Biases in processing of mood-congruent facial expressions in depression

Thomas Van Vleet[a], Alit Stark-Inbar[a,b], Michael M. Merzenich[a], Joshua T. Jordan[c], Deanna L. Wallace[d], Morgan B. Lee[d], Heather E. Dawes[d], Edward F. Chang[d,e,f,1], Mor Nahum[a,g,1,*]

[a] Department of R&D, Posit Science Corporation, San Francisco, CA, USA
[b] Department of Psychology, University of California, Berkeley, Berkeley, CA, USA
[c] Department of Psychiatry, University of California, San Francisco, San Francisco, CA, USA
[d] Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA, USA
[e] Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, USA
[f] Kavli Institute for Fundamental Neuroscience, University of California, San Francisco, San Francisco, CA, USA
[g] Faculty of Medicine, Hebrew University, Jerusalem, Israel

ABSTRACT

Cognitive models of depression suggest that depressed individuals exhibit a tendency to attribute negative meaning to neutral stimuli, and enhanced processing of mood-congruent stimuli. However, evidence thus far has been inconsistent. In this study, we sought to identify both differential interpretation of neutral information as well as emotion processing biases associated with depression. Fifty adult participants completed standardized mood-related questionnaires, a novel immediate mood scale questionnaire (IMS-12), and a novel task, Emotion Matcher, in which they were required to indicate whether pairs of emotional faces show the same expression or not. We found that overall success rate and reaction time on the Emotion Matcher task did not differ as a function of severity of depression. However, more depressed participants had significantly worse performance when presented with sad-neutral face pairs, as well as increased reaction times to happy-happy pairs. In addition, accuracy of the sad-neutral pairs was found to be significantly associated with depression severity in a regression model. Our study provides partial support for the mood-congruent hypothesis, revealing only a potential bias in interpretation of sad and neutral expressions, but not a general deficit in processing of facial expressions. The potential of such bias in serving as a predictor for depression should be further examined in future studies.

1. Introduction

Mood disorders, such as major depressive disorder (MDD), afflict a significant portion of the population and pose a major burden in total disability-adjusted years among midlife adults (Kessler et al., 2005). These disorders are associated with significant impairments in social, occupational, and educational functioning.

Cognitive models of depression (Beck, 1976) as well as Bower's mood congruency hypothesis (Bower, 1981) suggest that there are negative interpretations or response biases in depression, in addition to enhanced processing of mood-congruent (i.e. negative) stimuli. Specifically, they argue that the depressed individual is biased to attribute negative emotions to neutral stimuli over positive emotions. Furthermore, preferential processing of mood-congruent information in the environment may enhance recognition of facial expressions of negative emotions over expressions of positive emotions, relative to individuals without depression.

Studies examining these hypotheses have thus far provided support for the negative-interpretation bias but yielded mixed results regarding the processing of emotional stimuli (see recent reviews in Bourke et al., 2010; Foland-Ross and Gotlib, 2012). Specifically, several studies have reported that acutely depressed patients (Gollan et al., 2008; Leppanen et al., 2004), patients in remission, or those at high risk for depression onset (Maniglio et al., 2014) misclassified more neutral expressions as sad compared to non-depressed individuals, in line with the negative interpretation bias hypothesis. In contrast, findings regarding selective impairments in processing of emotional stimuli are inconsistent across studies. While some studies found differences in the processing of sad or happy expressions relative to other emotions (e.g. longer reaction times, RTs), other studies found no such differences (Leppanen et al.,

2004; Surguladze et al., 2004). For example, Milders et al. (2010) found no differences on an emotion matching task ('does the emotion of the 3rd face match that of the 1st or the 2nd?') but did find that depressed patients were more accurate and had longer reaction times for the sad faces on an emotion labeling task. Zwick and Wolkenstein (2017) found that acutely depressed patients recognized happy faces less accurately than did matched healthy controls. Two recent meta-analyses found evidence for impairments in emotion recognition in depression in all emotions except sadness (Dalili et al., 2015; Demenescu et al., 2010), but with small effect sizes.

The different paradigms used across various studies could contribute to the inconsistency in the pattern of results. For example, it has been suggested that difficulties are found when using a labeling but not a matching task (Milders et al., 2010). Moreover, stimulus duration may be important, as effects have often been reported for longer stimulus presentation durations (e.g. Gollan et al., 2008; Gotlib et al., 2004).

In the current study, we sought to identify both differential interpretation and emotion processing biases using a single novel paradigm, *Emotion Matcher*, which requires that participants decide whether pairs of faces of different individuals express the same or different emotions. Based on the negative interpretation bias, we hypothesized that accurate comparisons of sad and neutral expressions would become increasingly difficult in accordance with an increased severity of depression, yielding lower success rates relative to pairs of other emotions. Based on theories that address the processing of emotional stimuli, we hypothesized that more depressed participants would have longer RTs for pairs of faces involving sad expressions. We delivered the paradigm remotely, on a mobile device (tablet), to support a more accurate, real-life measurement of mood-related cognitive effects relative to other studies.

2. Methods

2.1. Subjects

A convenience sample of fifty ($n = 50$) participants completed the online battery of assessments (see details below). Participants were recruited from two sites: The Epilepsy Monitoring Unit (EMU) of the University of California in San Francisco (UCSF) Hospital and UC Berkeley. UCSF EMU participants were recruited as part of broader efforts to examine daily mood fluctuations while participants were hospitalized for seizure monitoring with electroencephalography (EEG) or electrocorticography (ECoG) (see Sani et al., 2018). These participants were enrolled in the study serially, as admitted to the EMU. UC Berkeley participants were recruited via the Research Participant Pool of the Psychology department and received course credit. The study was run under institutional review board approvals from UCSF and UC Berkeley. All participants gave written informed consent before engaging with the tasks and all study-related tasks were completed using a study-provided iPad-mini (Apple, Inc.). Participants did not receive monetary compensation for their participation.

2.2. Procedures

Following informed consent, participants were asked to login using a unique password-protected credential, and data was saved on a password-protected HIPAA-compliant server that was only accessible to study investigators through a web browser. Upon login, participants completed the following tasks:

(1) Patient Health Questionnaire, 9-item (PHQ-9; Kroenke et al., 2001): A standardized 9-item self-report questionnaire used to assess DSM-V-TR [17] symptoms of depression experienced in adults in the two weeks preceding administration. The total score of the PHQ-9 (range of 0–27) allows for classification of participants into depression severity categories: minimal depression (0–4), mild

depression (5–9), moderate depression (10–14), moderately severe depression (15–19) or severe depression (20 and over). The PHQ-9 is the most commonly administered self-report tool for depression, has good diagnostic and psychometric properties, and has been shown to be valid across numerous modes of administration (Fann et al., 2009).

(2) Generalized Anxiety Disorder, 7-item (GAD-7; Spitzer et al., 2006): A standardized, validated 7-item self-report questionnaire used to assess symptoms of anxiety experienced in the two weeks preceding administration. GAD-7 total scores range from 0–21, with the following cut-offs: minimal (0–4), mild (5–9), moderate (10–14), and severe anxiety (≥ 15).

(3) Immediate Mood Scale (IMS-12; Nahum et al., 2017): A novel 12-item measure developed to assess dynamic components of mood. Participants are asked to rate their current mood state on a continuum using 7-point Likert scales (e.g. happy-sad, distracted-focused, sleep-alert, fearful-fearless). For each item, an integer score between 1 to 7 was derived. The total score for this scale is the sum of the scores on all 12 items, where higher scores reflect more negative mood states. The scale has been recently used to help identify an amygdala-hippocampus sub-network that encodes variations in human mood (Kirkby et al., 2018).

(4) Emotion Matcher: This paradigm includes 48 trials (Fig. 1). Each trial was initiated with a central fixation appearing on the iPad screen for 500 ms. Following the fixation screen, two faces, each showing either neutral, happy, or sad expression appeared for 750 ms. Participants were required to hit the 'yes' button if the two faces showed the same facial expression (e.g. both showing sad expressions), or 'no' if the two faces showed different expressions. The yes/no buttons remained on the screen until the participant made his/her selection, but were disabled while the faces were presented. Auditory (sound) and visual (highlight) feedback then appeared on the screen for 500 ms, indicating either correct/incorrect response, followed by an inter-trial-interval (ITI) of 1750 ms until the next trial. On every trial, a pair of faces was pseudo-randomly selected from the following 6 potential pair options: happy-happy, happy-neutral, happy-sad, neutral-neutral, neutral-sad, or sad-sad. Each category pair appeared 8 times during the block of 48 trials. The order of presentation was randomized during the run, such that each participant received a different order of pairs. The stimulus set contained 43 female and 43 male faces from a database generated by Posit Science, each with 3 images, one for each emotional expression. Therefore, different faces were randomly selected for the pseudo-random pairs. Images were 250 × 280 pixels each. The specific identity of each face in a pair was not included in the analysis.

2.3. Data analysis

To analyze Emotion Matcher RT data, we first removed outliers with RT values of 3 absolute deviations from the median for at least one pair of the 6 pairs (Leys et al., 2013). We ended up with $N = 45$ participants (90%) with valid RT measures. Participants without valid RT measures did not differ significantly from participants with valid RT measures in terms of age, PHQ-9, or GAD-7 scores (all $p \geq .23$). Furthermore, participants without valid RT measures did not differ significantly from participants with valid RT measures in accuracy. Thus, invalid RTs were considered to be "missing at random". Analyzed data included accuracy and RT for each of the 6 emotion pairs for each participant. To test our hypotheses that a) increased depression severity is associated with reduced accuracy of sad and neutral expressions, and b) that increased depression severity is associated with slower RTs for sad expressions, we used a multivariate regression model with full information maximum likelihood (FIML). FIML is a gold standard approach to handle missing data (Enders, 2010). This allowed us to test the associations between PHQ-9/GAD-7 scores and each of the outcome measures while
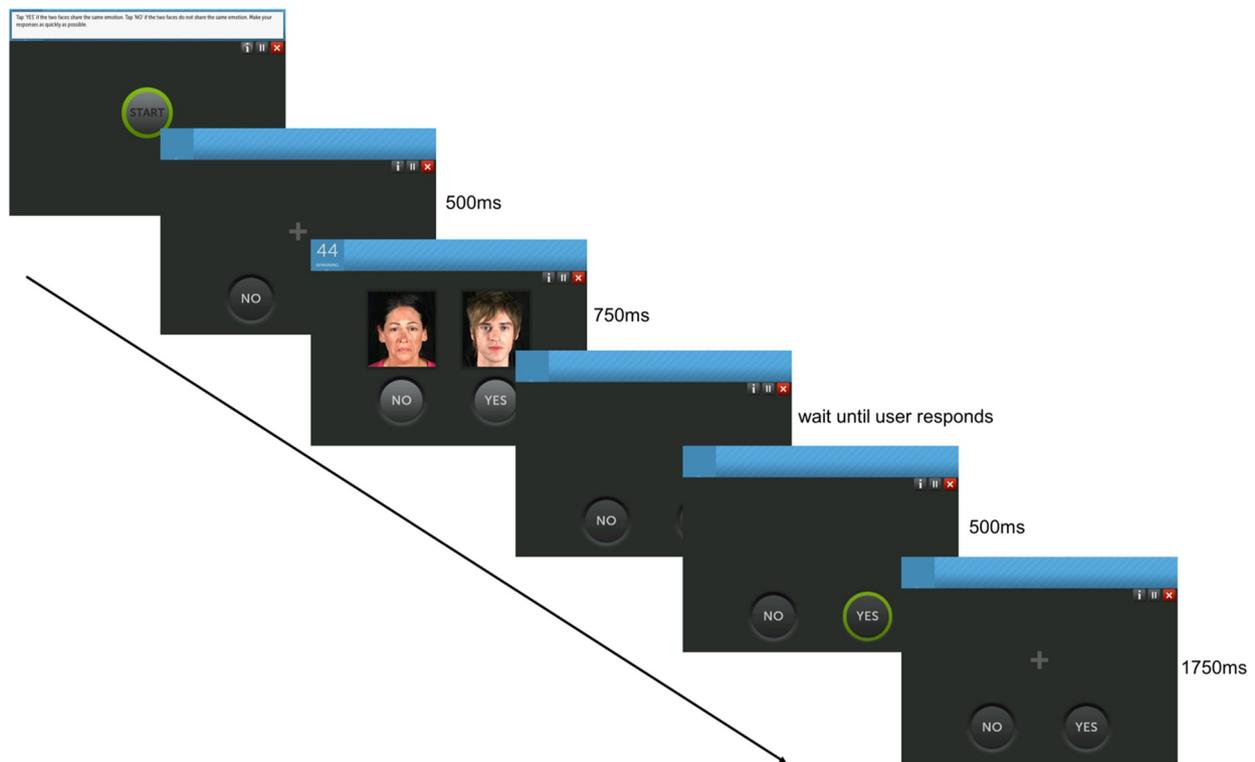
Fig. 1. The Emotion Matcher iPad Paradigm. The 'start' button is presented only at the beginning of the 48-trial block. Then, every trial starts with a 500 ms fixation cross, followed by stimulus presentation for 750 ms. The user should then select 'yes' or 'no' to indicate if the two faces show the same expression or not. Visual feedback is shown for 500 ms, accompanied by a short auditory sound, followed by an inter-trial-interval window of 1750 ms, until the next trial starts.

a) taking into consideration residual covariance between each of the outcome variables, and b) retaining all participants in analysis. All analyses were conducted in Stata v15.1 (StataCorp, 2017). Two-sided *p*-values < .05 were considered significant.

3. Results

3.1. Characterization of study sample

The age range of the 50 participants was 21–59 years (average: 35.1 ± 12.7 years). 27 of the participants (54%) were female. As per the PHQ-9 categories, 10 participants (20%) were classified as having minimal depression (PHQ-9 scores of 0–4), 11 (22%) were classified as having mild depression (PHQ-9 scores of 5–9), 16 (32%) as having moderate depression (PHQ- scores of 10–14), and 13 (26%) as having moderately severe to severe depression (PHQ-9 scores of 15 and over). Demographic information is presented in Table 1.

Age was significantly associated with PHQ-9 score as well as with Emotion Matcher overall RT, such that those who were older tended to have higher depression scores and had slower RT ($r = 0.38$, $p = .008$; $r = 0.29$, $p = .044$, respectively). Because of this, all regression and multivariate models were adjusted for age. In addition, depression (PHQ-9) and anxiety (GAD-7) scores were highly correlated ($r = 0.76$,

$p < .0001$), pointing to their high co-morbidity in this sample population.

3.2. Performance on Emotion Matcher task

Table 2 summarizes the results of the multivariate regression analysis performed for the Emotion Matcher task data.

3.2.1. Accuracy

Overall performance (accuracy: 0.78 ± 0.2) on the task was not associated with depression severity ($r = -0.19$, $p = .18$). In the multivariate regression model, we found that only the sad-neutral (S-N) pair accuracy was negatively associated with depression severity, such that as depression severity increased, accuracy on this pair decreased ($\beta = -0.29$, $p = .031$). Accuracy of all other pairs was not significantly associated with level of depression (Fig. 2, two top rows).

3.2.2. Reaction time (RT)

Average overall RT (across all pairs: 1550 ± 46 ms) was not associated with depression severity ($r = 0.23$, $p = .12$). In the multivariate regression model, only the happy-happy (H–H) pair was associated with depression severity, such that as severity of depression increased RT for this pair increased as well ($\beta = 0.29$, $p = .025$). RT for all other pairs was not associated with depression severity (Fig. 2, two bottom rows).

Similar results for accuracy were obtained using the GAD-7 anxiety scale, as GAD-7 total score was associated with sad-neutral accuracy ($\beta = -0.30$, $p = .015$) only. GAD-7 was not associated with any of the RT pairs (all $p \geq .278$).

3.3. A regression model predicting PHQ-9 scores

To examine the potential contribution of performance on the Emotion Matcher task to the diagnosis of depression, we conducted an

Table 1
Demographic Information of Study Participants.

Depression level (PHQ-9 Category)	PHQ-9 Score Range	N	Gender (% F)	Age (y)
Minimal	0–4	10	40%	32.3 ± 11.9
Mild	5–9	11	18%	27 ± 9.8
Moderate	10–14	16	56%	38 ± 13.8
Mod. Severe to Severe	15–27	13	84%	40.5 ± 11
TOTAL		50	54%	35.1 ± 12.7

Table 2

Results of the multivariate regression analysis for Emotion Matcher as a function of PHQ-9 total score[&]. Analyses for reaction time use Full Information Maximum Likelihood. Abbreviations: S–S (sad-sad), S−N (sad-neutral), N−N (neutral-neutral), H−H (happy-happy), H–S (happy-sad), H−N (happy-neutral).

Pair	Accuracy	β	$|z|$	p-value	95% CI	Reaction Time (RT)	β	$|z|$	p-value	95% CI
S–S	PHQ-9	.03	.23	.816	−.26, .33	PHQ-9	.14	.98	.33	−.14, .42
S–N	**PHQ-9**	**−.29**	**2.15**	**.031**	**−.55, −.03**	PHQ-9	.09	.60	.55	−.21, .39
N–N	PHQ-9	.08	.50	.62	−.22, .37	PHQ-9	.17	1.19	.23	−.11, .45
H–H	PHQ-9	.05	.32	.75	−.25, .35	**PHQ-9**	**.29**	**2.24**	**.025**	**.04, .55**
H–S	PHQ-9	.07	.49	.62	−.22, .36	PHQ-9	.14	.95	.34	−.15, .42
H–N	PHQ-9	.01	.10	.92	−.26, .29	PHQ-9	.24	1.69	.09	−.04, .52

[&] All analyses are adjusted for age.

ordinary least squares hierarchical multiple regression predicting PHQ-9 scores using GAD-7, IMS-12 and the sad-neutral (S-N) accuracy of Emotion Matcher. The IMS-12 scale was included as it was shown to significantly contribute to depression prediction in a previous study by our group (Nahum et al., 2017). The multiple regression showed that all three variables are independent predictors of PHQ-9 scores, with the accuracy of the S-N pair of the Emotion Matcher task adding about 3% of variance accounted above and beyond anxiety (GAD-7 total score)
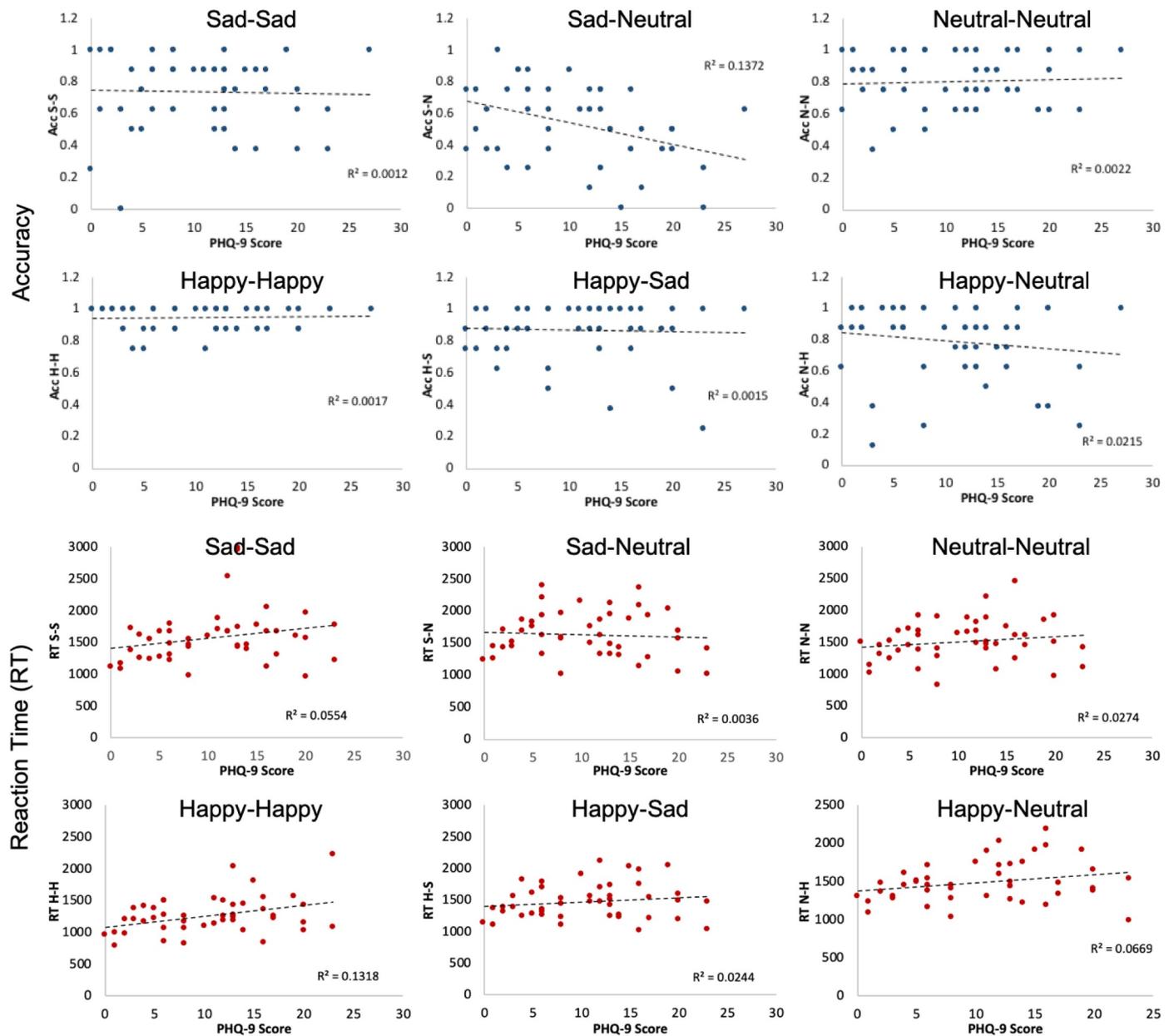

Fig. 2. Results of the Emotion Matcher paradigm as a function of PHQ-9 score. Accuracy (two top rows) and reaction time (two bottom rows) data on the Emotion Matcher paradigm. Each graph shows data for a different stimulus pair in the task. Higher PHQ-9 scores indicate higher level of depression. Abbreviations: S-S (sad-sad), S-N (sad-neutral), N−N (neutral-neutral), H−H (happy-happy), H-S (happy-sad), N−H (neutral-happy).

Table 3
Hierarchical linear regression models predicting PHQ-9 depression level.

Variable	B	t	p
GAD-7	.544	6.02	<.001
IMS-12	.375	4.37	<.001
ACC_SN	− 0.171	− 2.20	.034

Abbreviation: ACC_SN: Accuracy for the sad-neutral pair of the Emotion Matcher task.

and the IMS-12 total score ($\beta = -0.17$, $p = .034$). Variables remained significant even after including age as a covariate. By itself, accuracy of the S-N pair has a correlation (Spearman) of -0.38 ($p = .0064$). Full results of the hierarchical linear regression are given in Table 3.

4. Discussion

Our study examined performance on a novel speeded emotion discrimination task, *Emotion Matcher*, as a function of level of depression. The results indicate that discrimination accuracy between sad and neutral expressions decreased as the severity of depression increased. Sad-neutral (S-N) accuracy was also a significant predictor of the magnitude of depression in a regression model. Finally, while overall reaction time was not significantly associated with depression, reaction times for the happy-happy (H−H) pair increased as a function of depression severity.

4.1. Negative interpretation bias

Our results manifest negative emotion interpretation bias, with accuracy for correctly identifying sad-neutral pairs decreasing as a function of increased depression severity. Importantly, errors in interpreting neutral expressions in our study were only recorded in the context of sad expressions (sad-neutral pairs). There was no evidence of misclassification of neutral-neutral or neutral-happy pairs. On the other hand, longer RTs for happy-happy pairs for those with higher levels of depression may indicate a more general mood-congruent bias.

These results are consistent with those of several earlier studies that reported an increased tendency in depression to interpret neutral expressions as more negative (Gollan et al., 2008; Leppanen et al., 2004; Maniglio et al., 2014). A recent review by Bourke and colleagues indicates that there is "reasonably consistent evidence of misclassification of emotional expressions to more negative emotions in major depression" (Bourke et al., 2010). Indeed, Mangilio et al. (2014) found that individuals with affective temperaments who are more depressed interpreted neutral facial expressions more negatively, mostly misclassifying neutral facial expressions as sad.

Gollan et al. (2008) also found that participants with MDD tended to interpret neutral facial expressions as sad significantly more often than healthy participants. Interestingly, though, they also found that healthy controls interpreted neutral faces as happy significantly more often than depressed participants. Our study did not show this effect, as there was no association between depression severity and performance on the happy-neutral expressions.

Finally, Leppanen et al. (2004) found that depressed patients misclassified more neutral expressions as sad as well, but that they were also slower to respond to neutral expressions compared to sad or happy ones. We did not record any RT difference for neutral expressions in the current study, perhaps due to the different paradigm used here in which matching rather than identification was required.

4.2. Processing of emotional stimuli

Our results did not support a specific deficit in the processing of sad expressions with increased depression as there was no RT difference for sad-sad pairs, and only partially supported deficit in processing happy expressions (longer RTs for happy-happy emotion pairs). Moreover, they do not support an overall motor sluggishness or impaired motor function with increased levels of depression, as RT results were limited to one pair out of the six used in the study.

Results in the literature regarding potential deficits in processing of sad expressions have been thus far mixed (see reviews in Demenescu et al., 2010; Foland-Ross and Gotlib, 2012), with several studies reporting no differences in the processing of emotional stimuli in MDD (e.g. Leppanen et al., 2004; Maniglio et al., 2014; Milders et al., 2010; Surguladze et al., 2004), and others reporting differences in the processing of either sad or happy expressions or both (e.g. Gollan et al., 2008; Zwick and Wolkenstein, 2017). There is also no consensus in the literature on whether processing is enhanced or delayed for sad expression, or whether mood negatively affects the processing of happy expressions (Bourke et al., 2010). Two recent meta-analyses concluded that there are impairments in all emotions except sadness in MDD, with small effect sizes (Dalili et al., 2015; Demenescu et al., 2010).

We found only partial support for this deficit, in longer processing of happy-happy pairs with elevated depression levels. However, there was no evidence for longer processing times for any of the other pairs involving happy expression (happy-neutral and happy-sad) or reduced accuracy rates for any of these pairs. Several potential explanations may account for the discrepancies between the current study and the results of these meta-analyses. First, the use of a matching rather than a labeling paradigm could potentially tap into different cognitive operations. Indeed, Milders and colleagues (2010) used both a labelling and a matching paradigm, but only found differences on the labelling task (increased accuracy and longer RTs for sad expressions), suggesting that there are no abnormalities in discrimination and that the processing bias is not based on automatic processing or biased attention. Second, most studies that have found such deficits employed relatively long stimulus presentation durations (Gollan et al., 2008; Gotlib et al., 2004) and failed to replicate these effects with shorter presentation durations. In the current study, we employed a *medium-length* presentation duration of 750 ms, which may not be sufficiently long enough to identify such deficits if they exist. Third, some studies suggest that depression is associated with difficulties identifying subtle positive emotions (Joormann and Gotlib, 2006), rather than identifying the full range of emotional expression. We did not find evidence of this phenomenon. Finally, it is unclear whether these deficits are seen in a sub-clinical population or only in acutely depressed patients (Milders et al., 2010; Surguladze et al., 2004), remitted patients (LeMoult et al., 2009) or those at high risk for depression (Joormann et al., 2010).

4.3. Limitations of current study

Our study has several limitations that should be addressed in future studies. First, our sample was a convenience sample and did not include acutely depressed individuals, but rather individuals that were available to participate and self-reported their depression level. In addition, our sample size was small, and included only a small subset of severely depressed individuals. A follow-up study should carefully compare clinically depressed individuals versus healthy controls to further elaborate the range of performance on the Emotion Matcher task predictive of clinical depression. Moreover, the Emotion Matcher paradigm does not allow us to definitively conclude that the bias observed in individuals with more severe depression is due to the misinterpretation of neutral expressions as sad. A variation of the paradigm in which presentation time is titrated or a comparison of performance to that recorded by use of an explicit emotion labeling task may provide better clarity in future studies.

4.4. Clinical implications

The delivery of the Emotion Matcher paradigm, as well as the self-

report questionnaires remotely, on a tablet device, may provide a cost-effective, ecological method to assess depression severity and depression-related cognitive biases. The availability of mobile devices and their increasing use by psychiatric populations (Torous et al., 2014) may provide a novel and efficient means to assess symptoms in addition to existing methods, and even track changes in them over time, in ecological settings. Several recent studies have examined the potential use of mobile tools for assessment of depression severity and treatment of depression (e.g. Nahum et al., 2017; Torous et al., 2015), where assessment tools often include self-report mood questionnaires and scales. Here, we show the potential of employing a short cognitive paradigm for identifying implicit negative attribution biases in depression. Future research is needed in order to replicate and extend these results in order to determine the validity of this and additional tools for remote, ecological assessment of mood disorders severity.

Acknowledgments

Conflict of Interests Statement

Authors Thomas Van Vleet and Michael M. Merzenich are employees of Posit Science, Inc., which supported the development of the paradigm described in the manuscript. Authors Mor Nahum and Alit Stark-Inbar were employed at Posit Science when the study was conducted, but no longer have financial relations with Posit Science.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2019.02.076.

References

Beck, A.T., 1976. Cognitive Therapy and the Emotional Disorders. International University Press, New York.

Bourke, C., Douglas, K., Porter, R., 2010. Processing of facial emotion expression in major depression: a review. Aust. N. Z. J. Psychiatry 44 (8), 681–696.

Bower, G.H., 1981. Mood and memory. Am. Psychol. 36 (2), 129–148.

Dalili, M.N., Penton-Voak, I.S., Harmer, C.J., Munafo, M.R., 2015. Meta-analysis of emotion recognition deficits in major depressive disorder. Psychol. Med. 45 (6), 1135–1144.

Demenescu, L.R., Kortekaas, R., den Boer, J.A., Aleman, A., 2010. Impaired attribution of emotion to facial expressions in anxiety and major depression. PLoS One 5 (12), e15058.

Enders, C.K., 2010. Applied Missing Data Analysis. The Guilford Press, New York, NY.

Fann, J.R., Berry, D.L., Wolpin, S., Austin-Seymour, M., Bush, N., Halpenny, B., Lober, W.B., McCorkle, R., 2009. Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer. Psychooncology 18 (1), 14–22.

Foland-Ross, L.C., Gotlib, I.H., 2012. Cognitive and neural aspects of information processing in major depressive disorder: an integrative perspective. Front. Psychol. 3, 489.

Gollan, J.K., Pane, H.T., McCloskey, M.S., Coccaro, E.F., 2008. Identifying differences in biased affective information processing in major depression. Psychiatry Res. 159 (1–2), 18–24.

Gotlib, I.H., Krasnoperova, E., Yue, D.N., Joormann, J., 2004. Attentional biases for negative interpersonal stimuli in clinical depression. J. Abnorm. Psychol. 113 (1), 121–135.

Joormann, J., Gotlib, I.H., 2006. Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. J. Abnorm. Psychol. 115 (4), 705–714.

Joormann, J., Nee, D.E., Berman, M.G., Jonides, J., Gotlib, I.H., 2010. Interference resolution in major depression. Cogn. Affect. Behav. Neurosci. 10 (1), 21–33.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., Walters, E.E., 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch. Gen. Psychiatry 62 (6), 593–602.

Kirkby, L.A., Luongo, F.J., Lee, M.B., Nahum, M., Van Vleet, T.M., Rao, V.R., Dawes, H.E., Chang, E.F., Sohal, V.S., 2018. An amygdala subnetwork that encodes variation in human mood. Cell 175 (6), 1688–1700 e1614.

Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16 (9), 606–613.

LeMoult, J., Joormann, J., Sherdell, L., Wright, Y., Gotlib, I.H., 2009. Identification of emotional facial expressions following recovery from depression. J. Abnorm. Psychol. 118 (4), 828–833.

Leppanen, J.M., Milders, M., Bell, J.S., Terriere, E., Hietanen, J.K., 2004. Depression biases the recognition of emotionally neutral faces. Psychiatry Res. 128 (2), 123–133.

Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L., 2013. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J. Exp. Soc. Psychol. 49 (4), 764–766.

Maniglio, R., Gusciglio, F., Lofrese, V., Belvederi Murri, M., Tamburello, A., Innamorati, M., 2014. Biased processing of neutral facial expressions is associated with depressive symptoms and suicide ideation in individuals at risk for major depression due to affective temperaments. Compr. Psychiatry 55 (3), 518–525.

Milders, M., Bell, S., Platt, J., Serrano, R., Runcie, O., 2010. Stable expression recognition abnormalities in unipolar depression. Psychiatry Res. 179 (1), 38–42.

Nahum, M., Van Vleet, T.M., Sohal, V.S., Mirzabekov, J.J., Rao, V.R., Wallace, D.L., Lee, M.B., Dawes, H., Stark-Inbar, A., Jordan, J.T., Biagianti, B., Merzenich, M., Chang, E.F., 2017. Immediate Mood Scaler: tracking symptoms of depression and anxiety using a Novel Mobile Mood Scale. JMIR Mhealth Uhealth 5 (4), e44.

Sani, O.G., Yang, Y., Lee, M.B., Dawes, H.E., Chang, E.F., Shanechi, M.M., 2018. Mood variations decoded from multi-site intracranial human brain activity. Nat. Biotechnol. 36 (10), 954–961. https://doi.org/10.1038/nbt.4200.

Spitzer, R.L., Kroenke, K., Williams, J.B., Lowe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097.

StataCorp, 2017. StataCorp LLC College Station, TX.

Surguladze, S.A., Young, A.W., Senior, C., Brebion, G., Travis, M.J., Phillips, M.L., 2004. Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. Neuropsychology 18 (2), 212–218.

Torous, J., Friedman, R., Keshavan, M., 2014. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2 (1), e2.

Torous, J., Staples, P., Onnela, J.P., 2015. Realizing the potential of mobile mental health: new methods for new data in psychiatry. Curr. Psychiatry Rep. 17 (8), 602.

Zwick, J.C., Wolkenstein, L., 2017. Facial emotion recognition, theory of mind and the role of facial mimicry in depression. J. Affect. Disord. 210, 90–99.