Objective and subjective emotional face classification in non-clinical depression

Ben J. Jennings, Derrick A. Boateng, Tamera Choudhury, Noof Alotaibi & Survjit Cheeta

Centre for Clinical and Cognitive Neuroscience

College of Health, Medicine and Life Science

Brunel University London

Uxbridge, U.K.

Abstract

Background: Many previous studies highlighting a relationship between depression and emotional face recognition have relied on measures of classification accuracy to determine recognition deficits. However, the perception of emotions is also related arousal levels and valence, and more research is needed to determine how depression impacts these dimensions.

Aims: To compare performance on both an objective forced choice emotional recognition task and subjective emotional face valence rating task in participants with self-reported high depression.

Methods: Based on screening using the depression sub-scale of the DASS-42, 46 participants (23 males, 23 female) were in the high depression group (mean DASS-42 34±5) and 50 participants in the control groups (25 males, 25 females) with DASS-42 scores of either 0 or 1. All participants completed both a performance-based task (objective) as well as a rating task (subjective) of emotional facial expressions.

Results: The data indicate that difference in performance exist in classification accuracy between the groups, with depressed participants demonstrating reduced accuracy for anger, sadness and neutral facial expressions. Additionally differences in subjective ratings exist in the depressed group, but with the important caveat that these only relate to faces display positive emotional expressions.

Discussion: The limitations of relying solely on objective tasks where recognition accuracy is the main outcome measure are discussed as well as the data quantitatively demonstrating a reduced response in the depression group to positive stimuli. This study justifies the need for future studies using both objective and subjective measures to assess emotion classification deficits in depression.

Introduction

Depression is one of the most common psychiatric disorders and a leading cause of global disability (WHO, 2008). Much of the current research on increasing our understanding of depression and thus identifying more effective treatments has been led by the findings that depression is accompanied by two complementary information processing biases: an increased focus on negative events/information and a decreased focus on positive events/information (Willner et al., 2013). This cognitive vulnerability has been studied extensively using facial emotion recognition tasks, where a mixed picture has emerge, with some studies showing performance deficits (Csukly *et al*., 2009; Langenecker *et al*., 2005) and others not (Joormann & Gotlib, 2006; Kan et al., 2004), so several reviews and meta-analyses have aimed to reconcile these findings.

In order to reconcile these finding several reviews and meta-analyses have been published. In their review of 40 studies (Bourke *et al*., 2010), depression was found to be associated with a negative bias towards sad faces, ambiguous or neutral facial expressions were evaluated as more sad and less happy, and there was increased vigilance and attention towards sad facial expressions and away from happy expressions. There was also some evidence of reduced accuracy of happy and sad facial expressions, but findings were complicated by patient heterogeneity and different methodologies. A later meta-analysis of 22 studies found depression to be associated with reduced recognition accuracy of anger, disgust, fear, happiness and surprised with a small effect size, and no effect was found for sadness (Dalili *et al*., 2015). Most recently, a meta-analysis of 23 studies including 516 dysthymic/depressed participants and 614 euthymic controls participants assessed the role of range of variables including patient status (inpatient vs outpatient), type of stimulus (static vs, dynamic; morphed vs. unmorphed), stimulus duration, type of emotion, symptom severity, sex, method of diagnosis on emotional face recognition deficits. The main finding was that severity of diagnosis was a key moderating risk factor for emotional face recognition deficits (Krause *et al*., 2021).

The current evidence base for an association between depression and impaired emotional facial recognition is largely based on the published studies using the same outcome measures which is accuracy of recognition or classification, i.e. there is generally a correct response for

each presented trial of an emotional face. However, despite it being well-established that the perception of emotional stimuli are directly related to other dimensions, especially arousal (degree of excitement or motivation) and valence (pleasure) (Russell, 1994) the impact of low mood/depression on these dimensions has received less investigation.

Albeit limited studies are available which have employed subjective tasks to assess the perception of emotions, Deckert *et al*., (2019) used a Likert scale to assess arousal levels for emotional faces and words, and the variables that predicted arousal were emotion specific. For example, age and emotional intensity predicted arousal to negative faces, the sex of the participants and emotional intensity predicted arousal to positive faces and sex and poor impulse control predicted arousal to neutral faces. Furthermore, cluster analysis identified that the impact of mood on arousal levels to emotional faces resulted in three different groups. The first group was mostly females with high emotion regulation difficulties, scoring high on depression and anxiety (emotional difficulties disposition group) and higher arousal levels were seen in response to emotional faces, A second group consistently mostly of males with low emotion regulation difficulties, i.e. scoring low on depression and anxiety (lowest emotion regulation difficulties) reported low arousal to emotional faces. Finally, the third group also showed low arousal to emotional faces who this time were mostly females and scored the lowest for depression and anxiety.

Given the above limitations of previous research, the aim of the present study was to understand how mood may impact subjective experience of emotions. Therefore, the current study employed two tasks: an objective forced-choice task and a subjective valance rating task. The expectation is that as depressive symptoms predict poorer performance on recognition tasks, accuracies will be generally lower on the forced-choice task. Additionally, due to the negative outlook and bias that is associated with depression, in the subjective (valence) rating task participants scoring high on depression will rate the faces more negatively. Furthermore, as many previous studies have typically used the Ekman's six basic emotions (Ekman, 1992) or a subset of them of them, the aim of the present study was to assess a wider range of emotional stimuli.

Methods

Participants

A total of 481 participants were screened for depression as defined by the depression sub-scale of the DASS-42 (Lovibond and Lovibond, 1995). Participants were added to the depression group if they were identified as being in the "Extremely Severe" depression range (i.e., scoring 28 or above). Forty-six participants (23 males, 23 females) were subsequently identified and selected based on this criterion, they had an average age of 28±11 years (mean±SD) in the range 18 to 64 years. Their mean DASS-42 defined depression score was 34±5 in the range 28 to 42. A control group of 50 participants (25 males, 25 females) was identified, this group had an average age of 26±13 years in the range 18 to 71 years. Their DASS-42 scores were either 0 or 1. Ethical approval was obtained for the study from the College of Health, Medicine and Life Sciences Research Ethics Committee, Brunel University of London.

Materials

Face stimuli

In the forced-choice emotional face recognition task, stimuli were taken from the Karolinska Directed Emotional Faces (KDEF, Lundqvist, Flykt, & Ohman, 1998). In total 70 faces were chosen, 10 identities of each of the following emotions: afraid, anger, disgust, neutral, sad, happy and surprised, see figure 1a for examples. In the subjective valence rating task, images from the McGill Face Database (Schmidtmann *et al*., 2020) were employed. In total 26 stimuli were selected, based on previously published ratings, to cover a wide range of emotions, which were: alarmed, amused, contented, depressed, fearful, flirtatious, hostile, joking, panicked, perplexed, relaxed, suspicious, and terrified. For each stimulus condition both the male and female models were selected. See figure 1b for examples.

Figure 1a-b. *(a) Top row example images from the KDEF database: neutral, disgust and happy (left to right), (b) Bottom row example images from the McGill data base: contented, flirtatious and hostile (left to right).*

Depression questionnaire

To assess depression level the DASS-42 depression sub-scale employed (Lovibond and Lovibond, 1995).

Face classification task 1 – An objective Forced-choice task

A 7-alternative forced-choice classification task was performed. The 70 stimuli were presented in-turn on screen, in a random order per participant, for 500ms. Once the stimulus off-set the participants were tasked with clicking on the button whose label best described displayed facial emotion. The seven buttons were labelled: anger, disgust, fear, happy, neutral, sad, and surprised. Submitting a response initiated the next trial after an 800 ms inter-stimulus-interval. A practice block consisting of three trials was completed prior to the main testing block.

Face classification task 2 – A subjective valence rating task

A subjective valance rating task was also performed. The 26 stimuli were presented in-turn on screen, in a random order per participant, for 500ms. Once the stimulus off-set the participant was tasked with making a subjective rating of the valance of the face. This was achieved by making a mouse click on an on-screen horizontal valance scale that was label from Negative (on the left) through to positive (on the right), the neutral point was labelled in the middle of the scale. The width of the scale was 500 pixels, and the x-coordinate of the mouse click was recorded, hence a value of 1 corresponded to the most negative possible response and 500 the most positive, with 250 being the neutral point. Importantly no words, for example, the database emotion descriptions, were displayed at any time to the participants. Submitting a response initiated the next trial after an 800ms inter-stimulus-interval. A practice block consisting of three trials was completed prior to the main testing block.

Results

All data were analysed using Jamovi Version 2.3.28.0. For both tasks a repeated measures ANOVA with a single factor of emotion type (levels being, for example, fearful, sad, etc.) and a between subject factor of depression group (high vs low) were performed. Post hoc Welch's t-tests were performed and additionally effect sizes were calculated (Cohen, 1988). Throughout there were no examples of p-values becoming insignificant after a multiple comparison correction was applied.

Forced-choice Emotion Face Recognition task results

A significant main effect of emotion type was found ($F(6, 564)=170.7$, $p<.001$, $\eta^2_p=0.65$) along with a significant emotion type-depression group interaction ($F(6, 564)=2.7$, $p=.015$, $\eta^2_p=0.028$). The between subject effect (depression group vs controls) was also significant ($F(1, 94)=2.7$, $p=.015$, $\eta^2_p=0.028$).

Post hoc tests indicated significant differences between the depression and control groups for

the following three emotions : anger (t(79.1)=2.3, $p_{corrected}$=.017, d=0.48), sad (t(76.4)=2.2, $p_{corrected}$=.024, d=0.50) and neutral (t(74.0)=2.0, $p_{corrected}$=.033, d=0.42). No differences were found between the groups for the other four emotions: (fear: t(88.2)=-0.12, $p_{corrected}$=.82, disgust: t(91.0)=-0.54, $p_{corrected}$=1, d=0.11, happy: t(93.8)=0.34, $p_{corrected}$=.55, d=0.069, surprise: t(74.9)=1.62, $p_{corrected}$=.083., d=0.33, d=0.030 and). Results are presented in Figure 2.

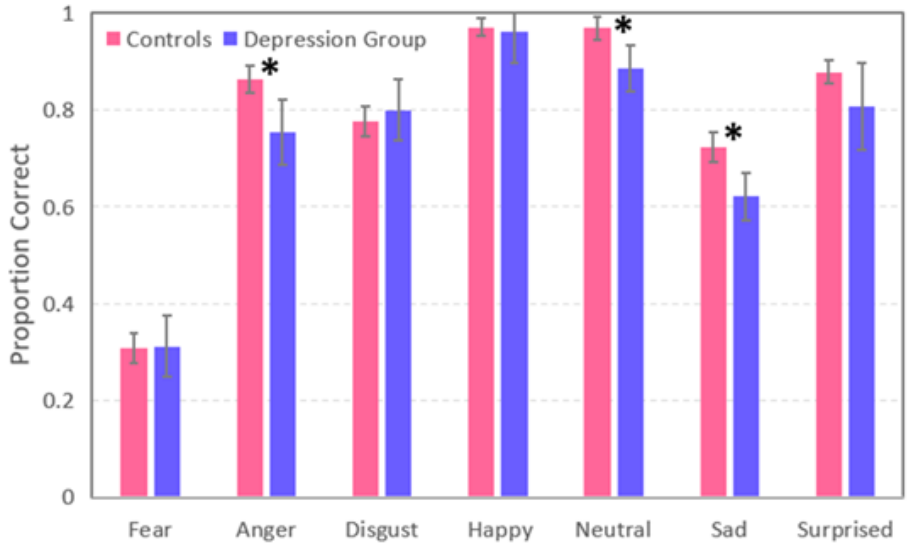

Figure 2. *Proportion correct for each emotion condition, the depression and control groups are plotted in light blue and magenta, respectively. Chance performance is at ~0.14 and the error bars represent ±2SE.*

Subjective valence ratings task results

A significant main effect emotion type F(12, 1716)=428.0, p<.0001, η^2_p=0.75) along with a significant emotion type-depression group interaction was found (F(12, 1716)=5.3, p=.0001, η^2_p =0.036). The between subject effect (depression group vs controls) was not significant F(1, 143)=1.4, p<.25, η^2_p=0.009).

Post hoc t-tests found significant differences in five of the thirteen emotions between the depression and control groups: relaxed (t(89.5)=2.4, $p_{corrected}$=.016, d=0.50), flirtatious (t(78.5)=2.5, $p_{corrected}$=.0.017, d=0.51), joking (t(83.0)=2.8, $p_{corrected}$=.006, d=0.57), contented (t(93.4)=2.2, $p_{corrected}$=.032, d=0.44) and amused (t(78.9)=3.3, $p_{corrected}$<.001, d=0.69). Of the

remaining eight emotions none were significant, all p-values were in the range: $.015 \leq p_{corrected} \leq 1$, with an average value of p=.89±0.30 (mean±SD). Additionally, the average effect size was in the range: $0.012 \leq d \leq 0.29$, with an average value of d=0.17±0.11 (mean±SD). Figure 3 shows the mean ratings per emotion, per group.

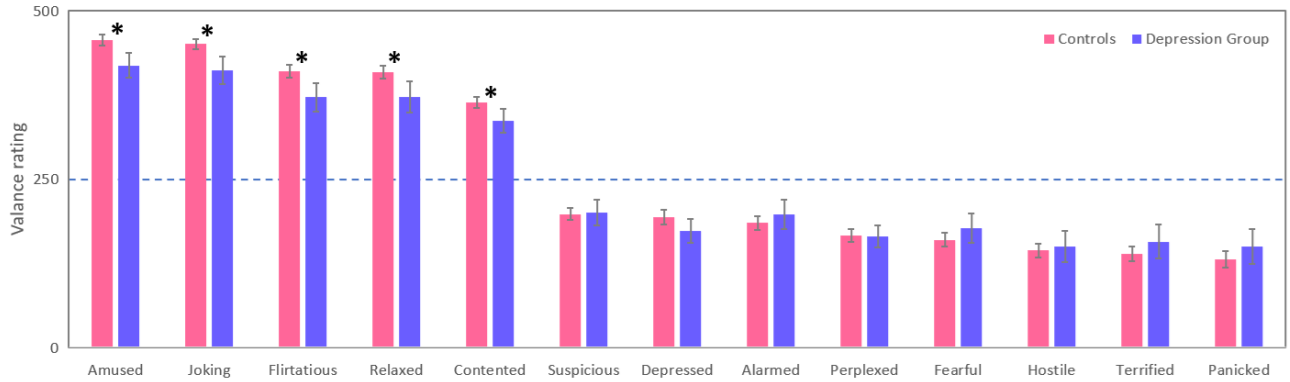

Figure 3. *Mean ratings for each emotion condition, the y-axis represented the valance with 1 being most negative, through to 500 (most positive), the horizontal dashed line represents a neutral rating at 250. The depression and control groups are plotted in light blue and magenta, respectively. Error bars represent ±2SE.*

Discussion

The forced-choice task indicates that accuracy for classifying emotional faces is not equal over emotional conditions, for example, consistent with previous research, the fear condition has much lower accuracy than the than the happy condition (Cheeta et al., 2021). The forced-choice task also indicates reduced accuracy for some specific emotion conditions (anger, neutral and sad) between the control and depression groups. This is consistence with other previous studies, for example, a reduced accuracy for classifying neutral faces in depression has ben demonstrated (Surguladze *et al*., 2004; Linden *et al*., 2011). Consistent with previous studies is also the inconsistency in the current findings, i.e., many previous studies are not in agreement as to exactly which conditions show performance differences. For example, other studies have found between group differences for the surprise condition that are absent in the current study (e.g., Gollan *et al*., 2010), and in a meta-analysis there was no effect on sadness (Dalili et al., 2015) which was seen in the present study. These differences between studies are

probably due to many factors, as discussed in the Introduction. Furthermore, the current study specifically the employed KDEF image database which is composed of relatively high-intensity stimuli which can serve to make the conditions easier to classify overall (Hess, Blairy & Kleck, 1997). Also, the current study screened for depression using the DASS-42, hence participants did not have an official clinical diagnosis of major depressive disorder. However overall, the results for this task are consistent with the broad literature that indicate the depression group will have a reduced ability to identify, and hence classify, facial emotional expressions.

In the subjective rating task the data indicated that with five out of the 13 of the emotional conditions there was a significant difference in valence rating, accompanied with a mean effect size of d=0.54, i.e., medium sized. All other conditions showed no differences. Considering the direction of this difference this is partly consistent with the prediction above that ratings would be more negative in the depression group. However, the data reveals that the conditions that were rated more negatively were all positive facial expressions (amused, joking, flirtatious, relaxed and contented). Hence, this result is consistent with findings that people with depression have a reduced or blunted response to positive events and stimuli (e.g., Rottenberg, Gross & Gotlib, 2005).

By employing both an objective (forced-choice task) and subjective (valence rating task) it is possible in the current study to draw additional conclusions beyond previous studies employing only a single objective forced-choice classification task. For example, the current data indicates that there is no difference in accuracy for classifying happy faces, if this was the only data available a reasonable conclusion would be that the two groups perceive happy faces equally, however by noting that all of the positive conditions in the rating task where given less positive ratings by the depression groups we can conclude that while the depression group can classify the happy faces with similar accuracy they are actually perceiving happy and positive faces differently to the non-depression group.

An additional issue is that subjective tasks are not vulnerable to is the circular nature of objective emotional face tasks. Here a set of images is selected by a researcher, validated by a focus group, then presented in a classification task. Participants are hence tasked with essentially agreeing or disagreeing with the focus groups in-order to be assigned a correct or

incorrect classification, this is a clear limitation if any conclusions are wanting to be drawn regarding how the faces are being perceived. It is recommended that future studies additionally include subjective arousal ratings in addition to valance ratings to give a fuller picture of potential deviations in depression, see for example the methods of Jennings, Yu and Kingdom (2017). It is also possible to employ, in subjective tasks, many more complex expressions, e.g., suspicion, that cover the full range of an emotion space beyond just the basic six Ekman conditions, or in many cases a subset of them, as the requirement of providing labels describing discrete emotions is removed.

Limitations

Among several limitations are the reliance of a self-report measure of depression and that the sample used did not have a clinical diagnosis. Furthermore, the study did not control for other co-morbidities, especially anxiety disorders which are common in those with mood disorders (Beck & Perkins, 2001; Etkin & Schatzberg, 2011). However, there is a huge body of existing evidence of emotional facial recognition deficits in depression, and the effect sizes reported in the present study are all medium to large. Furthermore, all effects remained significant even when controlling for multiple testing. However, studies are needed in clinically depressed samples to evaluate the extent to which the cognitive biases associated with depression are present on both subjective and objective measures of emotional face classification.

Conclusions

Several meta-analyses have confirmed that depression is associated with a bias in the classification of emotional faces (Bourke et al., 2010; Dalili et al., 2015; Krause et al., 2021). Based on the reported findings, the current study makes the methodological recommendation for future studies to not rely solely on objective tasks that measure performance, i.e., accuracy, and that by additionally including a subjective task more nuanced conclusions regarding how emotional faces are perceived, as well as how accurately they can be classified, can be obtained, this concept can be extended to other psychological domains beyond psychopathology.

Acknowledgments

This work was internally funded by Brunel University London via funds awarded by the Centre for Cognitive and Clinical Neuroscience.
Commercial relationships: none.
Corresponding author: Ben J. Jennings (ben.jennings@brunel.ac.uk).

References

Beck, R., & Perkins, T. S. (2001). Cognitive content-specificity for anxiety and depression: A meta-analysis. Cognitive Therapy and Research, 25(6), 651–663. https://doi.org/10.1023/A:1012911104891

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: A review. SAGE Publications. https://doi.org/10.3109/00048674.2010.496359

Csukly, G., Czobor, P., Szily, E., Takács, B., & Simon, L. (2009). Facial expression recognition in depressed subjects: The impact of intensity level and arousal dimension. The Journal of Nervous and Mental Disease, 197(2), 98–103. https://doi.org/10.1097/NMD.0b013e3181923f82

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafò, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. Psychological Medicine, 45(6), 1135–1144. https://doi.org/10.1017/S0033291714002591

Deckert, M., Schmoeger, M., Auff, E., & Willinger, U. (2020). Subjective emotional arousal: an explorative study on the role of gender, age, intensity, emotion regulation difficulties, depression and anxiety symptoms, and meta-emotion. Psychol Res. 84(7):1857-1876. doi: 10.1007/s00426-019-01197-z. Epub 2019 May 16. PMID: 31098662; PMCID: PMC7478944.

Ekman, P. (1992). An Argument for Basic Emotions. Cognition and Emotion. 6 (3/4): 169–200.

CiteSeerX 10.1.1.454.1984. doi:10.1080/02699939208411068.

Etkin, A., Schatzberg, A. F. (2011). Common abnormalities and disorder-specific compensation during implicit regulation of emotional processing in generalized anxiety and major depressive disorders. Am J Psychiatry. 168(9):968-78. doi:10.1176/appi.ajp.2011.10091290. Epub 2011 Jun 1. PMID: 21632648.

Flanagan, T.J., White, H., & Carter, B.G. (2011). Differential impairments in emotion face recognition in postpartum and nonpostpartum depressed women. J Affect Disord. 2011 Feb;128(3):314-8. doi: 10.1016/j.jad.2010.07.021. Epub 2010 Aug 25. PMID: 20800287.

Gollan, J.K., McCloskey, M., Hoxha, D., & Coccaro, E.F. (2010) How do depressed and healthy adults interpret nuanced facial expressions? J Abnorm Psychol. 2010 Nov;119(4):804-10. doi: 10.1037/a0020234. PMID: 20939654; PMCID: PMC3805828.

Hess, U., Blairy, S., & Kleck, R. (1997). The Intensity of Emotional Facial Expressions and Decoding Accuracy. Journal of Nonverbal Behavior. 21. 241-257. 10.1023/A:1024952730333.

Krause, F.C., Linardatos. E., Fresco. D.M., & Moore. M.T. (2021). Facial emotion recognition in major depressive disorder: A meta-analytic review. J Affect Disord. 2021 Oct 1;293:320-328. doi: 10.1016/j.jad.2021.06.053. PMID: 34229285; PMCID: PMC8457509.

Jennings, B. J., Yu, Y., & Kingdom, F. A. A. (2017). The role of spatial frequency in emotional face classification. Attention, Perception, & Psychophysics, 79(6), 1573–1577. https://doi.org/10.3758/s13414-017-1377-7

Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? biases in the identification of emotional facial expressions in depression and social phobia. Journal of Abnormal Psychology (1965), 115(4), 705–714. https://doi.org/10.1037/0021-843X.115.4.705

Kan, Y., Mimura, M., Kamijima, K., & Kawamura, M. (2004). Recognition of emotion from

moving facial and prosodic stimuli in depressed patients. Journal of Neurology, Neurosurgery and Psychiatry, 75(12), 1667–1671. https://doi.org/10.1136/jnnp.2004.036079

Langenecker, S. A., Bieliauskas, L. A., Rapport, L. J., Zubieta, J., Wilde, E. A., & Berent, S. (2005). Face emotion perception and executive functioning deficits in depression. Journal of Clinical and Experimental Neuropsychology, 27(3), 320–333. https://doi.org/10.1080/13803390490490515720

Linden S. C., Jackson M. C., Subramanian L., Healy D., and Linden D. E. J., Sad benefit in face working memory: An emotional bias of melancholic depression, Journal of Affective Disorders. (2011) 135, no. 1-3, 251–257, https://doi.org/10.1016/j.jad.2011.08.002, 2-s2.0-80055025457, 21872338.

Lovibond, P.F., & Lovibond, S.H. (1995). The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther. 33(3):335-43. doi: 10.1016/0005-7967(94)00075-u. PMID: 7726811.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Rottenberg, J., Gross, J.J., & Gotlib, I.H. (2005). Emotion context insensitivity in major depressive disorder. J Abnorm Psychol. 114(4):627-39. doi: 10.1037/0021-843X.114.4.627. PMID: 16351385.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. Psychological Bulletin, 115(1), 102–141. https://doi.org/10.1037/0033-2909.115.1.102

Schmidtmann, G., Jennings, B. J., Sandra, D. A., Pollock, J., & Gold, I. (2020). The McGill Face Database: Validation and Insights Into the Recognition of Facial Expressions of Complex

Mental States. Perception, 49(3), 310-329. https://doi.org/10.1177/0301006620901671

Surguladze S. A., Young A. W., Senior C., Brébion G., Travis M. J., and Phillips M. L., Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression, Neuropsychology. (2004) 18, no. 2, 212–218, https://doi.org/10.1037/0894-4105.18.2.212, 2-s2.0-1942438155.

WHO. (2008). The burden of disease: 2004 update. https://iris.who.int/bitstream/handle/10665/43942/9789241563710_eng.pdf?sequence=1. Retrieved 2025

Willner, P., Scheel-Krüger, J., & Belzung, C. (2013). The neurobiology of depression and antidepressant action. Neuroscience and Biobehavioral Reviews, 37(10), 2331–2371. https://doi.org/10.1016/j.neubiorev.2012.12.007