Negative recognition bias. These models indicated a poorer fit for predicting depressive symptoms compared to those with lower MSE values.

All MSE rankings of each model are noted down in Table 1.

[Table 1]

Incremental Predictive Validity

Given that the endorsement variables are among the most commonly used and strongest predictors of depressive symptoms in the SRET literature, the count of negative and positive words endorsed were entered as predictors in the base regression model. This base regression model yielded a R^2 value of 0.453, and an MSE value of 15.71 ($F(2, 168) = 69.52$, $p = 1.01E^{-22}$). Subsequently, each of the remaining variables were individually added to the base regression model, and the change in R^2 values was examined to assess whether additional variables offer incremental predictive validity beyond the count of endorsed words.

The results show that the matrix endorsement bias variables (Negative Matrix Endorsement Bias, Positive Matrix Endorsement Bias, and the Difference in Matrix Endorsement Bias, Proportion of Matrix Negative Words Endorsed) and Likert Endorsement Bias variables (Difference in Likert Endorsement Bias, Negative Likert Endorsement Bias, Proportion of Likert Negative Words Endorsed) showed the most significant changes in MSE values from the base regression model. A greater decrease in MSE values upon adding each measure to the baseline model suggests that these measures provide substantial improvements in predicting depressive symptoms. Specifically, Negative Matrix Endorsement Bias improved the model's performance, resulting in a reduction of the MSE value by 1.67 compared to the baseline model. In contrast, Drift rates to Positive and Negative words did not yield significant improvements, contributing only minor reductions in MSE values of 0.09 and 0.17 respectively. Similar trends were observed with recall variables, RT bias variables and recognition bias variables, which also did not significantly enhance the model beyond the count of endorsed words.

Please refer to Table 2 for the MSE values as predictors are iteratively added individually to the base model, and the corresponding change in MSE is observed.

[Table 2]

Most influential subset of variables

Considering the wide range of predictors in our dataset, it was valuable to identify a subset of variables that best explain the variability in depressive symptoms. To achieve this, we employed both forward step regression and backward step regression techniques. These methods iteratively add or remove predictors based on their contribution to the model's predictive power.

Four models were generated to identify the optimal subset of variables using different criterions: (1) Forward Stepwise Regression with AIC as criterion, (2) Backward Stepwise Regression with AIC as criterion, (3) Forward Stepwise Regression with BIC as criterion and (4) Backward Stepwise Regression with BIC as criterion.

The selected models of each method are presented in Table 3.

[Table 3]

The best model is (2) Backward Stepwise Regression with AIC as the criterion as the model has the smallest AIC value and the lowest MSE value. Although the BIC value for this model is not the smallest, it is comparable to the BIC values of the other models. The selected predictors for (2) include Endorsement Bias, Positive Likert Endorsement Ratio, Negative Matrix Endorsement Bias. Notably, Negative Matrix Endorsement Bias was also included in the final models of the other three methods. Taken together, these findings suggest that having a larger number of words may serve as a more robust predictor for depressive symptoms. The implications of these results will be further discussed in the Discussion section.

Comparison of different word lists

Regression analysis was conducted to assess the predictive validity of endorsement bias across four different word lists: the 23 overlapping words list, 40 words list from LeMoult's study, the 60 words list used in Dataset A and the 200 words list used in the matrix for Dataset A.

The results showed that the MSE value for the 200 words list was the lowest among all the word lists analysed. Specifically, the 200 words list yielded an MSE value of 12.88 for Negative Endorsement Bias while the MSE values for the 23 words, 40 words and 60 words were 16.61, 16.16 and 15.47, respectively. For Positive Endorsement Bias and Difference in Endorsement Bias measures, a similar trend was observed as the MSE value of the 200 words list was the highest compared to the other word lists.

[Table 4]

DISCUSSION

The SRET offers multiple measures for assessing SRP and its associations with depressive symptoms. In this study, we investigated several key measures derived from the SRET and its task variations. These measures included endorsement bias measures, RT measures, drift rates, recall measures, recognition measures, as well as matrix measures and Likert measures. The aim was to determine which of these measures had the highest predictive accuracy in predicting depressive symptoms across different samples.

The comparison of SRET measures across the three distinct datasets revealed a trend: while SRET measures demonstrated significant predictive utility within the hospital and healthy controls population (Dataset A), this association was not observed in the university populations (Datasets B and C). This finding suggests that the SRET may hold greater predictive value within populations already experiencing depressive symptoms, rather than in non-clinical samples. This is corroborated by previous research (Segal et al., 1988; Collins and Winer, 2023),[16,27] which found that individuals with major depressive disorder (MDD) exhibit greater negative SRP bias, possessing more negative self-schemas compared to non-depressed individuals. Although a F-test was conducted to compare the variance of depressive symptoms between datasets, the result was surprisingly non-significant. However, the observed mean and variance differences between the two populations suggest a difference that warrants further investigation. The findings imply that, with a larger sample size, our hypothesis might yield significant results. Consequently, subsequent analyses were conducted on Dataset A alone as it demonstrated more robust associations with depressive symptoms. This focus on Dataset A can be attributed to the small sample size of the university datasets. Future studies could explore the efficacy of the SRET in non-clinical samples with larger sample sizes, to further validate the use of the SRET beyond clinical populations.

Within Dataset A, the comparative analysis revealed that endorsement bias emerged as one of the strongest SRET measures in predicting depressive symptoms, ranking sixth overall. Notably, it performed better than the count of negative and positive words endorsed, and proportion of negative and positive words endorsed. This is corroborated by Dainer-Best et al. (2018),[12] whose research indicated that endorsement measures were among the strongest predictors of depressive symptoms. Given the multitude of calculation methods for endorsement data, this finding suggests that utilising endorsement bias may enhance the predictive power of endorsement measures in assessing depressive symptoms.

Although conventional RT bias measures significantly predicted depressive symptoms, our findings suggest that drift rates may be better predictors of depressive symptoms. Previous research indicates that depressed individuals often show a heightened attentional bias towards negative stimuli, which can be hypothesised to manifest as quicker responses to negative words due to increased salience for depressed individuals (Donaldson, Lam, and Matthews, 2007).[28] This is supported by Fritzsche et al. (2013) and Collins and Winer (2023),[15,16] who revealed that depressed individuals consistently exhibit slower reaction times than their non-depressed counterparts, indicating a prolonged decision-making process when processing self-referential adjectives. These findings complicate the use of RT as a reliable marker for depressive symptoms and underscores the complexity of assessing cognitive processes in MDD. The drift-diffusion model emerged as a better predictor than conventional SRET measures, such as RT bias, ranking at fourteenth place overall. Despite this, the analysis showed that the drift rates did not offer substantial incremental predictive value beyond endorsement measures. This highlights the need for further exploration and development of alternative models to improve the predictive validity of RT measures in assessing depressive symptoms. Consistent with findings from previous research (Dainer-Best et al., 2018),[12] our study did not find robust associations between recall measures and depressive symptoms. This suggests that recall performance may not be a reliable predictor of depressive symptoms across different populations and methodologies. Further investigations are warranted to better understand the role of recall measures in assessing depressive symptomatology.

In our current study, we expanded the conventional SRET task by incorporating additional components such as a recognition task, presenting the SRET in a matrix format and including a Likert scale for a more nuanced analysis of SRP. For the recognition task, the use of SDT to analyse participants' recognition of self-descriptive adjectives yielded more robust results as compared to simply calculating the proportion of correctly recognised and endorsed negative and positive words. Specifically, the response bias measure (c+ minus c-) was best in predicting depressive symptoms amongst the recognition measures and ranked ninth overall, suggesting that individuals' tendencies to both endorse and reject self-descriptive adjectives are crucial for predicting depressive symptoms. Hence, SDT offers a novel and promising approach to understanding SRP and its relationship to depressive symptoms (Macmillan, 1993).[29] Its suitability for analysing SRET data, particularly in the context of making self-referential judgements under conditions of uncertainty, makes it a valuable tool for advancing our understanding of depression. By utilising SDT, we can gain deeper insights into the cognitive biases that underlie depressive symptomatology, thereby enhancing both theoretical knowledge and clinical practice.

The present findings suggest that the matrix endorsement bias and the difference in Likert endorsement bias variables are the most robust SRET measures in terms of their predictive accuracy for depressive symptoms as they have the lowest MSE values and highest R^2 values, ranking the top amongst other SRET measures. Furthermore, when these measures were incrementally added to the base regression model, the MSE consistently decreased, and the R^2 value increased, indicating a significant improvement in the model's explanatory power to depressive symptoms. The study employed forward and backward stepwise regression models to determine the best subset of measures. Interestingly, Negative Matrix Endorsement Bias consistently appeared in all four models, regardless of whether AIC or BIC was used as a stopping criteria, suggesting its robust predictive power across different model selection techniques. Therefore, Negative Matrix Endorsement Bias, which is indicative of a depressotypic processing style (Duyser et al., 2020),[30] might be a reliable marker for screening for depressive symptoms. Consistent with our findings, the Endorsement Bias measures derived from the 200 words list demonstrated superior predictive validity, as evidenced by having the lowest MSE value compared to the other words lists.

These results have several implications to the SRET. First, expanding the word list in the SRET appears to be more beneficial for predicting depressive symptoms. Existing literature has well-documented that depressogenic patterns of self-schemas are associated with concurrent and prospective depressive symptoms in clinical samples of depressed adults (Gotlib et al., 2004; Fritzsche et al., 2010),[14,15] adolescents (Auerbach et al., 2015),[31] as well as in non-clinical samples of youth (Connolly et al., 2016; Liu et al., 2020).[19,32] By expanding the word list, our study may have captured a broader spectrum of self-referential thoughts and feelings, allowing for a richer representation of individuals' self-schemas, potentially enhancing the sensitivity of the SRET for screening for depressive symptoms (Derry & Kuiper, 1981)[6]. Furthermore, the expanded word list could help in identifying subtle nuances in self-referential processing that might be missed with a shorter list. Future research should explore the optimal length and content of word lists in SRET to maximise its predictive power. Clinically, using a more comprehensive word list in SRET could improve the accuracy of depressive symptom screening and better inform treatment strategies.

Next, presenting the SRET in a matrix format allowed participants to engage with a variety of self-descriptive adjectives simultaneously. The simultaneous presentation of the words might promote more effective encoding and retrieval of self-related information, potentially enhancing the sensitivity and accuracy of the task for screening depressive symptoms. Research by Reinitz and Hannigan (2001)[33] suggests that simultaneously-presented stimuli evoked a significantly larger effect on recall due to having stimuli in closer proximity with each other relative to sequentially-presented stimuli. This suggests that presenting stimuli together (simultaneously) can lead to better encoding and retrieval processes. Additionally, memory effects are enhanced. Additionally, a study by Bharti, Yadav, and Jaswal (2020)[34] also found that simultaneous presentation affects the organisation of stimuli in the visual working memory differently than sequential presentation, possibly leading to more efficient cognitive processing. Cognitive processes of memory are closely related to self-referential processing (Sajonz et al., 2010, Sui & Humphreys, 2015),[35,36] suggesting that the presentation of SRET stimuli in a matrix format might better facilitate the integration of self-referential information. Hence, the matrix format could still provide a richer and more accurate representation of self-schemas and may offer valuable insights for screening depressive symptoms. Future research could explore the neural underpinnings associated with sequential versus simultaneous presentation of the self–descriptive words to better understand these cognitive processes and their impact on self-referential encoding.

CONCLUSION

In summary, our study adds to the extensive literature on the SRET in depression by providing a head to head comparison of the predictive accuracy and incremental predictive value between SRET measures in the prediction of depressive symptoms. Our findings underscore the importance of expanding the word list and employing a matrix format for the simultaneous presentation of stimuli, as these methods enhance the sensitivity and accuracy of the SRET as a screening tool. The implications of this study are significant for both research and clinical practice. By refining the SRET with innovative measures and formats, we can improve early identifications and screening of depressive symptoms, leading to more targeted and effective interventions. Future research should continue to explore the neural and cognitive mechanisms underlying different presentation formats and further optimise the SRET by refining word selection to maximise its predictive accuracy. In conclusion, our study contributes to a deeper understanding of cognitive biases in depression and demonstrates the value of innovative approaches to the SRET. These findings offer valuable tools for improving mental health screening and assessment.

CONFLICTS OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

ESET: Writing of manuscript, performing statistical analyses of the different datasets, writing the code to perform statistical analyses, data cleaning, and creating the tables. HMT and JK: Provided guidance on machine learning analysis and scripts. KVF: Performed statistical analyses of the different datasets, data cleaning and refining the code for statistical analysis. CWH: Performed statistical analyses for different datasets, writing of scripts for statistical analysis, and contributed to writing of the manuscript. CT: Contributed to writing of manuscript. ZYT: Calculated drift rates for Dataset A. RJMO: Calculated response bias measures. JY: Performed data cleaning and calculation of LeMoult words, wrote the code for data cleaning of subset of words. ML: Wrote up justification for results in discussion section and collated supplementary materials. ART: Performed data cleaning of patient datasets. SKO, XYL: study PIs for university datasets, data collection of university datasets used in the analysis of the manuscript. JLK: Data collection of community datasets and data cleaning. NR, SYXT: data collection of patient datasets used in the analysis of the manuscript. GCYT: study PI for patient datasets, study conceptualisation and design, collection of data, supervision of analysis and writing of manuscript.

FUNDING

This research is supported by A*STAR Brain-Body Initiative (BBI) (#21718), IMH Research Seed Funding (642-2018), LKCMed-NUSMed-NHG Collaborative Mental Health Research Pilot Grant Call 2020 (MHRPG/2003) and FY21 PRENATAL / EARLY CHILDHOOD GRANT CALL (H22P0M0005).

ACKNOWLEDGEMENTS

We thank the study participants of the Xchange and CholDep studies.

REFERENCES

Allison, G. O., Kamath, R. A., Carrillo, V., Alqueza, K. L., Pagliaccio, D., Slavich, G. M., Shankman, S. A., & Auerbach, R. P. (2021). Self-referential Processing in Remitted Depression: An Event-Related Potential Study. *Biological Psychiatry Global Open Science*, *3*(1), 119–129. https://doi.org/10.1016/j.bpsgos.2021.12.005

Auerbach, R. P., Stanton, C. H., Proudfit, G. H., & Pizzagalli, D. A. (2015). Self-referential processing in depressed adolescents: A high-density event-related potential study. *Journal of Abnormal Psychology*, *124*(2), 233.

Beck: Depression: Clinical, experimental, and theoretical... - Google Scholar. (n.d.). Retrieved March 4, 2024, from https://scholar.google.com/scholar_lookup?author=A.+T.+Beck&publication_year=1967&journal=Depression:+clinical+experimental+and+theoretical+aspects

Bentley, S. V., Greenaway, K. H., & Haslam, S. A. (2017). An online paradigm for exploring the self-reference effect. *PLOS ONE*, *12*(5), e0176611. https://doi.org/10.1371/journal.pone.0176611

Bharti, A. K., Yadav, S. K., & Jaswal, S. (2020). Feature Binding of Sequentially Presented Stimuli in Visual Working Memory. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00033

Bradley, B., & Mathews, A. (1983). Negative self-schemata in clinical depression. *British Journal of Clinical Psychology*, *22*(3), 173–181. https://doi.org/10.1111/j.2044-8260.1983.tb00598.x

Castagna, P. J., Waters, A. C., Edgar, E. V., Budagzad-Jacobson, R., & Crowley, M. J. (2023). Catch the drift: Depressive symptoms track neural response during more efficient decision-making for negative self-referents. *Journal of Affective Disorders Reports*, *13*, 100593. https://doi.org/10.1016/j.jadr.2023.100593

Chodavadia, P., Teo, I., Poremski, D., Fung, D. S. S., & Finkelstein, E. A. (2023). Prevalence and economic burden of depression and anxiety symptoms among Singaporean adults: Results from a 2022 web panel. *BMC Psychiatry*, *23*(1), 104. https://doi.org/10.1186/s12888-023-04581-7

Collins, A. C., & Winer, E. S. (2023). Self-Referential Processing and Depression: A Systematic Review and Meta-Analysis. *Clinical Psychological Science*, 21677026231190390. https://doi.org/10.1177/21677026231190390

Connolly, S. L., Abramson, L. Y., & Alloy, L. B. (2016). Information processing biases concurrently and prospectively predict depressive symptoms in adolescents: Evidence from a self-referent encoding task. *Cognition and Emotion*, *30*(3), 550–560. https://doi.org/10.1080/02699931.2015.1010488

Dainer-Best, J., Lee, H. Y., Shumake, J. D., Yeager, D. S., & Beevers, C. G. (2018). Determining optimal parameters of the self-referent encoding task: A large-scale examination of self-referent cognition and depression. *Psychological Assessment*, *30*(11), 1527–1540. https://doi.org/10.1037/pas0000602

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, *90*(4), 286–297. https://doi.org/10.1037/0021-843X.90.4.286

Disner, S. G., Shumake, J. D., & Beevers, C. G. (2017). Self-referential schemas and attentional bias predict severity and naturalistic course of depression symptoms.

Cognition and Emotion, *31*(4), 632–644. https://doi.org/10.1080/02699931.2016.1146123

Donaldson, C., Lam, D., & Mathews, A. (2007a). Rumination and attention in major depression. *Behaviour Research and Therapy*, *45*(11), 2664–2678. https://doi.org/10.1016/j.brat.2007.07.002

Donaldson, C., Lam, D., & Mathews, A. (2007b). Rumination and attention in major depression. *Behaviour Research and Therapy*, *45*(11), 2664–2678. https://doi.org/10.1016/j.brat.2007.07.002

Duyser, F. A., Van Eijndhoven, P. F. P., Bergman, M. A., Collard, R. M., Schene, A. H., Tendolkar, I., & Vrijsen, J. N. (2020). Negative memory bias as a transdiagnostic cognitive marker for depression symptom severity. *Journal of Affective Disorders*, *274*, 1165–1172. https://doi.org/10.1016/j.jad.2020.05.156

El-Den, S., Chen, T. F., Gan, Y.-L., Wong, E., & O'Reilly, C. L. (2018). The psychometric properties of depression screening tools in primary healthcare settings: A systematic review. *Journal of Affective Disorders*, *225*, 503–522. https://doi.org/10.1016/j.jad.2017.08.060

Fritzsche, A., Watz, H., Magnussen, H., Tuinmann, G., Löwe, B., & von Leupoldt, A. (2013). Cognitive biases in patients with chronic obstructive pulmonary disease and depression – a pilot study. *British Journal of Health Psychology*, *18*(4), 827–843. https://doi.org/10.1111/bjhp.12025

Gotlib, I. H., Kasch, K. L., Traill, S., Joormann, J., Arnow, B. A., & Johnson, S. L. (2004). Coherence and Specificity of Information-Processing Biases in Depression and Social Phobia. *Journal of Abnormal Psychology*, *113*(3), 386–398. https://doi.org/10.1037/0021-843X.113.3.386

Hsu, T.-Y., Liu, T.-L., Cheng, P. Z., Lee, H.-C., Lane, T. J., & Duncan, N. W. (2021). Depressive rumination is correlated with brain responses during self-related processing. *Journal of Psychiatry and Neuroscience*, *46*(5), E518–E527. https://doi.org/10.1503/jpn.210052

Joormann, J., Dkane, M., & Gotlib, I. H. (2006). Adaptive and Maladaptive Components of Rumination? Diagnostic Specificity and Relation to Depressive Biases. *Behavior Therapy*, *37*(3), 269–280. https://doi.org/10.1016/j.beth.2006.01.002

Kiang, M., Farzan, F., Blumberger, D. M., Kutas, M., McKinnon, M. C., Kansal, V., Rajji, T. K., & Daskalakis, Z. J. (2017). Abnormal self-schema in semantic memory in major depressive disorder: Evidence from event-related brain potentials. *Biological Psychology*, *126*, 41–47. https://doi.org/10.1016/j.biopsycho.2017.04.003

LeMoult, J., Kircanski, K., Prasad, G., & Gotlib, I. H. (2017). Negative Self-Referential Processing Predicts the Recurrence of Major Depressive Episodes. *Clinical Psychological Science : A Journal of the Association for Psychological Science*, *5*(1), 174–181. https://doi.org/10.1177/2167702616654898

Li, J. (2017a). Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? *PLOS ONE*, *12*(8), e0183250. https://doi.org/10.1371/journal.pone.0183250

Li, J. (2017b). Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? *PLOS ONE*, *12*(8), e0183250. https://doi.org/10.1371/journal.pone.0183250

Liu, P., Vandemeer, M. R. J., Joanisse, M. F., Barch, D. M., Dozois, D. J. A., & Hayden, E. P. (2020). Depressogenic self-schemas are associated with smaller regional grey matter volume in never-depressed preadolescents. *NeuroImage: Clinical*, *28*, 102422. https://doi.org/10.1016/j.nicl.2020.102422

Lynn, S. K., & Barrett, L. F. (2014). "Utilizing" Signal Detection Theory. *Psychological Science*, *25*(9), 1663–1673. https://doi.org/10.1177/0956797614541991

Macmillan, N. A. (1993). Signal detection theory as data analysis method and psychological decision model. In *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 21–57). Lawrence Erlbaum Associates, Inc.

Maier, H. B., Neyazi, A., Bundies, G. L., Meyer-Bockenkamp, F., Bleich, S., Pathak, H., Ziert, Y., Neuhaus, B., Müller, F.-J., Pollmann, I., Illig, T., Mücke, S., Müller, M., Möller, B. K., Oeltze-Jafra, S., Kacprowski, T., Voges, J., Müntefering, F., Scheiber, J., … Frieling, H. (2024). Validation of the predictive value of BDNF -87 methylation for antidepressant treatment success in severely depressed patients—A randomized rater-blinded trial. *Trials*, *25*(1), 247. https://doi.org/10.1186/s13063-024-08061-5

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2013). *Introduction to Linear Regression Analysis*. John Wiley & Sons.

Myers, C. E., Interian, A., & Moustafa, A. A. (2022). A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.1039172

Nejad, A., Fossati, P., & Lemogne, C. (2013). Self-Referential Processing, Rumination, and Cortical Midline Structures in Major Depression. *Frontiers in Human Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00666

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage*, *31*(1), 440–457. https://doi.org/10.1016/j.neuroimage.2005.12.002

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Reinitz, M. T., & Hannigan, S. L. (2001). Effects of Simultaneous Stimulus Presentation and Attention Switching on Memory Conjunction Errors. *Journal of Memory and Language*, *44*(2), 206–219. https://doi.org/10.1006/jmla.2000.2727

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., & Manber, R. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, *54*(5), 573–583.

Sajonz, B., Kahnt, T., Margulies, D. S., Park, S. Q., Wittmann, A., Stoy, M., Ströhle, A., Heinz, A., Northoff, G., & Bermpohl, F. (2010). Delineating self-referential processing from episodic memory retrieval: Common and dissociable networks. *NeuroImage*, *50*(4), 1606–1617. https://doi.org/10.1016/j.neuroimage.2010.01.087

Segal, Z. V., Hood, J. E., Shaw, B. F., & Higgins, E. T. (1988). A structural analysis of the self-schema construct in major depression. *Cognitive Therapy and Research*, *12*(5), 471–485. https://doi.org/10.1007/BF01173414

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Sui, J., & Humphreys, G. W. (2015). The Integrative Self: How Self-Reference Integrates Perception and Memory. *Trends in Cognitive Sciences*, *19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Tan, G. C.-Y., Wang, Z., Tan, E. S. E., Ong, R. J. M., Ooi, P. E., Lee, D., Rane, N., Tey, S. Y. X., Chua, S. Y., Goh, N., Lam, G. W., Chakraborty, A., Yew, A. K. L., Ong, S. K., Kee, J. L., Lim, X. Y., Hashim, N., Lu, S. H., Meany, M., … Keppo, J. (2024). Transdiagnostic clustering of self-schema from self-referential judgements identifies subtypes of healthy personality and depression. *Frontiers in Neuroinformatics*, *17*. https://doi.org/10.3389/fninf.2023.1244347

Tyagi, K., Rane, C., Harshvardhan, & Manry, M. (2022). Chapter 4—Regression analysis. In R. Pandey, S. K. Khatri, N. kumar Singh, & P. Verma (Eds.), *Artificial Intelligence and Machine Learning for EDGE Computing* (pp. 53–63). Academic Press. https://doi.org/10.1016/B978-0-12-824054-0.00007-1

van Kleef, R. S., Kaushik, P., Besten, M., Marsman, J.-B. C., Bockting, C. L. H., van Vugt, M., Aleman, A., & van Tol, M.-J. (2023a). Understanding and predicting future relapse in depression from resting state functional connectivity and self-referential processing. *Journal of Psychiatric Research*, *165*, 305–314. https://doi.org/10.1016/j.jpsychires.2023.07.034

van Kleef, R. S., Kaushik, P., Besten, M., Marsman, J.-B. C., Bockting, C. L., van Vugt, M., Aleman, A., & van Tol, M.-J. (2023b). Understanding and predicting future relapse in depression from resting state functional connectivity and self-referential processing. *Journal of Psychiatric Research*, *165*, 305–314.

Young, K. D., Friedman, E. S., Collier, A., Berman, S. R., Feldmiller, J., Haggerty, A. E., Thase, M. E., & Siegle, G. J. (2020). Response to SSRI intervention and amygdala activity during self-referential processing in major depressive disorder. *NeuroImage: Clinical*, *28*, 102388. https://doi.org/10.1016/j.nicl.2020.102388

TABLES

TABLE 1 | Regression Analysis of Endorsement Bias for Full Word List with Depressive Symptoms for Dataset A

			95% CI				Model			
Variable	*B*	*SE*	LL	UL	*t*	*p*	R^2	MSE	*F(df)*	*p*
Endorsement Bias										
Proportion of Negative Words Endorsed (MSE rank: 7)	12.57	1.13	10.34	14.81	11.11	$4.75E^{-22}$	0.456	16.24	20.97 (8, 175)	$2.33E^{-20}$
Proportion of Positive Words Endorsed (MSE rank: 11)	-11.29	1.29	-13.83	-8.75	-8.78	$1.42E^{-15}$	0.356	19.22	13.83 (8, 175)	$3.31E^{-14}$
Negative Endorsement Bias† (MSE rank: 6)	14.41	1.22	11.99	16.83	11.77	$6.43E^{-24}$	0.482	15.47	23.27 (8, 175)	$3.73E^{-22}$
Positive Endorsement Bias† (MSE rank: 6)	-14.41	1.22	-16.83	-11.99	-11.77	$6.43E^{-24}$	0.482	15.47	23.27 (8, 175)	$3.73E^{-22}$
Difference in Endorsement Bias (MSE rank: 6)	7.21	0.61	6	8.41	11.77	$6.43E^{-24}$	0.482	15.47	23.27 (8, 175)	$3.73E^{-22}$
RT Bias										
Positive RT Bias (MSE rank: 21)	-16.00	8.33	-32.44	0.44	-1.92	.056	0.09	27.13	2.52 (8, 175)	.017
Negative RT Bias (MSE rank: 17)	-7.03	2.57	-12.10	-1.95	-2.73	.007	0.11	26.57	3.10 (8, 175)	.004
Drift Rates										
Drift rates to Negative words (*v*-) (MSE rank: 15)	2.76	0.4	1.98	3.54	6.96	$7.00E^{-11}$	0.284	20.88	9.69 (8, 171)	$3.91E^{-10}$
Drift rates to Positive words (*v*+) (MSE rank: 14)	-3.04	0.41	-3.86	-2.23	-7.36	$7.27E^{-12}$	0.302	20.35	10.59 (8, 171)	$4.92E^{-11}$
Recall Bias										
Positive Recall Bias (MSE rank: 22)	4.58	3.02	-1.38	10.53	1.52	.131	0.08	27.3727.37	2.31 (8, 174)	.028
Negative Recall Bias (MSE rank: 19)	8.18	3.22	1.82	14.55	2.54	.012	0.106	26.7426.74	2.95 (8, 174)	.006
Difference in Recall Bias (MSE rank: 23)	2.30	2.95	-3.53	8.13	0.78	.437	0.07	27.64	2.05 (8, 174)	.051
Recognition Bias										
Negative Recognition Bias (MSE rank: 20)	2.37	2.73	-3.03	7.76	0.87	0.39	0.057	27.08	2.51 (5, 166)	0.04
Positive Recognition Bias (MSE rank: 18)	-3.88	2.21	-8.25	0.48	-1.76	0.08	0.07	26.70	3.13 (5, 166)	0.02
Response bias for recognition of endorsed positive words (*c*+) (MSE rank: 16)	2.52	0.43	1.67	3.37	5.85	$2.41E^{-08}$	0.235	22.42	7.48 (8, 170)	$7.68E^{-8}$
Response bias for recognition of endorsed negative words (*c*-) (MSE rank: 13)	-3.09	0.42	-3.91	-2.26	-7.38	$6.62E^{-12}$	0.303	20.32	10.62 (8, 171)	$4.51E^{-11}$
c+ *minus c*- (MSE rank: 9)	2.32	0.25	1.83	2.82	9.28	$7.95E^{-17}$	0.390	17.89	15.53 (8, 170)	$1.18E^{-15}$
Matrix Endorsement Bias										

Proportion of Matrix Negative Words Endorsed (MSE rank: 5)	17.13	1.52	14.14	20.12	11.3	$1.68E^{-22}$	0.472	15.36	22 (8, 172)	$4.55E^{-21}$
Proportion of Matrix Positive Words Endorsed (MSE rank: 10)	-12.08	1.41	-14.87	-9.29	-8.54	$6.78E^{-15}$	0.354	18.79	13.49 (8, 172)	$7.71E^{-14}$
Positive Matrix Endorsement Bias† (MSE rank: 2)	-14.17	1.04	-16.22	-12.12	-13.62	$3.88E^{-29}$	0.558	12.88	30.99 (8, 172)	$1.74E^{-27}$
Negative Matrix Endorsement Bias† (MSE rank: 2)	14.17	1.04	12.12	16.22	13.62	$3.88E^{-29}$	0.558	12.88	30.99 (8, 172)	$1.74E^{-27}$
Difference in Matrix Endorsement Bias (MSE rank: 2)	7.09	0.52	6.06	8.11	13.62	$3.88E^{-29}$	0.558	12.88	30.99 (8, 172)	$1.74E^{-27}$
Likert Endorsement Bias										
Proportion of Likert Negative Words Endorsed (MSE rank: 8)	6.58	0.63	5.34	7.83	10.47	5.06^{E-20}	0.43	16.38	31.26 (5, 166)	$2.12E^{-19}$
Proportion of Likert Positive Words Endorsed (MSE rank: 12)	-4.71	0.6	-5.89	-3.53	-7.87	4.22^{E-13}	0.31	19.80	18.68 (5, 166)	$1.08E^{-12}$
Positive Likert Endorsement Bias (MSE rank: 4)	-5.45	0.45	-6.33	-4.57	-12.24	$3.54E^{-25}$	0.509	14.31	25.43 (8, 172)	$1.20E^{-23}$
Negative Likert Endorsement Bias (MSE rank: 3)	5.23	0.4	4.45	6.02	13.19	$6.78E^{-28}$	0.542	13.31	29.18 (8, 172)	$2.80E^{-26}$
Difference in Likert Endorsement Bias (MSE rank: 1)	2.87	0.21	2.45	3.28	13.67	$2.88E^{-29}$	0.559	12.83	31.18 (8, 172)	$1.30E^{-27}$

Note. This table presents the results of regression analysis for SRET measures on depressive symptoms. Standard errors were derived from the regression analysis conducted using RStudio. Statistical significance was determined using $p < .05$. However, for comparing R^2 values, the exact values were reported. Measures with statistically equivalent values are marked with a dagger (†) in the table. These measures include positive endorsement bias and negative endorsement bias, positive matrix endorsement bias and negative matrix endorsement bias. Please refer to the Methods section for details on how these measures were calculated.

TABLE 2 | Iterative Addition of Predictors to Base Regression Model and Corresponding R^2 Values

Predictor	R^2	ΔR^2	MSE	ΔMSE	*p*
Negative Matrix Endorsement Bias	0.511	0.058	14.04	-1.67	$8.43E^{-26}$
Positive Matrix Endorsement Bias	0.511	0.058	14.04	-1.67	$8.43E^{-26}$
Difference in Matrix Endorsement Bias	0.511	0.058	14.04	-1.67	$8.43E^{-26}$
Difference in Likert Endorsement Bias	0.505	0.053	14.2	-1.51	$2.20E^{-25}$
Negative Likert Endorsement Bias	0.502	0.049	14.3	-1.41	$3.96E^{-25}$
Proportion of Matrix Negative Words Endorsed	0.498	0.045	14.42	-1.29	$7.60E^{-25}$
Proportion of Likert Negative Words Endorsed	0.496	0.043	14.48	-1.23	$1.06E^{-24}$
Positive Likert Endorsement Bias	0.481	0.028	14.9	-0.81	$1.18E^{-23}$
Positive Recall Bias	0.464	0.011	15.4	-0.31	$1.83E^{-22}$
Drift rates to Negative Words	0.459	0.006	15.54	-0.17	$3.74E^{-22}$
Negative Recognition Bias	0.458	0.006	15.55	-0.16	$4.07E^{-22}$
Proportion of Matrix Positive Words Endorsed	0.458	0.005	15.57	-0.14	$4.56E^{-22}$
Proportion of Likert Positive Words Endorsed	0.457	0.004	15.6	-0.11	$5.34E^{-22}$
Negative Endorsement Bias	0.456	0.003	15.62	-0.09	$5.73E^{-22}$
Positive Endorsement Bias	0.456	0.003	15.62	-0.09	$5.73E^{-22}$
Difference in Endorsement Bias	0.456	0.003	15.62	-0.09	$5.73E^{-22}$
Drift rates to Positive Words	0.456	0.003	15.62	-0.09	$5.92E^{-22}$
Response bias for recognition of endorsed negative words (*c*-)	0.455	0.002	15.64	-0.07	$6.66E^{-22}$
c+ minus *c*-	0.455	0.002	15.65	-0.06	$6.69E^{-22}$
Difference in Recall Bias	0.453	0.001	15.69	-0.02	$8.63E^{-22}$
Negative RT Bias	0.453	0.001	15.7	-0.01	$8.71E^{-22}$
Positive RT Bias	0.453	0	15.7	-0.01	$9.08E^{-22}$
Positive Recognition Bias	0.453	0	15.71	0	$9.24E^{-22}$

Response bias for recognition of endorsed positive words (*c+*)	0.453	0	15.71	0	$9.34E^{-22}$
Proportion of Negative Words Endorsed	0.453	0	15.71	0	$1.01E^{-22}$
Proportion of Positive Words Endorsed	0.453	0	15.71	0	$1.01E^{-22}$

Note. This table presents the R^2 values and MSE as predictors are iteratively added individually to the base model, sorted from highest to lowest change in R^2 and MSE This provides insight into the corresponding change in R^2 and MSE observed with the addition of each predictor.

TABLE 3 | Final Backward Stepwise Regression and Forward Stepwise Regression Models

Model	Final Variables	AIC	BIC	R^2	MSE
(1) Forward Stepwise Regression (AIC)	**Negative Matrix Endorsement Bias**, Negative Likert Endorsement Ratio, Negative Recall Bias	939.81	955.52	0.53	13.46
(2) Backward Stepwise Regression with (AIC)	Negative Matrix Endorsement Ratio, Positive Matrix Endorsement Ratio, Positive Likert Endorsement Bias, Positive Likert Endorsement Ratio, **Negative Matrix Endorsement Bias**	939.64	961.63	0.52	12.79
(3) Forward Stepwise Regression with (BIC)	**Negative Matrix Endorsement Bias**, Negative Likert Endorsement Ratio	940.53	953.1	0.54	13.67
(4) Backward Stepwise Regression with (BIC)	Negative Matrix Endorsement Ratio, **Negative Matrix Endorsement Bias**	941.97	954.53	0.52	13.79

TABLE 4 | Effect of word selection on Endorsement Bias measures in Dataset A

Variable	*B*	*SE*	95% CI LL	95% CI UL	*t*	*p*	R^2	*MSE*	*F (df)*	*p*
Negative Endorsement Bias										
23 words list	14.68	1.36	12	17.36	10.81	$3.41E^{-21}$	0.444	16.61	19.96 (8, 175)	$1.54E^{-19}$
LeMoult word list (40 words)	15.11	1.35	12.44	17.78	11.18	$3.11E^{-22}$	0.459	16.16	21.196 (8, 175)	$1.55E^{-20}$
60 words list	14.41	1.22	11.99	16.83	11.77	$6.43E^{-24}$	0.482	15.47	23.27 (8, 175)	$3.73E^{-22}$
200 words list	14.17	1.04	12.12	16.22	13.62	$3.88E^{-29}$	0.558	12.88	30.99 (8, 172)	$1.74E^{-27}$
Positive Endorsement Bias										
23 words list	-0.48	0.06	-0.6	-0.36	-8.02	$1.70E^{-13}$	0.324	19.17	11.56 (8, 169)	$5.81E^{-12}$
LeMoult word list	-15.11	1.35	-17.78	-12.44	-11.18	$3.11E^{-22}$	0.459	16.16	21.196 (8, 175)	$1.55E^{-20}$
60 words list	-14.41	1.22	-16.83	-11.99	-11.77	$6.43E^{-24}$	0.482	15.47	23.27 (8, 175)	$3.73E^{-22}$
200 words list	-14.17	1.04	-16.22	-12.12	-13.62	$3.88E^{-29}$	0.558	12.88	30.99 (8, 172)	$1.74E^{-27}$
Difference in Endorsement Bias										
23 words list	0.47	0.06	0.36	0.58	8.13	$8.80E^{-14}$	0.329	19.03	11.83 (8, 169)	$3.16E^{-12}$
LeMoult word list	7.56	0.68	6.22	8.89	11.18	$3.11E^{-22}$	0.459	16.16	21.196 (8, 175)	$1.55E^{-20}$
60 words list	7.21	0.61	6	8.41	11.77	$6.43E^{-24}$	0.482	15.47	23.27 (8, 175)	$3.73E^{-22}$
200 words list	7.09	0.52	6.06	8.11	13.62	$3.88E^{-29}$	0.558	12.88	30.99 (8, 172)	$1.74E^{-27}$

Note. This table presents the R^2 values compared across the different word lists.