Bone, J. K., Lewis, G., Button, K. S., Duffy, L., Harmer, C. J., Munafò, M. R., ... Lewis, G. (2019). Variation in recognition of happy and sad facial expressions and self-reported depressive symptom severity: A prospective cohort study. *Journal of Affective Disorders*, *257*, 461-469. https://doi.org/10.1016/j.jad.2019.06.025

Journal of Affective Disorders 257 (2019) 461–469

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Variation in recognition of happy and sad facial expressions and self-reported depressive symptom severity: A prospective cohort study

Jessica K. Bone[a,*], Gemma Lewis[a], Katherine S. Button[b], Larisa Duffy[a], Catherine J. Harmer[c], Marcus R. Munafò[d,e], Ian S. Penton-Voak[d], Nicola J. Wiles[f], Glyn Lewis[a]

[a] Division of Psychiatry, Faculty of Brain Sciences, University College London, 6th Floor, Maple House, 149 Tottenham Court Road, London W1T 7NF, UK
[b] Department of Psychology, University of Bath, Bath, UK
[c] Department of Psychiatry, University of Oxford, Oxford, UK
[d] School of Experimental Psychology, University of Bristol, Bristol, UK
[e] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[f] Centre for Academic Mental Health, Population Health Sciences, University of Bristol, Bristol, UK

ARTICLE INFO

Keywords:
Depression
Facial expressions
Emotion recognition
Cognition
Cohort study

ABSTRACT

Objective: Cognitive theories suggest people with depression interpret self-referential social information negatively. However, it is unclear whether these biases precede or follow depression. We investigated whether facial expression recognition was associated with depressive symptoms cross-sectionally and longitudinally.
Methods: Prospective cohort study of people who had visited UK primary care in the past year reporting depressive symptoms (n = 509). Depressive symptoms were measured using the Patient Health Questionnaire (PHQ-9) at four time-points, 2 weeks apart. A computerised task assessed happy and sad facial expression recognition at three time-points (n = 505 at time 1). The unbiased hit rate measured ability to recognise emotions accounting for any general tendency to identify the emotion when it was not present.
Results: The sample included the full range of depressive symptom severity, with 45% meeting diagnostic criteria for depression. There was no evidence that happy or sad unbiased hit rates were associated with concurrent or subsequent depressive symptoms. There was weak evidence that, for every additional face incorrectly classified as happy, concurrent PHQ-9 scores reduced by 0.05 of a point (95% CI = -0.10 to 0.002, p = 0.06 after adjustment for confounders). This association was strongest for more ambiguous facial expressions (interaction term $p < 0.001$).
Limitations: This was an observational study with relatively short follow-up (6 weeks) and small changes in depressive symptoms and emotion recognition. Only 7% of invited patients consented to participate.
Conclusions: Reduced misclassifications of ambiguous faces as happy could be a state marker of depression, but was not associated with subsequent depressive symptoms. Future research should focus on the interpretation of ambiguous social information.

1. Introduction

Beck proposed that negative self-evaluations, beliefs, and memories play a key role in depression (Beck, 1967). Since then, evidence has supported increased negative social and emotional processing in depression (Beck, 2008; Elliott et al., 2011; Roiser et al., 2012). In recent theories of depression, negative processing biases are proposed to play a causal role in the development of depressive symptoms (Disner et al., 2011; Roiser et al., 2012).

Studies of information processing and depression traditionally used self-report questionnaires in which participants recorded hypothetical responses to events (Elliott et al., 2011). This restricts investigation to conscious thoughts and behaviours which may be susceptible to mood-congruent response biases (Colman et al., 2016). There has been an increasing focus on behavioural tasks that assess automatic cognitive processes, which may influence thoughts and behaviours without conscious awareness (Kahneman, 2011; Roiser et al., 2012).

One automatic process which may be important in depression is the interpretation of emotional facial expressions. Facial expressions represent what other people think of you, providing potentially ambiguous self-referential cues used to make inferences and decisions. For example, seeing a happy face may facilitate approach and

* Corresponding author.
E-mail address: jessica.bone.15@ucl.ac.uk (J.K. Bone).

https://doi.org/10.1016/j.jad.2019.06.025
Received 11 March 2019; Received in revised form 31 May 2019; Accepted 21 June 2019
Available online 22 June 2019

reinforcement of one's behaviour. Negative self-schema may influence the interpretation of facial expressions for example by making an ambiguous expression more likely to be negatively interpreted (Beck, 1979).

Some studies have found that people with depression are more likely to interpret neutral or ambiguous faces as sad (Bouhuys et al., 1999; Gollan et al., 2010; Gur et al., 1992; Hale, 1998; Lee et al., 2016; Leppänen et al., 2004) or more accurately identify sadness (Milders et al., 2010). One meta-analysis found that people with depression were worse at recognising all facial emotions except for sadness. However, the effect size was small and there was evidence of publication bias (Dalili et al., 2015).

In contrast, other studies have found that people with depression identify happy faces less accurately (Gur et al., 1992; Surguladze et al., 2004; Zwick and Wolkenstein, 2017), or are less likely to interpret neutral faces as happy than healthy controls (Gollan et al., 2010), with no differences for sad faces. Another finding is that healthy people have a positive processing bias, which may increase resilience, and is reduced in depression (Alloy and Abramson, 1979; Moore and Fresco, 2012). This is in line with the decreased positive emotionality model of depression, which states that a reduction in positive affect is specific to depression (Tellegen, 1985; Watson et al. 1988).

These inconsistent findings may arise from methodological limitations of prior studies. Most studies have used small samples, limiting statistical power (Button et al., 2013a), and case-control designs which are prone to selection biases unless cases and controls are selected from the same population (Schulz and Grimes, 2002). It is generally accepted that depression is a continuum, ranging from no symptoms to many (Hankin et al., 2005). However, there are very few studies which have examined biases in emotion recognition across the whole continuum of depressive symptom severity (Kohler et al., 2011; Lee et al., 2016).

Additionally, most previous studies have been cross-sectional, meaning there is little evidence on the direction of the association between emotion recognition and depression. According to the cognitive neuropsychological model of depression, changes in automatic emotional processing precede and have a causal role in depressive illness (Disner et al., 2011; Harmer et al., 2009; Roiser et al., 2012). One study found that greater recognition of happy faces during early antidepressant use (before symptomatic change) was associated with improved depressive symptoms six weeks later (Tranter et al., 2009). Additionally, increased incorrect classification of neutral faces as happy has been associated with depressive symptom remission (Leppänen et al., 2004). In a large study, slow identification of happiness was associated with the onset of depression during the following 8 years (Vrijen et al., 2016). In contrast, greater recognition of sadness has been associated with persistence of depression six months later (Bouhuys et al., 1999; Hale, 1998) and lower negative perceptual bias was associated with decreased depressive symptoms three months later (Münkler et al., 2015).

Other studies have found unchanged facial emotion recognition despite reduced depressive symptoms. In one small study, responses to sad facial expressions remained stable despite reductions in depressive symptom severity six months later (Milders et al., 2010). Over a longer period (2.5 years), negative biases in facial expression recognition were present regardless of recurrence-status for individuals with major depressive disorder (Ruhe et al., 2019). The inconsistent longitudinal evidence may be a result of small samples, with limited variation in depressive symptoms and low power.

Other methodological differences in facial expression recognition tasks may also cause conflicting findings. Some tasks involve different intensities of emotional expressions. Others present only prototypical emotional expressions which are easier to identify and may result in ceiling effects. A variety of measures can be taken from these tasks including hits, false alarms, reaction times and measures of accuracy or sensitivity. Responses can be analysed by emotion, averaged across positive and negative emotions, or used to create an index of positive or negative bias. Different approaches may give different results, and there is a tendency to interpret diverse findings as supporting the presence of emotional biases, without giving due weight to results which do not support their existence. It is therefore unsurprising that the literature is inconsistent, with few studies reproducing prior results.

In this study, we aimed to address these issues by using data from a large (n = 509) prospective cohort of people who had presented to UK primary care surgeries with depressive symptoms. Depressive symptoms ranged from mild to severe and participants were recruited from the same population, reducing selection bias. We tested concurrent and longitudinal associations between depressive symptoms and recognition of happy and sad facial expressions of varying emotional intensities.

2. Methods

2.1. Participants

Participants were recruited from General Practice (GP) surgeries in three United Kingdom sites (Bristol, Liverpool, York). GPs searched computerised records for those who had reported depressive symptoms during the last year (details in Supplementary Methods 1). Patients were aged 18–70 years but there were no restrictions on whether they were receiving antidepressants or psychological therapy. We excluded people who were diagnosed with bipolar disorder, psychosis or an eating disorder; had alcohol or substance use problems; were unable to complete study questionnaires; or were over 29 weeks pregnant (as these women would not have been able to complete the study).

Eligible patients (N = 7721) were sent information and 1470 (19%) replied. Of these, 821 (55%) were willing to be contacted, 23 (3%) of whom were ineligible. The remaining 798 were contacted to arrange interviews, and 563 (7%) consented. Data were collected at four time-points, each two weeks apart, at the participant's home or GP surgery. Assessments were completed by trained research assistants and clinical research nurses. At time one, 558 patients provided data (five could not be contacted), and at follow-ups two (two weeks), three (four weeks), and four (six weeks); 476 (85%), 443 (79%) and 430 (77%) respectively. Participants were not compensated for their participation in this study.

2.2. Ethics approval

Ethics approval was obtained from NRES Committee South West - Central Bristol and participants provided written informed consent after receiving a complete description of the study. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

2.3. Measures

The primary aim of the original study was to estimate a clinically important difference on commonly used self-administered questionnaires for depressive symptoms (unpublished). This was a secondary analysis.

2.3.1. Depressive symptoms

The Patient Health Questionnaire (PHQ-9; Spitzer et al., 1999) was completed at each time-point. It is a nine-item self-report measure of depressive symptoms, with higher scores indicating greater severity. Internal consistency was high at each time-point (Cronbach's alpha 0.89–0.92). The Beck Depression Inventory (BDI-II; Beck et al., 1961) was also used at each time-point and is a 21-item self-report measure of the severity of depressive symptoms. The PHQ-9 was our primary measure of depressive symptoms as it may be more responsive to

change in symptom severity (Kounali et al., 2016). However, as two measures of depressive symptoms were available, we repeated all analyses using depressive symptoms measured on the BDI-II.

At baseline, the self-administered computerised Clinical Interview Schedule Revised (CIS-R; Lewis et al., 1992) was used to assess whether participants met diagnostic criteria for depression according to the International Classification of Diseases (ICD-10). Diagnostic criteria was based on severity of depressive symptoms in the past week, with a cut-off score of 12. All fieldwork staff attended a training day organised by the lead study site (Bristol) and were trained in the use of the CIS-R.

2.3.2. Recognition of facial expressions of emotion

Participants performed a computerised task at times one, two, and three that required them to identify facial expressions in a six-alternative forced choice task (Attwood et al., 2017). They were asked to identify whether the emotional facial expression presented was happiness, sadness, fear, anger, disgust, or surprise. Facial stimuli were prototypes created by averaging across 12 adult male faces posing the same emotional expressions (details in Supplementary Methods 2). Fifteen intensities of each emotion were presented, created by morphing faces incrementally from an emotionally ambiguous prototype (0% emotion) to each emotional prototype (100% emotion) over 15 steps. As the intensity of expressions increased linearly from ambiguous to the full prototype, emotions became clearer and easier to identify.

In each trial, a cross appeared in the centre of the screen for 1500 to 2500 ms, followed by the face stimulus for 150 ms, and then a mask of visual noise for 250 ms (to prevent processing of after images). Six emotion labels then appeared on the screen in a circular formation until the participant responded by selecting one of them, at which point the next trial started. Label positions were randomly chosen for each participant and stayed the same throughout the session. Each emotion was presented once at each intensity in one block of 90 trials. The task took approximately six to eight minutes to complete and was delivered using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) on a laptop computer. Responses were made using a computer mouse.

We focussed on happy and sad facial expression recognition because we had clear hypotheses for the role of these emotions in depression. As in previous use of this task, trials displaying anger, fear, surprise and disgust were included in the design in order to make the discrimination task sufficiently challenging (Griffiths et al., 2015). We only tested recognition of happy and sad facial expressions to reduce the number of exploratory analyses and multiple comparisons in this study. For each emotion, correct responses were 'hits' and incorrect responses were 'misses'. Incorrect responses were also categorised as 'false alarms', the misidentification of an emotion when a different expression was presented. For example, responding 'happy' to a surprised facial expression was a happy false alarm.

2.3.3. Confounders

Demographic variables collected at baseline included sex, age, ethnic group, marital status, education, and negative life events. We created a binary education variable, split at the level which used to be compulsory in the UK (0 = school education up to 16 years, 1 = completed secondary school or university education). Negative life events were self-reported and categorised (0 = none, 1 = one or more). We also adjusted for self-reported antidepressant use at each time-point.

2.4. Primary analyses

The main analyses were performed using an 'unbiased hit rate' for each emotion. This balanced the ability to correctly identify an emotion (hits) with incorrect identifications when the emotion was not present (false alarms). By providing an overall measure of sensitivity that accounted for response bias (Wagner, 1993), this demonstrated ability to recognise emotions as opposed to a general tendency to identify the emotion regardless of whether it was present. Higher values of the unbiased hit rate indicated more accurate responses, or an increased ratio of hits to misses and false alarms. The formula for the unbiased hit rate took simultaneous account of whether a stimulus was correctly identified when it was presented and whether a response was correct when it was made (Wagner, 1993):

$$Unbiased\ hit\ rate = \frac{hits^2}{(hits + misses)\ x\ (hits + false\ alarms)}$$

2.4.1. Concurrent associations between emotion recognition and depressive symptoms

Multilevel linear regression models were used to calculate concurrent associations between emotion recognition (continuous exposure) and depressive symptoms (continuous outcome) across time-points one to three (when both facial recognition and depressive symptom data were available). Measures of emotion recognition (happy and sad unbiased hit rate) were included as exposures in separate regression models. As false alarm rates were not independent exposures, we could not enter measures of happy and sad facial recognition in the same model. An increase in happy false alarms, for example, could have led to a decrease in sad false alarms (Attwood et al., 2017).

2.4.2. Longitudinal associations between emotion recognition and depressive symptoms

Longitudinal associations between facial expression recognition (continuous exposure) and depressive symptoms (continuous outcome) were examined using multilevel linear models across time-points. We included emotion recognition at times one to three and depressive symptoms at times two to four. This model tested whether emotion recognition was associated with subsequent depressive symptoms whilst adjusting for depressive symptoms at baseline. This allowed us to include data from all time-points. Separate analyses were performed for happy and sad unbiased hit rates.

All multilevel models had a random effect for participant to allow for the clustering of responses within individuals over time-points and are presented before and after adjustment for confounders. Analyses were performed using Stata 14.

2.5. Secondary analyses

We repeated analyses using hits and false alarms as measures of emotion recognition, allowing examination of associations between depressive symptoms, accuracy and response bias separately. Hits and false alarms were included as exposure variables in the same model for each emotion, as done previously (Button et al., 2013b). This accounted for the general tendency to identify an emotion regardless of whether it was present (response bias; Wagner, 1993).

2.6. Exploratory analyses

We then performed unplanned post-hoc analyses where there was evidence of an association between emotion recognition and depressive symptoms, examining whether facial expression intensity influenced these associations. Depressive symptom severity was an exposure and the binary outcome was whether or not a false alarm occurred on each trial. This model included a level specifying the facial expression intensity on each trial. Multilevel logistic regression modelled concurrent associations between happy false alarms (binary outcome) and depressive symptoms and emotion intensity (continuous exposures). As the association between intensity and happy false alarms was not linear, a quadratic term was included in this model. We tested whether there was an interaction between emotion intensity and depressive symptoms for the probability of happy false alarms.

We also analysed associations between antidepressant use and

Table 1
Number of happy and sad hits (from a total of 15) and happy and sad false alarms (from a possible total of 75) according to sample characteristics at baseline.

	Participants	Hits				False alarms			
	n	Happy Mean (SD)	*P*	Sad Mean (SD)	*p*	Happy Mean (SD)	*p*	Sad Mean (SD)	*p*
Characteristic									
Age									
Under 50	254 (50%)	12.08 (2.07)	0.01	10.66 (2.41)	<0.001	5.29 (5.91)	0.002	6.66 (5.68)	0.14
50+	255 (50%)	11.58 (2.43)		8.90 (3.03)		6.96 (6.09)		7.37 (4.97)	
Gender									
Male	162 (32%)	11.73 (2.45)	0.50	9.76 (2.89)	0.92	7.32 (6.82)	0.002	6.53 (5.18)	0.17
Female	347 (68%)	11.87 (2.17)		9.79 (2.87)		5.57 (5.59)		7.24 (5.42)	
Ethnicity									
White	492 (97%)	11.85 (2.26)	0.14	9.79 (2.87)	0.05	6.15 (6.07)	0.93	7.01 (5.40)	0.88
Black	8 (2%)	10.75 (1.83)		7.50 (2.14)		5.63 (5.68)		8.00 (2.00)	
Asian	3 (<1%)	13.33 (0.58)		13.00 (1.73)		4.67 (2.08)		5.00 (6.24)	
Mixed	3 (<1%)	9.33 (2.31)		10.33 (3.06)		3.67 (3.06)		8.67 (4.51)	
Other	3 (<1%)	11.67 (3.05)		11.00 (1.73)		7.33 (11.02)		5.67 (5.03)	
Education									
Lower	197 (39%)	11.56 (2.44)	0.04	9.08 (3.13)	<0.001	6.89 (6.24)	0.03	7.03 (5.00)	0.96
Higher	312 (61%)	11.99 (2.14)		10.23 (2.61)		5.65 (5.90)		7.00 (5.56)	
Negative life events									
None	211 (42%)	11.96 (2.01)	0.53	9.69 (2.99)	0.32	6.48 (5.92)	0.28	7.29 (5.04)	0.41
One or more	287 (58%)	11.84 (2.30)		9.94 (2.70)		5.88 (6.15)		6.88 (5.61)	
Depression diagnosis									
No	276 (55%)	11.97 (2.06)	0.33	9.84 (2.72)	0.96	6.47 (6.19)	0.17	7.28 (5.47)	0.30
Yes	222 (45%)	11.78 (2.33)		9.83 (2.95)		5.71 (5.88)		6.77 (5.25)	
Currently taking antidepressants									
No	156 (31%)	11.65 (2.53)	0.18	9.61 (3.16)	0.35	5.97 (5.95)	0.67	7.36 (5.94)	0.34
Yes	348 (69%)	11.94 (2.11)		9.87 (2.73)		6.22 (6.11)		6.87 (5.07)	

Note. Emotion recognition data was present for $N = 505$ (99% of total sample). Negative life event and depression diagnosis were both missing for 11 participants (2%) and 5 participants were missing data on whether they were taking antidepressants at baseline (1%).

emotion recognition (Supplementary Materials).

3. Results

3.1. Descriptive statistics

We excluded participants (9% of total sample) who were missing baseline demographic data, or data on emotion recognition or depressive symptoms at all time-points. The final sample comprised 509 participants (68% female) aged between 18 and 71 years (mean 48.08, SD 12.71).

At baseline, 45% met diagnostic criteria for depression according to the ICD-10. The sample included the full range of depressive symptom severity with PHQ-9 scores ranging from 0 to 27 (mean 10.38, SD 6.62; Supplementary Figure 1) and BDI scores ranging from 0 to 58 (mean 20.32, SD 12.15; Supplementary Figure 2). Mean happy and sad hits and false alarms are shown in Table 1, according to demographic and clinical characteristics. Overall, participants made more happy than sad hits (mean difference = 1.66, 95% CI = 1.38 to 1.93, $p<0.001$), accurately identifying 12 of the 15 happy faces presented. Participants made more sad than happy false alarms (mean difference = −1.26, 95% CI = −2.03 to −0.49, $p = 0.001$), misidentifying seven sad faces versus six happy faces on average. Supplementary Table 1 shows depressive symptoms, unbiased hit rate, hits, and false alarms over time.

3.2. Primary analyses: unbiased hit rate

We found no evidence for associations between the happy or sad unbiased hit rate and concurrent depressive symptoms measured using the PHQ-9 and BDI-II (Table 2). There was also no evidence that emotion recognition was associated with depressive symptoms longitudinally. The happy and sad unbiased hit rates were not associated with subsequent PHQ-9 or BDI-II scores (Table 3).

Table 2
Concurrent associations between unbiased hit rate for happy and sad facial expressions (exposure variables) and depressive symptoms (outcome variable).

	Concurrent model 1 ($n = 446$)			Concurrent model 2: adjusted for confounders* ($n = 437$)		
	Coef	95% CI	*p*	Coef	95% CI	*p*
Change in PHQ-9 score for a one-unit increase in unbiased hit rate						
Happy	1.37	−0.16 to 2.90	0.08	1.31	−0.25 to 2.86	0.10
Sad	0.23	−1.47 to 1.93	0.79	−0.50	−2.26 to 1.26	0.58
Change in BDI-II score for a one-unit increase in unbiased hit rate						
Happy	0.78	−1.79 to 3.35	0.55	0.71	−1.89 to 3.31	0.59
Sad	−0.95	−3.82 to 1.92	0.52	−1.55	−4.50 to 1.39	0.30

Note. The coefficient is the unstandardized regression coefficient, representing the change in depressive symptoms for each unit (one point) increase in unbiased hit rate. Concurrent associations use the exposures and outcome measured at times 1–3.

* Confounders were age, sex, ethnic group, education, marital status, negative life events, and concurrent antidepressant use.

3.3. Secondary analyses: hits and false alarms

3.3.1. Concurrent associations

Concurrent associations between happy hits and depressive symptoms are shown in Table 4. The main analyses used the PHQ-9. There was no evidence that happy hits, alone or adjusted for false alarms, were associated with concurrent depressive symptoms. There was weak evidence for an association between happy false alarms (incorrectly classifying other emotions as happy) and depressive symptoms. For every additional happy false alarm, PHQ-9 scores reduced by 0.06 points (95% CI = −0.11 to −0.004, $p = 0.04$). As shown in Fig. 1, depressive symptoms decreased as the number of happy false alarms increased. After adjustment for confounders, any evidence for an

Table 3
Longitudinal associations between unbiased hit rate for happy and sad facial expressions (exposure variables) and depressive symptoms (outcome variable).

	Longitudinal model 1 (n = 446)			Longitudinal model 2: adjusted for confounders* (n = 437)		
	Coef	95% CI	p	Coef	95% CI	p
Change in PHQ-9 score for a one-unit increase in unbiased hit rate						
Happy	0.59	−0.90 to 2.08	0.44	0.33	−1.05 to 1.71	0.64
Sad	−0.15	−1.79 to 1.50	0.86	−0.63	−2.18 to 0.92	0.43
Change in BDI-II score for a one-unit increase in unbiased hit rate						
Happy	1.31	−1.20 to 3.81	0.31	0.67	−1.59 to 2.93	0.56
Sad	1.18	−1.60 to 3.96	0.41	1.07	−1.47 to 3.61	0.41

Note. The coefficient is the unstandardized regression coefficient, representing the change in depressive symptoms for each unit (one point) increase in unbiased hit rate. Longitudinal period-lagged associations use unbiased hit rate for happy and sad facial expressions at times 1–3 (exposure variables) and depressive symptoms at times 2–4 (outcome variable).

* Confounders were baseline depressive symptoms, age, sex, ethnic group, education, marital status, negative life events, and concurrent antidepressant use.

association between happy false alarms and PHQ-9 scores remained very weak (coef = −0.05, 95% CI = −0.10 to 0.002, p = 0.06).

There was no evidence that sad hits or sad false alarms were associated with concurrent depressive symptoms (Table 4). As shown in Fig. 1, depressive symptoms were similar across the range of sad false alarms made by participants.

When these analyses were repeated using BDI-II score, no evidence was found for any associations between happy and sad facial expression recognition and depressive symptoms (Table 4). There was no evidence for an association between happy false alarms and BDI-II score (coef = −0.003, 95% CI = −0.09 to 0.09, p = 0.95 after adjustment for confounders). We found no evidence of an interaction between depression diagnoses and happy or sad facial expression recognition on depressive symptoms (measured using either the PHQ-9 or BDI-II; all $p > 0.10$).

3.3.2. Longitudinal associations

There was no evidence of longitudinal associations between facial expression recognition (hits and false alarms) and depressive symptoms measured using the PHQ-9 and BDI-II (see Supplementary Table 2).

3.4. Exploratory analyses

3.4.1. Associations with emotion intensity

Analyses so far have investigated emotion recognition across all intensities of facial expressions. In these exploratory analyses we examine whether the intensity of each facial expression is associated with the probability of happy false alarms. When both intensity and PHQ-9 score were included as exposures, increased depressive symptoms were associated with decreased odds of happy false alarms (OR = 0.98, 95% CI = 0.96 to 0.99, p = 0.002 after adjustment for confounders). For each 1-unit increase in emotion intensity, the odds of a happy false alarm decreased by 29% (OR = 0.71, 95% CI = 0.69 to 0.73, $p < 0.001$ after adjustment for confounders).

There was also evidence for an interaction between emotion intensity and depressive symptoms on the probability of making a happy false alarm (interaction term after adjustment for confounders $p < 0.001$). Emotion intensity was analysed continuously but Fig. 2 shows high versus low intensities to illustrate this interaction. At higher emotion intensities (when facial expressions were clearer and identification relatively easy) the probability of making happy false alarms did

Table 4
Concurrent associations between hits and false alarms (exposure variables) and depressive symptoms (outcome variable). Separate models were conducted for happy and sad exposure variables.

		Model 1: Hits (n = 509)			Model 2: False alarms (n = 509)			Model 3: False alarms adjusted for hits (n = 509)			Model 4: Additionally adjusted for confounders[a] (n = 498)		
		Coef	95% CI	p	Coef	95% CI	p	Coef	95% CI	p	Coef	95% CI	p
	Change in PHQ-9 scores for a one-unit increase in facial expression recognition												
Happy	Hits	−0.06	−0.18 to 0.05	0.26	–	–	–	−0.04	−0.16 to 0.07	0.46	−0.05	−0.17 to 0.07	0.40
	False alarms	–	–	–	−0.06	−0.11 to −0.004	0.04	−0.05	−0.10 to −0.001	0.05	−0.05	−0.10 to 0.002	0.06
Sad	Hits	0.02	−0.08 to 0.12	0.71	–	–	–	0.02	−0.08 to 0.12	0.76	−0.003	−0.11 to 0.10	0.95
	False alarms	–	–	–	0.01	−0.04 to 0.06	0.71	0.01	−0.04 to 0.06	0.76	0.03	−0.03 to 0.08	0.34
	Change in BDI-II scores for a one-unit increase in facial expression recognition												
Happy	Hits	−0.06	−0.25 to 0.13	0.52	–	–	–	−0.06	−0.25 to 0.13	0.53	−0.06	−0.25 to 0.14	0.57
	False alarms	–	–	–	−0.01	−0.10 to 0.08	0.83	−0.005	−0.09 to 0.08	0.92	−0.003	−0.09 to 0.09	0.95
Sad	Hits	0.01	−0.16 to 0.18	0.91	–	–	–	−0.01	−0.18 to 0.17	0.93	−0.003	−0.18 to 0.18	0.97
	False alarms	–	–	–	0.04	−0.04 to 0.13	0.34	0.04	−0.05 to 0.13	0.34	0.07	−0.02 to 0.16	0.13

Note. The coefficient is the unstandardized regression coefficient, representing the change in depressive symptoms for each unit (one point) increase in hits or false alarms.

[a] Confounders were age, sex, ethnic group, education, marital status, negative life events, and concurrent antidepressant use.

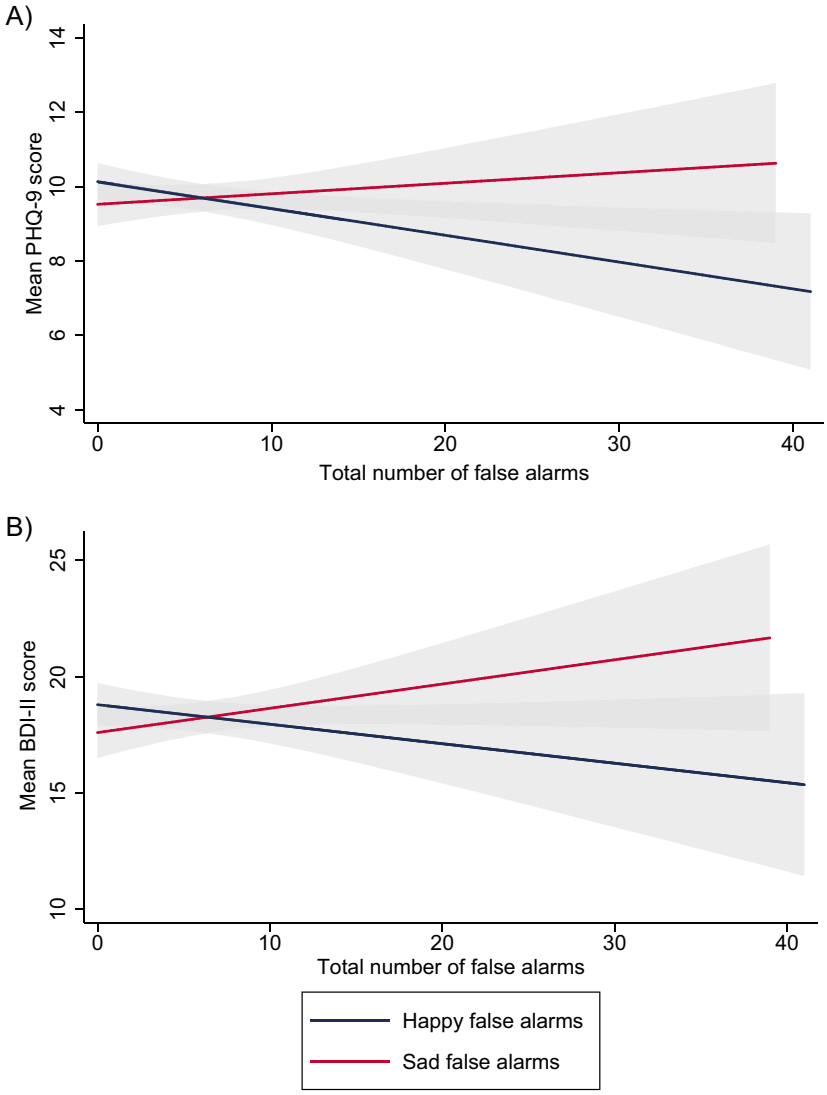

Fig. 1. Association between number of happy and sad false alarms and depressive symptoms. Shading represents the 95% confidence intervals. A) Shows depressive symptoms measured using the PHQ-9 B) shows depressive symptoms measured using the BDI-II.

not differ according to depressive symptoms. However, at lower emotion intensities (when facial expressions were more ambiguous and identification harder) the probability of making a happy false alarm decreased as depressive symptoms increased. This pattern held over both measures of depressive symptoms as, despite the lack of evidence for associations between BDI-II score and happy false alarms, there was evidence for an interaction between emotion intensity and BDI-II score on the probability of making a happy false alarm (interaction term after adjustment for confounders $p < 0.001$).

3.4.2. Associations with antidepressant use

Antidepressant use was relatively stable over time (between 69% and 71% of participants reported taking antidepressants across timepoints) and self-reported adherence was high (87%−90% took them every day). Participants taking antidepressants had more severe depressive symptoms (mean difference = 2.50, 95% CI = 1.82 to 3.18, $p < 0.001$). There was no evidence for associations between antidepressant use and happy or sad face recognition (Supplementary Table 3).

4. Discussion

There was no evidence that the happy or sad unbiased hit rate (a measure of participants' ability to recognise emotions versus the general tendency to identify the emotion regardless of whether it was present) was associated with concurrent depressive symptoms. After breaking the unbiased hit rate down into its constituent parts, there was no evidence of an association between happy hits, sad hits, or sad false alarms and depressive symptoms. We found weak evidence of an association between happy false alarms and depressive symptoms.

Overall, as happy false alarms increased, depressive symptoms measured using the PHQ-9 decreased. However, this finding was not replicated with the BDI-II. This may be due to the small size of the effects, especially as the PHQ-9 coefficient lies within the confidence interval for the BDI-II coefficient. Other explanations could be that this is a chance finding, particularly when taken in light of the lack of evidence for other associations between facial expression recognition and depressive symptoms. Alternatively, the PHQ-9 and BDI-II may measure different aspects of depression which are related to different biases in emotional processing.

In exploratory analyses, there was evidence that both the PHQ-9 and BDI-II measures of depressive symptoms were associated with fewer happy false alarms at lower emotion intensities, when faces were ambiguous. Although the conclusion is speculative, based on exploratory analyses, it is plausible as one would expect any cognitive biases to act more powerfully with ambiguous stimuli. Therefore,

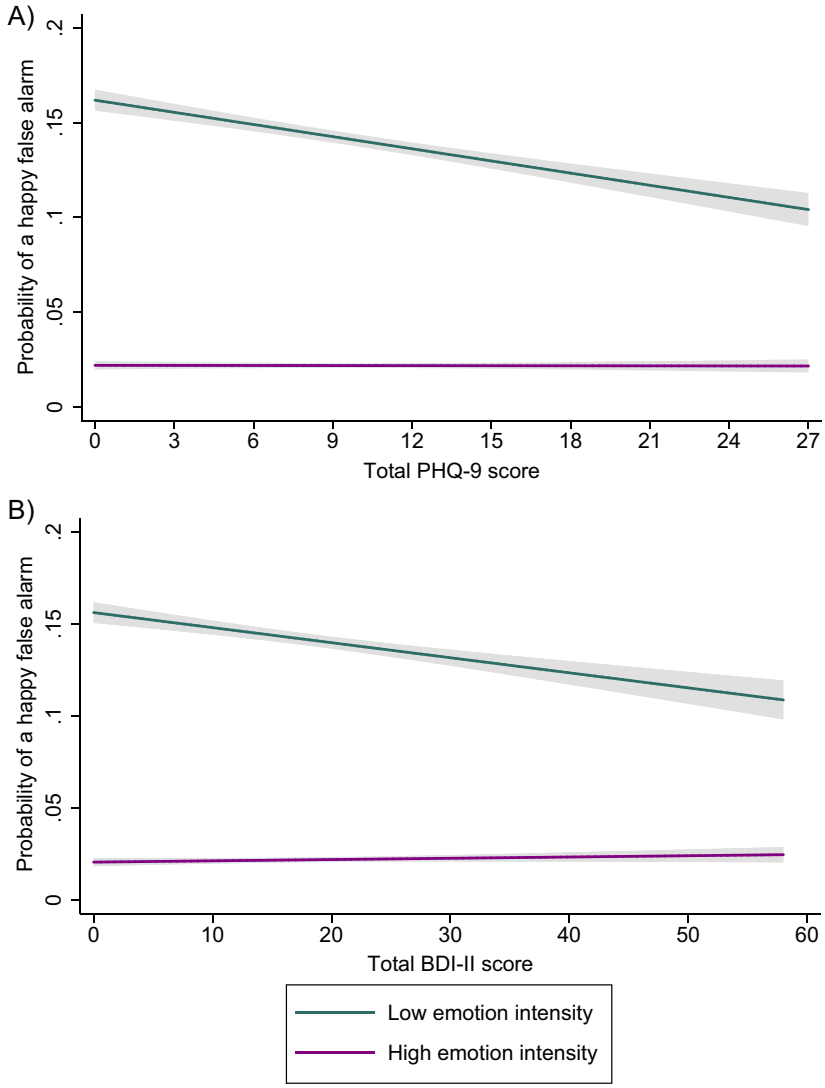

Fig. 2. Interactions between emotion intensity and depressive symptoms on the probability of making happy false alarms. The mean probability of a happy false alarm on each trial is shown for lower emotion intensities (steps 1–7 of the morphed faces) versus higher emotion intensities (steps 8–15). Shading represents the 95% confidence intervals. A) Shows depressive symptoms measured using the PHQ-9 B) shows depressive symptoms measured using the BDI-II.

reduced positive biases in interpreting ambiguous social information could be more important for concurrent depressive symptoms than increased negative biases.

The size of the effects in our study should be considered. For every additional face incorrectly classified as happy, concurrent PHQ-9 scores reduced by 0.05 of a point. To see a reduction of one point on the PHQ-9, it would require an additional 20 happy false alarms. On average, participants made 6 happy false alarms. However, the evidence for an interaction between facial expression intensity and depressive symptoms suggests that the effect would be larger for more ambiguous faces.

Longitudinally, there was no evidence for associations between recognition of happy or sad faces and subsequent depressive symptoms, suggesting that changes in facial expression recognition may not precede changes in depressive symptoms. Emotion recognition biases may not play a causal role in depression. However, this was an observational study, where changes in depressive symptoms and emotion recognition were modest and follow-up was relatively short.

4.1. Strengths and limitations

To our knowledge, this is the largest study investigating the association between emotion recognition and depression. The sample size should increase the reproducibility of our results. However, one meta-analysis indicated that a case-control study would require approximately 1230 participants to detect differences in emotion recognition (Dalili et al., 2015). Our sample may not have been adequately powered to show such small effects, although using severity of depressive symptoms continuously in analyses should have increased the power.

We sampled the full range of depressive symptom severity, and a relatively large proportion of the sample had low PHQ-9 scores (55% did not meet the diagnostic criteria for depression). Recruiting our sample from one population reduced the risk of selection bias relative to case-control studies that select cases and controls from different populations. We found no evidence that associations between happy and sad face recognition and depressive symptoms differed according to whether people currently met diagnostic criteria for depression. Future research should analyse depressive symptoms continuously as emotion recognition processes may not differ on this basis.

The prospective design and repeated measures meant our analyses took account of changes over time. Using multilevel models allowed data from all time-points to be included in the same model, increasing statistical power and precision of our estimates (Diggle, 1998). However, each time-point was two weeks apart, so the longitudinal element of the study consisted of a maximum of six weeks follow-up. Further

prospective studies of facial expression recognition and subsequent depressive symptoms should investigate whether these findings are replicated over longer follow-up periods.

The use of different intensities of facial expressions in our task make it particularly sensitive. Short presentations of ambiguous emotional faces mimicked the fleeting unclear expressions often observed in real life. Our findings suggest that the inclusion and analysis of ambiguous facial expressions in this task may be especially relevant in depression. Additionally, using a mask of visual noise ensured automatic processing of facial expressions which may involve different mechanisms to slower more effortful processing (Kahneman, 2011).

Our study has several limitations. The sample excluded people with depression who had not visited their GP, which might introduce selection biases and would have affected generalisability. We also had a low response rate making the sample unrepresentative of everyone with depressive symptoms presenting to GPs in the UK. However, the inclusion of participants did not depend on emotion recognition, so is unlikely to have biased any associations between emotion recognition and depressive symptoms. Additionally, we did not control for symptoms of anxiety in our analyses, despite the high comorbidity between anxiety and depression.

Although six emotions were included to make the task sufficiently difficult, it meant that participants saw more negative than positive facial expressions overall, which could have affected associations with depressive symptoms. We only assessed recognition of two of the six emotions presented in the facial expression recognition task. However, we had clear hypotheses for the presence of biases in happy and sad face recognition as these have most commonly been linked to depression (Bourke et al., 2010; Dalili et al., 2015; Surguladze et al., 2004). Additionally, examining two emotions lowered the probability of type I errors by reducing multiple comparisons (Curran-Everett, 2000).

A further limitation of this study was the necessity of multiple comparisons. Our primary analyses used the unbiased hit rate as the primary outcome. As planned, we then studied hits and false alarms separately in secondary analyses. The final stage of our analysis was exploratory and examined facial expression intensity. We did not take account of the multiple comparisons using a Bonferroni correction because they are often too conservative, increasing the risk of type 2 errors, particularly in exploratory analyses (Perneger, 1998; Rothman, 1990; Streiner & Norman, 2011). However, it is important to be cautious about drawing definitive conclusions given the number of comparisons we conducted and the increased probability of a chance finding, alongside the lack of evidence for our primary analysis.

We found no evidence for associations between antidepressant use and processing of happy or sad faces. This was an observational study, in which most participants had been using antidepressants over varying time periods. Even though our findings do not support the hypothesis that biases in emotion recognition play a role in antidepressant function (Harmer et al., 2009; Roiser et al., 2012), experimental studies provide much stronger evidence for associations between antidepressant use and facial emotion recognition.

4.2. Existing literature

Facial expression recognition tasks have been used in various forms with healthy individuals and people with depression or anxiety. Task performance varies across studies, likely due to different presentation timings, emotion intensities, and the outcomes tested. Compared to healthy individuals, participants in this study performed similarly in terms of hits but made more false alarms (cf. Bamford et al., 2015; Button et al., 2013b; Harmer et al., 2003a; Harmer et al., 2003b; Tranter et al., 2009). It is surprising that happy false alarms were more common in this sample than in studies with healthy individuals, given the negative association found between happy false alarms and depressive symptoms. Individuals in this study may have been worse at performing the task overall. They were, on average, approximately 25 years older than people in previous studies (Bamford et al., 2015; Button et al., 2013b). Minimal false alarm data is available from previous studies (as noted in a meta-analysis by Dalili et al., 2015), making it hard to compare our findings to other studies of participants with depressive symptoms.

Our lack of evidence for increased negative processing of facial expressions in depression is inconsistent with many smaller studies (Dalili et al., 2015; Gollan et al., 2010; Lee et al., 2016; Leppänen et al., 2004). Previous results may be due to selection bias or type 1 errors. The risk of type 1 errors is high when small samples are used (Button et al., 2013a), particularly with facial recognition tasks that are often analysed in several ways. These tasks are subject to errors in interpretation and vulnerable to multiple testing, with the tendency to select results which support the presence of cognitive biases in depression. There is also evidence of publication bias in previous research (Dalili et al., 2015). Additionally, not all studies have included or explicitly examined ambiguous faces, which may be particularly important in depression.

Our findings suggest that reduced misclassifications of ambiguous faces as happy could be a state marker of depression. Positive biases in interpreting ambiguous social information could therefore be more important for depressive symptoms than negative biases. This supports some previous evidence on cognitive biases in depression. One study found that people with depression were less likely to interpret neutral faces as happy than healthy controls, with no differences for sad faces (Gollan et al., 2010). In a longitudinal study, patients in remission showed an increased number of incorrect classifications of neutral faces as happy, indicating an increased positive bias with reduced depressive symptom severity (Leppänen et al., 2004). Additionally, our previous study of this sample assessed recall of positive and negative personality characteristics (Lewis et al., 2017). As recall of positive words increased, severity of depressive symptoms decreased but recall of negative words was not associated with depressive symptoms. In both studies, we have used a large sample and analysed depressive symptoms continuously to suggest that depression may be characterised by reduced positive rather than increased negative processing.

4.3. Implications

In line with Beck's model, reduced positive interpretations of social information relevant to the self could affect people's evaluations of social situations, reduce positive affect, and encourage social withdrawal (Beck, 2008). Our findings also fit with the idea that individuals with depression have reductions specifically in positive affect, as described in the decreased positive emotionality model of depression (Tellegen, 1985; Watson et al. 1988). With depression the leading cause of disability worldwide (World Health Organisation, 2017), understanding the mechanisms underlying depressive symptoms is important.

We would speculate that the interpretation of ambiguous social information, in this case ambiguous facial expressions, could be most relevant. Future research in this area should focus on the interpretation of ambiguous information. This is particularly important for tasks such as facial expression recognition, in which ceiling effects may occur when full intensities of emotion are presented. Neuroimaging research has attempted to elucidate the mechanisms underlying depression, but with inconsistent findings (Groenewold et al., 2013). The reduction in happy false alarms for ambiguous facial expressions could represent a target for imaging studies, allowing more detailed investigation of the cognitive mechanisms associated with depression.

Author contributions

GlL, GeL, KB, LD, CH, MM, IPV, and NW assisted with the design and completion of the original cohort study. GlL, GeL, and JB conceptualised this study. JB analysed the data and drafted the manuscript, with input from GeL and GlL. MM and IPV provided advice on

acquisition and analysis of the data. GlL provided senior supervision. All authors read and revised the whole report.

Acknowledgements

We would like to thank the participants who took part in the study, and the general practice surgeries for their help with recruitment. We thank the research team: researchers, research nurses and administrative staff. We would also like to thank the co-applicants on the PANDA programme and Professor Jonathan Roiser for his advice on our statistical analyses and an earlier version of this manuscript.

Funding

This article summarises independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research, Reference Number RP-PG-0610-10048. This study was also supported by the NIHR Biomedical Research Centre and the University Hospitals Bristol NHS Foundation Trust and the University of Bristol and University College London. CJH was supported by the NIHR Oxford Health Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funding source had no role in study design, data collection, data analysis, interpretation or writing of the report. The corresponding author had full access to all data used in the study and had final responsibility for the decision to submit for publication.

Declaration of interest

CJH has received consultancy fees from Lundbeck, Servier, Johnson and Johnson, and P1vital. The remaining authors have no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2019.06.025.

References

Alloy, L.B., Abramson, L.Y., 1979. Judgment of contingency in depressed and nondepressed students: sadder but wiser? J. Exper. Psychol. Gen. 108, 441–485.
Attwood, A.S., Easey, K.E., Dalili, M.N., Skinner, A.L., Woods, A., Crick, L., Ilett, E., Penton-Voak, I., Munafò, M.R., 2017. State anxiety and emotional face recognition in healthy volunteers. R. Soc. Open Sci. 4, 160855.
Bamford, S., Penton-voak, I., Pinkney, V., Baldwin, D.S., Munafò, M.R., Garner, M., 2015. Early effects of duloxetine on emotion recognition in healthy volunteers. J. Psychopharmacol. 29, 634–641.
Beck, A., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. Arch. Gen. Psychiatry 4, 561–571.
Beck, A.T., 1967. Depression: Clinical, Experimental, and Theoretical Aspects. Harper & Row, New York.
Beck, A.T., 1979. Cognitive therapy of depression. Guilford press.
Beck, A.T., 2008. The evolution of the cognitive model of depression and its neurobiological correlates. Am. J. Psychiatry 165, 969–977.
Bouhuys, A.L., Geerts, E., Gordijn, M.C.M., 1999. Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study. J. Nerv. Ment. Dis. 187, 595–602.
Bourke, C., Douglas, K., Porter, R., 2010. Processing of facial emotion expression in major depression: a review. Aust. N. Z. J. Psychiatry 44, 681–696.
Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013a. Confidence and precision increase with high statistical power. Nat. Rev. Neurosci. 14, 585.
Button, K.S., Lewis, G., Penton-Voak, I., Munafò, M., 2013b. Social anxiety is associated with general but not specific biases in emotion recognition. Psychiatry Res. 210, 199–207.
Colman, I., Kingsbury, M., Garad, Y., et al., 2016. Consistency in adult reporting of adverse childhood experiences. Psychol. Med. 46, 543–549.
Curran-Everett, D., 2000. Multiple comparisons: philosophies and illustrations. Am. J. Physiol. Regulat. Integr. Comp. Physiol. 279, R1–R8.
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, 1135–1144.
Diggle, P.J., 1998. Dealing with missing values in longitudinal studies. In: Everitt, B.S., Dunn, G. (Eds.), Statistical Analysis of Medical Data: New Developments. Arnold, London.
Disner, S.G., Beevers, C.G., Haigh, E.A., Beck, A.T., 2011. Neural mechanisms of the cognitive model of depression. Nat. Rev. Neurosci. 12, 467–477.
Elliott, R., Zahn, R., Deakin, J., Anderson, I., 2011. Affective cognition and its disruption in mood disorders. Neuropsychopharmacology 36, 153–182.
Gollan, J.K., McCloskey, M., Hoxha, D., Coccaro, E.F., 2010. How do depressed and healthy adults interpret nuanced facial expressions? J. Abnorm. Psychol. 119, 804.
Griffiths, S., Jarrold, C., Penton-Voak, I.S., Munafò, M.R., 2015. Feedback training induces a bias for detecting happiness or fear in facial expressions that generalises to a novel task. Psychiatry Res. 230, 951–957.
Groenewold, N.A., Opmeer, E.M., de Jonge, P., Aleman, A., Costafreda, S.G., 2013. Emotional valence modulates brain functional abnormalities in depression: evidence from a meta-analysis of fMRI studies. Neurosci. Biobehav. Rev. 37, 152–163.
Gur, R.C., Erwin, R.J., Gur, R.E., Zwil, A.S., Heimberg, C., Kraemer, H.C., 1992. Facial emotion discrimination: II. Behavioral findings in depression. Psychiatry Res. 42, 241–251.
Hale, W.W., 1998. Judgment of facial expressions and depression persistence. Psychiatry Res. 80, 265–274.
Hankin, B.L., Fraley, R.C., Lahey, B.B., Waldman, I.D., 2005. Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. J. Abnorm. Psychol. 114, 96–110.
Harmer, C.J., Bhagwagar, Z., Perrett, D.I., Völlm, B.A., Cowen, P.J., Goodwin, G.M., 2003a. Acute SSRI administration affects the processing of social cues in healthy volunteers. Neuropsychopharmacol. 28, 148–152.
Harmer, C.J., Hill, S.A., Taylor, M.J., Cowen, P.J., Goodwin, G.M., 2003b. Toward a neuropsychological theory of antidepressant drug action: increase in positive emotional bias after potentiation of norepinephrine activity. Am. J. Psychiatry 160, 990–992.
Harmer, C.J., Goodwin, G.M., Cowen, P.J., 2009. Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. Br. J. Psychiatry 195, 102–108.
Kahneman, D., 2011. Thinking Fast and Slow. Farrar, Straus and Giroux, New York.
Kohler, C.G., Hoffman, L.J., Eastman, L.B., Healey, K., Moberg, P.J., 2011. Facial emotion perception in depression and bipolar disorder: a quantitative review. Psychiatry Res 188, 303–309.
Kounali, D.Z., Button, K.S., Lewis, G., Ades, A.E., 2016. The relative responsiveness of test instruments can be estimated using a meta-analytic approach: an illustration with treatments for depression. J. Clin. Epidemiol. 77, 68–77.
Lee, J.S., Mathews, A., Shergill, S., Yiend, J., 2016. Magnitude of negative interpretation bias depends on severity of depression. Behav. Res. Ther. 83, 26–34.
Leppänen, J.M., Milders, M., Bell, J.S., Terriere, E., Hietanen, J.K., 2004. Depression biases the recognition of emotionally neutral faces. Psychiatry Res. 128, 123–133.
Lewis, G., Kounali, D.Z., Button, K.S., Duffy, L., Wiles, N.J., Munafò, M.R., Harmer, C.J., 2017. Variation in the recall of socially rewarding information and depressive symptom severity: a prospective cohort study. Acta Psychiatr. Scand. 135, 489–498.
Lewis, G., Pelosi, A.J., Araya, R., Dunn, G., 1992. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. Psychol. Med. 22, 465–486.
Milders, M., Bell, S., Platt, J., Serrano, R., Runcie, O., 2010. Stable expression recognition abnormalities in unipolar depression. Psychiatry Res. 179, 38–42.
Moore, M.T., Fresco, D.M., 2012. Depressive realism: a meta-analytic review. Clin. Psychol. Rev. 32, 496–509.
Münkler, P., Rothkirch, M., Dalati, Y., Schmack, K., Sterzer, P., 2015. Biased recognition of facial affect in patients with major depressive disorder reflects clinical state. PLoS ONE 10, e0129863.
Perneger, T.V., 1998. What's wrong with Bonferroni adjustments. BMJ 316, 1236–1238.
Roiser, J.P., Elliott, R., Sahakian, B.J., 2012. Cognitive mechanisms of treatment in depression. Neuropsychopharmacology 37, 117–136.
Rothman, K.J., 1990. No adjustments are needed for multiple comparisons. Epidemiol. 1, 43–46.
Ruhe, H.G., Mocking, R.J., Figueroa, C.A., Seeverens, P.W., Ikani, N., Tyborowska, A., Browning, M., Vrijsen, J.N., Harmer, C.J., Schene, A.H., 2019. Emotional biases and recurrence in Major Depressive Disorder. Results of 2.5 years follow-up of drug-free cohort vulnerable for recurrence. Frontiers in Psychiatry 10, 145.
Schulz, K.F., Grimes, D.A., 2002. Case-control studies: research in reverse. Lancet 359, 431–434.
Spitzer, R., Kroenke, K., Williams, J., Patient Health Questionnaire Primary Care Study Group, 1999. Validation and utility of a self-report version of PRIME-MD: the patient health questionnaire primary care study. JAMA 282, 1737–1744.
Streiner, D.L., Norman, G.R., 2011. Correction for multiple testing: is there a resolution? Chest 140, 16–18.
Surguladze, S.A., Young, A.W., Senior, C., Brébion, 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, 212–218.
Tellegen, A., 1985. Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In: Tuma, A.H., Maser, J.D. (Eds.), Anxiety and the anxiety disorders. Edbaum, Hillsdale, NJ, pp. 681–706.
Tranter, R., Bell, D., Gutting, P., Harmer, C., Healy, D., Anderson, I.M., 2009. The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. J. Affect Disord. 118, 87–93.
Vrijen, C., Hartman, C.A., Oldehinkel, A.J., 2016. Slow identification of facial happiness in early adolescence predicts onset of depression during 8 years of follow-up. Eur. Child Adolesc. Psychiatry 25, 1255–1266.
Wagner, H., 1993. On measuring performance in category judgment studies of nonverbal behaviour. J. Nonverbal. Behav. 17, 3–28.
Watson, D., Clark, L.A., Carey, G., 1988. Positive and negative affectivity and their relation to anxiety and depressive disorders. J. Abnorm. Psychol. 97, 346–353.
World Health Organization, 2017. Depression and other common mental disorders: global health estimates. World Health Organization, pp. 1–24.
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.