PHD

A neurocognitive investigation of the role of reinforcement learning in updating dysfunctional self-schema in depression: A putative mechanism for antidepressant action?

(Alternative Format Thesis)

Hobbs, Catherine

Award date:
2022

Awarding institution:
University of Bath

PHD

A neurocognitive investigation of the role of reinforcement learning in updating dysfunctional self-schema in depression: A putative mechanism for antidepressant action?

(Alternative Format Thesis)

Hobbs, Catherine

Award date:
2022

Awarding institution:
University of Bath

A neurocognitive investigation of the role of reinforcement learning in updating dysfunctional self-schema in depression: A putative mechanism for antidepressant action?

Catherine Hobbs

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Psychology

February 2022

Copyright Notice

Declaration of any previous submission of the work

The material presented here for examination for the award of a higher degree by research has not been [redacted] for another degree.

[redacted]

Declaration of authorship

I am the author of this thesis, and the work described therein was carried out by myself personally. Individual chapters are accompanied by an authorship statement indicating my contribution to the work.

[redacted]

Conflict of Interest Statement

The candidate declares no conflict of interest.

Table of Contents

Acknowledgements

I would firstly like to thank my funder, the GW4 Biomed MRC DTP, who have been hugely supportive in providing flexible funding supplements that enabled two additional chapters in this thesis outside of my core funding. This thesis was supported in part by grant MR/N0137941/1 for the GW4 BIOMED MRC DTP, awarded to the Universities of Bath, Bristol, Cardiff and Exeter from the Medical Research Council (MRC)/UKRI.

I would also like to say a huge thank you to my supervisory team. To my primary supervisor, Katherine Button, thank you for pushing me to complete this PhD to the best of my ability even when I was faltering and for all the helpful career chats. To Jie Sui, thank you for providing a cognitive perspective to my work. To Marcus Munafò, thank you for your careful reading of my drafts. And finally, to David Kessler, thank you for providing me with medical supervision and for your considerate approach to supporting me in tackling difficult conversations.

I am also very grateful for the support of Susannah Murphy and Catherine Harmer who hosted me for a placement at the Psychopharmacology and Emotion Research Laboratory at the Department of Psychiatry, University of Oxford. The study that I completed under their supervision was integral to this PhD. I would also like to thank Janina Hoffman, Michael Moutoussis, and Julian Faraway for their support with statistical analyses.

Thank you as well to the research support team at the Department of Psychology, University of Bath. I would particularly like to thank Caroline Ransford for her help sourcing equipment, and Susie Martin, for all her technical support within the psychology labs. Thank you to the wellbeing service at the University of Bath, particularly Sam Williams whose support helped me finish this thesis in the final few months.

I would also like to thank all the participants that have taken part in my research. This thesis would not have been possible without their help.

Finally, I would like to express my love and gratitude to my family. Thank you to my husband James, who moved across the country to allow me to complete this PhD. I truly would not have been able to complete this thesis without his kind words cheering me up when I was on the verge of quitting. I am eternally grateful for his unfaltering love and support. Thank you to my dog Harry for always forcing me to take a break and walk in fresh air. He has been the bright spot throughout all this work, and I don't know how I would have managed without having him available for a cuddle. Thank you to my parents for all their

emotional and financial support throughout my extended time at university. They have provided me with a template for hard work throughout their lives. Finally, thank you to my grandparents, particularly my Nanny Kate who always asks if ‘I have finished university yet’ – with the submission of this thesis I’m proud to say I almost have.

Abstract

Depression is one of the leading mental health problems experienced worldwide. Whilst treatments are available for depression, individual treatment response is varied. The intersection between self, emotion, and reward, referred to in combination as self-referential affective processing, may be a key cognitive vulnerability in depression and a sensitive target for intervention. I aimed to: (1) evaluate whether depression is associated with differences in processing of self-related information, and whether this varies depending on the emotional valence or rewarding nature of information; (2) understand whether antidepressants are associated with change in self-referential affective processing.

In chapter 1, I addressed my first aim by conducting a cross-sectional study where adults experiencing varying depression (n = 144) completed cognitive tasks measuring self, emotion, and reward processing, occurring independently and in combination. Depression was most reliably associated with self-referential processing occurring in interaction with reward and emotion. Participants with greater depression were worse at learning positive versus negative social evaluations about the self in a reinforcement learning task ($\beta = 0.13$, 95% CI: 0.06, 0.20, $p < .001$). In chapter 2, I validated these findings using an individual participant dataset collated from studies previously conducted within this research group (n = 552) and a second independent dataset of participants recruited online (n = 807). Replicating chapter 1, participants with greater depression were again worse at learning positive versus negative social evaluations about the self (dataset 1: $\beta = 0.27$, 95% CI: 0.20, 0.35, $p < .001$; dataset 2: $\beta = 0.10$, 95% CI: 0.02, 0.17, $p = 0.009$). Treatments that increase learning of positive social evaluations about the self may therefore be helpful in addressing depressive symptoms.

In the second half of my thesis, I subsequently investigated the effect of antidepressants on self-referential affective processing. In chapter 3 I conducted a systematic review, narrative synthesis (k = 82), and meta-analysis (k = 28) of behavioural evidence for change in emotional processing following antidepressant administration. I did not find reliable evidence that current behavioural measures of emotional processing were altered following antidepressant administration (positive: SMD = 0.08, 95% CI = -0.05, 0.22, $p = 0.230$; negative: SMD = -0.04, 95% CI = -0.12, 0.04, $p = 0.341$). In chapters 4 and 5, I therefore investigated whether change in self-referential affective processing may be a more sensitive

measure of early antidepressant action. I did not find evidence that learning social evaluations about the self differed in healthy volunteers randomised to acute citalopram versus placebo (chapter 4: β = 1.95, -2.25, 6.16), or that this was associated with change in depression over the first eight weeks of antidepressant treatment in a prospective observational cohort of primary care patients (chapter 5: β = 0.15, 95% CI: -0.10, 0.39, p = 0.239). However, acute administration of an antidepressant increased positive social behaviours (chapter 4; β = 20%, 95% CI: 2%, 37%, p = 0.030). Additionally, exploratory analyses also indicated that antidepressants increased positive affective biases towards familiar others (chapter 4: β = 4.06, 95% CI: 0.88, 7.24), which weakly predicted a reduction in a secondary measure of depression, the BDI-II, in primary care patients (chapter 5: β = 0.24, 95% CI: -0.01, 0.49, p = 0.069). Exploratory analyses in chapter 5 also indicated stronger evidence that increased learning of positive versus negative evaluations about both the self (β = 0.34, 95% CI: 0.14, 0.55, p = 0.002) and a friend (β = 0.36, 95% CI: 0.16, 0.56, p = 0.001) were associated with a reduction in anxiety.

Due to weak evidence for the role of self-referential affective processing in antidepressant treatment found in chapters 4-5, as well as chapter 3 highlighting the need for novel psychological measures of antidepressant effects, in chapter 6 I evaluated a different area of cognition: optimistic belief updating. Whereas healthy participants updated their beliefs regarding negative life events more following good versus bad news, individuals experiencing depression lacked this bias (β = 0.71, β 95% CI: 0.24, 1.18, p = 0.004). In line with my previous findings of reduced learning of positive social evaluations, treatments attempting to increase the accommodation of positive information may be beneficial in remediating negative beliefs associated with depression.

Overall, whilst depression was reliably associated with reduced self-referential affective biases, this did not appear to be important in early antidepressant treatment. However, I identified potential areas of interest for future research on psychopharmacological antidepressant effects. Firstly, I found evidence that antidepressants may operate by increasing sensitivity to positive information about familiar others and increasing positive social behaviours. This may treat depression by increasing enjoyment and engagement in social interactions. Secondly, exploratory analyses suggested that change in affective biases may be important in remediating anxiety rather than depression. It is therefore possible that antidepressants operate in part by remediating threat-related biases that maintain anxiety symptoms.

General Introduction

Depression

Depression is a relatively heterogenous disorder, characterised by a wide range of symptoms (Fried & Nesse, 2015). According to DSM-V diagnostic criteria, core symptoms of major depressive disorder include a depressed mood (e.g., feeling sad, empty, or hopeless) and anhedonia – a loss of interest or pleasure in almost all activities. Other symptoms include a significant change in weight or appetite, disruptions to sleep, change in speed of speech or movement, fatigue, difficulties with concentration, and suicidal ideation. To meet diagnostic criteria individuals must be experiencing five or more symptoms, at least one of which is a core symptom, for the majority of time over a continuous 2-week period (American Psychiatric Association, 2013).

Depression is the leading mental health problem worldwide (Vos et al., 2015), affecting 4.4% of the global population at a given point (World Health Organization, 2017). A large proportion of individuals also experience depression and anxiety symptoms at a subclinical level (Macrory, 2016). Depression is the main cause of disability globally (World Health Organization, 2017) and is projected to be the principal cause of disease burden worldwide by 2030 (World Health Organization, 2011). Experiencing depression places individuals at increased risk of experiencing a chronic physical disorder (Scott et al., 2007), and a reduced life expectancy (Laursen, Musliner, Benros, Vestergaard, & Munk-Olsen, 2016).

Whilst effective treatments are available for depression (Churchill et al., 2002; Cipriani et al., 2018), individual response varies (Warden, Rush, Trivedi, Fava, & Wisniewski, 2007). Understanding the cognitive processes maintaining depression may allow us to improve the effectiveness of current treatments by identifying sensitive targets for intervention. In this thesis I examine the role of self, emotion, and rewarding processing, referred to in combination as self-referential affective processing, as a potential source of cognitive biases maintaining negative self-schema and depressive symptoms. Additionally, I investigate self-referential affective processing as a putative mechanism for early antidepressant action.

Self-Referential Affective Processing in Depression

Self-Referential Processing

Across the general population information related to the self is preferentially processed (Cunningham & Turk, 2017). People show greater recall, attention, and learning for

information about the self (Bargh, 1982; Sui & Humphreys, 2015; Sui, He, & Humphreys, 2012; Symons & Johnson, 1997). Reflecting the specialised nature of the self, a network in the brain has been associated with processing of self-related information. Activity in the default mode network, comprised of the medial prefrontal cortex, posterior cingulate cortex, and inferior parietal lobe, is heightened when thinking about the self (Davey, Pujol, & Harrison, 2016). Self-prioritisation effects are believed to be adaptive in ensuring information integral to pursuit of future goals and potential threat to the self is attended to (Cunningham & Turk, 2017). Additionally, due to the strength and consistency of self-referential biases, it has been suggested that the self may integrate information processing across cognitive domains, leading to greater elaboration of self-relevant information (Sui & Humphreys, 2015). Self-referential processing is therefore believed to be an integral cognitive domain across the population.

Self-prioritisation effects are also believed to play a key role in depression (Northoff, 2007). Individuals experiencing depression show a heightened focus on the self and increased cognitive processing of self-related stimuli (Northoff, 2007). Behaviourally, this is demonstrable in an increase in the use of first-person singular pronouns (Edwards & Holtzman, 2017). Neurologically, individuals with depression show a hyperconnectivity within the default mode network (Whitfield-Gabrieli & Ford, 2012), particularly the ventral medial prefrontal cortex, considered the 'home' of internal self-representations (Lemogne, Delaveau, Freton, Guionnet, & Fossati, 2012). Whereas healthy participants demonstrate supressed activity in the default mode network during cognitively demanding tasks, individuals with depression show a hyperactivity (Sheline et al., 2009). Depression is therefore believed to be associated with an impaired ability to shift attention from an internal self-referential focus to the external environment (Sheline et al., 2009). This may limit the ability of individuals with depression to update representations of the self in response to environmental feedback, perpetuating maladaptive views of the self. Attempts to understand the causal and maintenance factors of depression would therefore benefit from understanding the precise role of self-processing (Davey & Harrison, 2022).

Emotional Processing

Abnormalities in self-processing in individuals experiencing depression are thought to differ depending on the emotional valence of incoming information, with effects being heightened for negative relative to positive self-referential stimuli. The cognitive theory of depression argues that due to maladaptive early experiences, individuals with depression hold heavily

entrenched negative sets of beliefs and views of the self, termed negative self-schema. When activated by social stressors in later life, negative self-schema promote automatic processing of negative information about the self. Additionally, deliberative cognitive processes are weakened, preventing re-evaluation of these automatic negative self-referential biases. Over time repeated activation of negative self-schema leads to lower thresholds for future activation, leading negative self-referential processing to become the dominant viewpoint (Beck, 1976, 2008; Beck & Dozois, 2011). More recently, the cognitive neuropsychological model has reformulated this theory but has continued to argue for the causal role of negative self-schema in core depressive symptoms. Rather than negative self-schema being caused by early negative experiences, this theory proposes that negative affective biases occur due to alterations in monoamine transmission. Over time, repeated automatic integration of negative information with the self because of negative affective biases, leads to the development of negative self-schema. This in turn, reinforces negative automatic processing in a cyclical process (Roiser, Elliott, & Sahakian, 2012).

Empirical support for the presence and importance of negative self-schema in depression has been found. Following negative life events, individuals with greater levels of depression report more tightly interconnected negative perceptions of the self (Disner, Shumake, & Beevers, 2017). Negative self-schema also precede and predict the onset of depression, suggesting a possible causal role (Evans, Heron, Lewis, Araya, & Wolke, 2005). Cognitive Behavioural Therapy (CBT), developed based on the cognitive theory of depression, has been found to be effective at treating depression, suggesting a possible mechanistic role of negative self-schema (Butler, Chapman, Forman, & Beck, 2006). Addressing negative self-schema is therefore believed to be central to successful treatment of depression.

Evidence has also been found for the importance of negative self-referential affective biases in depression. Individuals with depression show greater recall of negative information (Dalgleish & Watts, 1990), enhanced attention towards negative stimuli (Peckham, McHugh, & Otto, 2010), and are more likely to interpret ambiguous stimuli negatively (Everaert, Podina, & Koster, 2017). Notably, these effects are enhanced for self-relevant information. Greater recall for negative stimuli, and reduced recall for positive stimuli, is more likely to be observed in depression when stimuli is encoded in reference to the self (Gaddy & Ingram, 2014). Furthermore, increased recall of positive words about the self has been found to predict reductions in depression, suggesting a possible causal role (Lewis et al., 2017). Higher levels of depression have also been linked to greater attentional avoidance of positive

information, but only when information is processed in reference to the self rather than others (Ji, Grafton, & MacLeod, 2017). Similarly, negative attentional biases in depression are believed to be heightened for self-related information (Mogg & Bradley, 2005). Interpretation biases have also only been observed when individuals experiencing low mood were primed to think about themselves versus others (Hertel & El-Messidi, 2006). Meta-analysis has confirmed that studies using self-relevant information report greater effects sizes for the presence of interpretation biases in depression (Everaert et al., 2017). The combination of a heightened focus on the self and negatively biased emotional processing may therefore be a key vulnerability for depression.

Reward Processing

In addition to biased emotional processing, depression has also been associated with differences in reward processing. An increasing body of evidence indicates that individuals with depression show a reduced sensitivity to rewarding information (Eshel & Roiser, 2010; Halahakoon et al., 2020). Individuals experiencing depression show impaired learning from positive feedback in probabilistic learning tasks (Huys, Pizzagalli, Bogdan, & Dayan, 2013), and are less willing to expend cognitive or physical effort for reward (Horne, Topp, & Quigley, 2021). Functional magnetic resonance imaging (fMRI) studies have also indicated that individuals with depression show reduced activation in the striatum, an area of the brain critical to reward systems, when anticipating reward or receiving rewarding feedback (Keren et al., 2018). Similarly, depression has been associated with blunted neural responses following reward in event related potential (ERP) studies (Brush, Ehmann, Hajcak, Selby, & Alderman, 2018). Research has also reported some evidence of an increased sensitivity to punishing information in depression (Eshel & Roiser, 2010), although impairments appear to be more reliably observed for processing of reward versus punishment (McFarland & Klein, 2009; Mukherjee, Filipowicz, Vo, Satterthwaite, & Kable, 2020). Differences in reward processing are believed to underly a core symptom of depression, anhedonia – a reduced interest or pleasure in activities previously enjoyed (Kaya & McCabe, 2019). Clarifying the conditions in which abnormalities in reward processing are observed may therefore allow us to develop more targeted interventions to address such core symptoms of depression.

Reward processing is thought to be modulated by the self-relevance of information (Northoff & Hayes, 2011). Self-relevant information induces neural activity in areas of the brain activated during reward processing, including the ventral medial prefrontal cortex, ventral striatum, and ventral tegmental area (Northoff & Hayes, 2011). Preferential processing of

self-related information may be driven by the self being intrinsically rewarding (Northoff & Hayes, 2011). However, this may be weakened in depression, where the self holds negative rather than positive associations (Woolfolk, Novalany, Gara, Allen, & Polino, 1995). In keeping with this possibility, patients with depression show reduced activation of both reward and self-related areas of the brain in response to positive stimuli (Northoff, 2007; Northoff & Hayes, 2011). At a behavioural level, depression is associated with difficulties in updating self-beliefs in a positive direction based on reward. In a study examining probabilistic reversal learning, participants experiencing depression initially successfully learnt to engage in negative self-referential thoughts based on rewarding feedback. However, when this rule was reversed so that engaging in negative self-reference was associated with a loss, participants experiencing depression showed an impaired ability to update their behaviour (Iijima, Takano, Boddez, Raes, & Tanno, 2017). Similarly, participants with depression continue to choose to engage in negative versus positive memories even when this response results in economic loss (Takano, Van Grieken, & Raes, 2019). Engaging with negative self-referential information may therefore be rewarding to individuals with depression. Hyposensitivity to reward may cause an inability to integrate positive information with the self, resulting in reduced cognitive biases towards positive information observed in depression. Likewise, increased sensitivity to punishing feedback may enhance preferential processing of negative self-referential feedback, reinforcing negative views of the self. Self-referential affective processing, combining self, emotion, and reward, may therefore represent a key abnormality in information processing underlying depression.

Social Environment

The relationship between self, emotion, and reward processing is especially relevant for social stimuli. The self is a social construct, shaped by our interpretations of others' beliefs about us (Shrauger & Schoeneman, 1979). Whereas healthy individuals tend to find social stimuli rewarding (Bhanji & Delgado, 2014) and preferentially incorporate positive social evaluations into their self-concept (Korn, Prehn, Park, Walter, & Heekeren, 2012), individuals with depression report reduced reward from social interactions (Barkus & Badcock, 2019) and show a greater tendency to engage in negative social feedback. Individuals with greater depression are more likely to believe that others will evaluate them negatively (Moritz & Roberts, 2018), preferentially seek self-verifying negative feedback (Giesler, Josephs, & Swann, 1996; Hames, Hagan, & Joiner, 2013), rate social interactions more negatively (Hoehn-Hyde, Schlottmann, & Rush, 1982; Setterfield, Walsh, Frey, &

McCabe, 2016), and are more sensitive to social rejection (Fossati, Hinfray, Fall, Lemogne, & Rotge, 2019). At a behavioural level, individuals with depression are less likely to demonstrate prosocial behaviours and are more likely to demonstrate hostile or withdrawn behaviours (Girard et al., 2014; Hames et al., 2013; Pulcu et al., 2015). Additionally, recent research has indicated that depression is associated with impaired learning from social feedback, which is linked to social anhedonia (Frey & McCabe, 2020a; Frey, Frank, & McCabe, 2021). Being able to accurately interpret and learn from social feedback is important in maintaining successful social interactions and developing positive interpersonal relationships (Weightman, Air, & Baune, 2014). Inability to learn from social reward in depression may also contribute to a reduced ability to incorporate positive information into the self-concept. Simultaneously, heightened sensitivity towards negative social stimuli is likely to reinforce negative concepts of the self. Repeated exposure to negative social experiences as a result of self-referential affective biases may also evoke socially withdrawn behaviours, increasing the likelihood of reciprocal negative behaviours from others, and reinforcing negative expectations of social interactions in a vicious cycle (Lewinsohn, 1975). Addressing self-referential affective learning biases occurring within social contexts may be important when treating depression.

Self-Referential Affective Processing: A Potential Mechanism for Early Antidepressant Action?

As I have outlined above, self-referential affective biases may be a key factor maintaining depression. Treatments that can address self-referential affective biases may be beneficial in addressing negative self-schema and reducing depressive symptoms. In this PhD I explore the role of antidepressants in remediating self-referential affective biases associated with depression.

Antidepressant Treatment

Within the UK most patients experiencing depression are managed within primary care settings using antidepressants (Kendrick et al., 2009; McManus, Bebbington, Jenkins, & Brugha, 2016). National Institute for Health and Care Excellence (NICE) guidelines recommend antidepressants for treating moderate to severe depression, or those with long-term subthreshold depressive symptoms (NICE, 2009). Antidepressant prescriptions have risen over time, with seventy-two million antidepressants being prescribed in 2019, a 25% increase since 2015 (Lalji, McGrogan, & Bailey, 2021).

Most antidepressants currently prescribed were developed based on the monoamine hypothesis, that reduced levels of the monoamines serotonin and norepinephrine are causally implicated in depression (Hirschfeld, 2000). Antidepressants are designed to increase levels of these neurotransmitters, although classes of antidepressants differ in their specific mechanisms of action.

Monoamine oxidase inhibitors (MAOIs) were the first class of antidepressants developed. However, they are no longer typically used in clinical practice due to potentially serious side effects and drug interactions (Ciraulo, Shader & Greenblatt, 2011). In 2018 MAOIs accounted for only 0.1% of antidepressant prescriptions (Bogowicz, Curtis, Walker, Cowen, Geddes & Goldacre, 2021). MAOIs operate by inhibiting monoamine oxidase enzymes which metabolize monoamines including serotonin and norepinephrine, preventing their breakdown (Ciraulo, Shader & Greenblatt, 2011).

The most common classes of antidepressants currently prescribed are Selective Serotonin Reuptake Inhibitors (SSRIs), Serotonin-Norepinephrine Reuptake Inhibitors (SNRIs), Tricyclic Antidepressants (TCAs), and Noradrenaline and specific serotonergic antidepressants (NASSAs). SSRIs such as citalopram, sertraline, and fluoxetine, are recommended as the first-line antidepressant for adults in the UK, accounting for 53.9% of prescriptions (Bogowicz et al, 2021). SSRIs operate by blocking the reuptake of 5-HT (5-hydroxytryptamine; serotonin) into presynaptic terminals, increasing levels of serotonin (Ciraulo, Shader & Greenblatt, 2011). SNRIs, such as venlafaxine and duloxetine, are offered as an alternative to SSRIs, making up 9.5% of antidepressant prescriptions (Bogowicz et al, 2021). SNRIs increase levels of both serotonin and norepinephrine by blocking their reuptake (Ciraulo, Shader & Greenblatt, 2011). TCAs also operate by preventing the reuptake of serotonin and norepinephrine. However, TCAs block additional receptors such as histamine (H1) and muscarinic (M1) receptors, contributing to more severe side effects than SSRIs and SNRIs (Ciraulo, Shader & Greenblatt, 2011). Despite this, TCAs account for 21.9% of antidepressant prescriptions (Bogowicz et al, 2021). NASSAs, such as Mirtazapine, are another class of antidepressants offered as an alternative to SSRIs. Prescriptions of NASSAs has increased over time; Mirtazapine accounted for only 0.5% of prescriptions in 1998 versus 12.6% in 2018 (Bogowicz et al, 2021). NASSAs block specific serotonin receptors including 5-$HT_{2A,}$ 5-HT_{2c} and 5-HT_3 to increase levels of serotonin. Additionally, NASSAs also increase levels of norepinephrine by blocking α_2 presynaptic autoreceptors (Ciraulo, Shader & Greenblatt, 2011).

Most recently an isomer of ketamine, esketamine, was licensed as a therapy for treatment-resistant depression to be used in conjunction with SSRIs and SNRIs. However, at the time of writing it is not recommended for treatment in the UK due to concerns over its cost-effectiveness and efficacy compared to currently available treatments (Mahase, 2020). Esketamine is a N-methyl-D-aspartate (NDMA) receptor antagonist, although the exact mechanisms of action on depression are currently debated (Miller, 2020).

Antidepressants are generally regarded to be effective at treating depression, with moderate effects on reducing depression symptoms (Cipriani et al., 2018, although see Hengartner & Plöder, 2018; McCormack & Korownyk, 2018; Munkholm, Paludan-Müller, & Boesen, 2019 for alternative viewpoints). However, whilst antidepressants seem to have some effect at a group level, treatment response appears to vary within individuals. In the largest study of sequenced treatment for depression conducted to date (the STAR*D trial), two-thirds of patients did not remit after 12 weeks of initial treatment with citalopram (Warden et al., 2007). Ultimately, approximately 30% of patients did not remit following 4 stages of treatment (Warden et al., 2007). Antidepressant selection at a patient level has therefore been described as imprecise, following a trial-and-error approach (Chekroud & Krystal, 2015).

Exacerbating this issue is the delay in therapeutic action of antidepressants. Whilst some reductions in depression symptoms are apparent from 2 weeks after beginning treatment, the eventual outcome of antidepressant response appears to be unrelated to early changes in mood (De Vries et al., 2019). Current NICE guidelines recommend altering treatment if there is inadequate response only after three to four weeks (2009). Patients are therefore currently exposed to long periods of potentially ineffective treatment before an appropriate treatment can be found. As antidepressants are associated with a number of serious side effects, including an increased risk of suicide in early treatment (Healy & Whitaker, 2003; Tiihonen et al., 2006), reducing unnecessary treatment before therapeutic response is essential. Identifying mechanisms of antidepressant action may allow identification of early indicators of treatment response, allowing clinicians to predict whether patients are likely to benefit from an antidepressant at an earlier timepoint and subsequently minimise unnecessary treatment.

Change in affective processing as a mechanism of antidepressant action

One potential psychopharmacological predictor of antidepressant response that has generated considerable research is change in affective processing. The cognitive neuropsychological

model proposes that rather than influencing mood directly, antidepressants operate by remediating negative affective biases associated with depression, increasing sensitivity to positive relative to negative information. It is believed that this occurs early in treatment, but accounting for the delay in the therapeutic action of antidepressants, an improvement in mood is produced by individuals interacting with their social environment with these remediated affective biases. Repeated interactions over time with increased positive affective biases allows patients to relearn environmental associations from a more positive perspective, subsequently improving mood (Godlewska & Harmer, 2021; Harmer, Hill, Taylor, Cowen, & Goodwin, 2003; Roiser et al., 2012). Change in affective biases may therefore act as an early biomarker of treatment response, allowing patients' treatment to be altered earlier than currently possible, if no change in affective processing is observed.

Consistent with this theory, evidence for change in neural activity underlying affective processing has been reported following antidepressant administration. A meta-analysis of nine fMRI and positron emission tomography (PET) studies found that neural activity in response to emotional stimuli in brain regions associated with major depression were normalised following antidepressant treatment (Delaveau et al., 2011). Activity in the amygdala, hippocampus, and orbitofrontal cortex was reduced in response to negative stimuli, whereas activity in the fusiform gyrus and dorsolateral prefrontal cortex was increased in response to positive stimuli (Delaveau et al., 2011). The authors interpreted these findings as suggesting that antidepressants may decrease sensitivity to negative environmental information and facilitate positive emotional processing (Delaveau et al., 2011). Similar conclusions were drawn by a later meta-analysis examining sixty fMRI studies on the effect of antidepressants on brain activity underlying emotion and reward processing (Ma, 2015).

To date, a systematic review or meta-analysis on behavioural evidence of change in affective processing has not been conducted. However, a narrative review examining the influence of serotonin manipulations (including both depletion and augmentation) on emotional processing reported evidence of change in attention, facial emotional recognition, memory, attitudes, and decision making consistent with the cognitive neuropsychological model (Merens, Willem Van der Does, & Spinhoven, 2007). Although, the relationship of change in emotional processing with improvement in mood was unclear (Merens et al., 2007). Other narrative reviews have identified evidence of change in a range of affective biases including attention, recall, and facial emotion recognition, following both acute and longer term

administration of antidepressants in healthy and clinical samples (Godlewska, 2019; Godlewska & Harmer, 2021; Pringle, Browning, Cowen, & Harmer, 2011; Warren, Pringle, & Harmer, 2015). However, recently the reliability of findings across studies has been questioned (Ahmed et al., 2021). Additionally, studies have been criticised for lacking ecological validity (Ahmed et al., 2021). Most studies have been conducted on short-term doses in healthy volunteers under laboratory conditions. Whilst some studies have recruited individuals experiencing depression, they are typically recruited based on meeting diagnostic criteria. As most individuals receiving antidepressant treatment in primary care typically experience more mild forms of depression (Kendrick et al., 2009), it is possible that effects reported to date are unrepresentative of patients in clinical practice.

To date, a small number of studies have examined change in affective processing in relation to depression symptoms in primary care patients receiving antidepressant treatment. In a prospective cohort study, an increase in the recognition of happiness two weeks after commencing treatment was related to an improvement in well-being at six weeks (Tranter et al., 2009). Similarly, in a study of veterans aged 55 years and older researchers found weak evidence that increased recognition of facial emotions from baseline to one week of antidepressant treatment predicted remission at eight weeks (Shiroma, Thuras, Johns, & Lim, 2014). Researchers have also developed a predictive algorithm based on change in facial emotion recognition and depression severity from baseline to one week of treatment. The algorithm predicted response to citalopram with 77% accuracy (Browning et al., 2019). However, attempts to tailor treatment based on this algorithm in primary care had limited effectiveness. In a randomised controlled-trial, patients assigned to the algorithm arm versus treatment as usual showed no differences in reduction of depressive symptoms at 8 weeks, although a greater reduction in anxiety was found (Browning et al., 2021). Subsequently, whilst there is some promising evidence that change in affective processing may be associated with treatment response, findings are somewhat preliminary at present.

Self-referential processing

As outlined previously there is extensive evidence that affective biases associated with depression are heightened when information is self-relevant. Research has also indicated that antidepressants normalise activity in the default mode network, an area of the brain associated with self-referential processing (Posner et al., 2013; Van Wingen et al., 2014). Increasing positive affective biases through antidepressant treatment may increase integration of positive information with the self, addressing negative self-schema. Change in self-

referential affective biases may therefore be a particularly sensitive measure of early antidepressant mechanisms. However, there is a currently a limited understanding of the importance of self-referential affective biases in antidepressants. Whilst antidepressants have been found to increase endorsement and recall of positive characteristics about the self (Harmer et al., 2009; Harmer, Shelley, Cowen, & Goodwin, 2004), the cognitive tasks employed to measure affective processing related to the self do not typically include a control referential condition, limiting conclusions as to the specificity of effects to the self. Further research is required to understand the relative importance of self-referential affective biases in antidepressant action and as a potential indicator of treatment response.

Social Environment

The cognitive neuropsychological model suggests that interaction with the social environment with remediated affective biases is integral to antidepressant response (Godlewska & Harmer, 2021). To date, there is little research directly examining this aspect of the model. However, a large body of evidence has documented a positive relationship between levels of serotonin and sensitivity to social cues (Kiser, SteemerS, Branchi, & Homberg, 2012). For example, in a recent study depletion of serotonin was found to impair learning from social reward (Frey & McCabe, 2020b). However, the influence of serotonin on processing of social cues is thought to depend in part on perceptions of the social environment (Kiser et al., 2012). In a positive social environment, higher levels of serotonin are thought to produce favourable outcomes such as a greater empathy and more positive social interactions. Conversely, in negative social environments high serotonin levels are thought to contribute to unfavourable outcomes such as reduced bonding and cooperation (Kiser et al., 2012). Increasing positive perceptions of the social environment through change in affective biases may explain observed increases in prosocial behaviours following antidepressant administration (Bond, 2005; Crockett, Clark, Hauser, & Robbins, 2010; Young, Moskowitz, & aan het Rot, 2014). In keeping with previous evidence of increases in positive affective biases (Harmer et al., 2009, 2004; Tse & Bond, 2002), antidepressant treatment is likely to strengthen positive learning during social interactions. Repeated social interactions with remediated positive affective biases may reinforce engagement in future social interactions, potentially addressing issues of social withdrawal characteristic of depression.

Aims

Self-referential affective biases, particularly those occurring in a social context, may be a key area of information processing underlying negative self-schema and depressive symptoms. Change in self-referential affective processing may therefore be a potential mechanism for early antidepressant action. In this thesis I therefore aimed:

1. To evaluate whether depression is associated with differences in processing of self-related information, and whether this varies depending on the emotional valence or rewarding nature of information
2. To evaluate whether administration of an antidepressant is associated with change in self-referential affective processing.

Overview of Chapters

The five main papers included in this thesis aim to investigate the association between self-referential affective processing and depression (chapters 1 and 2), and to understand the role of self-referential affective processing in early antidepressant treatment (chapters 3, 4, and 5). I have also included an additional paper in the format of a registered report in which I investigate the role of an additional aspect of self-referential affective processing in depression, optimistic belief updating (chapter 6).

Chapter 1: Self processing in relation to emotion and reward processing in depression

The intersection between self, emotion, and reward processing may play a central role in maintaining depression. In the first paper of my thesis, I aimed to understand the relationship between self, emotion, and reward processing occurring independently and in interaction in relation to depression using three cognitive tasks. To understand the linear relationship between self-referential affective processing and depression severity I recruited participants from the community experiencing varying levels of depression (n = 144). I planned to use the findings from this study to inform the later stages of my PhD. Specifically, I aimed to identify a sensitive and reliable measure of self-referential affective processing to use as a potential indicator of early antidepressant action (Chapters 4 and 5). In preparation for a later study examining change in self-referential affective processing over the first eight weeks of antidepressant treatment (Chapter 5), I also examined the stability of associations between self-referential affective learning and depression over two testing sessions to establish the test-retest reliability of tasks where this was unknown.

Chapter 2: Individual differences in social evaluation learning and their association with psychopathology

In chapter 1 I found evidence of a linear relationship between reduced learning of positive evaluations about the self and increased depression severity. In chapter 2 I assessed the reliability of this association in a mega-analysis of ten studies previously conducted within this research group (n = 552). I further validated these findings in a large independent dataset of participants with varying levels of social anxiety and depression recruited online (n = 807). As self-referential affective learning within social interactions has previously been associated with social anxiety symptoms (Button, Browning, Munafò, & Lewis, 2012; Button et al., 2015), I also investigated the potential unifying and differential patterns of learning underlying biased social evaluation learning in depression and social anxiety to understand potential transdiagnostic mechanisms and targets for intervention.

Chapter 3: The cognitive neuropsychological theory of antidepressants: a systematic review and meta-analysis of behavioural evidence

In chapter 3 I aimed to establish the current evidence base for change in emotional processing following antidepressant administration. From this, I aimed to identify current gaps in research that could be addressed in the further chapters of this PhD. Chapter 3 therefore reports findings from a systematic review and meta-analysis of evidence for the cognitive neuropsychological theory of antidepressants. I chose to focus on behavioural evidence for the cognitive neuropsychological theory in this review as cognitive behavioural measures of emotional processing would be the most feasible to implement in primary care settings for future chapters of this thesis, and previous systematic reviews have focused on evidence of alterations in neural activity.

Chapter 4: The effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers

From systematic review and meta-analysis of literature surrounding antidepressant effects on behavioural evidence of emotional processing reported in chapter 3, I established that at present there are not reliable findings for these effects amongst cognitive tasks currently used. Chapters 1 and 2 indicated that self-referential affective biases were associated with depression. Therefore, in this chapter I used the cognitive tasks employed in chapter 1 and adapted a widely used cognitive task of emotional processing identified in chapter 3, to understand the acute effects of antidepressants on self-referential affective processing. This was a preliminary study to examine the potential mechanistic effects of antidepressants on

self-referential affective processing. I therefore conducted a double-blind randomised controlled trial examining differences in task performance between healthy volunteers administered an acute dose of citalopram versus placebo.

Chapter 5: The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care

The systematic review of literature in chapter 3 indicated that most research on the effect of antidepressants on emotional processing has been conducted in healthy samples with acute doses. In this chapter I built upon our findings in chapter 4, which examined an acute dose of citalopram in healthy volunteers, to examine the relationship between self-referential affective learning and depression over the first eight weeks of antidepressant treatment in a prospective cohort study of patients recruited from primary care. Due to adaptations made to the study design during the COVID-19 pandemic, we focused on one cognitive task that provided the strongest and most reliable effects across chapters: the social evaluation learning task, which measures self-referential affective learning within social contexts. In this chapter I therefore examine social evaluation learning as a potential mechanism of early antidepressant effects by examining associations between change in learning of positive and negative social evaluations and change in depression and anxiety symptoms.

Chapter 6: Is depression associated with reduced optimistic belief updating?

In chapter 6 I report an additional paper that examines another potential cognitive mechanism underlying depression, optimistic belief updating. In chapters 1 and 2 I report evidence of a reliable association between self-referential affective learning and depression severity. However, in chapters 4 and 5 I found only weak evidence for the importance of change in self-referential affective learning in antidepressant treatment. Systematic review of literature reported in chapter 3 identified the need for novel measures of psychopharmacological antidepressant effects. In chapter 6, I therefore focus on a novel area of cognition in this thesis – optimistic belief updating. Previous research has indicated that healthy individuals update their beliefs about the likelihood of experiencing a negative life event in the future more following good versus bad news (Sharot, Korn, & Dolan, 2011). However, individuals with depression lack this optimistic belief updating bias (Garrett et al., 2014; Korn, Sharot, Walter, Heekeren, & Dolan, 2014). I aimed to replicate this effect and to build upon previous findings to evaluate positive life events. This may help inform potential identification of future psychopharmacological measures of early antidepressant action.

References

Ahmed, N., Bone, J. K., Lewis, G., Freemantle, N., Harmer, C. J., Duffy, L., & Lewis, G. (2021). The effect of sertraline on emotional processing: Secondary analyses of the PANDA randomised controlled trial. *Psychological Medicine*, 1–8. https://doi.org/10.1017/S0033291720004985

American Psychiatric Association. (2013). Major Depressive Disorder. In *Diagnostic and Statistical Manual of Mental Disorders (DSM-V)* (5th ed.). https://doi.org/https://doi.org/10.1176/appi.books.9780890425596.dsm04

Bargh, J. A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.43.3.425

Barkus, E., & Badcock, J. C. (2019). A transdiagnostic perspective on social anhedonia. *Frontiers in Psychiatry*, *10*, 216. https://doi.org/10.3389/FPSYT.2019.00216/BIBTEX

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International University Press.

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry*, *165*(8), 969–977. https://doi.org/10.1176/appi.ajp.2008.08050721

Beck, A. T., & Dozois, D. J. A. (2011). Cognitive therapy: Current status and future directions. *Annual Review of Medicine*, *62*, 397–409. https://doi.org/10.1146/annurev-med-052209-100032

Bhanji, J. P., & Delgado, M. R. (2014). The social brain and reward: Social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(1), 61–73. https://doi.org/10.1002/wcs.1266

Bogowicz, P., Curtis, H. J., Walker, A. J., Cowen, P., Geddes, J., & Goldacre, B. (2021). Trends and variation in antidepressant prescribing in English primary care: a retrospective longitudinal study. *BJGP open*, *5*(4).

Bond, A. J. (2005). Antidepressant treatments and human aggression. *European Journal of Pharmacology*, *526*(1–3), 218–225. https://doi.org/10.1016/j.ejphar.2005.09.033

Browning, M., Bilderbeck, A. C., Dias, R., Dourish, C. T., Kingslake, J., Deckert, J., …

Dawson, G. R. (2021). The clinical effectiveness of using a predictive algorithm to guide antidepressant treatment in primary care (PReDicT): an open-label, randomised controlled trial. *Neuropsychopharmacology*, *46*(7), 1307–1314. https://doi.org/10.1038/s41386-021-00981-z

Browning, M., Kingslake, J., Dourish, C. T., Goodwin, G. M., Harmer, C. J., & Dawson, G. R. (2019). Predicting treatment response to antidepressant medication using early changes in emotional processing. *European Neuropsychopharmacology*, *29*(1), 66–75. https://doi.org/10.1016/j.euroneuro.2018.11.1102

Brush, C. J., Ehmann, P. J., Hajcak, G., Selby, E. A., & Alderman, B. L. (2018). Using Multilevel Modeling to Examine Blunted Neural Responses to Reward in Major Depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(12), 1032–1039. https://doi.org/10.1016/j.bpsc.2018.04.003

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review*, *26*(1), 17–31. https://doi.org/10.1016/j.cpr.2005.07.003

Button, K. S., Browning, M., Munafò, M. R., & Lewis, G. (2012). Social inference and social anxiety: Evidence of a fear-congruent self-referential learning bias. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(4), 1082–1087. https://doi.org/10.1016/j.jbtep.2012.05.004

Button, K. S., Kounali, D., Stapinski, L., Rapee, R. M., Lewis, G., & Munafò, M. R. (2015). Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. *PLoS ONE*, *10*(4), e0119456. https://doi.org/10.1371/journal.pone.0119456

Chekroud, A. M., & Krystal, J. H. (2015). Personalised pharmacotherapy: an interim solution for antidepressant treatment? *BMJ*, *350*, h2502. https://doi.org/10.1136/bmj.h2502

Churchill, R., Hunot, V., Corney, R., Knapp, M., McGuire, H., Tylee, A., & Wessely, S. (2002). A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Healthy Technology Assessment*, *5*(35), 1–73.

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., …

Geddes, J. R. (2018). Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *FOCUS*, *16*(4), 420–429. https://doi.org/10.1176/appi.focus.16407

Ciraulo, D. A., Shader, R. I., & Greenblatt, D. J. (2011). Clinical pharmacology and therapeutics of antidepressants. In *Pharmacotherapy of Depression* (pp. 33–124). Humana Press. https://doi.org/10.1007/978-1-60327-435-7_2

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(40), 17433–17438. https://doi.org/10.1073/PNAS.1009396107/-/DCSUPPLEMENTAL

Cunningham, S. J., & Turk, D. J. (2017). A review of self-processing biases in cognition. *Quarterly Journal of Experimental Psychology*, *70*(6), 987–995. https://doi.org/10.1080/17470218.2016.1276609

Dalgleish, T., & Watts, F. N. (1990). Biases of attention and memory in disorders of anxiety and depression. *Clinical Psychology Review*, *10*(5), 589–604. https://doi.org/10.1016/0272-7358(90)90098-U

Davey, C. G., & Harrison, B. J. (2022). The self on its axis: a framework for understanding depression. *Translational Psychiatry*, *12*(1), 1–9. https://doi.org/10.1038/s41398-022-01790-8

Davey, C. G., Pujol, J., & Harrison, B. J. (2016). Mapping the self in the brain's default mode network. *NeuroImage*, *132*, 390–397. https://doi.org/10.1016/j.neuroimage.2016.02.022

De Vries, Y. A., Roest, A. M., Bos, E. H., Burgerhof, J. G. M., Van Loo, H. M., & De Jonge, P. (2019). Predicting antidepressant response by monitoring early improvement of individual symptoms of depression: Individual patient data meta-analysis. *British Journal of Psychiatry*, *214*, 4–10. https://doi.org/10.1192/bjp.2018.122

Delaveau, P., Jabourian, M., Lemogne, C., Guionnet, S., Bergouignan, L., & Fossati, P. (2011). Brain effects of antidepressants in major depression: A meta-analysis of emotional processing studies. *Journal of Affective Disorders*, *130*(1–2), 66–74. https://doi.org/10.1016/j.jad.2010.09.032

Disner, S. G., Shumake, J. D., & Beevers, C. G. (2017). Self-referential schemas and attentional bias predict severity and naturalistic course of depression symptoms. *Cognition and Emotion*, *31*(4), 632–644. https://doi.org/10.1080/02699931.2016.1146123

Edwards, T., & Holtzman, N. S. (2017). A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality*, *68*, 63–68. https://doi.org/10.1016/j.jrp.2017.02.005

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, *68*(2), 118–124. https://doi.org/10.1016/j.biopsych.2010.01.027

Evans, J., Heron, J., Lewis, G., Araya, R., & Wolke, D. (2005). Negative self-schemas and the onset of depression in women: longitudinal study. *The British Journal of Psychiatry*, *186*(4), 302–307. https://doi.org/10.1192/BJP.186.4.302

Everaert, J., Podina, I. R., & Koster, E. H. W. (2017). A comprehensive meta-analysis of interpretation biases in depression. *Clinical Psychology Review*, *58*, 33–48. https://doi.org/10.1016/j.cpr.2017.09.005

Fossati, P., Hinfray, S., Fall, A., Lemogne, C., & Rotge, J. (2019). Emotional And Cognitive Consequences of Social Rejection: An Entry Door to Major Depression. In B. T. Baune & C. Harmer (Eds.), *Cognitive Dimensions of Major Depressive Disorder* (1st ed., pp. 211–221). Oxford University Press.

Frey, A. L., Frank, M. J., & McCabe, C. (2021). Social reinforcement learning as a predictor of real-life experiences in individuals with high and low depressive symptomatology. *Psychological Medicine*, *51*(3), 408–415. https://doi.org/10.1017/S0033291719003222

Frey, A. L., & McCabe, C. (2020a). Impaired social learning predicts reduced real-life motivation in individuals with depression: A computational fMRI study. *Journal of Affective Disorders*, *263*, 698–706. https://doi.org/10.1016/J.JAD.2019.11.049

Frey, A. L., & McCabe, C. (2020b). Effects of serotonin and dopamine depletion on neural prediction computations during social learning. *Neuropsychopharmacology 2020 45:9*, *45*(9), 1431–1437. https://doi.org/10.1038/s41386-020-0678-z

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR∗D study. *Journal of Affective*

Disorders, *172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010

Gaddy, M. A., & Ingram, R. E. (2014). A meta-analytic review of mood-congruent implicit memory in depressed mood. *Clinical Psychology Review*, *34*(5), 402–416. https://doi.org/10.1016/j.cpr.2014.06.001

Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the rose tinted glasses: Neural substrates of unbiased belief updating in depression. *Frontiers in Human Neuroscience*, *8*, 1–9. https://doi.org/10.3389/fnhum.2014.00639

Giesler, R. B., Josephs, R. A., & Swann, W. B. (1996). Self-verification in clinical depression: The desire for negative evaluation. *Journal of Abnormal Psychology*, *105*(3), 358. https://doi.org/10.1037/0021-843X.105.3.358

Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., & Rosenwald, D. P. (2014). Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses. *Image and Vision Computing*, *32*(10), 641–647. https://doi.org/10.1016/j.imavis.2013.12.007

Godlewska, B. R. (2019). Cognitive neuropsychological theory: Reconciliation of psychological and biological approaches for depression. *Pharmacology and Therapeutics*, *197*, 38–51. https://doi.org/10.1016/j.pharmthera.2018.12.010

Godlewska, B. R., & Harmer, C. J. (2021). Cognitive neuropsychological theory of antidepressant action: a modern-day approach to depression and its treatment. *Psychopharmacology*, *238*(5), 1265–1278. https://doi.org/10.1007/s00213-019-05448-0

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-Processing Behavior in Depressed Participants Relative to Healthy Volunteers: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *77*(12), 1286–1295. https://doi.org/10.1001/JAMAPSYCHIATRY.2020.2139

Hames, J. L., Hagan, C. R., & Joiner, T. E. (2013). Interpersonal processes in depression. *Annual Review of Clinical Psychology*, *9*, 355–377. https://doi.org/10.1146/annurev-clinpsy-050212-185553

Harmer, C. J., Hill, S. A., Taylor, M. J., Cowen, P. J., & Goodwin, G. M. (2003). Toward a neuropsychological theory of antidepressant drug action: Increase in positive emotional bias after potentiation of norepinephrine activity. *American Journal of Psychiatry*,

160(5), 990–992. https://doi.org/10.1176/appi.ajp.160.5.990

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., … Cowen, P. J. (2009). Effect of acute antidepressant administration on negative affective bias in depressed patients. *American Journal of Psychiatry*, *166*(10), 1178–1184. https://doi.org/10.1176/appi.ajp.2009.09020149

Harmer, C. J., Shelley, N. C., Cowen, P. J., & Goodwin, G. M. (2004). Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *American Journal of Psychiatry*, *161*(7), 1256–1263. https://doi.org/10.1176/appi.ajp.161.7.1256

Healy, D., & Whitaker, C. (2003). Antidepressants and suicide: Risk-benefit conundrums. *Journal of Psychiatry and Neuroscience*, *28*(5), 331–337.

Hengartner, M. P., & Plöder, M. (2018). Statistically significant antidepressant-placebo differences on subjective symptom-rating scales do not prove that the drugs work: Effect size and method bias matter! *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00517

Hertel, P. T., & El-Messidi, L. (2006). Am I Blue? Depressed Mood and the Consequences of Self-Focus for the Interpretation and Recall of Ambiguous Words. *Behavior Therapy*, *37*(3), 259–268. https://doi.org/10.1016/j.beth.2006.01.003

Hirschfeld, R. (2000). History and evolution of the monoamine hypothesis of depression. *Journal of Clinical Psychiatry*, *61*(6), 4–6.

Hoehn-Hyde, D., Schlottmann, R. S., & Rush, A. J. (1982). Perception of social interactions in depressed psychiatric patients. *Journal of Consulting and Clinical Psychology*, *50*(2), 209–212. https://doi.org/10.1037/0022-006X.50.2.209

Horne, S. J., Topp, T. E., & Quigley, L. (2021). Depression and the willingness to expend cognitive and physical effort for rewards: A systematic review. *Clinical Psychology Review*, *88*, 102065. https://doi.org/10.1016/j.cpr.2021.102065

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 1–16. https://doi.org/10.1186/2045-5380-3-12

Iijima, Y., Takano, K., Boddez, Y., Raes, F., & Tanno, Y. (2017). Stuttering thoughts:

Negative self-referent thinking is less sensitive to aversive outcomes in people with higher levels of depressive symptoms. *Frontiers in Psychology*, *8*, 1333. https://doi.org/10.3389/fpsyg.2017.01333

Ji, J. L., Grafton, B., & MacLeod, C. (2017). Referential focus moderates depression-linked attentional avoidance of positive information. *Behaviour Research and Therapy*, *93*, 47–54. https://doi.org/10.1016/j.brat.2017.03.004

Kaya, S., & McCabe, C. (2019). Can Understanding Reward Help Illuminate Anhedonia? *Current Behavioral Neuroscience Reports*, *6*(4), 236–242. https://doi.org/10.1007/s40473-019-00186-1

Kendrick, T., Dowrick, C., McBride, A., Howe, A., Clarke, P., Maisey, S., … Smith, P. W. (2009). Management of depression in UK general practice in relation to scores on depression severity questionnaires: Analysis of medical record data. *BMJ (Online)*, *338*(7697), b750. https://doi.org/10.1136/bmj.b750

Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., … Stringaris, A. (2018). Reward processing in depression: A conceptual and meta-analytic review across fMRI and EEG studies. *American Journal of Psychiatry*, *175*(11), 1111–1120. https://doi.org/10.1176/appi.ajp.2018.17101124

Kiser, D., SteemerS, B., Branchi, I., & Homberg, J. R. (2012). The reciprocal interaction between serotonin and social behaviour. *Neuroscience and Biobehavioral Reviews*, *36*(2), 786–798. https://doi.org/10.1016/j.neubiorev.2011.12.009

Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *Journal of Neuroscience*, *32*(47), 16832–16844. https://doi.org/10.1523/JNEUROSCI.3016-12.2012

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*(3), 579–592. https://doi.org/10.1017/S0033291713001074

Lalji, H. M., McGrogan, A., & Bailey, S. J. (2021). An analysis of antidepressant prescribing trends in England 2015–2019. *Journal of Affective Disorders Reports*, *6*, 100205. https://doi.org/10.1016/J.JADR.2021.100205

Laursen, T. M., Musliner, K. L., Benros, M. E., Vestergaard, M., & Munk-Olsen, T. (2016).

Mortality and life expectancy in persons with severe unipolar depression. *Journal of Affective Disorders*, *193*, 203–207. https://doi.org/10.1016/j.jad.2015.12.067

Lemogne, C., Delaveau, P., Freton, M., Guionnet, S., & Fossati, P. (2012). Medial prefrontal cortex and the self in major depression. *Journal of Affective Disorders*, *136*(1–2), e1–e11. https://doi.org/10.1016/J.JAD.2010.11.034

Lewinsohn, P. M. (1975). The behavioral study and treatment of depression. In *Progress in behavior modification* (Vol. 1, pp. 19–64). Elsevier.

Lewis, G., Kounali, D.-Z., Button, K. S., Duffy, L., Wiles, N. J., Munafò, M. R., … Lewis, G. (2017). Variation in the recall of socially rewarding information and depressive symptom severity: a prospective cohort study. *Acta Psychiatrica Scandinavica*, *135*(5), 489–498. https://doi.org/10.1111/ACPS.12729

Ma, Y. (2015). Neuropsychological mechanism underlying antidepressant effect: A systematic meta-analysis. *Molecular Psychiatry*, *20*(3), 311–319. https://doi.org/10.1038/mp.2014.24

Macrory, I. (2016). *Measuring National Well-being: Life in the UK, 2016*. Office for National Statistics.

Mahase, E. (2020). Esketamine for treatment resistant depression is not recommended by NICE. *BMJ (Clinical Research Ed.)*, *368*, 329. https://doi.org/10.1136/bmj.m329

McCormack, J., & Korownyk, C. (2018). Effectiveness of antidepressants. *BMJ (Online)*, *360*. https://doi.org/10.1136/bmj.k1073

McFarland, B. R., & Klein, D. N. (2009). Emotional reactivity in depression: diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance. *Depression and Anxiety*, *26*(2), 117–122. https://doi.org/10.1002/DA.20513

McManus, S., Bebbington, P., Jenkins, R., & Brugha, T. (2016). *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. A survey carried out for NHS Digital by NatCen Social Research and the Department of Health Sciences, University of Leicester*.

Merens, W., Willem Van der Does, A. J., & Spinhoven, P. (2007). The effects of serotonin manipulations on emotional information processing and mood. *Journal of Affective*

Disorders, *103*(1–3), 43–62. https://doi.org/10.1016/j.jad.2007.01.032

Miller, J. J. (2020). Ketamine/esketamine: Putative mechanisms of action. *Current Psychiatry*, *19*(1), 32–36.

Mogg, K., & Bradley, B. P. (2005). Attentional Bias in Generalized Anxiety Disorder Versus Depressive Disorder. *Cognitive Therapy and Research*, *29*(1), 29–45. https://doi.org/10.1007/s10608-005-1646-y

Moritz, D., & Roberts, J. E. (2018). Self-Other Agreement and Metaperception Accuracy Across the Big Five: Examining the Roles of Depression and Self-Esteem. *Journal of Personality*, *86*(2), 296–307. https://doi.org/10.1111/jopy.12313

Mukherjee, D., Filipowicz, A. L. S., Vo, K., Satterthwaite, T. D., & Kable, J. W. (2020). Reward and punishment reversal-learning in major depressive disorder. *Journal of Abnormal Psychology*, *129*(8), 810–823. https://doi.org/10.1037/abn0000641

Munkholm, K., Paludan-Müller, A. S., & Boesen, K. (2019). Considering the methodological limitations in the evidence base of antidepressants for depression: A reanalysis of a network meta-analysis. *BMJ Open*, *9*(6). https://doi.org/10.1136/bmjopen-2018-024886

Northoff, G. (2007). Psychopathology and pathophysiology of the self in depression - Neuropsychiatric hypothesis. *Journal of Affective Disorders*, *104*(1–3), 1–14. https://doi.org/10.1016/j.jad.2007.02.012

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*(11), 1019–1025. https://doi.org/10.1016/j.biopsych.2010.12.014

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety*, *27*(12), 1135–1142. https://doi.org/10.1002/DA.20755

Posner, J., Hellerstein, D. J., Gat, I., Mechling, A., Klahr, K., Wang, Z., … Peterson, B. S. (2013). Antidepressants Normalize the Default Mode Network in Patients With Dysthymia. *JAMA Psychiatry*, *70*(4), 373–382. https://doi.org/10.1001/JAMAPSYCHIATRY.2013.455

Pringle, A., Browning, M., Cowen, P. J., & Harmer, C. J. (2011). A cognitive neuropsychological model of antidepressant drug action. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(7), 1586–1592.

https://doi.org/10.1016/j.pnpbp.2010.07.022

Pulcu, E., Thomas, E. J., Trotter, P. D., McFarquhar, M., Juhasz, G., Sahakian, B. J., … Elliott, R. (2015). Social-economical decision making in current and remitted major depression. *Psychological Medicine*, *45*(6), 1301–1313. https://doi.org/10.1017/S0033291714002414

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*, *37*(1), 117–136. https://doi.org/10.1038/npp.2011.183

Scott, K. M., Bruffaerts, R., Tsang, A., Ormel, J., Alonso, J., Angermeyer, M. C., … De Graaf, R. (2007). Depression–anxiety relationships with chronic physical conditions: results from the World Mental Health Surveys. *Journal of Affective Disorders*, *103*(1–3), 113–120.

Setterfield, M., Walsh, M., Frey, A. L., & McCabe, C. (2016). Increased social anhedonia and reduced helping behaviour in young people with high depressive symptomatology. *Journal of Affective Disorders*, *205*, 372–377. https://doi.org/10.1016/j.jad.2016.08.020

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. https://doi.org/10.1038/nn.2949

Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., … Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1942–1947. https://doi.org/10.1073/pnas.0812686106

Shiroma, P. R., Thuras, P., Johns, B., & Lim, K. O. (2014). Emotion recognition processing as early predictor of response to 8-week citalopram treatment in late-life depression. *International Journal of Geriatric Psychiatry*, *29*(11), 1132–1139. https://doi.org/10.1002/gps.4104

Shrauger, J. S., & Schoeneman, T. J. (1979). Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, *86*(3), 549. https://doi.org/10.1037/0033-2909.86.3.549

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

from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1105. https://doi.org/10.1037/a0029792

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

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

Takano, K., Van Grieken, J., & Raes, F. (2019). Difficulty in updating positive beliefs about negative cognition is associated with increased depressed mood. *Journal of Behavior Therapy and Experimental Psychiatry*, *64*, 22–30. https://doi.org/10.1016/j.jbtep.2019.02.001

The National Institute for Health and Care Excellence. (2009). The NICE Guideline on the Treatment and Depression the Treatment and Management of Depression. In *NICE guidelines [CG90]*.

Tiihonen, J., Lönnqvist, J., Wahlbeck, K., Klaukka, T., Tanskanen, A., & Haukka, J. (2006). Antidepressants and the risk of suicide, attempted suicide, and overall mortality in a nationwide cohort. *Archives of General Psychiatry*, *63*(12). https://doi.org/10.1001/archpsyc.63.12.1358

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. *Journal of Affective Disorders*, *118*(1–3), 87–93. https://doi.org/10.1016/j.jad.2009.01.028

Tse, W. S., & Bond, A. J. (2002). Serotonergic intervention affects both social dominance and affiliative behaviour. *Psychopharmacology*, *161*(3), 324–330. https://doi.org/10.1007/s00213-002-1049-7

Van Wingen, G. A., Tendolkar, I., Urner, M., van Marle, H. J., Denys, D., Verkes, R. J., & Fernández, G. (2014). Short-term antidepressant administration reduces default mode and task-positive network connectivity in healthy individuals during rest. *NeuroImage*, *88*, 47–53. https://doi.org/10.1016/j.neuroimage.2013.11.022

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., … Dicker, D. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, *386*(9995), 743–800.

Warden, D., Rush, A. J., Trivedi, M. H., Fava, M., & Wisniewski, S. R. (2007). The STAR*D project results: A comprehensive review of findings. *Current Psychiatry Reports*, *9*(6), 449–459. https://doi.org/10.1007/s11920-007-0061-3

Warren, M. B., Pringle, A., & Harmer, C. J. (2015). A neurocognitive model for understanding treatment action in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1677). https://doi.org/10.1098/rstb.2014.0213

Weightman, M. J., Air, T. M., & Baune, B. T. (2014). A review of the role of social cognition in major depressive disorder. *Frontiers in Psychiatry*, *5*, 179. https://doi.org/10.3389/fpsyt.2014.00179

Whitfield-Gabrieli, S., & Ford, J. M. (2012). Default mode network activity and connectivity in psychopathology. *Annual Review of Clinical Psychology*, *8*. https://doi.org/10.1146/annurev-clinpsy-032511-143049

Woolfolk, R. L., Novalany, J., Gara, M. A., Allen, L. A., & Polino, M. (1995). Self-Complexity, Self-Evaluation, and Depression: An Examination of Form and Content Within the Self-Schema. *Journal of Personality and Social Psychology*, *68*(6), 1108–1120. https://doi.org/10.1037/0022-3514.68.6.1108

World Health Organization. (2011). Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level. *EB130/9*. https://doi.org/https://apps.who.int/gb/ebwha/pdf_files/eb130/b130_9-en.pdf

World Health Organization. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

Young, S. N., Moskowitz, D. S., & aan het Rot, M. (2014). Possible role of more positive social behaviour in the clinical effect of antidepressant drugs. *Journal of Psychiatry and*

Neuroscience, *39*(1), 60–65. https://doi.org/10.1503/jpn.130165

This declaration concerns the article entitled:			
Self processing in relation to emotion and reward processing in depression			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Hobbs, C., Sui, J., Kessler, D., Munafò, M. R., & Button, K. S. (2021). Self-processing in relation to emotion and reward processing in depression. *Psychological Medicine*, 1-13.		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to the formulation of ideas (50%), design of methodology (50%), and experimental work including creation of study materials (50%), participant recruitment (100%), data collection (100%), data cleaning, analysis, visualisation, and archiving (100%), interpretation of results (70%). The candidate predominantly executed the presentation of the data in journal format (90%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	[redacted]	**Date**	05/01/2022

Chapter 1: Self processing in relation to emotion and reward processing in depression

Catherine Hobbs[a], Jie Sui[b], David Kessler[c], Marcus R Munafò[d e f], Katherine S Button[a]

[a] Department of Psychology, University of Bath

[b] School of Psychology, University of Aberdeen

[c] Population Health Sciences, University of Bristol

[d] School of Psychological Science, University of Bristol

[e] MRC Integrative Epidemiology Unit at the University of Bristol

[f] National Institute of Health Research Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol

Chapter Rationale

Depression is characterised by a heightened internal focus on the self (Sheline et al., 2009), enhanced processing of negative information (Everaert, Duyck, & Koster, 2014), and impaired processing of reward (Huys, Pizzagalli, Bogdan, & Dayan, 2013; Keren et al., 2018). Differences in processing of self-related information are believed to be associated with emotion and reward cognition (Northoff, 2007; Northoff & Hayes, 2011). The interaction between self, emotion, and reward processing may therefore be a key cognitive vulnerability in depression. However, the precise nature of the relationship between these areas of processing in association with depression is currently unclear. In this first chapter of my thesis, I examined the role of the self in emotion and reward processing occurring separately and in interaction using three cognitive tasks in individuals experiencing varying levels of depression (n = 144). I planned to use these findings to inform the later chapters of my PhD by identifying a sensitive and reliable measure of differences in self-related cognition associated with depression to use as a potential indicator of early antidepressant action (Chapters 4 and 5). Additionally, I examined the stability of association between self-processing and depression in one of the cognitive tasks that currently lacks information on test-retest reliability, to inform its suitability in a longitudinal study of early antidepressant action (Chapter 5).

Open Science Statement

This study was pre-registered on the Open Science Framework (https://osf.io/34ma2), where study materials are also available (https://osf.io/7syc8/). Study data are available in the University of Bath Research Data Archive (https://doi.org/10.15125/BATH-00924; Hobbs, Sui, Kessler, Munafò, & Button, 2020).

Abstract

Background: Depression is characterised by a heightened self-focus, which is believed to be associated with differences in emotion and reward processing. However, the precise relationship between these cognitive domains is not well understood. We examined the role of self-reference in emotion and reward processing, separately and in combination, in relation to depression.

Methods: Equal numbers of adults were recruited with none, mild, or moderate to severe levels of depression (n = 144). Participants completed self-report depression measures (PHQ-9, BDI-II), and three cognitive tasks, measuring self, emotion, and reward processing, separately and in combination.

Results: When self processing was measured independently of emotion and reward, in a simple associative learning task, there was little association with depression. However, when self and emotion processing occurred in combination in a self-esteem go/no-go task, depression was associated with an increased positive other bias (*b* = 3.51, 95% CI: 1.24, 5.79). When the self was processed in relation to emotion and reward, in a social evaluation learning task, depression was associated with reduced positive self biases (*b* = 0.11, 95% CI: 0.05, 0.17).

Conclusions: Depression was associated with enhanced positive implicit associations with others, and reduced positive learning about the self, culminating in reduced self-favouring biases. However, when self, emotion, and reward processing occurred independently there was little evidence of an association with depression. Treatments targeting reduced positive self-biases may provide more sensitive targets for therapeutic intervention and potential biomarkers of treatment responses, allowing the development of more effective interventions.

Introduction

Depression is a highly prevalent mental health problem worldwide (World Health Organization, 2017), and is projected to be the leading cause of disease burden globally by 2030 (World Health Organization, 2011). Treatments for depression are moderately effective (Cipriani et al., 2018; Cuijpers, Andersson, Donker, & Van Straten, 2011), but individual response varies (Maslej, Furukawa, Cipriani, Andrews, & Mulsant, 2020). Understanding the cognitive processes maintaining depression may allow us to develop sensitive targets for therapeutic intervention. In this study we explored the role of self processing in depression, in relation to emotion and reward processing.

Self Processing

Across the general population, people show greater attention, recall and learning of self-related stimuli, often referred to as the self-prioritisation effect (Cunningham & Turk, 2017; Sui, He, & Humphreys, 2012; Sui & Humphreys, 2015a). However, individuals experiencing depression exhibit a heighted focus on the self, and difficulty disengaging from an internal self-referential focus (Northoff, 2007; Sheline et al., 2009). Paradoxically, this heightened internal self-referential focus may prevent individuals from associating novel stimuli with internal representations of the self (Sui, Ohrling, & Humphreys, 2016). This concept has previously been demonstrated in a study where following negative mood induction participants were worse at associating arbitrarily assigned neutral shapes with the self (Sui et al., 2016). Individuals experiencing depression may subsequently be limited in their ability to update their self-concept from environmental feedback, perpetuating maladaptive views of the self.

The strength and consistency of self-prioritisation effects has led to proposals of the self being an integrative hub through which incoming stimuli is processed (Sui & Humphreys, 2015a). Targeting abnormalities in self-referential processing in depression may have wider implications for other cognitive domains implicated in depression. This is likely to include emotion (Ma & Han, 2010) and reward (Northoff & Hayes, 2011), as they are fundamental behavioural drivers and neurally overlap in the medial prefrontal cortex.

Self and Emotional Processing

Negative perceptions of the self are believed to play a causal role in the development of depression. According to Beck's cognitive theory, individuals experiencing depression develop negative views of the self as an internalised reaction to repeated adverse social

experiences. When activated by stressful life events these negative self-schema dominate information processing, promoting automatic processing of negative information about the self (Beck, 2008). Supportive of this theory, emotional biases are more likely to be observed in depression when stimuli is processed in reference to the self (Gaddy & Ingram, 2014; Hertel & El-Messidi, 2006; Ji, Grafton, & MacLeod, 2017). Altering negative information processing in relation to the self is therefore a key target for therapeutic interventions for depression.

Self and Reward Processing

Depression is also associated with a hyposensitivity to reward and hypersensitivity to punishment (Eshel & Roiser, 2010). There is evidence to suggest that this is linked to self processing. Self-relevant information induces activity in areas of the brain also activated during reward processing, such as the ventral medial prefrontal cortex, ventral striatum, and ventral tegmental area (Northoff & Hayes, 2011). Differences in self-processing in depression may be at least partially driven differences in reward processing. In keeping with this theory, individuals with greater depression were found to continue to selectively engage in negative thoughts about the self even when this resulted in economic loss (Takano, Van Grieken, & Raes, 2019). Targeting reward processing in depression may have wider effects on self processing and vice versa.

Self, Emotion and Reward Processing

The interaction between self, emotion and reward processing may be a key combination of cognitive processes maintaining depression. Patients with depression show reduced activation of both reward and self-related areas of the brain when processing positive stimuli (Northoff, 2007; Northoff & Hayes, 2011). Reduced self-referential processing of positive information has also been identified as the most robust predictor of low approach motivation and reward responsivity (Hsu et al., 2020). Increased sensitivity to punishing feedback may sustain preferential processing of negative information about the self, reinforcing negative self-schema. Likewise, reduced sensitivity to positive feedback may reduce the ability to learn positive information about the self. The intersection between self, emotion and reward may therefore be the most effective target for cognitive treatments for depression.

Aims and Hypotheses

We explored the role of the self in relation to emotion and reward processing associated with varying levels of depressive symptoms. In contrast to previous studies that focused on either

of these independent cognitive process (based on self, reward, or emotion) or interactions between any two components, we used three cognitive tasks to examine relationships between these processes and depressive symptoms, not only as distinct cognitive processes but also how they functionally interact.

To examine self, reward and emotion processing occurring independently we used associative learning tasks where participants paired neutral shapes with self-relevant, emotionally valenced, and varying degrees of reward, in three separate tasks. Based on previous research (Sui et al., 2016), we predicted that increased depression severity would be associated with worse performance when associating shapes with the self. Similarly based on evidence of impaired affective processing in depression (Dalgleish & Watts, 1990; Dalili, Penton-Voak, Harmer, & Munafò, 2015; Eshel & Roiser, 2010), we predicted that depression would be associated with worse performance when associating shapes with positive and rewarding stimuli.

To examine self, reward and emotion processing occurring in interaction we used a social reinforcement learning task where participants learnt when the computer liked themselves and others. Based on previous evidence (Hobbs et al., 2019), we hypothesised that increasing depression severity would be associated with worse learning of the self being 'liked'.

We also included a self-esteem go/no-go task due to its ability to integrate self and emotion processing. Participants rapidly categorised emotional and referential words, with greater discriminative accuracy believed to reflect existing implicit associations. An implicit negative self-esteem would therefore be reflected by greater discriminative accuracy when categorising self-referential and negative stimuli. However, due to mixed findings regarding the role of response inhibition in depression (Lewis, Button, Pearson, Munafò, & Lewis, 2020), and no previous use of this task within our research group we made no hypotheses regarding this task.

Methods

This study was pre-registered on the Open Science Framework (https://osf.io/34ma2), where study materials are also available. Study data are available in the University of Bath Research Data Archive (https://doi.org/10.15125/BATH-00924 ; Hobbs et al., 2020).

Participants

We recruited participants aged 18 to 65, fluent in English, with normal or corrected-to-normal vision, through campus advertising at the University of Bath. As depression severity is positively skewed (Tomitaka, Kawasaki, & Furukawa, 2015), to ensure balanced levels of depression we screened participants using the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001). We recruited an equal number of participants with no depression (PHQ-9 < 5), mild depression (PHQ-9 5-9), and moderate to severe depression (PHQ-9 ≥ 10).

Procedure

Participants completed two testing sessions, on average eight days apart (SD 3). At each session participants completed a social evaluation learning task, allowing measurement of test-retest reliability. To reduce fatigue effects associated with reaction time tasks, participants completed a go/no-go task at session one and an associative learning task at session two. At each session participants completed self-report measures of mood after the cognitive tasks.

Materials

Cognitive Tasks

To personalise tasks, prior to testing participants provided the first names of themselves, a friend, and a stranger.

Associative Learning Task

We used three simple associative learning tasks to measure how self, emotion and reward processing are independently associated with depression (Stolte, Humphreys, Yankouskaya, & Sui, 2017; Sui & Humphreys, 2015b). In each task, participants learnt to associate stimuli related to the relevant area of processing (e.g., Self: names of the self, a friend, and a stranger; Emotion: happy, neutral, and sad faces; Reward: £9, £3, £1), with abstract shapes. These tasks were completed sequentially in a counterbalanced order.

At the start of each task participants were instructed to learn randomly assigned stimuli-shape pairings. Two blocks of 60 trials were completed per task. In each trial a fixation point was displayed for 200 ms, followed by a stimuli-shape pairing presented for 100 ms (self, reward) or 150 ms (valence task only due to greater visual stimuli complexity). Participants pressed the 'n' or 'm' keys to indicate whether the presented pairings matched with the learnt association (Figure 1a). Key assignment to 'matching' or 'non-matching' responses was randomised for each participant but consistent across tasks. A response limit of 1100 ms was applied. Feedback was presented for each trial for 500 ms ("correct" / "incorrect" / "too slow"). At the end of each block participants were informed of their accuracy. For the reward task only, participants received a monetary reward based on the proportion of correct trials per reward stimuli.

Accuracy and reaction times (ms) were recorded. Prioritisation of stimuli is indicated by faster reaction times and/or higher accuracy.

Self-Esteem Go/No-Go Task

To measure how self and emotion processing occurring in interaction are associated with depression, we used a self-esteem go/no-go task. This task is proposed to measure implicit self-esteem (Gregg & Sedikides, 2010).

Participants were asked to categorise characteristics as positive (e.g., 'charming', 'smart') or negative (e.g., 'cruel', 'boring'), and referential worlds as related to the self, specified to participants as 'me' (e.g., participants' first name, 'me', 'I'), or others, specified to participants as 'not-me' (e.g., 'they', 'them', 'others'). In the training phase participants categorised words according to single categories (e.g., positive, negative, me, not-me), with 20 trials per condition. In the test phase, participants categorised words belonging to paired categories (e.g., positive OR me, positive OR not-me, negative OR me, negative OR not-me). There were 16 practice trials and 48 test trials for each paired combination of categories. An equal number of trials for stimuli relating to each condition was presented per block. A response timeout of 600 ms was applied. Block order was randomised.

At the beginning of each block the condition(s) by which words should be categorised was presented at the top of the screen and remained in place throughout the block. In each trial a word belonging to any of the conditions (e.g., positive, negative, me or not-me) was presented at the centre of the screen for 600 ms. Participants were asked to press the spacebar if the presented word related to the specified category (a 'go' response) or to refrain from

pressing the spacebar if the word did not relate to the specified category (a 'no-go' response) (Figure 1b).

We categorised responses in test trials according to hits (a 'go' response when the stimuli belonged to the specified categories) and false alarms (a 'go' response when the stimuli did not belong to the specified categories). Responses to both referential and valence stimuli were included. For example, if the specified categories were 'positive OR me' a trial was considered a hit if a 'go' response was given upon presentation of a positive characteristic *or* a self-referential word.

Discriminative accuracy (d') for each referential-emotion block was calculated by applying z-score transformations and subtracting hits from false alarms. Greater d' values indicate greater accuracy, suggesting stronger associations between paired-categories.

Social Evaluation Learning Task

To measure self, emotion, and reward learning occurring simultaneously we used a reinforcement learning task within a social context (Button, Karwatowska, Kounali, Munafò, & Attwood, 2016; Button, Browning, Munafò, & Lewis, 2012; Button et al., 2015). Participants learnt how much the computer 'liked' the self, a friend, and a stranger based on feedback to a forced choice selection between positive and negative social evaluation pairings (Figure 1c). A response time limit was not imposed. Participants learnt two rules based on the probability of the positive evaluations being 'correct' ('Like' 60-80%, 'Dislike' 20-40%). The number of errors made before reaching the criterion of eight consecutive rule-congruent responses were recorded. Bias scores were calculated by subtracting errors to criterion made when learning the dislike rule from the like rule. A positive value indicates a negative bias, as fewer errors were made learning the dislike rule compared to the like rule. We also calculated participants cumulative accuracy across trials in each condition-rule block to visualise learning curves.

After completing each rule block participants were also asked to provide a global rating of how much the computer liked the person, ranging from 'Complete Dislike' (0) to 'Complete Like' (10).

Participants completed all referential-conditions and rules. Order of referential-condition, and nested within this rule, was randomised. All participants completed 24 trials per referential-condition rule block.

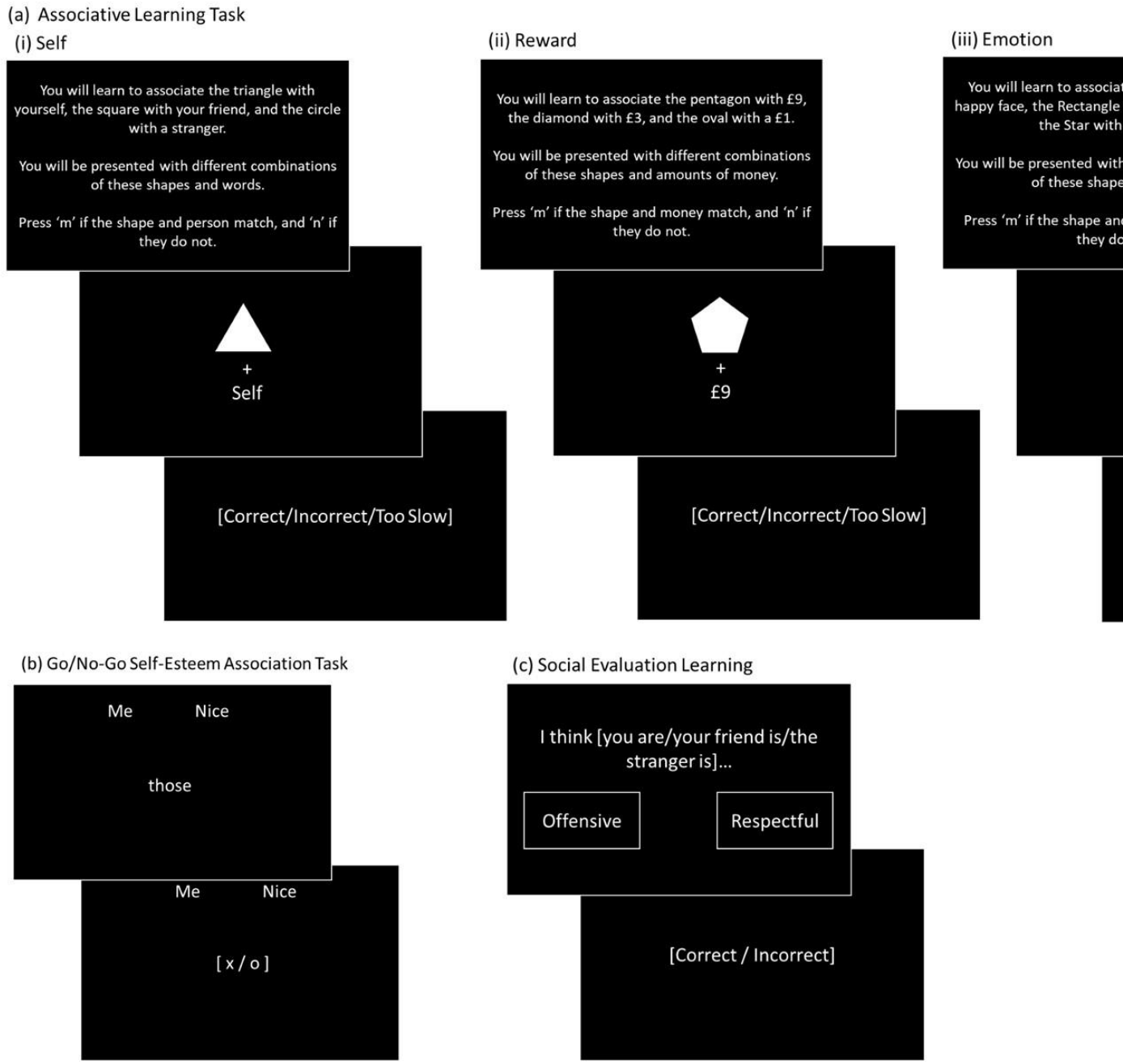


Figure 1

Cognitive Task Procedures

(a) Associative Learning Tasks: Example of an introduction, trial, and feedback for each type for each type of task (self, reward, emotion). In the introduction of each task participants were instructed to associate specified randomly assigned shape and stimuli pairings. In each trial participants were presented with a random combination of these shape-stimuli pairings and were asked to use the 'n' and 'm' keys to indicate whether these matched with the pairings they had previously learnt. In these examples, the 'm' key indicates a 'matching' response, and the 'n' key indicates a 'non-matching'

response, however key assignment was randomised for each participant. Following each trial, feedback was given indicating if the participant was correct, incorrect, or too slow (> 1100 ms). Each of these examples demonstrate a 'matching' trial, where the presented shape-stimuli match with the pairings specified in the introduction. A 'matching' response would therefore be correct, in this example the 'm' key, whereas an 'non-matching' response would be incorrect, in this example the 'n' key.

(b) Go/No-Go Self-Esteem Association Task: Example of a trial and feedback for the Self-Positive condition. The conditions that words should be categorised according to (in this instance Me or Nice) were presented at the top of the screen throughout the block. In each trial a word was presented at the centre of the screen. Participants were asked to press the spacebar if the word belonged to a specified category (a 'go' response) or to refrain from pressing the spacebar if the word did not belong to the specified category (a 'no-go' response). Feedback (correct indicated by a green circle, or incorrect indicated by a red cross) was given for each response. In this example, a 'no-go' response would be considered a correct rejection and a 'go' response would be considered a false alarm, as the stimuli ('those') does not belong to the Me or Positive categories.

(c) Social Evaluation Learning Task: Example of a trial and feedback. Participants were asked to select the word that they felt reflected the computers' opinion of the person being learnt about (self, friend, or stranger), and were given feedback on their response. The proportion of trials deemed correct upon selection of the positive word was manipulated to reflect learning of two different rules: positive 'like' 60-80%, negative 'dislike' 20-40%.

Self-Report Measures

We measured depression severity using the Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001) and Beck Depression Inventory (BDI-II) (Beck, Steer, & Brown, 1996). The PHQ-9 and BDI-II are self-administered questionnaires of the experience of depression within the previous two weeks. The PHQ-9 consists of nine items relating to the DSM-IV diagnostic criteria with scores ranging from 0-27, whereas the BDI-II consists of 21 items with scores ranging from 0-63 and has a greater focus on cognitive symptoms. Both measures demonstrate good psychometric properties (Cameron, Crawford, Lawton, & Reid, 2008; Wang & Gorenstein, 2013), and are widely used in clinical practice (Kendrick et al., 2009).

We also identified whether participants met ICD-10 criteria for a primary diagnosis of a Major Depressive Episode (MDE) using the Clinical Interview Schedule-Revised (CIS-R; Lewis, Pelosi, Araya, & Dunn, 1992). The CIS-R is a fully structured self-administered computerised assessment that provides ICD-10 diagnoses of common mental health disorders. It has previously been used in large scale epidemiological studies within the general population.

As social anxiety has previously been associated with performance on the Social Evaluation Learning task (Button et al., 2015), we also measured social anxiety using the Brief Fear of Negative Evaluation Scale (Leary, 1983). To characterise the clinical profile of our sample we collected additional self-report measures of mental health and cognition. We measured anxiety using the Generalised Anxiety Disorder Scale (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006), anxiety relating to positive social feedback using the Fear of Positive Evaluation Scale (Weeks, Heimberg, & Rodebaugh, 2008), self-esteem using the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), and self-schema using the Dysfunctional Attitude Scale (DAS; Weissman & Beck, 1978). Finally, we measured change in state mood before and after completion of the cognitive tasks using the Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988).

Statistical Analyses

Sample Size Calculation

A priori power calculations indicated that 144 participants would be required to provide greater than 80% power at an alpha level of 0.05 to detect previously observed effect sizes ($\eta^2 = 0.05$) for the relationship between bias scores in the self condition in the Social Evaluation Learning task and depression severity (Button et al., 2016, 2012, 2015; Hobbs et al., 2019),

and greater than 99% power to detect previously observed effect sizes for the relationship between reaction times when matching shapes with the 'self' on the Associative Learning Task and depression severity ($\eta^2 = 0.17$) (Sui & Button, 2017).

Data Exclusion

Data was excluded according to a priori criteria as specified in our pre-registration.

For the associative learning task, trials with reaction times less than 200 ms (0.8%) and trials with no response (8%) were excluded. We included matching and non-matching trials in our analyses. For reaction time data we used both correct and incorrect responses.

We excluded 36 (25%) participants from the Go/No-Go Self-Esteem analyses due to a pattern of response indicating non-compliance (discrimination scores lower than 5 and/or bias scores less than 12 or greater than 36). As the exclusion rate was high, we repeated the main analyses for this task with all participants included as a sensitivity analysis.

Due to a technical error, data for the social evaluation learning task was unavailable in the second session for one participant.

Statistical Models

All analyses were conducted in R 3.6.

To aid interpretation we have provided both standardised (β) and unstandardised (b) regression coefficients.

We first assessed whether task performance differed across conditions using mixed-effects linear regression models. Separate models were used for each task, and for each measure of performance. Subject was entered as a random effect to account for within-subject effects. Task performance measures were entered as the outcome, and conditions as predictors.

Whilst the associative learning task and go/no-go task have previously been evidenced to have acceptable levels of reliability (Stolte et al., 2017; Williams & Kaufmann, 2012), the reliability of the social evaluation learning task is yet to be tested. We calculated intraclass correlation coefficients for bias scores in the social evaluation learning task, using two-way mixed-effects models to calculate absolute agreement and consistency as recommended for cognitive-behavioural measures (Parsons, Kruijt, & Fox, 2019).

We used linear regression models to assess the relationship between task performance and depression. In all models, task outcomes were entered as separate predictors according to

condition (e.g., in the self associative learning task accuracy model, accuracy in the self, friend and stranger condition were entered as separate predictors). We used depression as the outcome in these models, rather than a predictor as is typical in psychiatric experimental models, in preparation for future work using the cognitive task outcomes as predictors of change in depression severity. Separate models were conducted for each task outcome with PHQ-9 or BDI-II scores used as continuous outcomes. As the social evaluation learning task was completed in two sessions, we used mixed-effects linear regression models with session included as an additional predictor and subject as a random effect.

To examine the reliability of our findings for individuals meeting diagnostic criteria for depression, we repeated the primary analyses for each task using logistic regression models. Primary diagnosis of major depressive episode derived from the CIS-R was used as a binary outcome (diagnostic criteria met/not met). As the CIS-R was only completed at session 1, for tasks with multiple timepoints data from session 1 was used.

Full details of models are provided in the supplementary materials.

Results

Participant Characteristics

We recruited 144 participants, all of whom provided data for analysis. To demonstrate variability across depression severity, participant characteristics grouped according to PHQ-9 clinical cut-offs are presented in Table 1. The PHQ-9 and BDI-II showed excellent test-retest reliability between sessions (PHQ-9: ICC 0.94 (95% CI 0.89, 0.96), BDI-II: ICC: 0.96 (95% CI: 0.94, 0.97)), and strongly correlated (r = 0.90, 95% CI: 0.88, 0.92).

Table 1

Participant Characteristics according to Depression Severity

	PHQ-9 Depression Severity		
	None (<5)	**Mild (5-9)**	**Moderate to Severe (≥10)**
N (%)	48 (33)	56 (39)	40 (28)
Age, M (SD)	23.4 (7.3)	22.6 (7.9)	20.9 (3.1)
Gender, N (%)			
Male	11 (23)	11 (19)	9 (22)
Female	37 (77)	44 (79)	31 (78)
Other	0 (0)	1 (2)	0 (0)
Ethnicity, N (%)			
White	33 (69)	30 (54)	33 (85)
Black	0 (0)	2 (4)	1 (3)
Asian	11 (23)	18 (32)	3 (7)
Mixed	4 (8)	4 (7)	2 (5)
Other	0 (0)	2 (3)	0 (0)
Employment, N (%)			
Student	42 (88)	50 (89)	36 (90)
Employed	5 (10)	6 (11)	3 (8)
Other	1 (2)	0 (0)	1 (2)
CIS-R Primary Diagnosis Major Depressive Episode, N (%)	0 (0)	9 (16)	26 (65)[a]
Current Treatment, N (%)			
Psychological Therapy	0 (0)	3 (5)	5 (13)
Antidepressants	0 (0)	2 (4)	7 (18)
PHQ-9, M (SD)	2.5 (1.2)	6.9 (1.4)	15.0 (4.0)
BDI-II, M (SD)	4.6 (3.6)	13.1 (5.6)	27.2 (10.5)
BFNE, M (SD)	34.3 (10.2)	38.8 (9.1)	45.9 (8.2)
GAD-7, M (SD)	2.1 (2.1)	5.4 (3.0)	10.9 (4.1)
FPE, M (SD)	23.2 (11.1)	26.8 (13.5)	36.5 (14.2)

DAS-24, M (SD)	90.3 (17.8)	94.9 (18.5)	108.3 (15.5)
RSES, M (SD)	13.6 (1.9)	12.9 (2.5)	12.7 (2.1)
PANAS Positive Change, M (SD)	-1.5 (3.2)	-1.9 (3.3)	-1.9 (4.2)
PANAS Negative Change, M (SD)	-0.7 (2.1)	-0.7 (2.2)	-1.1 (4.0)

[a] Participants who met criteria for a primary diagnosis of a MDE within this group had higher PHQ-9 (M 16.21, SD 4.35) and BDI-II scores (M 31.88, SD 10.42), compared to those that did not have a primary diagnosis of a MDE (PHQ-9: M 12.00, SD 1.83, BDI-II: 19.57, SD 5.95).

CIS-R = Clinical Interview Schedule Revised, PHQ-9 = Patient Health Questionnaire, BDI-II = Beck Depression Inventory, BFNE = Brief Fear of Negative Evaluation, GAD-7 = Generalised Anxiety Questionnaire, BFNE = Brief Fear of Negative Evaluation Scale, FPE = Fear of Positive Evaluation Scale, DAS-24 = Dysfunctional Attitudes Scale, RSES = Rosenberg Self-Esteem Scale, PANAS = Positive and Negative Affect Schedule.

Note: All data presented in this table were collected at the first testing session. PANAS change scores reflect differences in scores from pre- to post-completion of the cognitive tasks.

Associative Learning Task

Hypothesis: Depression will be associated with reduced learning of self, highly rewarding and positive stimuli as indicated by reduced accuracy and greater reaction times.

Self

Consistent with prior evidence of self-prioritisation (Sui et al., 2012), participants on average showed the highest level of accuracy and fastest reaction times when matching shapes with the name of the self versus a friend or stranger (Supplementary Tables S1 and S2). We found no evidence to support our hypothesis; ability to associate shapes with the self, a friend or a stranger was not associated with depression (Table 2).

Reward

Likewise consistent with previous evidence of prioritisation of higher levels of reward (Sui & Humphreys, 2015b), participants on average were more accurate and faster when matching shapes with the highest level of reward (Supplementary Tables S1 and S2).

We found some evidence that increased accuracy when matching shapes with the medium level of reward was associated with greater depression. For every 1% increase in accuracy when matching shapes with '£3', PHQ-9 and BDI-II scores increased by 0.10 (b 95% CI: 0.02, 0.19, p = 0.021) and 0.24 (b 95% CI: 0.05, 0.43, p = 0.012) points respectively.

There was also weak evidence that decreasing accuracy when matching shapes with the high level of reward was associated with increased BDI-II scores (b = -0.19, b 95% CI: -0.37, 0.00, p = 0.051). However, confidence intervals overlapped with the null and there was little evidence of a similar relationship for PHQ-9 scores. We therefore found only weak support for our hypothesis.

No association was observed between accuracy when matching shapes with the low level of reward (£1) and depression severity (Table 2). We also found no relationship between reaction times and depression for this task (Table 2).

Emotion

Consistent with previous evidence of prioritisation of positive stimuli (Stolte et al., 2017), participants on average were more accurate and faster at matching shapes with happy faces (Supplementary Tables S1 and S2). However, in contrast with our hypothesis, accuracy or reaction times were not associated with depression (Table 2).

Table 2

Results from linear regression models examining the association between accuracy and reaction times for each task condition (predictors) in the associative learning task with depression (Outcome: PHQ-9/BDI-II)

Task	Stimuli	PHQ-9					BDI-II				
		b	*b* 95% CI	β	β 95% CI	p	*b*	*b* 95% CI	β	β 95% CI	p
Accuracy (%)											
Self	Intercept	11.44	3.38, 19.49	0.00	-0.16, 0.16	0.006	14.51	-2.78, 31.79	0.00	-0.17, 0.17	0.099
	Self	-0.06	-0.17, 0.05	-0.12	-0.33, 0.10	0.288	-0.15	-0.39, 0.10	-0.13	-0.35, 0.08	0.231
	Friend	-0.04	-0.15, 0.06	-0.09	-0.32, 0.13	0.414	0.03	-0.20, 0.26	0.03	-0.20, 0.26	0.790
	Stranger	0.05	-0.04, 0.15	0.13	-0.10, 0.36	0.279	0.11	-0.09, 0.32	0.13	-0.10, 0.36	0.275
Reward	Intercept	6.07	1.08, 11.06	0.00	-0.16, 0.16	0.018	8.59	-2.02, 19.20	0.00	-0.16, 0.16	0.112

	High (£9)	-0.06	-0.15, 0.03	-0.18	-0.43, 0.07	0.166	-0.19	-0.37, 0.00	-0.25	-0.50, 0.00	0.051
	Medium (£3)	0.10	0.02, 0.19	0.30	0.05, 0.56	0.021	0.24	0.05, 0.43	0.33	-0.07, 0.58	0.012
	Low (£1)	-0.03	-0.10, 0.04	-0.10	-0.31, 0.11	0.366	0.02	-0.13, 0.16	0.02	-0.18, 0.23	0.814
Emotion	Intercept	6.05	1.05, 11.04	0.00	-0.17, 0.17	0.018	10.72	0.05, 21.39	0.00	-0.17, 0.17	0.049
	Happy	-0.02	-0.09, 0.05	-0.06	-0.27, 0.15	0.588	-0.05	-0.21, 0.11	-0.06	-0.28, 0.15	0.547
	Neutral	0.03	-0.05, 0.11	0.08	-0.15, 0.32	0.498	0.06	-0.11, 0.23	0.08	-0.15, 0.32	0.489
	Sad	0.01	-0.07, 0.08	0.02	-0.22, 0.25	0.881	0.04	-0.13, 0.20	0.05	-0.18, 0.28	0.668
Reaction Times (ms)											

Self	Intercept	11.50	2.64, 20.37	0.00	-0.17, 0.17	0.011	24.45	5.50, 43.40	0.00	-0.16, 0.16	0.012
	Self	0.00	-0.02, 0.03	0.05	-0.27, 0.37	0.755	0.00	-0.05, 0.06	0.01	-0.31, 0.34	0.929
	Friend	-0.01	-0.04, 0.01	-0.20	-0.60, 0.19	0.317	-0.04	-0.09, 0.02	-0.24	-0.64, 0.15	0.277
	Stranger	0.00	-0.03, 0.03	0.04	-0.39, 0.48	0.846	0.02	-0.05, 0.08	0.11	-0.32, 0.55	0.610
Reward	Intercept	4.53	-2.38, 11.44	0.00	-0.17, 0.17	0.197	7.89	-6.88, 22.65	0.00	-0.17, 0.17	0.293
	High (£9)	0.01	-0.01, 0.04	0.22	-0.15, 0.59	0.245	0.03	-0.01, 0.08	0.26	-0.11, 0.63	0.168
	Medium (£3)	-0.01	-0.04, 0.02	-0.19	-0.65, 0.27	0.422	-0.01	-0.07, 0.04	-0.09	-0.55, 0.36	0.685
	Low (£1)	0.00	-0.02, 0.02	0.02	-0.36, 0.40	0.933	-0.01	-0.06, 0.03	-0.11	-0.49, 0.26	0.549

Emotion	Intercept	7.51	1.64, 13.37	0.00	-0.17, 0.17	0.013	14.33	1.75, 26.91	0.00	-0.17, 0.17	0.026
	Happy	0.01	-0.01, 0.02	0.10	-0.28, 0.48	0.614	0.00	-0.04, 0.04	0.02	-0.36, 0.41	0.898
	Neutral	0.00	-0.03, 0.02	-0.06	-0.55, 0.44	0.824	0.00	-0.05, 0.05	0.00	-0.49, 0.50	0.990
	Sad	0.00	-0.03, 0.02	-0.07	-0.57, 0.44	0.793	0.00	-0.05, 0.04	-0.04	-0.55, 0.46	0.867

b = unstandardised regression coefficients, β = standardised regression coefficients

Self-Esteem Go/No-Go Task

Due to previous mixed findings for the role of response inhibition in depression we made no hypothesis regarding this task, our findings should therefore be considered exploratory.

We found strong evidence of an interaction between referential condition and emotion on discriminative accuracy in the self-esteem Go/No-Go Task (b = 0.79, b 95% CI: 0.61, 0.97, β = 1.31, β 95% CI: 1.01, 1.61, p < .001; Supplementary Table S3). On average, participants showed a positive bias towards the self with greater discriminative accuracy for positive (M 1.40, SD 0.56) versus negative (M 1.0, SD 0.52) associations with the self. The opposite pattern was observed when associating words with the other (positive: M 0.71, SD 0.48, negative: M 1.12, SD 0.62).

We found consistent evidence that discriminative accuracy in the 'other' condition was associated with depression severity. Increased discriminative accuracy when associating positive words with others was associated with greater depression severity using both the PHQ-9 (b = 3.51, b 95% CI: 1.24, 5.79, β = 0.30, β 95% CI: 1.24, 5.79, p = 0.003) and BDI-II (b = 6.78, b 95% CI: 1.93, 11.64, β = 0.28, β 95% CI: 0.08, 0.47, p = 0.007). Conversely, increased discriminative accuracy when associating negative words with others was associated with lower PHQ-9 (b = -2.46, b 95% CI: -4.24, -0.68, β = -0.27, β 95% CI: -0.46, -0.07, p = 0.007), and BDI-II scores (b = -5.13, b 95% CI: -8.92, -1.34, β = -0.27, β 95% CI: -0.46, - 0.07, p = 0.008). Individuals with greater depression therefore showed both a greater positive bias, and a reduced negative bias, when processing information about others.

Evidence for an association between discriminative accuracy in the self condition and depression was less consistent. Increased discriminative accuracy when associating positive words with the self was associated with a decrease in PHQ-9 scores (b = -2.47, b 95% CI: -4.54, -0.39, β = -0.24, β 95% CI: -0.44, -0.04, p = 0.020). Although an effect in the same direction was observed for BDI-II scores, confidence intervals overlapped substantially with the null (b = -3.20, b 95% CI: -7.62, 1.23, β = -0.15, β 95% CI: -0.36, 0.06, p = 0.155). There was little evidence of an association between discriminative accuracy when associating negative words with the self with either the PHQ-9 (b = -0.59, b 95% CI: -2.57, 1.39, β = -0.05, β 95% CI: -0.24, 0.13, p = 0.553) or BDI-II (b = 0.81, b 95% CI: -5.03, 3.41, β = -0.04, β 95% CI: -0.22, 0.15, p = 0.704).

As we excluded a large proportion of participants (25%) in these analyses due to a priori criteria indicating non-compliance with the task, we repeated these analyses including all participants as a sensitivity analysis. We no longer found evidence for an association between discriminative accuracy in the other-negative condition and PHQ-9 severity, as confidence intervals overlapped with the null. However, the results described above persisted for all other associations (Supplementary Table S4).

Social Evaluation Learning

Hypothesis: Depression will be associated with reduced positive biases when learning about the self, driven by a greater number of errors before learning the positive 'like' rule.

Bias Scores

Participants on average were most positively biased when learning about the friend, making 2.07 fewer errors learning positive relative to negative evaluations (b 95% CI: -2.93, -1.21, β = -0.35, β 95% CI: -0.49, -0.20, $p < .001$), compared to when learning about the self. Participants displayed similar levels of bias when learning about the self and stranger (b = -0.44, b 95% CI: -1.31, 0.42, β = -0.07, β 95% CI: -0.22, 0.07, p = 0.318). The estimated agreement and consistency for bias scores across test sessions was ICC = 0.41 (95% CI: 0.29, 0.52).

In support of our hypothesis, bias scores when learning about the self were associated with depression severity. For every additional error learning the positive relative to the negative rule, PHQ-9 scores increased by 0.11 points (b 95% CI: 0.05, 0.17, $p < .001$) and BDI-II scores increased by 0.23 points (b 95% CI: 0.12, 0.34, $p < .001$). Effects were specific to learning about the self; bias scores when learning about the friend or a stranger were not associated with depression (Figure 2a; Table 3).

We also conducted additional exploratory analyses to examine whether the relationship between self bias scores and depression symptoms was consistent across sessions. We found little evidence of an interaction suggesting that the relationship did not vary over the two sessions (PHQ-9 b = 0.04, b 95% CI: -0.04, 0.11, β = 0.02, β 95% CI: -0.02, 0.06, p = 0.377; BDI-II b = 0.07, b 95% CI: -0.07, 0.21, β = 0.02, β 95% CI = -0.02, 0.06, p = 0.315).

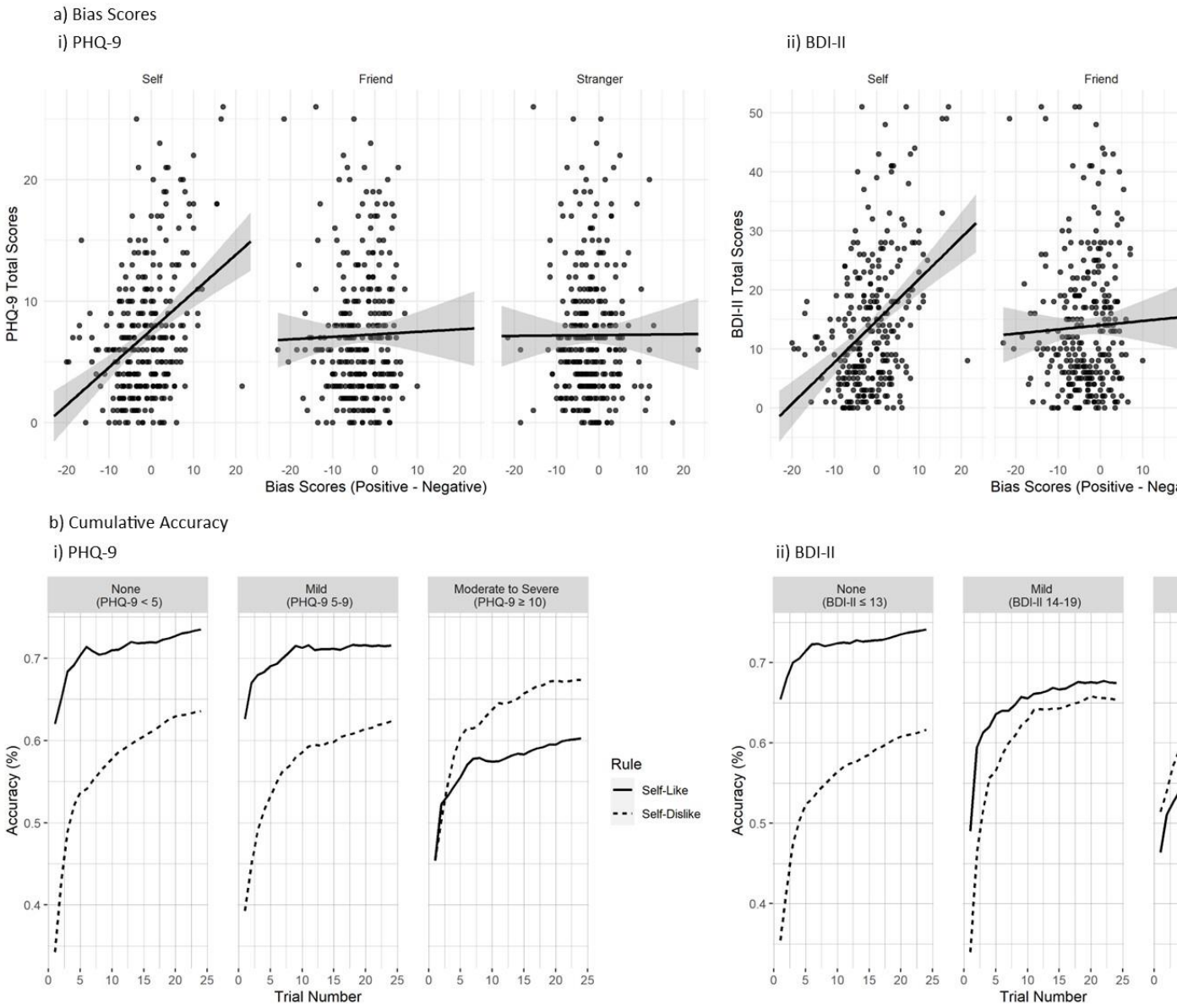


Figure 2

(a) Relationship between bias scores in the self, friend, and stranger conditions in the social evaluation learning task with (i) PHQ-9 and (ii) BDI-II scores.

(b) Learning curves in the self condition in the social evaluation learning task based on cumulative accuracy with depression severity grouped according to (i) PHQ-9 clinical cut-offs and (ii) BDI-II clinical cut-offs.

Errors to Criterion

To investigate whether the relationship between bias scores and depression severity was driven by worse learning of the positive rule, or better learning of the negative rule, we examined the relationship between errors to criterion in each referential-rule condition and depression.

Participants overall were positively biased, making greater errors learning the negative versus positive rules (b = 1.45, b 95% CI: 0.82, 2.07, p < .001; Supplementary Table S5) The greater bias scores in the friend condition, as outlined above, was driven by participants making both fewer errors learning the positive rule (M 5.39, SD 3.76) and greater errors learning the negative rule (M 8.90, SD 4.24), compared to the self (positive M 6.50, SD 4.22; negative M 7.95 SD 4.28) and stranger (positive M 6.34 SD 3.90, negative M 8.23 SD 3.97) conditions.

We found consistent evidence to support our hypothesis that depression would be associated with a greater number of errors when learning the self-positive rule. For every additional error before learning the self-positive rule, PHQ-9 scores increased by 0.17 points (b 95% CI: 0.08, 0.26, p < .001) and BDI-II scores increased by 0.31 points (b 95% CI: 0.15, 0.47, p < .001).

We also found weak evidence that worse learning of the friend being disliked was associated with greater PHQ-9 scores, and better learning of the self being disliked was associated with reduced BDI-II scores (Table 3). However, confidence intervals were relatively wide, and these effects were not observed in the alternative depression measure for each, suggesting unreliable effects.

Errors to criterion when learning that a friend was liked, or either rule about the stranger, were not associated with PHQ-9 or BDI-II scores (Table 3).

Cumulative Accuracy

Figure 2b demonstrates the cumulative mean accuracy over the 24 learning trials for the positive ‘like’ and negative ‘dislike’ rules about the self in participants grouped according to none, mild, and moderate to severe levels of depression on the PHQ-9 and BDI-II. In keeping with our findings for errors to criterion, participants with moderate to severe levels of depression demonstrated impaired learning of the self-like rule as indicated by lower levels of mean accuracy both initially and cumulatively across trials.

Global Ratings

After each rule we asked participants to provide a global rating of how much the computer 'liked' the person.

Demonstrating understanding of each rule, participants gave lower global ratings following completion of the negative versus positive rules (b = -2.67, 95% CI: -2.85, -2.49, p < .001).

Additionally, participants showed slightly increased perceptions of the friend being liked compared to the self (b = 0.32, 95% CI: 0.14, 0.50, p = 0.001), but gave similar global ratings in the self and stranger conditions (b = 0.09, b 95% CI: -0.10, 0.27, p = 0.354). Full results are available in supplementary Table S5.

Consistent with our findings for errors to criterion, increased perceptions of being liked after completing the self-positive rule were associated with lower depression severity (Table 3). We also found weak evidence that greater global ratings in the stranger-positive condition was associated with greater PHQ-9 scores, however there was little evidence of this association with BDI-II scores (Table 3).

Social Anxiety

The effects outlined above persisted when social anxiety was taken into account, suggesting an independent relationship between social evaluation learning and depression (Supplementary Table S6).

Table 3

Results from mixed-effect linear regression models examining the relationship between social evaluation learning task outcomes (predictors) and depression (Outcome: PHQ-9/BDI-II)

	PHQ-9				**BDI-II**			
	b	***b* 95% CI**	**β**	**95% CI**	**p**	***b***	***b* 95% CI**	**β**	**95% CI**	**p**
Bias Scores										
Intercept	8.54	7.47, 9.60	0.00	-0.15, 0.15	< .001	15.18	13.06, 17.30	0.00	-0.16, 0.15	< .001
Self	0.11	0.05, 0.17	0.13	0.06, 0.20	< .001	0.23	0.12, 0.34	0.13	0.07, 0.19	< .001
Friend	-0.03	-0.09, 0.03	-0.04	-0.11, 0.01	0.259	0.01	-0.10, 0.11	0.00	-0.05, 0.06	0.898
Stranger	-0.01	-0.08, 0.05	-0.01	-0.08, 0.05	0.731	0.00	-0.12, 0.11	0.00	-0.06, 0.05	0.943
Session	-0.88	-1.29, -0.46	-0.08	-0.12, -0.04	< .001	-0.73	-1.47, 0.02	-0.03	-0.06, 0.00	0.057
Errors to Criterion										
Intercept	7.45	5.91, 8.99	0.00	-0.15, 0.15	< .001	13.79	10.84, 16.73	0.00	-0.15, 0.15	< .001
Self-Positive	0.17	0.08, 0.26	0.13	0.06, 0.20	< .001	0.31	0.15, 0.47	0.12	0.06, 0.18	< .001

Self-Negative	-0.05	-0.13, 0.04	-0.04	-0.10, 0.03	0.264	-0.17	-0.32, -0.02	-0.06	-0.12, -0.01	0.031
Friend-Positive	0.03	-0.05, 0.16	0.02	-0.04, 0.08	0.492	0.01	-0.14, 0.16	0.00	-0.05, 0.05	0.916
Friend-Negative	0.08	0.05, 0.16	0.06	0.00, 0.12	0.038	-0.01	-0.15, 0.13	0.00	-0.06, 0.05	0.867
Stranger-Positive	-0.05	-0.13, 0.04	-0.03	-0.09, 0.03	0.294	0.02	-0.14, 0.17	0.01	-0.05, 0.06	0.840
Stranger-Negative	-0.03	-0.12, 0.06	-0.02	-0.09, 0.04	0.475	0.04	-0.13, 0.20	0.01	-0.04, 0.07	0.659
Session	-0.87	-1.29, -0.45	-0.08	-0.12, -0.04	< .001	-0.73	-1.49, 0.03	-0.03	-0.07, 0.00	0.062
Global Ratings										
Intercept	9.24	6.45, 12.02	0.00	-0.16, 0.16	< .001	17.77	12.48, 23.06	0.00	-0.16, 0.15	<.001
Self-Positive	-0.52	-0.82, -0.22	-0.12	-0.19, -0.05	0.001	-0.73	-1.29, -0.17	-0.08	-0.14, -0.02	0.012
Self-Negative	0.13	-0.17, 0.44	0.03	-0.04, 0.10	0.398	0.03	-0.54, 0.60	0.00	-0.06, 0.07	0.925
Friend-Positive	-0.04	-0.35, 0.28	-0.01	-0.07, 0.06	0.806	0.08	-0.51, 0.67	0.01	-0.05, 0.07	0.796

Friend-Negative	0.23	-0.04, 0.51	0.05	-0.01, 0.12	0.094	0.34	-0.16, 0.85	0.04	-0.02, 0.09	0.186
Stranger-Positive	0.32	0.03, 0.62	0.07	0.01, 0.14	0.033	-0.17	-0.72, 0.39	-0.02	-0.08, 0.04	0.554
Stranger-Negative	-0.16	-0.45, 0.13	-0.04	-0.10, 0.03	0.272	0.20	-0.34, 0.74	0.02	-0.04, 0.08	0.465
Session	-0.91	-1.33, -0.49	-0.08	-0.12, -0.04	< .001	-0.78	-1.57, 0.00	-0.03	-0.07, 0.00	0.052

b = unstandardised regression coefficients, β = standardised regression coefficients

Reliability of findings with clinical diagnosis of depression

To examine whether our findings were valid for participants meeting clinical diagnostic criteria for depression, we repeated the primary analyses for each task using logistic regression models with primary diagnosis of major depressive episode, derived from the CIS-R, as a binary outcome. The primary effects of each task were replicated; increased positive biases towards others in the self-esteem go/no-go task and reduced positive biases towards the self in the social evaluation learning task, were associated with an increased odds of meeting diagnostic criteria for a major depressive episode. Full details are available in supplementary materials (Supplementary Tables S7-S9).

Adjusting for Age and Gender

The results of our primary analyses were consistent when we adjusted for age and gender (Supplementary Table S10).

Discussion

Depression is characterised by differences in processing self-related information, which are believed to be related to emotion and reward cognition. However, the precise relationship between these areas of processing is not yet well understood. In this study we examined the role of the self in emotion and reward processing, separately and in interaction, in individuals experiencing varying levels of depression. Healthy individuals typically show enhanced positive perceptions of the self, relative to others (De Jong, 2002). We found that when the self was processed in relation to emotion and reward, this self-favouring bias was reduced in individuals with greater depression severity. However, when self, emotion and reward processing occurred independently there was little evidence of an association with depression.

Using a social evaluation learning task, we found evidence of an interaction between self, emotion, and reward processing with depression. During social interactions, healthy individuals preferentially incorporate positive evaluations into their self-concept (Korn, Prehn, Park, Walter, & Heekeren, 2012). In support of our pre-registered hypothesis, we found that participants with greater depression showed a reduced positive self-bias when learning social evaluations. Participants with greater depression made a greater number of errors before learning that they were 'liked' and gave lower global ratings of being liked. Depression was therefore consistently associated with a reduced ability to learn positive, socially rewarding information about the self.

Using a go/no-go task, we found that individuals with greater depression severity showed increased sensitivity to positive words in relation to others, and decreased sensitivity to negative words. However, in keeping with previous research using response inhibition tasks we found only weak evidence of an association between implicit self-esteem and depression (De Jong, Sportel, De Hullu, & Nauta, 2012; Franck, De Raedt, & De Houwer, 2008; Van Tuijl, De Jong, Sportel, De Hullu, & Nauta, 2014). Depression was therefore characterised by increased positive 'other-esteem', but not by an increased negative self-esteem. Our research adds to evidence suggesting that individuals with depression tend to perceive others more positively (Kuiper, Derry, & MacDonald, 1982). Depression has previously been theorised to originate from discrepancies between internal self-representations, and representations of the ideal self (Higgins, 1987). Enhanced positive perceptions of others may increase discrepancies between views of the actual and idealised self, perpetuating depressive symptoms. Alternatively, our findings of a weak association between implicit self-esteem and

depression may reflect debate over the construct validity of implicit association tests (Hahn, Judd, Hirsh, & Blair, 2014), or questions over the extent to which affective response inhibition are associated with depression severity (Lewis et al., 2020).

When the self was processed independently of emotion or reward, within an associative learning task, we did not find evidence of changes in self-prioritisation with greater depression severity. This contrasts with previous findings of reduced self-prioritisation following negative mood induction (Sui et al., 2016). Whilst temporary, sudden changes in state mood may inhibit self-prioritisation in the absence of emotional processing, this does not seem to apply to low trait mood. We also found no evidence that depression was associated with differences in learning emotional associations when processed independently of the self. There were some indications of differences in reward learning associated with depression. Although, in contrast to our expectations this was only observed for medium levels of reward. It is possible that depression alters sensitivity to reward, with greater value being placed on lower levels of reward. However, confidence intervals were relatively wide for this effect. Further research replicating these results is therefore required in order to understand their importance.

A substantial body of research suggests that healthy individuals hold relatively enhanced perceptions of the self versus others (Kuiper et al., 1982), and typically rate their abilities as better-than-average (Zell, Strickhouser, Sedikides, & Alicke, 2019). These positive self-biases are believed to be beneficial for mental health in increasing self-esteem and confidence (Button et al., 2015). Our results indicate that when processed independently of emotion, at least at a 'cold' perceptual level as in the associative learning task, self-referential processing is similar irrespective of depression severity. However, differences were observed when integrating positive and negative information with the self and others. Overall, depression was characterised by a reduction in self-favouring biases. Individuals with greater depression showed both greater implicit positive perceptions of others, and impaired learning of positive associations with the self. Depression may be driven by other-favouring biases strengthened by reduced learning of positive information about the self. In combination, reduced positive perceptions of the self and enhanced positive perceptions of others are likely to maintain negative views of the self.

Clinical Implications

Acknowledging that much of the work in therapy already implies self-reference, our findings suggest that it may be beneficial to explicitly manipulate referential focus and target biases in emotion and reward processing in relation to the self. Social evaluation learning in particular may be an important target for intervention. Depression is associated with poorer quality social interactions (Teo, Choi, & Valenstein, 2013), and social withdrawal (Hirschfeld et al., 2000). Our findings suggest that individuals with depression show a stable pattern of reduced learning of positive evaluations about the self. Reduced positive self-biases in social interactions are likely to maintain negative perceptions of the self, reinforcing social withdrawal and increasing the likelihood of poor social relationships, subsequently maintaining depression symptoms (Lewinsohn, Mischel, Chaplin, & Barton, 1980). Social evaluation learning provides an important and potentially reversible target for therapeutic intervention that can address impairments in social functioning, negative perceptions of the self, and wider depressive symptoms. It is also possible that social evaluation learning may be a transdiagnostic mechanism. Future research examining latent mental health traits would allow us to understand the importance of social evaluation learning across mental health disorders.

Additionally, we found evidence that the relationship between biased learning about the self and depression was consistent across testing sessions. Change in social evaluation learning may therefore be a viable predictor of change in depressive symptoms. Individual treatment response for depression is varied (Maslej et al., 2020). It is currently difficult to predict which treatments are effective at an individual level (Simon & Perlis, 2010). Exacerbating these difficulties are the long time periods between commencing treatment and improvement in mood (Uher et al., 2011). Identifying markers of therapeutic change would be beneficial in allowing identification of effective treatments at an earlier timepoint. Further research examining changes in learning positive evaluations about the self as a potential predictor of treatment response would be beneficial.

Limitations

We recruited participants based on depression severity to gain a balanced range of depression. However, in the time between screening and testing, depression severity on average decreased potentially weakening our effects. In-depth analysis of larger samples representative of the spectrum of individuals with depression would be fruitful to further characterise changes in self-referential processing and to replicate the current findings.

Although, our results were replicated for individuals meeting diagnostic criteria for depression, suggesting that our results are reliable for greater severities of depression.

Additionally, whilst our sample was representative of the range of depressive symptoms experienced in the general population it was limited in its demographic diversity. Participants were predominantly young, students and female. While this may be an ideal sample to investigate the role of self biases in depression, given the worrying increase of depression in this population at a developmentally sensitive time where self-identity and peer relations are evolving (Blakemore & Mills, 2014; Royal College of Psychiatrists, 2011), future studies should investigate whether these finding generalise across the wider population and test whether the strength of the associations alter across adulthood.

Whilst we found evidence of a consistent relationship between biased learning about the self and depression in the social evaluation learning task, bias scores themselves showed limited reliability between test sessions. Further development of this task to improve reliability would be beneficial.

Finally, this was a cross-sectional study examining the association between self, emotion, and reward processing with depression. We are therefore unable to comment on the causal role of self-processing in relation to emotion and reward. Future research examining the longitudinal relationship between self-processing and depression would provide insight into the potential causal role of reduced positive self-biases. Additionally, manipulating self-referential affective processing through cognitive bias modification would help us understand the importance of this cognitive style in maintaining depression symptoms.

Conclusion

Overall, our findings suggest that depression is characterised by enhanced positive implicit associations with others, and reduced positive learning about the self, culminating in reduced self-favouring biases observed in healthy individuals. We also found some evidence of altered sensitivity to reward in individuals with greater depression severity using a simple associative learning task, although this effect requires further replication. Treatments targeting reduced positive self-biases may provide more sensitive targets for therapeutic intervention and potential biomarkers of treatment responses, allowing the development of more effective interventions.

References

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry*, *165*(8), 969–977. https://doi.org/10.1176/appi.ajp.2008.08050721

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for the Beck depression inventory-II. *San Antonio, TX: Psychological Corporation*.

Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Button, K. S., Browning, M., Munafò, M. R., & Lewis, G. (2012). Social inference and social anxiety: Evidence of a fear-congruent self-referential learning bias. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(4), 1082–1087. https://doi.org/10.1016/j.jbtep.2012.05.004

Button, K. S., Karwatowska, L., Kounali, D., Munafò, M. R., & Attwood, A. S. (2016). Acute anxiety and social inference: An experimental manipulation with 7.5% carbon dioxide inhalation. *Journal of Psychopharmacology*, *30*(10), 1036–1046. https://doi.org/10.1177/0269881116653105

Button, K. S., Kounali, D., Stapinski, L., Rapee, R. M., Lewis, G., & Munafò, M. R. (2015). Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. *PLoS ONE*, *10*(4), e0119456. https://doi.org/10.1371/journal.pone.0119456

Cameron, I. M., Crawford, J. R., Lawton, K., & Reid, I. C. (2008). Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *British Journal of General Practice*, *58*(546), 32–36. https://doi.org/10.3399/bjgp08X263794

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., … Geddes, J. R. (2018). Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *FOCUS*, *16*(4), 420–429. https://doi.org/10.1176/appi.focus.16407

Cuijpers, P., Andersson, G., Donker, T., & Van Straten, A. (2011). Psychological treatment

of depression: Results of a series of meta-analyses. *Nordic Journal of Psychiatry*, *65*(6), 354–364. https://doi.org/10.3109/08039488.2011.596570

Cunningham, S. J., & Turk, D. J. (2017). A review of self-processing biases in cognition. *Quarterly Journal of Experimental Psychology*, *70*(6), 987–995. https://doi.org/10.1080/17470218.2016.1276609

Dalgleish, T., & Watts, F. N. (1990). Biases of attention and memory in disorders of anxiety and depression. *Clinical Psychology Review*, *10*(5), 589–604. https://doi.org/10.1016/0272-7358(90)90098-U

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

De Jong, P. J. (2002). Implicit self-esteem and social anxiety: Differential self-favouring effects in high and low anxious individuals. *Behaviour Research and Therapy*, *40*(5), 501–508. https://doi.org/10.1016/S0005-7967(01)00022-5

De Jong, P. J., Sportel, B. E., De Hullu, E., & Nauta, M. H. (2012). Co-occurrence of social anxiety and depression symptoms in adolescence: Differential links with implicit and explicit self-esteem? *Psychological Medicine*, *42*(3), 475. https://doi.org/10.1017/S0033291711001358

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, *68*(2), 118–124. https://doi.org/10.1016/j.biopsych.2010.01.027

Everaert, J., Duyck, W., & Koster, E. H. W. (2014). Attention, interpretation, and memory biases in subclinical depression: A proof-of-principle test of the combined cognitive biases hypothesis. *Emotion*, *14*(2), 331–340.

Franck, E., De Raedt, R., & De Houwer, J. (2008). Activation of latent self-schemas as a cognitive vulnerability factor for depression: The potential role of implicit self-esteem. *Cognition and Emotion*, *22*(8), 1588–1599. https://doi.org/10.1080/02699930801921271

Gaddy, M. A., & Ingram, R. E. (2014). A meta-analytic review of mood-congruent implicit memory in depressed mood. *Clinical Psychology Review*, *34*(5), 402–416. https://doi.org/10.1016/j.cpr.2014.06.001

Gregg, A. P., & Sedikides, C. (2010). Narcissistic fragility: Rethinking its links to explicit

and implicit self-esteem. *Self and Identity*, *9*(2), 142–161. https://doi.org/10.1080/15298860902815451

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*(3), 1369–1392. https://doi.org/10.1037/a0035028

Hertel, P. T., & El-Messidi, L. (2006). Am I Blue? Depressed Mood and the Consequences of Self-Focus for the Interpretation and Recall of Ambiguous Words. *Behavior Therapy*, *37*(3), 259–268. https://doi.org/10.1016/j.beth.2006.01.003

Higgins, E. T. (1987). Self-Discrepancy: A Theory Relating Self and Affect. *Psychological Review*, *94*(3), 319. https://doi.org/10.1037/0033-295X.94.3.319

Hirschfeld, R., Montgomery, S., Keller, M., Kasper, S., Schatzberg, A., Moller, H.-J., … Bourgeois, M. (2000). Social functioning in depression: a review. *Journal of Clinical Psychiatry*, *61*(4), 268–275.

Hobbs, C., Faraway, J., Kessler, D., Munafò, M. R., Sui, J., & Button, K. S. (2019). Social evaluation learning in social anxiety disorder and depression: a mega-analysis presented at the Summer Meeting of the Bristish Association for Psychopharmacology. *Journal of Psychopharmacology*, *Supplement*(8), A33.

Hobbs, C., Sui, J., Kessler, D., Munafò, M. R., & Button, K. S. (2020). *Dataset for 'Self processing in relation to emotion and reward processing in depression*. University of Bath Research Data. https://doi.org/https://doi.org/10.15125/BATH-00924

Hsu, K. J., McNamara, M. E., Shumake, J., Stewart, R. A., Labrada, J., Alario, A., … Beevers, C. G. (2020). Neurocognitive predictors of self-reported reward responsivity and approach motivation in depression: A data-driven approach. *Depression and Anxiety*, *37*, 682–697. https://doi.org/10.1002/da.23042

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 1–16. https://doi.org/10.1186/2045-5380-3-12

Ji, J. L., Grafton, B., & MacLeod, C. (2017). Referential focus moderates depression-linked attentional avoidance of positive information. *Behaviour Research and Therapy*, *93*, 47–54. https://doi.org/10.1016/j.brat.2017.03.004

Kendrick, T., Dowrick, C., McBride, A., Howe, A., Clarke, P., Maisey, S., … Smith, P. W. (2009). Management of depression in UK general practice in relation to scores on depression severity questionnaires: Analysis of medical record data. *BMJ (Online)*, *338*(7697), b750. https://doi.org/10.1136/bmj.b750

Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., … Stringaris, A. (2018). Reward processing in depression: A conceptual and meta-analytic review across fMRI and EEG studies. *American Journal of Psychiatry*, *175*(11), 1111–1120. https://doi.org/10.1176/appi.ajp.2018.17101124

Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *Journal of Neuroscience*, *32*(47), 16832–16844. https://doi.org/10.1523/JNEUROSCI.3016-12.2012

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613.

Kuiper, N. A., Derry, P. A., & MacDonald, M. R. (1982). Self-reference and person perception in depression: A social cognition perspective. In *Integrations of clinical and social psychology* (pp. 79–2103).

Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, *9*(3), 371–375. https://doi.org/10.1177/0146167283093007

Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology*, *89*(2), 203. https://doi.org/10.1037/0021-843X.89.2.203

Lewis, G., Button, K. S., Pearson, R. M., Munafò, M. R., & Lewis, G. (2020). Inhibitory control of positive and negative information and adolescent depressive symptoms: a population-based cohort study. *Psychological Medicine*, 1–11. https://doi.org/https://doi.org/10.1017/S0033291720002469

Lewis, G., Pelosi, A., Araya, R., & Dunn, G. (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine*, *22*(2), 465–486. https://doi.org/10.1017/S0033291700030415

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit

positive association theory of self-advantage during implicit face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 619–633. https://doi.org/10.1037/a0015797

Maslej, M. M., Furukawa, T. A., Cipriani, A., Andrews, P. W., & Mulsant, B. H. (2020). Individual differences in response to antidepressants: a meta-analysis of placebo-controlled randomized clinical trials. *JAMA Psychiatry*, *77*(6), 607–617. https://doi.org/10.1001/jamapsychiatry.2019.4815

Northoff, G. (2007). Psychopathology and pathophysiology of the self in depression - Neuropsychiatric hypothesis. *Journal of Affective Disorders*, *104*(1–3), 1–14. https://doi.org/10.1016/j.jad.2007.02.012

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*(11), 1019–1025. https://doi.org/10.1016/j.biopsych.2010.12.014

Parsons, S., Kruijt, A.-W., & Fox, E. (2019). Psychological Science Needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in Methods and Practices in Psychological Science*, *2*(4), 378–395. https://doi.org/10.1177/2515245919879695

Rosenberg, M. (1965). Rosenberg self esteem scale. *Personality and Individual Differences*. https://doi.org/10.1007/s12671-015-0407-6

Royal College of Psychiatrists. (2011). *Mental health of students in higher education*. https://www.rcpsych.ac.uk/docs/default-source/improving-care/better-mh-policy/college-reports/college-report-cr166.pdf?sfvrsn=d5fa2c24_2

Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., … Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1942–1947. https://doi.org/10.1073/pnas.0812686106

Simon, G. E., & Perlis, R. H. (2010). Personalized medicine for depression: Can we match patients with treatments? *American Journal of Psychiatry*, *167*(12), 1445–1455. https://doi.org/10.1176/appi.ajp.2010.09111680

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder. *Archives of Internal Medicine*, *166*(10), 1092–

1097. https://doi.org/10.1001/archinte.166.10.1092

Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, *70*(6), 1011–1022. https://doi.org/10.1080/17470218.2015.1101477

Sui, J., & Button, K. S. (2017). Associative Learning and Depression. *Unpublished Data*.

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1105. https://doi.org/10.1037/a0029792

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

Sui, J., & Humphreys, G. W. (2015b). The interaction between self-bias and reward: Evidence for common and distinct processes. *Quarterly Journal of Experimental Psychology*, *68*(10), 1952–1964. https://doi.org/10.1080/17470218.2015.1023207

Sui, J., Ohrling, E., & Humphreys, G. W. (2016). Negative mood disrupts self- and reward-biases in perceptual matching. *Quarterly Journal of Experimental Psychology*, *69*(7), 1438–1448. https://doi.org/10.1080/17470218.2015.1122069

Takano, K., Van Grieken, J., & Raes, F. (2019). Difficulty in updating positive beliefs about negative cognition is associated with increased depressed mood. *Journal of Behavior Therapy and Experimental Psychiatry*, *64*, 22–30. https://doi.org/10.1016/j.jbtep.2019.02.001

Teo, A. R., Choi, H. J., & Valenstein, M. (2013). Social Relationships and Depression: Ten-Year Follow-Up from a Nationally Representative Study. *PLoS ONE*, *8*(4), e62396. https://doi.org/10.1371/journal.pone.0062396

Tomitaka, S., Kawasaki, Y., & Furukawa, T. (2015). A distribution model of the responses to each depressive symptom item in a general population: A cross-sectional study. *BMJ Open*, *5*(9), e008599. https://doi.org/10.1136/bmjopen-2015-008599

Uher, R., Mors, O., Rietschel, M., Rajewska-Rager, A., Petrovic, A., Zobel, A., … McGuffin, P. (2011). Early and Delayed Onset of Response to Antidepressants in Individual

Trajectories of Change During Treatment of Major Depression. *The Journal of Clinical Psychiatry*, *72*(11), 1586–1592. https://doi.org/10.4088/jcp.10m06419

Van Tuijl, L. A., De Jong, P. J., Sportel, B. E., De Hullu, E., & Nauta, M. H. (2014). Implicit and explicit self-esteem and their reciprocal relationship with symptoms of depression and social anxiety: A longitudinal study in adolescents. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(1), 113–121. https://doi.org/10.1016/j.jbtep.2013.09.007

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Brazilian Journal of Psychiatry*, *35*(4), 416–431. https://doi.org/10.1590/1516-4446-2012-1048

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063. https://doi.org/10.1037//0022-3514.54.6.1063

Weeks, J. W., Heimberg, R. G., & Rodebaugh, T. L. (2008). The Fear of Positive Evaluation Scale: Assessing a proposed cognitive component of social anxiety. *Journal of Anxiety Disorders*, *22*(1), 44–55. https://doi.org/10.1016/j.janxdis.2007.08.002

Weissman, A. N., & Beck, A. T. (1978). Development and Validation of the Dysfunctional Attitude Scale: A Preliminary Investigation. *Acceptance and Commitment Therapy*, *54*.

Williams, B. J., & Kaufmann, L. M. (2012). Reliability of the Go/No Go Association Task. *Journal of Experimental Social Psychology*, *48*(4), 879–891. https://doi.org/10.1016/j.jesp.2012.03.001

World Health Organization. (2011). Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level. *EB130/9*. https://doi.org/https://apps.who.int/gb/ebwha/pdf_files/eb130/b130_9-en.pdf

World Health Organization. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2019). The Better-Than-Average Effect in Comparative Self- Evaluation: A Comprehensive Review and Meta-Analysis.

Psychological Bulletin, *146*(2), 118. https://doi.org/10.1037/bul0000218

This declaration concerns the article entitled:			
Individual differences in social evaluation learning and their association with psychopathology			
Publication status (tick one) **Draft manuscript** [X] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [ ]			
Publication details (reference)	Hobbs, C., Hoffmann, J., Hopkins, A., Moutoussis, M., Button, K.S. (2022). Individual differences in social evaluation learning and their association with psychopathology. *Manuscript in preparation.*		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to the formulation of ideas (25%), design of methodology (25%), and experimental work including creation of study materials (25%), participant recruitment (50%), data collection (50%), data cleaning, analysis, visualisation, and archiving (100%), interpretation of results (75%). The candidate predominantly executed the presentation of the data in journal format (90%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	05/01/2022

Chapter 2: Individual differences in social evaluation learning and their association with psychopathology

Catherine Hobbs [a], Janina Hoffmann [a], Alexandra Hopkins [b c], Michael Moutoussis [b c], Katherine S Button [a]

[a] Department of Psychology, University of Bath

[b] Max-Planck – UCL Center for Computational Psychiatry and Ageing Research, University College London

[c] Wellcome Trust Centre for Human Neuroimaging, University College London

Chapter Rationale

In chapter 1 I found evidence that participants with greater levels of depression showed a reduced positive self-bias when learning social evaluations. Participants with greater depression made a greater number of errors before learning that they were 'liked'. In chapter 2 I aimed to validate these findings in a larger sample of individual participant data from ten studies previously conducted within this research group (n = 552) and a large independent dataset of participants recruited online (n = 807). Additionally, as previous research has reported evidence of increased learning of negative evaluations in individuals with greater social anxiety [1,2] I investigated the potential unifying and differential mechanisms underlying the relationship between social evaluation learning with depression and social anxiety

Open Science Statement

This study was pre-registered on Open Science Framework (https://osf.io/ke3d5), where study materials are also publicly available (https://osf.io/utyw5/). Dataset 2 is openly accessible via the University of Bath research data archive (https://doi.org/10.15125/BATH-01108). Dataset 1 is available upon request as participants did not provide informed consent to publish data as open access.

Abstract

Introduction: Depression and social anxiety are characterised by negative views of the self and social difficulties. Differences in learning of social evaluations may drive these impairments. Whereas healthy individuals are more sensitive to positive versus negative social feedback, individuals with depression and social anxiety are thought to lack this positive bias. We therefore investigated patterns of social evaluation learning associated with depression and social anxiety symptoms.

Method: Using two individual datasets to assess the reliability of our findings (n = 552; n = 807), we evaluated the association between learning of social evaluations on a computer-based reinforcement task and depression (PHQ-9) and social anxiety (BFNE) questionnaires, within structural equations models. We examined overall biases and specific patterns of learning in depression and social anxiety.

Results: Participants that were better at learning negative versus positive social evaluations about the self had greater depression and social anxiety. We found reliable evidence that for social anxiety this was driven by an increased learning of negative evaluations. In contrast, depression was associated with both increased learning of negative evaluations and decreased learning of positive evaluations. However, for depression this relationship was attenuated between datasets. Whilst we found strong evidence of this association in lab-based data (dataset 1), statistical evidence of this association was weakened in participants recruited online (dataset 2).

Discussion: In line with cognitive theories highlighting the role of threat-related biases, social anxiety was characterised by a heightened sensitivity to negative social feedback. In contrast, depression was associated with both heightened sensitivity to negative feedback as well as decreased sensitivity to positive feedback. These findings may reflect negative self-schema characteristic of depression as well as impaired reward processing. Treatments targeting these processes within each disorder may be beneficial.

Introduction

Depression and social anxiety are highly co-morbid disorders that significantly contribute to the global burden of disease [3,4]. Both disorders are characterised by negative views of the self [5,6] and widespread social difficulties [7,8]. Developing a detailed understanding of the nature of social cognitive biases sustaining negative self-beliefs in depression and social anxiety would inform our knowledge of the mechanisms underlying these disorders and help to identify potential targets for intervention.

Social interactions are dynamic and complex, with social behaviours (e.g., a smile, praise) and social inferences (e.g., "this person likes me") dependent on both the actual feedback received and subjective interpretation of the feedback. Perceptions of the self are largely informed by how we believe others view us [9]. During social interactions, healthy individuals preferentially update their self-concept in response to positive versus negative feedback [10]. These positive social biases are believed to be beneficial for mental health in increasing confidence and self-esteem [11]. In contrast, individuals with depression and social anxiety lack positive social biases, tending to believe that others judge them negatively [12,13] and underestimating their social skills [14,15]. Greater sensitivity to negative relative to positive information in social interactions is likely to maintain maladaptive views of the self characteristic of social anxiety and depression.

However, it is possible that different mechanisms underlie social evaluation learning biases in depression and social anxiety. According to cognitive models, individuals experiencing social anxiety construct negative mental images of how they appear to others based on maladaptive beliefs of their social skills (e.g., "I am stupid") and other's expectations (e.g., "They will criticise me"). These self-beliefs contribute to enhanced sensitivity to socially threatening information which then reinforce negative expectations in a cyclical process [6,16]. In keeping with this theory, individuals experiencing social anxiety show greater sensitivity to negative facial emotions [17,18]. Although research has also documented decreased biases towards positive social information[19–21], in line with cognitive theories of social anxiety we may expect socially anxious individuals to predominantly demonstrate an increased learning of negative social evaluations about the self.

Whilst cognitive theories of depression propose a similar causal role for negative self-beliefs [5], depression also appears to be characterised by a reduced sensitivity to reward. Individuals experiencing depression show impaired learning from positive feedback in probabilistic

learning tasks [22] and are less willing to expend effort for reward [23]. Depression is also associated with reduced activation in the striatum, an area of the brain critical to reward systems, when anticipating reward or receiving rewarding feedback [24]. Although some research has also reported evidence of an increased sensitivity to punishing information in depression [25], impairments appear to be more reliably observed for reward [26,27]. Within social contexts, individuals with depression report social anhedonia – a loss of interest or pleasure from social interactions [28]. We may therefore expect depression to be characterised by reduced learning of positive evaluations about the self.

Identifying specific maladaptive patterns of learning associated with social anxiety and depression is important in furthering our causal understanding of these disorders. Through understanding the nature of underlying mechanisms, we may be able to develop more effective treatments by addressing causal factors rather than symptoms. A transdiagnostic approach in understanding and treating depression and anxiety has previously been advocated [29]. However, if different social cognitive mechanisms underly depression and anxiety it may be more effective to tailor individual treatments for each disorder. It is therefore important that we understand both the unifying and differential factors underlying social evaluation learning in depression and social anxiety.

We aimed to examine learning of social evaluations in individuals experiencing varying levels of social anxiety and depression. Using two independent datasets to establish the reliability of our findings, we firstly evaluated the extent to which social anxiety and depression are associated with an overall bias in learning social evaluations. We tested the hypothesis that depression and social anxiety would be associated with better learning of negative relative to positive social evaluations about the self. We then evaluated the specific patterns underlying biased learning associated with each disorder, testing the hypothesis that social anxiety would be associated with better learning of a negative 'dislike' rule about the self, whereas depression would be associated with impaired learning of a positive 'like' rule about the self.

Methods

This study was pre-registered on Open Science Framework (https://osf.io/ke3d5), where study materials are also publicly available (https://osf.io/utyw5/). Dataset 2 is openly accessible via the University of Bath research data archive (https://doi.org/10.15125/BATH-01108). Dataset 1 is available upon request as participants did not provide informed consent to publish data as open access.

Procedure

We created two datasets for this study. Firstly, we collated individual participant data from ten studies previously collected within our research group (dataset 1). Where studies included multiple testing sessions, we included only data from session one to reduce potential practice effects. Additionally, for two studies that used pharmacological manipulations we used only data from control conditions. To assess the reliability of our findings from dataset 1, we collected a second independent dataset using the online recruitment platform Prolific (dataset 2). In both datasets participants completed a social evaluation learning task and self-report measures of mood. For dataset 1 participants completed the study procedure within psychology labs under the supervision of researchers. For dataset 2 participants independently completed the study procedure remotely using online survey [30] and cognitive task software [31]. Dataset characteristics are summarised in Table 1.

Subjects

All studies within dataset 1 recruited participants from the local community, although inclusion criteria and recruitment methods varied (Table 1). In four studies participants were recruited based on levels of social anxiety or depression symptoms. The remaining studies recruited healthy volunteers.

For dataset 2, we recruited participants aged 18 to 65 who were current residents of the United Kingdom using the online recruitment platform Prolific [32]. We restricted participants to individuals who spoke English as a first language and had no literacy difficulties to ensure comprehension of the task. We also restricted participants to individuals that had completed 5 or more prolific studies with a ≥ 98% acceptance rate to ensure data quality.

Materials

Social Evaluation Learning

We measured learning of social evaluations using a two-alternative forced choice reinforcement computer-based task. Participants were asked to learn whether the computer liked the self and a fictional other based on the computers' feedback to selection of a positive or negative social evaluative word pairs (Figure 1). Feedback contingencies were manipulated to reflect learning of three rules, positive ('like'), negative ('dislike') and neutral, based on the probability of 'correct' feedback following selection of a positive evaluation (positive: 60-80%, negative: 20-40%, neutral: 50%). Characteristics of the task according to each dataset are outlined in Table 1.

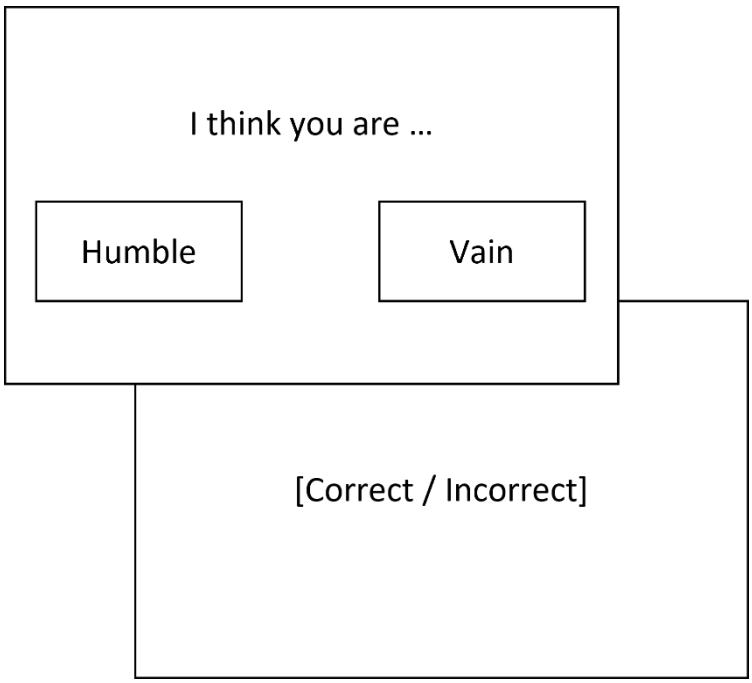


Figure 1

Example of a trial in the self-referential condition on the social evaluation learning task. Participants were asked to learn how much the computer liked them based on feedback to selection of positive or negative evaluative words. The proportion of trials with 'correct' feedback after selection of the positive word was manipulated to reflect learning of three possible rules: positive 'like' 60-80%, negative 'dislike' 20-40%, and neutral 50%.

Self-Report Measure of Mood and Personality

We measured depression severity using the PHQ-9, a nine-item measure of DSM-IV depression symptoms within the previous two weeks [33]. The PHQ-9 has good psychometric properties and is widely used in both clinical settings and the general population [34,35].

We measured social anxiety using the Brief Fear of Negative Evaluation Scale (BFNE) [36]. The BFNE is a 12 item measure of fear or worries in social contexts, and has been validated for use in community and clinical samples [37,38].

Data Analysis

We measured participants' ability to learn the positive and negative rules by calculating the average number of errors made before eight-consecutive rule-congruent responses (errors to criterion) for each referential-condition rule. Greater errors to criterion indicate worse learning. From this we then calculated overall bias scores by subtracting errors to criterion in the negative rule from the positive rule for each referential condition. Greater bias scores indicate relatively better learning of the negative relative to the positive rule. As the number of trials varied across studies in dataset 1 from 20 - 32, we used only the first 20 trials for each referential condition-rule block across datasets.

We firstly used mixed-effects linear regression models to evaluate differences in learning of social evaluations by referential condition and rule, irrespective of depression and anxiety. Errors to criterion was entered as the outcome, and referential condition, rule, and an interaction term between referential condition and rule were entered as predictors. We conducted individual models for each dataset. We included subject clustered within study as a random effect for dataset 1, and subject as a random effect for dataset 2.

Hypothesis: depression and social anxiety will be associated with better learning of negative relative to positive social evaluations about the self.

We then conducted multivariate linear regressions within a structural equation modelling framework to assess the relationship between bias scores and depression and social anxiety severity. Bias scores for the self and other condition were entered as exogenous (predictor) variables, and total PHQ-9 and BFNE scores as individual endogenous (outcome) variables. This model allowed us to examine the individual relationships between social evaluation learning with both depression and social anxiety whilst accounting for symptomatic comorbidity.

Hypothesis: Social anxiety will be associated with better learning of a negative 'dislike' rule about the self, whereas depression will be associated with impaired learning of a positive 'like' rule about the self.

To examine the mechanisms underlying the relationship between social evaluation learning, depression, and social anxiety, we then conducted additional multivariate regression models examining errors to criterion in each referential condition-rule block. Errors to criterion in each of the referential condition-rule combinations were entered as exogenous variables, and PHQ-9 and BFNE scores were entered as endogenous variables in a structural equation model.

Separate statistical models were created for dataset 1 and 2 for all analyses outlined above. In structural equation models for dataset 1 we included study as a random effect to account for potential clustering of data within studies.

Table 1

Characteristics of study populations and the social evaluation learning task from dataset 1 by study and dataset 2

				Social Evaluation Learning Task				
Dataset	**Recruitment**	**N**	**Inclusion Criteria**	**Referential Conditions**	**Feedback Contingencies**	**Trials per Block**	**Total Trials**	**Test Sessions ***
Dataset 1								
SEL1 [1]	Low (BFNE < 32) or High (BFNE > 43) Social Anxiety	102	Female, aged 18-50, no current antidepressant medication.	Self	Positive (80%), Negative (20%)	32	64	1
SEL2 [2]	Low (BFNE < 32), Moderate (BFNE 32-45) or High (BFNE > 45) Social Anxiety	100	Aged 18-50, no current psychiatric medication, English as a first language.	Self, Other	Positive (80%), Neutral (50%), Negative (20%)	32	192	1
SEL3 [39]	Low (BFNE ≤ 26) or High (BFNE ≥ 50) Social Anxiety	48	Aged 18-50, English as a first language.	Self, Other	Positive (80%), Negative (20%)	32	128	1
SEL4[a] [40]	Community	48	Aged 18-50, English as a first language, no significant current or previous medical or psychiatric illness.	Self, Other	Positive (80%), Negative (20%)	32	128	1[a]

SEL5 [41]	Community	53	Aged 18-50, English as a first language.	Self, Other	Positive (80%), Negative (20%)	32	192	1
SEL6 [42]	Community	17	Aged ≥ 16, English as a first language.	Self, Other	Positive (60%, 70%, 80%), Negative (20%, 30%, 40%)	20	240	1
SEL7 [43]	Community	8	Aged 18-35, normal or corrected-to-normal vision, right-handed.	Self, Other	Positive (80%), Neutral (50%), Negative (20%)	32	192	3
SEL8 [43]	Community	11	Aged 18-35, normal or corrected-to-normal vision, right-handed.	Self, Other	Positive (80%), Neutral (50%), Negative (20%)	32	192	3
SEL9 [44]	None (PHQ-9 < 5), Mild (PHQ-9 5-9), or Moderate to Severe (PHQ-9 ≥ 10) Depression	144	Aged 18-65, normal or corrected-to-normal vision, fluent in English.	Self, Friend[b], Stranger	Positive (60%, 80%), Negative (20%, 40%)	24	288	2
SEL10[a] [45]	Community	21	Aged 18-45, normal or corrected-to-normal vision, fluent in English, no significant current or previous medical or psychiatric	Self, Friend[b], Stranger	Positive (60%, 80%), Negative (20%, 40%)	24	288	1

			illness, no current or past drug or alcohol dependency, no current use of psychoactive medication, no current pregnancy or breastfeeding, ≤ 5 cigarettes per day, ≤ 6 caffeinated drinks per day, lactose intolerance, no recreational psychoactive drug use within previous 3 months.					
Dataset 2								
	Online (Prolific)	807	Aged 18-65, current resident in the UK, English as a first-language, no literacy difficulties, ≥ 5 Prolific studies completed with ≥ 98%	Self, Other	Positive (80%), Neutral (50%), Negative (20%)	32	192	1

	acceptance rate.

[*] To reduce potential order effects where studies had multiple data collection sessions we included only data from the first testing session.

[a] As these studies used pharmacological manipulations, only data from the control (placebo) conditions were included in dataset 1.

[b] Data from the 'friend' condition was not included in dataset 1.

Results

Sample

Dataset 1 included data from 552 participants collated across ten studies (Table 1). For dataset 2, we recruited 1062 individuals using the online recruitment platform Prolific [32]. Of these, 836 (79%) participants completed the full study procedure, 5 (0.6%) of whom were excluded for failing attention checks. A further 24 (3%) participants were excluded due to technical errors, leaving a final sample of 807.

Participants in dataset 2 were on average older, and a lower proportion were female and students. Depression and social anxiety severity were similar between datasets (Table 2).

Table 2

Sample Characteristics

	Dataset 1	**Dataset 2**
N	552	807
Age, M (SD)	23.2 (6.8)	35.1 (12.6)
Female, N (%)	405 (80)	455 (57)
Ethnicity, N (%)		
White	312 (75)	698 (87)
Ethnic Minority	102 (25)	109 (13)
Occupation, N (%)		
Employed	92 (20)	524 (65)
Student	348 (77)	122 (15)
Other	11 (3)	161 (20)
PHQ-9[a]		
M (SD)	6.0 (5.0)	6.8 (5.3)
None (< 5), N (%)	244 (48)	338 (42)
Mild (5-9), N (%)	159 (32)	265 (33)
Moderate to Severe (≥ 10), N (%)	101 (20)	204 (25)
BFNE		
M (SD)	37.8 (11.4)	37.9 (11.0)

[a] PHQ-9 was not collected for one study within dataset 1 (n = 48)

PHQ-9 = Patient Health Questionnaire; BFNE = Brief Fear of Negative Evaluation

Social Evaluation Learning

On average, participants in both datasets demonstrated a positive learning style, making a greater number of errors learning the negative relative to the positive rule. This effect was greatest when learning about the self (Tables 3-4).

Table 3

Mean errors to criterion and bias scores (positive – negative) in the social evaluation learning task by dataset. Greater bias scores indicate better learning of the negative rule relative to the positive rule.

	Dataset 1 (n = 552)		**Dataset 2 (n = 807)**	
	Self	Other	Self	Other
Errors to Criterion, M (SD)				
Positive	4.15 (3.53)	4.19 (3.53)	3.08 (3.47)	3.28 (3.71)
Negative	6.53 (4.32)	5.92 (3.87)	6.44 (5.36)	5.64 (4.71)
Bias Scores, M (SD)	-2.37 (5.64)	-1.73 (4.92)	-3.35 (6.73)	-2.36 (6.11)

Table 4

Results from mixed-effects linear regression models examining differences in errors to criterion (outcome) by condition and rule (predictors)

	Dataset 1 (n = 552)			**Dataset 2 (n = 807)**		
	β	*95% CI*	*p*	*β*	*95% CI*	*p*
Intercept	-0.33	-0.53, -0.12	< .001	-0.33	-0.40, -0.26	< .001
Condition	0.06	-0.05, 0.17	0.303	0.04	-0.05, 0.13	0.363
Rule	0.60	0.49, 0.70	< .001	0.73	0.64, 0.82	< .001
Condition x Rule	-0.16	-0.32, -0.01	0.041	-0.21	-0.34, -0.09	0.001

Reference categories: Condition – Self, Rule – Positive

Association with Depression and Social Anxiety

Visual inspection of linear relationships indicated that extent of positively biased learning about the self declined as depression and social anxiety increased. However, this was weakened in dataset 2, particularly for depression (Figure 2).

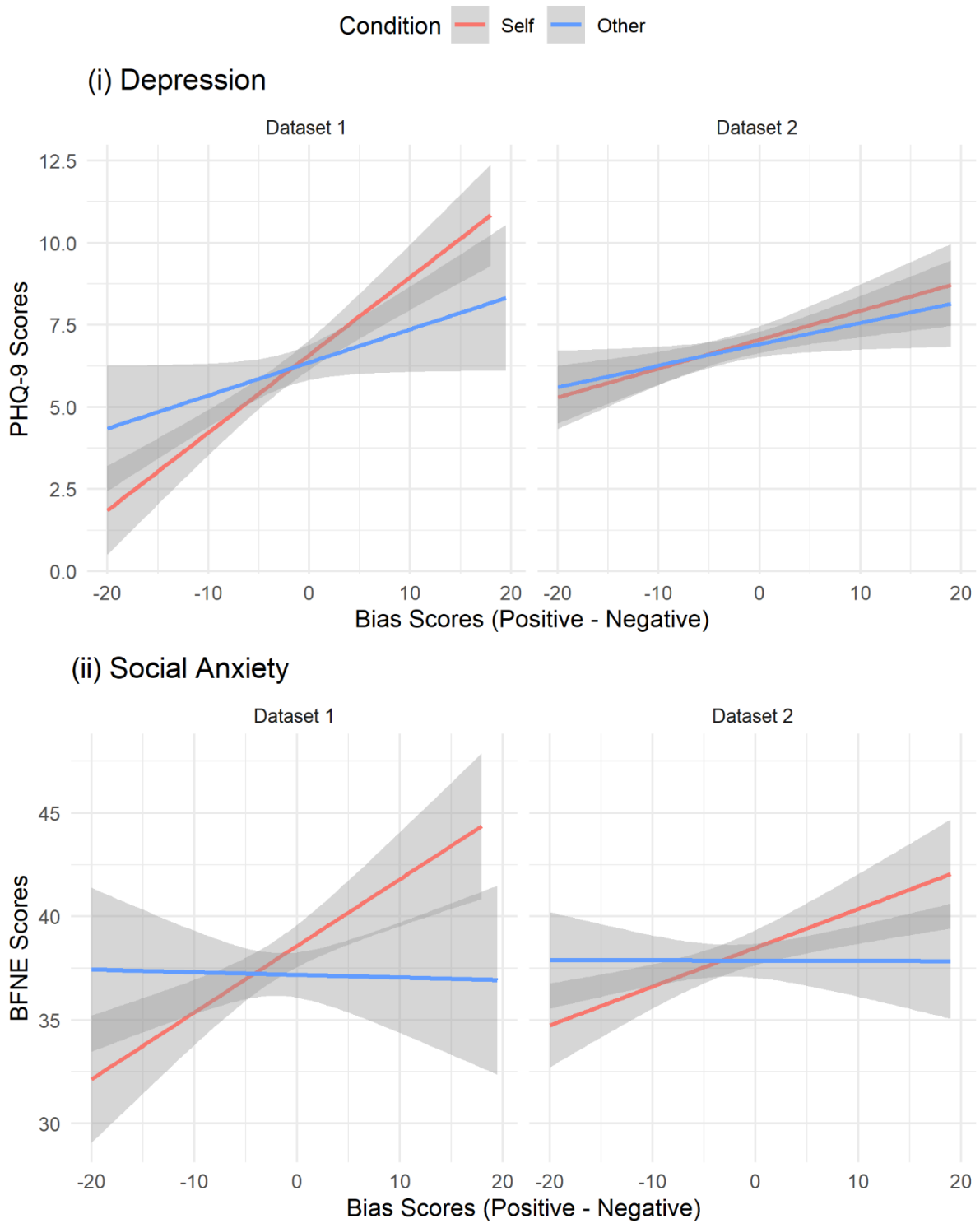


Figure 2

Relationship between biased learning on the social evaluation learning task and (i) depression, (ii) social anxiety. Greater bias scores indicate relatively better learning of the negative versus positive rule.

In keeping with visual inspection of the data, structural equation modelling indicated strong evidence that bias scores when learning about the self were associated with depression and social anxiety. In dataset 1 greater self-bias scores, indicating relatively better learning of the negative versus positive rule, were associated with increased symptoms of depression (β = 0.27, 95% CI: 0.20, 0.35, p < .001) and social anxiety (β = 0.20, 95% CI: 0.12, 0.29, p < .001) (Figure 3). These effects were specific to learning about the self; we found little evidence of a relationship between learning about others and depression or social anxiety (Table 5).

These results were replicated in dataset 2, although effects were attenuated (depression: β = 0.10, 95% CI: 0.02, 0.17, p = 0.009; social anxiety: β = 0.13, 95% CI: 0.06, 0.21, p < .001; Table 5, Figure 3).

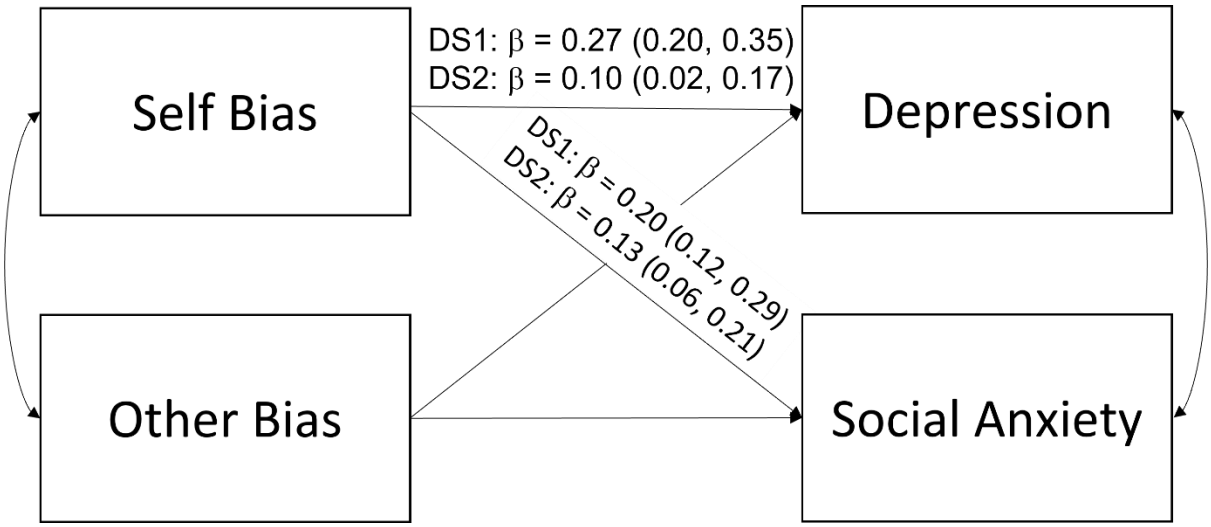


Figure 3

Structural Equation Model Diagram illustrating the relationship between self-bias scores in the social evaluation learning task with depression (PHQ-9 scores) and social anxiety (BFNE scores) in dataset 1 (DS1) and dataset 2 (DS2). Greater bias scores, indicating relatively better learning of the negative versus positive rule, when learning about the self were associated with increased depression and social anxiety severity. We did not find evidence of an association between biased learning about the other and depression or social anxiety.

To examine whether these associations were driven by better learning of the positive rule, or worse learning of the negative rule, we examined the relationship between depression and social anxiety with errors to criterion in each of the referential condition-rule combinations (Table 5).

In dataset 1 we found strong evidence that worse learning of the positive rule (β = 0.32, 95% CI: 0.25, 0.39, p < .001) and better learning of the negative rule (β = -0.09, 95% CI: -0.13, -0.06, p < .001) were associated with greater depression. In dataset 2, whilst these effects were observed in the same direction they were attenuated and overlapped with the null to a small extent (Self-Positive: β = 0.06, 95% CI: - 0.01, 0.13, p = 0.098; Self-Negative: β = -0.07, -0.14, 0.01, p = 0.080).

Across both datasets, better learning of the negative rule about the self was associated with greater social anxiety (dataset 1: β = -0.16, 95% CI: -0.21, -0.11, p < .001; dataset 2: β = -0.13, 95% CI: -0.21, -0.06, p < .001). We did not find evidence that learning of the positive rule about the self was associated with social anxiety in either dataset (dataset 1: β = 0.12, 95% CI: -0.06, 0.30, p = 0.206; dataset 2: β = 0.03, -0.04, 0.10, p = 0.398).

Table 5

Results from structural equation models reporting regression pathways between learning in the social evaluation learning task and depression (PHQ-9) and social anxiety (BFNE), and correlations between depression (PHQ-9) and social anxiety (BFNE) by dataset

	Dataset 1 (n = 450)[a]			**Dataset 2 (n = 807)**		
Bias Scores						
Regression	*β*	*95% CI*	*p*	*β*	*95% CI*	*p*
PHQ-9 ←						
Self	0.27	0.20, 0.35	< .001	0.10	0.02, 0.17	0.009
Other	0.03	-0.09, 0.14	0.655	0.04	-0.03, 0.11	0.300
BFNE ←						
Self	0.20	0.12, 0.29	< .001	0.13	0.06, 0.21	< .001
Other	-0.05	-0.16, 0.06	0.352	-0.05	-0.12, 0.02	0.177
Correlation	*r*	95% CI	*p*	*r*	95% CI	*p*
PHQ-9 ↔ BFNE	0.40	0.31, 0.49	< .001	0.41	0.35, 0.46	< .001
Errors to Criterion						
Regression	*β*	*95% CI*	*p*	*β*	*95% CI*	*p*
PHQ-9 ←						
Self-Positive	0.32	0.25, 0.39	< .001	0.06	-0.01, 0.13	0.098
Self-Negative	-0.09	-0.13, -0.06	< .001	-0.07	-0.14, 0.01	0.080
Other-Positive	0.02	-0.12, 0.16	0.772	-0.01	-0.07, 0.08	0.888
Other-Negative	-0.01	-0.08, 0.06	0.817	-0.05	-0.12, 0.03	0.216
BFNE ←						
Self-Positive	0.12	-0.06, 0.30	0.206	0.03	-0.04, 0.10	0.398
Self-Negative	-0.16	-0.21, -0.11	< .001	-0.13	-0.21, -0.06	< .001
Other-Positive	-0.01	-0.10, 0.08	0.845	-0.05	-0.12, 0.02	0.169
Other-Negative	0.06	-0.06, 0.18	0.302	0.03	-0.05, 0.10	0.484
Correlation	*r*	95% CI	*P*	*r*	95% CI	*p*
PHQ-9 ↔ BFNE	0.41	0.30, 0.52	< .001	0.41	0.35, 0.46	< .001

[a] One study within dataset 1 (n = 102) did not include an other condition and was therefore excluded from these models

PHQ-9 = Patient Health Questionnaire; BFNE = Brief Fear of Negative Evaluation Scale. β = standardised path coefficient

Discussion

Depression and social anxiety are associated with differences in sensitivity to positive and negative social feedback, which are believed to reinforce negative views of the self [6,46,47]. However, the pattern of learning of social evaluations associated with each disorder is unclear. We therefore investigated learning of positive and negative social evaluations in individuals experiencing varying levels of depression and social anxiety.

Our findings support previous work in suggesting that positive self-biases are protective for mental health [10,48]. On average participants displayed a positive social evaluation learning style, making relatively fewer errors learning a positive 'like' rule versus a negative 'dislike' rule. However, as symptoms of depression and social anxiety increased this pattern was reversed. Social anxiety and depression were independently associated with relatively better learning of the negative relative to the positive rule. These effects were specific to learning about the self; we did not find any evidence of an association between learning about others with either social anxiety or depression. Our findings suggest that depression and social anxiety are characterised by a greater sensitivity to negative relative to positive social feedback. The specificity of effects to the self suggest that social evaluation learning may be important in maintaining negative views of the self characteristic of both disorders.

However, we found evidence that different patterns of learning underpinned reduced positive biases in depression and social anxiety. In both datasets we found strong evidence that social anxiety was associated with better learning of the negative 'dislike' rule about the self. Conversely, we did not find evidence of an association between social anxiety and learning of the positive self 'like' rule. Our findings are consistent with traditional cognitive theories in suggesting that social anxiety is partially driven by an increased sensitivity to socially threatening information which reinforce negative beliefs about the self [6]. However, our finding that individuals with social anxiety did not show impairments in learning positive evaluations about the self are inconsistent with evidence of blunted neural response to social reward [49] and self-reported fears of positive evaluations [50]. It is possible that fears of positive stimuli are dependent on the context in which they occur, whereas fear of negative stimuli is more universal. For instance, receiving positive social feedback in a private situation, similar to the context of this cognitive task, may be less anxiety provoking than a group setting [51]. Conversely, receiving negative social feedback may be anxiety provoking regardless of whether this occurs in private or public social contexts.

In contrast to our findings for social anxiety, we found evidence that depression was associated with differences in both learning of being liked and disliked. In keeping with growing evidence of social anhedonia [28] and impaired reward-processing in depression [52–54] , we found that depression was associated with worse learning of the positive 'like' rule about the self. Additionally, supportive of previous evidence of negative self-referential biases in depression [55–57], we found that participants with greater depression were better at learning the negative self 'dislike' rule. Social evaluation learning may play a dual role in maintaining depression by both reinforcing negative views of the self, as well as perpetuating reduced reward from social interactions, reinforcing social anhedonia. Repeated exposure to negative social experiences through social evaluation learning biases may promote socially avoidant behaviours, increasing the probability of reciprocal negative behaviours from others, and contributing to relationship difficulties and a lack of social support commonly associated with depression [58].

However, the relationship between social evaluation learning and depression was less reliable across datasets compared to our findings for social anxiety. Whilst we found strong evidence of this association in individual participant data collated from ten previous studies conducted by this research group (dataset 1), statistical evidence of this association was weakened in a large independent dataset of participants recruited online (dataset 2). Distribution of PHQ-9 scores was similar across datasets suggesting that this attenuation was not driven by differences in depression severity. Our findings may instead be explained by differences in data collection methods. Whereas participants in dataset 1 completed the study in labs supervised by researchers, participants in dataset 2 completed measures independently online. Although cognitive testing across lab and online settings are believed to produce comparable results [59], concerns have been raised over the quality of data collected remotely [60]. As individuals with depression also experience wider cognitive impairments in memory and attention [61] these issues may be particularly relevant within this study. Further research evaluating the suitability of administering cognitive tasks online in depressed samples would be helpful in contextualising our results.

Clinical Implications

Our findings suggest that learning within social interactions may be a potential beneficial target for treatments for both depression and social anxiety. However, our results also suggest that tailoring treatments to target specific maladaptive patterns of learning underlying each disorder may be more effective than a universal treatment. Whereas treatments for social

anxiety may be most advantageous in focusing on addressing heightened sensitivity to negative social feedback, treatments for depression should also attempt to address reduced sensitivity to socially rewarding information. Additionally, our findings that positive learning biases towards a fictional 'other' were maintained in individuals with high social anxiety and depression support a technique commonly used in cognitive therapies; asking patients to challenge their maladaptive beliefs by considering a scenario from a third-person perspective [62]. Our findings suggest that cognitive treatments for depression and social anxiety should continue to use this technique to attempt to address maladaptive social cognitive biases.

Strengths and Limitations

We took a rigorous approach, using two independent datasets to assess the reliability of our findings. We also used relatively large samples in both datasets to ensure that we were adequately powered to detect small effects. Additionally, we used a computer-based task to understand complex social phenomena within a controlled environment.

However, whilst we found strong evidence of a cross-sectional relationship between social evaluation learning, depression, and social anxiety our study design does not allow us to comment on the causality of this association. It is possible that social cognitive processes may be symptomatic of these disorders rather than playing a causal role. Future research examining longitudinal relationships would be beneficial in helping to identify the importance of social evaluation in depression and social anxiety.

We used the BFNE as a measure of social anxiety. As this measure focuses on fear of negative evaluations it is likely to have amplified our observed associations between social anxiety and negative social evaluation learning. However, fear of negative evaluations are considered to be core to social anxiety [6] and the BFNE correlates highly with other social anxiety measures [63]. Additional research using alternative measures of social anxiety would be useful in validating our findings.

Conclusions

Across two independent datasets we found that participants with greater social anxiety and depression demonstrated biased learning of social evaluations about the self. Although this relationship was stronger for social anxiety, our findings suggest that maladaptive social evaluation learning may represent a transdiagnostic vulnerability for social anxiety and depression. However, specific patterns of learning underlying this bias differed. Whereas social anxiety was predominantly characterised by better learning of negative evaluations,

depression was characterised by a relative balance between enhanced learning of negative evaluations as well as reduced learning of positive evaluations. Treatments targeting these processes within each disorder may be beneficial.

References

1. Button, K. S., Browning, M., Munafò, M. R. & Lewis, G. Social inference and social anxiety: Evidence of a fear-congruent self-referential learning bias. *J. Behav. Ther. Exp. Psychiatry* **43**, 1082–1087 (2012).

2. Button, K. S. *et al.* Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. *PLoS One* **10**, e0119456 (2015).

3. Ohayon, M. M. & Schatzberg, A. F. Social phobia and depression: Prevalence and comorbidity. *J. Psychosom. Res.* **68**, 235–243 (2010).

4. Abbafati, C. *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1204–1222 (2020).

5. Beck, A. T. Cognitive models of depression. *J. Cogn. Psychother.* **1**, 5–37 (1987).

6. Clark, D. M. & Wells, A. A Cognitive Model of Social Phobia. in *Social phobia: Diagnosis, assessment, and Treatment* (ed. Heimberg, R. G.) 69–73 (The Guilford Press, 1995).

7. Hirschfeld, R. M. A. *et al.* Social functioning in depression: A review. *J. Clin. Psychiatry* **61**, 268–275 (2000).

8. Angélico, A. P., Crippa, J. A. de S. & Loureiro, S. R. Social anxiety disorder and social skills: A critical review of the literature. *Int. J. Behav. Consult. Ther.* **7**, 16 (2013).

9. Shrauger, J. S. & Schoeneman, T. J. Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychol. Bull.* **86**, 549 (1979).

10. Korn, C. W., Prehn, K., Park, S. Q., Walter, H. & Heekeren, H. R. Positively biased processing of self-relevant social feedback. *J. Neurosci.* **32**, 16832–16844 (2012).

11. Blaine, B. & Crocker, J. Self-Esteem and Self-Serving Biases in Reactions to Positive and Negative Events: An Integrative Review. in *Self-Esteem* (ed. Baumeister, R. F.) 55–85 (Springer, Boston, MA, 1993). doi:10.1007/978-1-4684-8956-9_4.

12. Christensen, P. N., Stein, M. B. & Means-Christensen, A. Social anxiety and interpersonal perception: A social relations model analysis. *Behav. Res. Ther.* **41**,

1355–1371 (2003).

13. Moritz, D. & Roberts, J. E. Self-Other Agreement and Metaperception Accuracy Across the Big Five: Examining the Roles of Depression and Self-Esteem. *J. Pers.* **86**, 296–307 (2018).

14. van Niekerk, R. E. *et al.* The Role of Cognitive Factors in Childhood Social Anxiety: Social Threat Thoughts and Social Skills Perception. *Cognit. Ther. Res.* **41**, 489–497 (2017).

15. Gable, S. L. & Shean, G. D. Perceived social competence and depression. *J. Soc. Pers. Relat.* **17**, 139–150 (2000).

16. Stopa, L. Why is the self important in understanding and treating social phobia? *Cogn. Behav. Ther.* **38**, 48–54 (2009).

17. Gilboa-Schechtman, E., Foa, E. B. & Amir, N. Attentional biases for facial expressions in social phobia: The face-in-the-crowd paradigm. *Cogn. Emot.* **13**, 305–318 (1999).

18. Brook, D. W., Brook, J. S., Zhang, C., Cohen, P. & Whiteman, M. Increased amygdala activation to angry and contemptuous faces in generalized social phobia. *Arch. Gen. Psychiatry* **59**, 1027–1034 (2002).

19. Chen, J., Short, M. & Kemps, E. Interpretation bias in social anxiety: A systematic review and meta-analysis. *J. Affect. Disord.* **276**, 1119–1130 (2020).

20. Campbell, D. W. *et al.* Happy but not so approachable: The social judgments of individuals with generalized social phobia. *Depress. Anxiety* **26**, 419–424 (2009).

21. Romano, M., Tran, E. & Moscovitch, D. A. Social anxiety is associated with impaired memory for imagined social events with positive outcomes. *Cogn. Emot.* **34**, 700–712 (2020).

22. Huys, Q. J., Pizzagalli, D. A., Bogdan, R. & Dayan, P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol. Mood Anxiety Disord.* **3**, 1–16 (2013).

23. Horne, S. J., Topp, T. E. & Quigley, L. Depression and the willingness to expend cognitive and physical effort for rewards: A systematic review. *Clin. Psychol. Rev.* **88**, 102065 (2021).

24. Keren, H. *et al.* Reward processing in depression: A conceptual and meta-analytic review across fMRI and EEG studies. *Am. J. Psychiatry* **175**, 1111–1120 (2018).

25. Eshel, N. & Roiser, J. P. Reward and punishment processing in depression. *Biol. Psychiatry* **68**, 118–124 (2010).

26. McFarland, B. R. & Klein, D. N. Emotional reactivity in depression: diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance. *Depress. Anxiety* **26**, 117–122 (2009).

27. Mukherjee, D., Filipowicz, A. L. S., Vo, K., Satterthwaite, T. D. & Kable, J. W. Reward and punishment reversal-learning in major depressive disorder. *J. Abnorm. Psychol.* **129**, 810–823 (2020).

28. Barkus, E. & Badcock, J. C. A transdiagnostic perspective on social anhedonia. *Front. Psychiatry* **10**, 216 (2019).

29. Craske, M. G. Transdiagnostic treatment for anxiety and depression. *Depress. Anxiety* **29**, 749–753 (2012).

30. Qualtrics. Qualtrics. (2020).

31. Inquisit. Inquisit 6. www.millisecond.com (2020).

32. Prolific. Prolific. https://www.prolific.co (2021).

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

34. Löwe, B., Kroenke, K., Herzog, W. & Gräfe, K. Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J. Affect. Disord.* **81**, 61–66 (2004).

35. Martin, A., Rief, W., Klaiberg, A. & Braehler, E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *Gen. Hosp. Psychiatry* **28**, 71–77 (2006).

36. Leary, M. R. A Brief Version of the Fear of Negative Evaluation Scale. *Personal. Soc. Psychol. Bull.* **9**, 371–375 (1983).

37. Duke, D., Krishnan, M., Faith, M. & Storch, E. A. The psychometric properties of the Brief Fear of Negative Evaluation Scale. *J. Anxiety Disord.* **20**, 807–817 (2006).

38. Collins, K. A., Westra, H. A., Dozois, D. J. A. & Stewart, S. H. The validity of the brief version of the Fear of Negative Evaluation Scale. *J. Anxiety Disord.* **19**, 345–359 (2005).

39. Button, K. S. Identifying the Neural Substrates of Learning Social Evaluation in Social Anxiety. *Unpubl. Data*.

40. Button, K. S., Karwatowska, L., Kounali, D., Munafò, M. R. & Attwood, A. S. Acute anxiety and social inference: An experimental manipulation with 7.5% carbon dioxide inhalation. *J. Psychopharmacol.* **30**, 1036–1046 (2016).

41. Button, K. S. Self-referential learning and self-esteem. *Unpubl. Data*.

42. Brennan-Craddock, A. & Button, K. S. Self-referential processing in Depression. *Unpubl. Data*.

43. Brennan-Craddock, A., Xin, W., Sui, J. & Button, K. S. The Emotional Self. *Unpubl. Data*.

44. Hobbs, C., Sui, J., Kessler, D., Munafò, M. R. & Button, K. S. Self-processing in relation to emotion and reward processing in depression. *Psychol. Med.* 1–13 (2021) doi:10.1017/s0033291721003597.

45. Hobbs, C. *et al.* Effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers. *BJPsych Open* **6**, (2020).

46. Hofmann, S. G. Cognitive factors that maintain social anxiety disorder: A comprehensive model and its treatment implications. *Cogn. Behav. Ther.* **36**, 193–209 (2007).

47. Weightman, M. J., Air, T. M. & Baune, B. T. A review of the role of social cognition in major depressive disorder. *Front. Psychiatry* **5**, 179 (2014).

48. Blaine, B. & Crocker, J. Self-Esteem and Self-Serving Biases in Reactions to Positive and Negative Events: An Integrative Review. in *Self-Esteem* 55–85 (Springer, Boston, MA, 1993). doi:10.1007/978-1-4684-8956-9_4.

49. Cremers, H. R., Veer, I. M., Spinhoven, P., Rombouts, S. A. R. B. & Roelofs, K. Neural sensitivity to social reward and punishment anticipation in social anxiety disorder. *Front. Behav. Neurosci.* **8**, 439 (2015).

50. Weeks, J. W., Heimberg, R. G., Rodebaugh, T. L. & Norton, P. J. Exploring the relationship between fear of positive evaluation and social anxiety. *J. Anxiety Disord.* **22**, 386–400 (2008).

51. Fredrick, J. W. & Luebbe, A. M. Fear of positive evaluation and social anxiety: A systematic review of trait-based findings. *Journal of Affective Disorders* vol. 265 157–168 (2020).

52. Pegg, S. *et al.* Blunted Social Reward Responsiveness Moderates the Effect of Lifetime Social Stress Exposure on Depressive Symptoms. *Front. Behav. Neurosci.* **13**, 178 (2019).

53. Ait Oumeziane, B., Jones, O. & Foti, D. Neural Sensitivity to Social and Monetary Reward in Depression: Clarifying General and Domain-Specific Deficits. *Front. Behav. Neurosci.* **13**, 199 (2019).

54. Fussner, L. M., Mancini, K. J. & Luebbe, A. M. Depression and Approach Motivation: Differential Relations to Monetary, Social, and Food Reward. *J. Psychopathol. Behav. Assess.* **40**, 117–129 (2018).

55. Mogg, K. & Bradley, B. P. Attentional Bias in Generalized Anxiety Disorder Versus Depressive Disorder. *Cognit. Ther. Res.* **29**, 29–45 (2005).

56. Everaert, J., Podina, I. R. & Koster, E. H. W. A comprehensive meta-analysis of interpretation biases in depression. *Clin. Psychol. Rev.* **58**, 33–48 (2017).

57. Gaddy, M. A. & Ingram, R. E. A meta-analytic review of mood-congruent implicit memory in depressed mood. *Clin. Psychol. Rev.* **34**, 402–416 (2014).

58. Lewinsohn, P. M., Mischel, W., Chaplin, W. & Barton, R. Social competence and depression: The role of illusory self-perceptions. *J. Abnorm. Psychol.* **89**, 203 (1980).

59. Backx, R., Skirrow, C., Dente, P., Barnett, J. H. & Cormack, F. K. Comparing web-based and lab-based cognitive assessment using the cambridge neuropsychological test automated battery: A within-subjects counterbalanced study. *J. Med. Internet Res.* **22**, e16792 (2020).

60. Gagné, N. & Franzen, L. How to run behavioral experiments online: Best practice suggestions for cognitive psychology and neuroscience [Preprint]. *PsyArXiv* 1–40 doi:10.31234/OSF.IO/NT67J.

61. Marazziti, D., Consoli, G., Picchetti, M., Carlini, M. & Faravelli, L. Cognitive impairment in major depression. *European Journal of Pharmacology* vol. 626 83–86 (2010).

62. Wallace-Hadrill, S. M. A. & Kamboj, S. K. The impact of perspective change as a cognitive reappraisal strategy on affect: A systematic review. *Front. Psychol.* **7**, 1715 (2016).

63. Carleton, R. N., Collimore, K. C. & Asmundson, G. J. G. Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *J. Anxiety Disord.* **21**, 131–141 (2007).

This declaration concerns the article entitled:			
The cognitive neuropsychological theory of antidepressants: a systematic review and meta-analysis of behavioural evidence			
Publication status (tick one) **Draft manuscript** [X] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [ ]			
Publication details (reference)	Hobbs, C., Bone, J., Ruzickova, T., Sui, J., Kessler, D., Munafò, M.R., & Button, K. S. (2022). The cognitive neuropsychological theory of antidepressants: a systematic review and meta-analysis of behavioural evidence. *Manuscript in preparation.*		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to the formulation of ideas (80%) and design of methodology (80%). The candidate conducted the search, extracted data, and evaluated risk of bias as first reviewer (100%). The candidate analysed and visualised the data (100%), and interpreted findings (90%). The candidate predominantly executed the presentation of the data in journal format (90%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	05/01/2022

Chapter 3: The cognitive neuropsychological theory of antidepressants: a systematic review and meta-analysis of behavioural evidence

Catherine Hobbs [a], Jessica Bone [b], Tereza Ruzickova [c], Jie Sui [d], David Kessler [e], Marcus R Munafò [f g h], Katherine S Button [a]

[a] Department of Psychology, University of Bath

[b] Institute of Epidemiology and Health, University College London

[c] Department of Psychiatry, University of Oxford

d School of Psychology, University of Aberdeen

[e] Population Health Sciences, University of Bristol

[f] School of Psychological Science, University of Bristol

[g] MRC Integrative Epidemiology Unit at the University of Bristol

[h] National Institute of Health Research Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol

Chapter Rationale

In chapters 1 and 2 I found that depression was associated with a loss of positive self-referential affective biases. Participants with greater depression made a greater number of errors learning a positive 'like' rule versus a negative 'dislike' rule about the self. Increasing learning of positive social evaluations about the self may be beneficial in remediating negative self-schema and subsequent depressive symptoms. Social evaluation learning may therefore provide an important and potentially reversible target for therapeutic intervention. The cognitive neuropsychological model proposes that antidepressants operate by increasing sensitivity to positive information and decreasing sensitivity to negative information [1,2]. In line with this theory, we would expect antidepressants to remediate social evaluation learning biases associated with depression by increasing learning of positive evaluations and decreasing learning of negative evaluations. In chapter 3 I conducted a systematic review and meta-analysis of evidence for the cognitive neuropsychological theory of antidepressants in preparation for future chapters examining the effect of antidepressants on self-referential

affective processing. I aimed to establish the current evidence base for change in emotional processing following antidepressant administration, and to identify current gaps in research that could be addressed in the further chapters of this PhD. As behavioural measures of emotional processing (e.g., reaction times or accuracy in cognitive tasks) would be the most feasible to implement in primary care settings in future chapters, and previous meta-analyses have focused on neural activity, I chose to focus on behavioural evidence for the cognitive neuropsychological theory in this review.

Open Science Statement

This review was pre-registered on PROSPERO (registration number: 95621). The summary data that support the findings of this study will be made openly available in the University of Bath Research Data Archive prior to publication.

Key Points

Question: Are antidepressants associated with increased processing of positive information and reduced processing of negative information on behavioural measures of emotional processing?

Findings: We did not find evidence of a reliable association between antidepressant administration and processing of positive or negative information.

Meaning: Our findings do not support the theory that antidepressants operate by remediating negative emotional biases associated with depression.

Abstract

Importance: Antidepressants are proposed to operate by remediating negative emotional biases associated with depression; increasing sensitivity to positive relative to negative information. Previous meta-analyses on neural activity have found supportive evidence. However, a meta-analysis has not been conducted for behavioural measures.

Objective: To examine whether antidepressants are associated with behavioural measures of positive and negative emotional processing.

Data Sources: We searched PubMed, PsychInfo and Embase databases for studies that examined the effect of antidepressants on behavioural measures of emotional processing published between January 1st 1998 and 6th February 2020.

Study Selection: Studies that administered an antidepressant medication to a human sample and included a quantitative behavioural measure of emotional processing.

Data Extraction and Synthesis: We conducted a narrative synthesis of findings grouped according to cognitive domain. We used summary statistics for antidepressant and control conditions to calculate standardised mean differences (SMDs) and conducted random-effects multi-level meta-analyses with effects clustered within study for positive and negative emotions. We used moderator analyses to understand if effects varied by cognitive domain.

Main Outcome Measures: Behavioural measures of emotional processing.

Results: Of 9299 records identified, 82 studies met inclusion criteria for narrative synthesis. Of these, 28 provided 75 effects for positive emotions (antidepressant n = 625, control n = 612) and 76 effects for negative emotions (antidepressant n = 626, control n = 613) for meta-analysis. Participants' mean age was 28.5 years and 58% were female. We did not find

evidence of a difference between antidepressant and control conditions in processing of positive (SMD = 0.08, 95% CI = -0.05, 0.22, p = 0.230) or negative stimuli (SMD = -0.04, 95% CI = -0.12, 0.04, p = 0.341). Effects did not vary by cognitive domain (Positive: F(6, 69) = 0.983, p = 0.443, Negative: F(6, 70) = 0.999, p = 0.433).

Conclusions and Relevance: We did not find evidence of a reliable association between antidepressant administration and altered processing of positive or negative information. Our findings do not support the theory that antidepressants operate by remediating negative emotional biases associated with depression. However, we integrated evidence across a broad range of samples and antidepressants. Specific effects on emotional processing may have been obscured.

Introduction

Depression is among the leading causes of global disability across the lifespan [3]. Individuals experiencing depression are at an increased risk of chronic physical disorders [4] and have a reduced life expectancy [5]. Depression also has a substantial economic burden; in the USA alone depression is associated with costs of $210.5 billion [6]. Effective treatments for depression are therefore essential.

Antidepressant medication is a first-line treatment for depression. Whilst antidepressants are effective in reducing depressive symptoms [7], approximately half of patients do not respond to the first antidepressant prescribed [8]. Sequential strategies can be used to address treatment resistance, including increasing dosages, switching antidepressants, and augmentation with additional treatments [9,10]. However, response cannot be reliably identified early in treatment. Current NICE guidelines recommend altering treatment only if improvement has not occurred in 3-4 weeks [10]. It is also not currently possible to predict which treatment strategy may be most beneficial. Patients are therefore exposed to long periods of potentially ineffective treatment with a trial-and-error approach. Exposure to side effects in the absence of therapeutic benefits during this period may contribute to high discontinuation rates [11,12]. To address these issues, research has focused on identifying the therapeutic mechanisms of early antidepressant action. By understanding these early mechanisms, it may be possible to identify potential biomarkers of treatment response, allowing treatment to be altered at an earlier period if necessary [13].

One potential psychopharmacological mechanism of early antidepressant action that has generated substantial research is change in emotional processing. Depression is associated with increased sensitivity to negative relative to positive information across a range of cognitive domains, including attention, memory, and emotion recognition [14–16]. The cognitive neuropsychological model of depression suggests that these negative biases are causal in the development and maintenance of depression. According to this model negative emotional biases are driven by alterations in monoamine transmission caused by environmental and genetic factors. Repeated automatic processing of negative environmental stimuli over time leads to the development of entrenched and persistent negative sets of expectations and beliefs about the self. In turn, these negative self-schema reinforce automatic negative emotional biases and produce deficits in top-down cognitive control biases, further maintaining depressive symptoms [2]. Within this framework, antidepressants are believed to operate by remediating negative emotional biases early in treatment, increasing sensitivity to

positive information and decreasing sensitivity to negative information. Accounting for the delay in the therapeutic action of antidepressants, an improvement in mood is only produced when individuals have interacted with their social environment with remediated emotional biases, allowing them to relearn associations in positive contexts [1,2,17,18].

Several reviews have been conducted on the cognitive neuropsychological model of antidepressant action, reporting generally supportive findings. In a narrative review of 25 studies assessing the influence of serotonin manipulations, change in emotional processing was observed across a variety of cognitive domains [19]. Additionally, several narrative reviews have reported evidence of change in emotional processing in behavioural and brain imaging studies following antidepressant administration [18,20–22]. Two meta-analyses have also been conducted on change in brain activity associated with antidepressant response. From nine studies, Delaveau et al (2011) [23] reported that antidepressant treatment was found to normalise brain activity in response to emotional stimuli. In 2015, Ma [24] found that antidepressants increased activity in response to positive stimuli and decreased activity to negative stimuli in the emotional brain network.

However, to date a systematic review on behavioural evidence for the cognitive neuropsychological theory of antidepressants has not been conducted. As cognition cannot be measured directly, research studies rely on neural (e.g., blood flow, brain waves) or behavioural (e.g., accuracy, reaction times) indicators of emotional processing. In this review, we aimed to systematically evaluate behavioural evidence of change in emotional processing following antidepressant administration. We took a broad approach, integrating evidence for both conventional and experimental antidepressant medications across various cognitive domains of emotional processing. We evaluated the evidence in reference to the predictions of the cognitive neuropsychological model of antidepressants, namely that administration of antidepressants should increase processing of positive emotional stimuli and decrease processing of negative stimuli, in both healthy volunteers and individuals experiencing depression, and that these changes should be detectable even with acute doses of antidepressants [1,2].

Method

This review was pre-registered on PROSPERO (reference: 95621)

Search Strategy

We searched PubMed, PsycInfo, and Embase databases for articles published between January 1st 1998 and 6th February 2020. We used three main search terms combined using the Boolean operator 'and': (1) antidepressant OR SSRI OR SNRI OR NDRI OR MAOI, (2) neurocognitive OR cognitive OR emotional processing OR psychological OR neurological OR neuropsychological OR affective processing, (3) mechanism OR action OR module OR effect

Inclusion and Exclusion Criteria

To be eligible for inclusion, studies must have administered an antidepressant medication (including experimental treatments) to a human sample and included a quantitative behavioural measure of emotional processing and self-report measures of mood. Studies were required to include a comparison condition of either placebo or no antidepressant. We excluded conference proceedings and presentations, and articles not available in English.

Study Screening and Selection

The references of studies identified in the electronic search were uploaded to the systematic review software Covidence, and duplications were removed. Articles were first assessed for eligibility based on titles and abstracts and were then re-assessed based on the full text. References of included articles were hand-searched for additional eligible studies, which then underwent the screening procedure outlined previously. Studies were screened independently by two reviewers (CH and JB). Disagreements were resolved by a third reviewer (KSB).

Data Extraction and Study Quality Assessment

Data were extracted independently by CH and JB using a standardised data collection form. Information extracted included the study design, sample, antidepressant administered, comparator condition, task procedures, and a summary of the effect of antidepressants on emotional processing. Study quality was assessed using an adapted Cochrane risk of bias tool. Disagreements in quality ratings were resolved by a third reviewer (KSB).

Narrative Synthesis

All studies identified as eligible were included in a narrative synthesis. A broad number of cognitive tasks were employed across studies to examine change in emotional processing. This review aimed to be broad, therefore results relating to all cognitive tasks measuring

emotional processing and reporting a behavioural outcome measure were summarised grouped according to cognitive domain.

Meta-Analysis

To be included in the meta-analysis section of this study appropriate data were required to be reported in the article or available upon request. As within-subject studies typically did not report data that would allow for calculation of effect sizes (e.g., within-subject correlations) we focused on standardised mean differences between experimental and control conditions. We calculated Hedge's g, using means and standard deviations, to represent differences in emotional processing between experimental and control conditions. Positive effect sizes represent increased processing in the experimental versus control conditions. Negative effect sizes represent decreased processing in the experimental versus control conditions. Hedges' g is comparable to Cohen's d but is more suitable for smaller and unbalanced sample sizes [25]. As positive and negative emotions are processed using distinct neural pathways [26], we calculated separate effect sizes for positive and negative emotions. Where data were reported for individual emotions (e.g., sadness, anger, disgust) we calculated the average mean and standard deviation.

There was substantial statistical dependency in the reported results of studies eligible for inclusion in the meta-analysis. To account for studies testing multiple antidepressants against a single control, we combined the antidepressant groups into a single group as recommended in Cochrane guidelines [27]. Additional dependency arose from studies using multiple emotional processing tasks to measure separate cognitive domains. Some studies also used multiple groups with different psychiatric characteristics (e.g., healthy controls versus depression) to assess the effect of antidepressants versus a control condition on emotional processing. To account for these dependencies, we conducted a restricted maximum-likelihood multi-level meta-analysis with effects clustered within study entered as a random effect using the metafor package [28] in R version 4.0.5. We applied the Hartung-Knapp-Sidik-Jonkman method due to relatively lower inflation rates for smaller number of studies, unequal sample sizes, and between-study heterogeneity (20). We examined heterogeneity by calculating Higgin's and Thompson's I^2 using the dmetar package [29]. Values of 25%, 50% and 75% indicate low, moderate, and substantial heterogeneity respectively [25]. To assess the potential presence of publication bias we conducted an Egger's test for small study bias by regressing the standard error of the effect sizes on the estimates calculated in our primary

meta-analyses [30,31]. To assess possible variations in effects according to cognitive domain, we repeated our primary analyses with cognitive domain added as a moderator effect.

Most studies also reported multiple outcome measures per cognitive task. However, to reduce the statistical complexity of the model in relation to the relatively small number of studies included, we used the outcome measures most commonly reported per cognitive domain across studies in our primary analyses. To account for potential bias in the selection of these outcomes we conducted a sensitivity analysis using the second outcome measure most commonly reported.

Results

Study Selection

A total of 82 studies published between 2003 and 2019 met inclusion criteria and were included in our narrative synthesis. Of these, 28 studies provided data for meta-analysis. A flow chart illustrating the review process is shown in Figure 1.

Study Characteristics

Study characteristics are outlined in Table 1. Most studies recruited healthy volunteers (k = 50), or individuals experiencing depression (k = 19). The most common antidepressants examined were selective serotonin reuptake inhibitors (k = 33), followed by noradrenaline reuptake inhibitors (k = 13), selective noradrenaline reuptake inhibitors (k = 6), and atypical antidepressants (k = 9). Ten studies were observational, testing mixed types of antidepressants and dosages, 20 studies examined experimental antidepressant treatments, and one study examined the herbal remedy St John's Wort. Most studies examined the effect of an acute dose of an antidepressant (k = 43); the duration in the remaining studies ranged from 1 to 18 weeks. A further detailed description is available in Supplementary Materials.

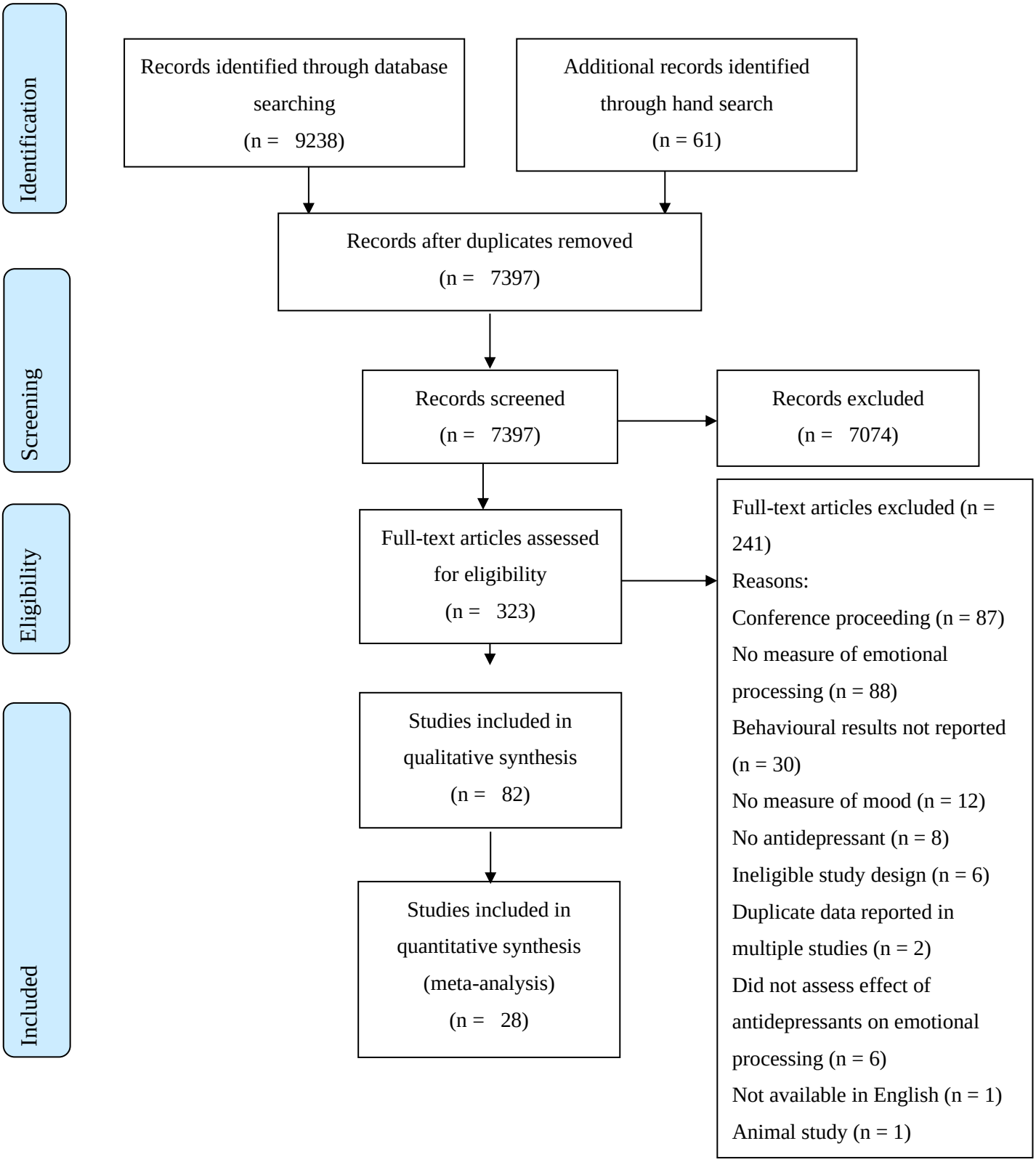


Figure 1

PRISMA flow diagram illustrating search results

Table 1

Study Characteristics

			Experimental Condition				Comparator Condition				
Author	Year	Drug	Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	Drug Assignment
Anderson[32] *	2011	Mixed	Mixed	Mixed	16 17	Current Depression Remitted Depression	No Drug No Drug	14 82	Current Depression Remitted Depression		Observatio
Arce[33]	2008	Escitalopram	5-10 mg/day [a]	3 weeks	13	Healthy	Placebo	13	Healthy		Experimer
Arnone[34]	2009	Mirtazapine	15 mg	Acute	15	Healthy	Placebo	15	Healthy		Experimer
Bamford[35]	2015	Duloxetine	30-60 mg/day [a]	2 weeks	19	Healthy	Placebo	19	Healthy		Experimer
Bershad[36]	2018	Buprenorphine	0.2 mg	Acute	38	Volunteers with varying levels of anxiety / depression	Placebo	38	Volunteers with varying levels of anxiety / depression		Experimer
Bhagwagar[37]	2004	Citalopram	10 mg	Acute	10 10	Healthy Remitted Depression	Placebo	10 10	Healthy Remitted Depression		Experimer
Browning[38]	2007	Citalopram	20 mg	Acute	16	Healthy	Placebo	15	Healthy		Experimer
Bruhl[39]	2010	Citalopram Reboxetine	40 mg 8 mg	Acute Acute	16 14	Healthy	Placebo	15	Healthy		Experimer
Bukala[40]	2019	Scopolamine	1.5 mg	Acute	16	Healthy	Placebo	17	Healthy		Experimer

			Experimental Condition				Comparator Condition				
Author	Year	Drug	Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	Drug Assignment
Capitao[41]	2015	Fluoxetine	20 mg	Acute	17	Healthy	Placebo	18	Healthy		Experimer
Carvalho[42]	2006	Bupropion Sertraline	150 mg 50 mg	Acute Acute	32 32	Healthy Healthy	Placebo	35	Healthy		Experimer
Cerit[43]	2015	ARA290 / Epo	2 mg	Acute	18	Healthy	Placebo	18	Healthy		Experimer
Chandra[44]	2010	Aprepitant	125 mg	Acute	20	Healthy	Placebo	18	Healthy		Experimer
Constant[45] *	2005	Sertraline	50 mg/day [a]	7 weeks	20	Current Depression	No Drug	20	Current Depression	X	Experimer
Cooper[46]	2015	Tianeptine	12.5 mg	Acute	20	Healthy	Placebo	20	Healthy		Experimer

			Experimental Condition				Comparator Condition				
Author	Year	Drug	Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	Drug Assignment
Cremers[47]	2016	Escitalopram	0.375 mg/kg	Acute	14 17	Healthy Intermittent Explosive Disorder	Placebo Placebo	14 17	Healthy Intermittent Explosive Disorder		Experimer
Dannlowski[48]	2006	Mixed	Mixed [a]	Mixed [a]	22	Current Depression	Mixed	22	Current Depression	X	Observatio
Di Simplicio[49]	2012	Citalopram	20 mg/day	1 week	18	Healthy	Placebo	16	Healthy		Experimer
Di Simplicio[50]	2014	Citalopram	20 mg/day	1 week	16	High Neuroticism	Placebo	17	High Neuroticism		Experimer
Duschek[51] *	2014	Mixed	Mixed	Mixed	8	Females with fibromyalgia syndrome	No Drug	19	Females with fibromyalgia syndrome	X	Observatio
Fales[52] *	2009	Escitalopram [a]	10-20 mg/day [a]	8 weeks	23	Current Depression	No Drug	23	Current Depression		Observatio
Fu[53] *	2015	Duloxetine	60-120 mg/day [a]	12 weeks	24	Current Depression	No Drug	32	Current Depression		Experimer
Furey[54]	2015	Scopolamine	4.0 µg/kg	3 sessions 15 mins	16 15	Current Depression Healthy	Placebo Placebo	16 15	Current Depression Healthy		Experimer
Gibbs[55]	2013	Reboxetine	4 mg	Acute	57	Healthy	Placebo	62	Healthy		Experimer
Gibbs[56]	2014	Reboxetine	4 mg	Acute	38	Healthy	Placebo	34	Healthy		Experimer
Gorka[57]	2019	Mixed [b]	Mixed	12 weeks	28	Current Depression or Anxiety	No Drug	59	Current Depression or Anxiety		Observatio
Greer[58]	2014	Duloxetine	30-120 mg/day [a]	12 weeks	21	Current Depression	No Drug	21	Current Depression		Experimer
Grimm[59]	2018	Psilocybin	0.16 mg / kg	Acute	18	Healthy	Placebo	18	Healthy		Experimer
Harmer[60]	2003a	Citalopram	10 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimer
Harmer[1]	2003b	Reboxetine	4 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimer

		Experimental Condition				Comparator Condition				
Author	Year	Drug	Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	Drug Assignment
Harmer[61]	2004	Citalopram	20 mg/day	1 week	14	Healthy	Placebo	14	Healthy		Experimer
		Reboxetine	4 mg b.i.d./day	1 week	14	Healthy					
Harmer[62]	2006	Citalopram	20 mg/day	1 week	12	Healthy	Placebo	12	Healthy		Experimer
Harmer[63]	2008	Duloxetine	60 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimer
Harmer[64]	2009	Reboxetine	4 mg	Acute	15	Current Depression	Placebo	18	Current Depression		Experimer
					16	Healthy		15	Healthy		
Harmer[65]	2011	Agomelatine	25 mg/day	1 week	16	Healthy	Placebo	16	Healthy		Experimer
			50 mg/day	1 week	16	Healthy					
Harmer[66]	2013	Citalopram		Acute	18	Healthy	Placebo	18	Healthy		Experimer
		GSK424887		Acute	18	Healthy					
Kamboj[67]	2006	Scopolamine [c]	0.6 mg	Acute	16	Healthy	Placebo	16	Healthy		Experimer
Kerestes[68]	2009	Citalopram	20 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimer
		Reboxetine	4 mg b.i.d.	Acute	12	Healthy					
Kometer[69]	2012	Psilocybin [d]	215 ug/kg	Acute	17	Healthy	Placebo	17	Healthy		Experimer

			Experimental Condition				Comparator Condition				
Author	Year	Drug	Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	Drug Assignment
Komulainen[70]	2016	Mirtazapine	15 mg	Acute	15	Healthy	No Drug	15	Healthy		Experimen
Komulainen[71]	2018	Escitalopram	10 mg/day	1 week	17	Current Depression	Placebo	15	Current Depression		Experimen
Kraehenmann[72]	2015	Psilocybin	0.16 mg/kg	Acute	25	Healthy	Placebo	25	Healthy		Experimen
Labuschagne[73]	2010	Citalopram	20 mg	Acute	14	Healthy	Placebo	14	Healthy		Experimen
Lemogne[74]	2010	Mixed [a]	Mixed	6-14 weeks	8	Current Depression	Mixed antidepressants, ≤ 1 week treatment	8	Current Depression		Observatio
Lochner[75]	2012	Escitalopram	20 mg	Acute	19	Healthy	Placebo	18	Healthy		Experimen
					9	OCD, chronic antidepressants	Placebo	9	OCD, chronic antidepressants	X	
					5	OCD, no chronic antidepressants	Placebo	5	OCD, no chronic antidepressants		
Ma[76]	2015	Citalopram	30 mg	Acute	46	Healthy	Placebo	46	Healthy		Experimen
Matthews[77]	2010	Escitalopram	5-10 mg/day [a]	3 weeks	15	Healthy	Placebo	15	Healthy		Experimen
McCabe[78]	2009	Aprepitant	125 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimen
Merens (study 1)[79]	2008	Mixed	Varied	Min. 4 weeks	19	Remitted Depression	No Drug	20	Healthy		Observatio
Merens (study 2)[80]	2008	Mixed	Varied	Min. 4 weeks	20	Remitted Depression	No Drug	21	Healthy		Observatio
Miskowiak[81]	2007a	Reboxetine	4 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimen
Miskowiak[82]	2007 b	Epo	40,000 IU/mL	Acute	12	Healthy	Placebo	12	Healthy		Experimen

			Experimental Condition				Comparator Condition				
Author	Year	Drug	Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	Drug Assignment
Miskowiak[83]	2008	Epo	40,000 IU/mL	Acute	12	Healthy	Placebo	12	Healthy		Experimen
Miskowiak[84]	2009	Epo	40,000 IU/ml	Acute	9	Current Depression	Placebo	8	Current Depression	X	Experimen
Miskowiak[85]	2010	Epo	40,000 IU/mL	Acute	10	Current Depression	Placebo	9	Current Depression	X	Experimen
Mogg[86] *	2004	Paroxetine/Citalopram	20 mg/day	4 weeks	16	Current GAD	No Drug	19	Current GAD		Observatio
Murphy[87]	2009a	Citalopram	20 mg	Acute	13	Healthy	Placebo	13	Healthy		Experimen
Murphy[88]	2009 b	Citalopram Reboxetine	20 mg/day 4 mg b.i.d./day	1 week 1 week	14 14	Healthy Healthy	Placebo	14	Healthy		Experimen
Murrough[89] *	2015	Ketamine hydrochloride	0.5 mg kg−1	Acute	18	Treatment resistant MDD	No Drug	18	Treatment resistant MDD		Experimen
Naudin[90] *	2014	Mixed	Mixed [a]	Mixed	22	Current Depression	No Drug	41	Healthy		Observatio
Norbury[91]	2008	Reboxetine	4 mg b.i.d./day	1 week	12	Healthy	Placebo	12	Healthy		Experimen
Norbury[92]	2009	Citalopram	20 mg/day	7-10 days	16	Healthy	Placebo	12	Healthy		Experimen
Outhred[93]	2014	Escitalopram	20 mg	Acute	36	Healthy	Placebo	36	Healthy		Experimen
Papadatou-Pastou[94]	2012	Reboxetine	4 mg	Acute	12	Healthy	Placebo	12	Healthy		Experimen
Pringle[95]	2011	Aprepitant	125 mg/day	1 week	16	Healthy	Placebo	16	Healthy		Experimen
Pringle[17]	2012	Memantine	10 mg	Acute	16	Healthy	Placebo	15	Healthy		Experimen

Author	Year	Drug	Experimental Condition				Comparator Condition				Drug Assignment
			Dose	Duration	N	Sample	Comparator	N	Sample	Other Treatment	
Reed[96]	2018	Ketamine	0.5 mg/kg	Acute	33 26	Current Depression Healthy	Placebo	33 26	Current Depression Healthy		Experimer
Robertson[97]	2007	Bupropion	150-450mg/day [a]	8 weeks	8	Current Depression	No Drug	10	Current Depression		Observatio
Scheidegger[98]	2016	S-Ketamine	0.12-0.25 mg/kg/hour [a]	Acute	23	Healthy	No Drug	23	Healthy		Experimer
Shiroma[99]	2014	Citalopram	10 mg/day	1 week	27	Veterans aged 55+ Current Depression	No Drug	30	Veterans aged 55+ Current Depression		Experimer
Simmons[100]	2009	Escitalopram	5-10 mg/day [a]	3 weeks	15	Healthy	Placebo	15	Healthy		Experimer
Stein[101]	2012	Citalopram Reboxetine	20 mg/day 4 mg b.i.d./day	1 week 1 week	13 12	Healthy Healthy	Placebo	14	Healthy		Experimer
Sterpenich[102]	2019	Ketamine Hydrochloride	0.5 mg/kg	Acute	10	Treatment Resistant Depression	No Drug	10	Treatment Resistant Depression	X	Experimer
Tendolkar[103]	2011	Duloxetine	60 mg/day	2 weeks	18	Healthy	Placebo	18	Healthy		Experimer
Tranter[104]	2009	Citalopram Reboxetine	20 mg/day 4 mg b.i.d./day	2 weeks 6 weeks 2 weeks 6 weeks	32 39 27 30	Current Depression Current Depression	No Drug No Drug	57 51	Current Depression Current Depression		Experimer
Van Marle[105]	2011	Duloxetine	60 mg/day	2 weeks	19	Healthy	Placebo	19	Healthy		Experimer
Walsh[106] *	2018	Bupropion	150 mg 1-2 times/day [a]	2 weeks 6 weeks	43 41	Current Depression	No Drug	46	Current Depression		Experimer
Walsh[107]	2018 b	Bupropion	150 mg	Acute	20	Healthy	Placebo	20	Healthy		Experimer

Author	Year	Drug	Experimental Condition				Comparator Condition			Other Treatment	Drug Assignment
			Dose	Duration	N	Sample	Comparator	N	Sample		
Wang[108] *	2012	Fluoxetine	20-80 mg [a]	8 weeks	18	Current Depression	No Drug	18	Current Depression		Experimer
Warren[109]	2019	St John's Wort	300 mg 3 times/day	1 week	23	Healthy	Placebo	24	Healthy		Experimer
Wells[110]	2014	Mixed	Mixed	Mixed	21	Current Depression	No Drug	26	Current Depression	X	Observatio
							No Drug	47	Healthy	X	
Windischberger[111]	2010	Citalopram	20 mg	10 days	18	Healthy	Placebo	18	Healthy		Experimer
		Escitalopram	10 mg	10 days	18	Healthy					

FERT = Facial Emotion Recognition Task, RT = Reaction Time

* Study also included healthy volunteer group who did not receive any medication. As we are interested in the effects of the intervention we have restricted analyses to those reporting the most suitable comparison between experimental and comparator conditions.

[a] Clarifications of drug condition:

- Abel et al (2003) bolus dose of 0.23 mg/kg over 0–5 min, followed by an infusion (controlled by an I-Med pump system) of 0.5 mg/kg from 5 to 45 (± 15) min (maximum 1 h)
- Arce et al (2008) 5mg/day for 3 days, followed by 10 mg/day for 18 days
- Arce et al (2012) 20 mg starting dose that could be increased to 40 mg after 4 weeks.
- Bamford et al (2015) 30 mg first 3 days, 60 mg remaining 11 days
- Constant et al (2005) as stated except for 1 subject who was given sertraline 75 mg from the beginning of the study and another who was given sertraline 75 mg after 3 weeks of treatment
- Dannlowski et al (2006) At time 1, six patients were unmedicated (three were drug naïve and three had a wash-out phase). The others were taking mirtazapine 15–45 mg (n = 7), citalopram 20–40 mg (n = 4), venlafaxine 150–225 mg (n = 3), sertraline 50 mg (n = 1) and reboxetine 8 mg (n = 1).At time 2, three patients were still drug-naïve. The others were taking mirtazapine 15–60 mg (n = 8), citalopram 20–60 mg (n = 4), venlafaxine 150–300 mg (n = 4), nefadazone 500 mg (n = 1), reboxetine 8 mg (n = 1) and venlafaxine 300 mg combined with mirtazapine 30 mg (n = 1). Patients were tested at intake and on average after 7 weeks of psychoanalytic-interactional group therapy (M: 50 days; SD: 13).
- Fales et al (2009) Main drug was escitalopram but 5 participants treated with other medications (n = 4 Sertraline, of which n = 3 150 mg / day and n = 1 100 mg / day; n = 2 Paroxetine 20 mg/day). For participants receiving Escitalopram they were administered 10 mg per day, increased based on clinical response, end doses were 20 mg (n = 9) and 10 mg (n = 8).
- Fu et al (2015) 60 mg 8 weeks, then optional increase to 120 mg if no remission by 8 weeks
- Grady et al (20130 20mg/h 2 hours prior to tests, maintenance dose during testing of 8 mg/h (50 mg in total)
- Greer et al (2014) 30 mg 4 days, then 60 mg per day, increased to max of 120 mg (mean 90 mg/d (±28.5) at follow-up))
- Hornboll et al (2018) 20 mg /hour for 2 hours, 8 mg /hour maintenance during testing
- Lemogne et al (2010) At session 1 (≤ 1 week of treatment) 4 patients were taking a selective serotoninreuptake inhibitor (SSRI), 2 a serotonin–norepinephrine reuptake inhibitor (SNRI), and 2 a tricyclic antidepressant. At session 2 (6-14 weeks of treatment), SSRI had been switched for a SNRI for 1 patient.

- Matthews et al (2010) 5 mg/day 3 days, 10 mg/day 18 days
- Naudin et al (2014) 20 patients were treated with escitalopram, 1 patient was treated with venlafaxine, and one patient was treated with paroxetine.
- Robertson et al (2007) 150mg, then increased after 7 days to 300 mg, based on clinical response increased to 450 mg
- Scheidegger et al (2016) 0.12 mg/kg 15 mins prior fMRI, continuous infusion 0.25 mg/kg/h during scanning period
- Simmons et al (2009) 5mg/day for 3 days, followed by 10 mg/day for 18 days
- Walsh et al (2018) 150 mg once per day for 7-10 days, then 150 mg 2 x per day for remaining study duration
- Wang et al (2012) 20 mg initially, final dose ranged from 20-80 mg

[b] This study also included a group of participants receiving cognitive behavioural therapy. We have not included this groups in our analyses.

[c] This study also included a group of participants receiving 2 mg of the benzodiazepine, Lorazepam. We have not included this group in our analyses.

[d] This study also included experimental conditions of 50 mg ketaserin, and 50 mg of ketaserin with 215 ug/kg Psilocybin.

Details of other treatments (indicated by a 'X' in the Other Treatment column):

- Constant et al (2005) If the subjects were taking other psychotropic drugs (e.g., benzodiazepines), the dose of these drugs was constant during this 7-week period. No other psychotropic drug was allowed during the 7-week period. No other antidepressants for 2 months prior to study were allowed.
- Dannlowski et al (2006) Patients received 7 weeks of psychoanalytic-interactional group therapy
- Duschek et al (2014) nine participants (33.3%) were using nonopioid analgesics, two participants (7.4%) opiates, and one participant (3.7%) anxiolytics.
- Lochner (2012) Chronic antidepressant use
- Miskowiak et al (2009) All patients but two were on antidepressant medication. No change in medication was made in the 2 weeks prior to or during the study. Epo group: any antidepressant n = 7, SSRI n = 4, Dual Action = 2, MAOI n = 1, Antipsychotic n = 2, Benzodiazepine n = 3, No meds n = 1. Placebo group: any antidepressant n = 5, SSRI: n = 3, Dual Action = 1, MAOI n = 2, Antipsychotic n = 1, Lithium n = 1, Benzodiazepine n = 1, Thyroxin n = 1, no meds n = 1
- Miskowiak et al (2010) All patients but two were on antidepressant medication. No change in medication was made in the 2 weeks prior to or during the study.
- Sterpernich et al (2019) participants were taking a range of medication including antidepressants, antipsychotics, mood stabilisers and benzodiazepines. Participants were on stable doses for at least 6 weeks before participation
- Wells et al (2014) MDD patients on antidepressants were taking other medications (lamotrigine, trazodone, bupropion, lithium).

Study Quality

Risk of bias judgements are summarised in Figure 2, and individual judgements per study are available in Supplementary Figure S1.

Studies were for the most part at low risk of bias relating to the appropriateness of comparators, incomplete outcome data, and blinding. Blinding of outcome assessment was low for all studies as the outcome of interest was emotional processing that was predominantly measured using computerised tasks. However, risk of bias for selective reporting, random sequence generation and allocation concealment were mainly unclear. For selective reporting, this was predominantly driven by a lack of clear pre-registration. Of studies that implemented randomisation, most did not provide details on the method used to randomly assign participants or how group allocation was concealed. Approximately half of studies were also at high risk of 'other' bias, mainly due to reported financial conflicts of interest.

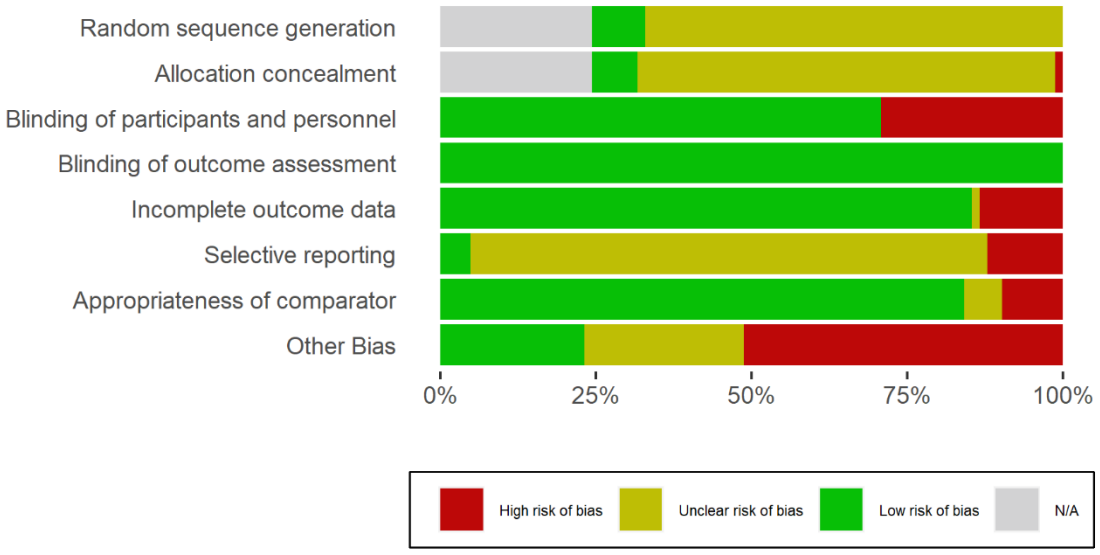


Figure 2

Risk of bias summary graph

Results of Narrative Synthesis

Narrative synthesis of findings grouped according to cognitive domain are reported in full in Supplementary Materials. Overall, effects were mixed. Whilst there was some evidence of effects consistent with the cognitive neuropsychological model, studies also reported null effects, or effects occurring in the opposite direction to those hypothesised. Within individual studies there was some evidence of differential effects according to sample [e.g., 37], antidepressant [e.g., 61], and dose [e.g., 65]. However, other studies did not find evidence of these differential effects [e.g., 47,64,68,104]. Additionally, when the evidence base is examined as a whole (Supplementary Tables S1-S14) varying directions of effects are observed across samples and antidepressants.

Results of Meta-Analysis

Of the 82 studies included in the narrative synthesis 28 were eligible for meta-analysis, providing 75 effects of positive emotions and 76 effects of negative emotions.

Positive

Across studies and effects, we did not find evidence of a difference in processing of positive stimuli between antidepressant and comparator conditions (SMD = 0.08, 95% CI = -0.05, 0.22, p = 0.230; Figure 3). Although, heterogeneity was moderate suggesting some inconsistency (Q (74) = 157.99, p < .001). This was primarily due to variation in effects reported within (I^2 Level 2 = 41.84%) rather than between studies (I^2 Level 3 = 10.31%). We also found evidence of small study bias (β = 1.24, 95% CI = -2.28, -0.21, p = 0.019), suggesting possible publication bias.

We did not find evidence that patterns of effects varied according to the cognitive domain examined (F(6, 69) = 0.983, p = 0.443). Confidence intervals for estimated SMDs per cognitive domain overlapped with the null (Table 2). However, there was weak evidence of an increase in accuracy for recognition of positive facial emotions in antidepressant versus control conditions (SMD = 0.20, 95% CI = -0.03, 0.44, p = 0.086). Sensitivity analyses of secondary outcomes were similar to our primary analyses (Supplementary Table S15). However, for facial emotion recognition we found no evidence of an association when using reaction times rather than accuracy (SMD = -0.04, 95% CI = -0.21, 0.12, p = 0.604).

Negative

We did not find evidence of an overall difference in processing of negative stimuli between antidepressant and comparator conditions (SMD = -0.04, 95% CI = -0.12, 0.04, p = 0.341; Figure 4). There was no evidence of substantial heterogeneity between or within studies suggesting that effects were consistent (Q(75) = 85.80, p = 0.185; between-study I^2 = 0.00%, within-study I^2 = 5.39%). There was also no evidence of small study bias (β = -0.32, 95% CI = -1.12, 0.48, p = 0.428).

Additionally, we did not find evidence that patterns of effects varied according to the cognitive domain examined (F(6, 70) = 0.999, p = 0.433). Across all cognitive domains confidence intervals for estimated SMDs overlapped with the null (Table 2). Sensitivity analyses of secondary outcomes did not show a deviation in findings from our primary analysis (Supplementary Table S15).

Table 2

Results from multilevel meta-analyses examining differences in processing of positive and negative stimuli in participants administered antidepressants compared to a control condition (placebo or no drug) with estimated standardised mean differences (Hedge's g) overall and by cognitive domain

	k		***n***		**SMD**	**SE**	**95% CI**	***p***
	Studies	**Effects**	**Antidepressant**	**Comparator**				
Positive [a]	28	75	625	612	0.08	0.07	-0.05, 0.22	0.230
Cognitive Domains [b]								
Word Categorisation (RT)	15	16	244	238	0.19	0.13	-0.08, 0.46	0.159
Dot Probe (RT Vigilance)	9	9	180	152	0.07	0.17	-0.27, 0.41	0.679
FERT (Accuracy)	18	20	340	388	0.20	0.12	-0.03, 0.44	0.086
Word Recall	12	13	223	204	0.00	0.15	-0.2	0.997

(Accuracy)							9, 0.29	
Word Recognition (Accuracy)	13	14	238	218	-0.09	0.14	-0.37, 0.19	0.505
Emotional Stroop (RT)	3	3	40	51	-0.12	0.31	-0.74, 0.50	0.700
Negative [c]	28	76	626	613	-0.04	0.04	-0.12, 0.04	0.341
Cognitive Domains [d]								
Word Categorisation (RT)	15	16	244	238	0.09	0.09	-0.10, 0.27	0.371
Dot Probe (RT Vigilance)	10	10	199	172	0.09	0.11	-0.13, 0.31	0.427
FERT (Accuracy)	18	20	340	388	-0.07	0.08	-0.23,	0.396

Word Recall (Accuracy)	12	13	223	204	-0.12	0.10	0.09 - 0.32, 0.09	0.255
Word Recognition (Accuracy)	13	14	239	218	-0.14	0.10	0.34 - 0.05	0.148
Emotional Stroop (RT)	3	3	40	52	-0.13	0.22	0.57 - 0.31	0.553

[a] Test for heterogeneity: Q (74) = 157.99, p < .001), σ Level 3 = 0.03, σ Level 2 = 0.11, I^2 Level 3 = 10.31%, I^2 Level 2 = 41.84%

[b] Test of moderators: F(6, 69) = 0.983, p = 0.443

[c] Test for heterogeneity: Q(75) = 85.80, p = 0.185, σ Level 3 = 0.00, σ Level 2 = 0.01, I^2 Level 3 = 0.00%, I^2 Level 2 = 5.39%

[d] Test of moderators: F(6, 70) = 0.999, p = 0.433

FERT = Facial Emotion Recognition Task, SMD = Standardised Mean Difference calculated using Hedge's g, SE = Standard Error, CI = Confidence Interval, RT = Reaction Times, RT Vigilance = Vigilance scores calculated by subtracting reaction times in congruent trials (probe appears in same position as emotional stimuli) from incongruent trials (probe appears in opposite position from the emotional stimuli)

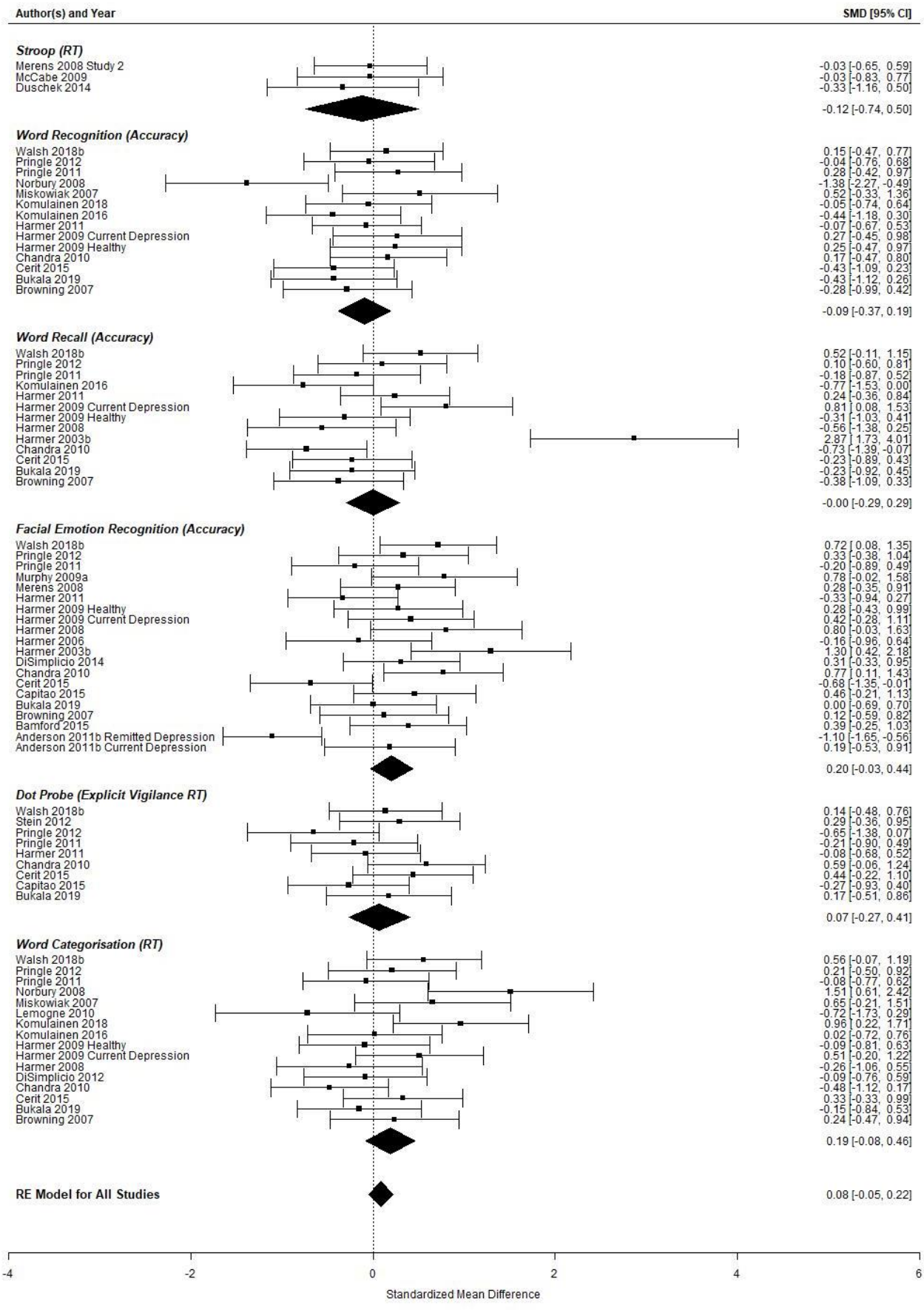

Author(s) and Year
SMD [95% CI]
Stroop (RT)
Merens 2008 Study 2 -0.03 [-0.65, 0.59]
McCabe 2009 -0.03 [-0.83, 0.77]
Duschek 2014 -0.33 [-1.16, 0.50]
-0.12 [-0.74, 0.50]
Word Recognition (Accuracy)
Walsh 2018b 0.15 [-0.47, 0.77]
Pringle 2012 -0.04 [-0.76, 0.68]
Pringle 2011 0.28 [-0.42, 0.97]
Norbury 2008 -1.38 [-2.27, -0.49]
Miskowiak 2007 0.52 [-0.33, 1.36]
Komulainen 2018 -0.05 [-0.74, 0.64]
Komulainen 2016 -0.44 [-1.18, 0.30]
Harmer 2011 -0.07 [-0.67, 0.53]
Harmer 2009 Current Depression 0.27 [-0.45, 0.98]
Harmer 2009 Healthy 0.25 [-0.47, 0.97]
Chandra 2010 0.17 [-0.47, 0.80]
Cerit 2015 -0.43 [-1.09, 0.23]
Bukala 2019 -0.43 [-1.12, 0.26]
Browning 2007 -0.28 [-0.99, 0.42]
-0.09 [-0.37, 0.19]
Word Recall (Accuracy)
Walsh 2018b 0.52 [-0.11, 1.15]
Pringle 2012 0.10 [-0.60, 0.81]
Pringle 2011 -0.18 [-0.87, 0.52]
Komulainen 2016 -0.77 [-1.53, 0.00]
Harmer 2011 0.24 [-0.36, 0.84]
Harmer 2009 Current Depression 0.81 [0.08, 1.53]
Harmer 2009 Healthy -0.31 [-1.03, 0.41]
Harmer 2008 -0.56 [-1.38, 0.25]
Harmer 2003b 2.87 [1.73, 4.01]
Chandra 2010 -0.73 [-1.39, -0.07]
Cerit 2015 -0.23 [-0.89, 0.43]
Bukala 2019 -0.23 [-0.92, 0.45]
Browning 2007 -0.38 [-1.09, 0.33]
-0.00 [-0.29, 0.29]
Facial Emotion Recognition (Accuracy)
Walsh 2018b 0.72 [0.08, 1.35]
Pringle 2012 0.33 [-0.38, 1.04]
Pringle 2011 -0.20 [-0.89, 0.49]
Murphy 2009a 0.78 [-0.02, 1.58]
Merens 2008 0.28 [-0.35, 0.91]
Harmer 2011 -0.33 [-0.94, 0.27]
Harmer 2009 Healthy 0.28 [-0.43, 0.99]
Harmer 2009 Current Depression 0.42 [-0.28, 1.11]
Harmer 2008 0.80 [-0.03, 1.63]
Harmer 2006 -0.16 [-0.96, 0.64]
Harmer 2003b 1.30 [0.42, 2.18]
DiSimplicio 2014 0.31 [-0.33, 0.95]
Chandra 2010 0.77 [0.11, 1.43]
Cerit 2015 -0.68 [-1.35, -0.01]
Capitao 2015 0.46 [-0.21, 1.13]
Bukala 2019 0.00 [-0.69, 0.70]
Browning 2007 0.12 [-0.59, 0.82]
Bamford 2015 0.39 [-0.25, 1.03]
Anderson 2011b Remitted Depression -1.10 [-1.65, -0.56]
Anderson 2011b Current Depression 0.19 [-0.53, 0.91]
0.20 [-0.03, 0.44]
Dot Probe (Explicit Vigilance RT)
Walsh 2018b 0.14 [-0.48, 0.76]
Stein 2012 0.29 [-0.36, 0.95]
Pringle 2012 -0.65 [-1.38, 0.07]
Pringle 2011 -0.21 [-0.90, 0.49]
Harmer 2011 -0.08 [-0.68, 0.52]
Chandra 2010 0.59 [-0.06, 1.24]
Cerit 2015 0.44 [-0.22, 1.10]
Capitao 2015 -0.27 [-0.93, 0.40]
Bukala 2019 0.17 [-0.51, 0.86]
0.07 [-0.27, 0.41]
Word Categorisation (RT)
Walsh 2018b 0.56 [-0.07, 1.19]
Pringle 2012 0.21 [-0.50, 0.92]
Pringle 2011 -0.08 [-0.77, 0.62]
Norbury 2008 1.51 [0.61, 2.42]
Miskowiak 2007 0.65 [-0.21, 1.51]
Lemogne 2010 -0.72 [-1.73, 0.29]
Komulainen 2018 0.96 [0.22, 1.71]
Komulainen 2016 0.02 [-0.72, 0.76]
Harmer 2009 Healthy -0.09 [-0.81, 0.63]
Harmer 2009 Current Depression 0.51 [-0.20, 1.22]
Harmer 2008 -0.26 [-1.06, 0.55]
DiSimplicio 2012 -0.09 [-0.76, 0.59]
Chandra 2010 -0.48 [-1.12, 0.17]
Cerit 2015 0.33 [-0.33, 0.99]
Bukala 2019 -0.15 [-0.84, 0.53]
Browning 2007 0.24 [-0.47, 0.94]
0.19 [-0.08, 0.46]
RE Model for All Studies
0.08 [-0.05, 0.22]
-4
-2
0
2
4
6
Standardized Mean Difference

Figure 3

Standardised mean differences (Hedge's g) effect sizes and 95% CIs for the effect of antidepressants on emotional processing of positive stimuli grouped by cognitive domain. Positive effect sizes represent an increase in emotional processing (increased accuracy, reduced reaction times), whereas negative effect sizes represent a decrease in emotional processing (decreased accuracy, increased reaction times).

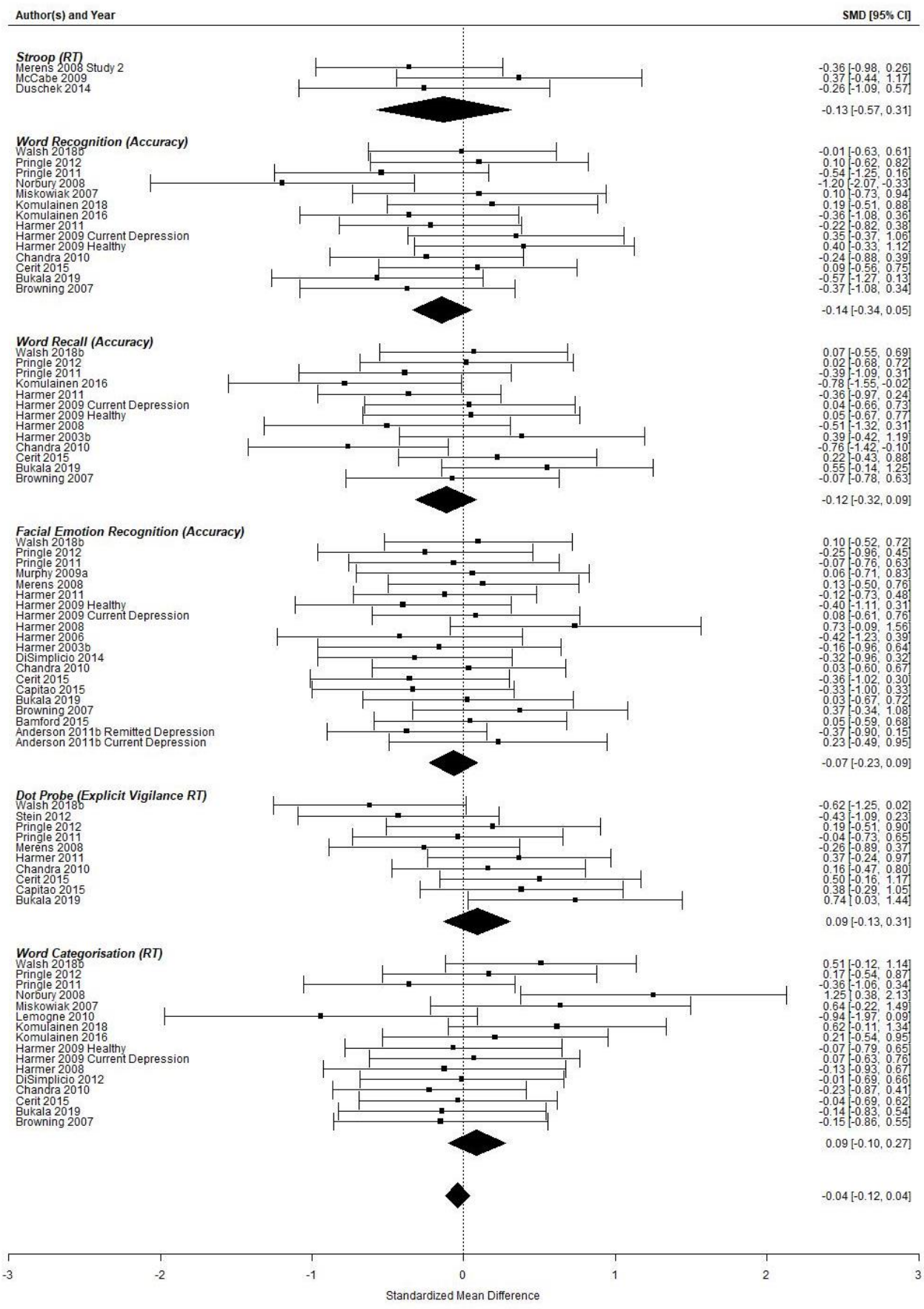

Author(s) and Year
SMD [95% CI]
Stroop (RT)
Merens 2008 Study 2 -0.36 [-0.98, 0.26]
McCabe 2009 0.37 [-0.44, 1.17]
Duschek 2014 -0.26 [-1.09, 0.57]
-0.13 [-0.57, 0.31]
Word Recognition (Accuracy)
Walsh 2018b -0.01 [-0.63, 0.61]
Pringle 2012 0.10 [-0.62, 0.82]
Pringle 2011 -0.54 [-1.25, 0.16]
Norbury 2008 -1.20 [-2.07, -0.33]
Miskowiak 2007 0.10 [-0.73, 0.94]
Komulainen 2018 0.19 [-0.51, 0.88]
Komulainen 2016 -0.36 [-1.08, 0.36]
Harmer 2011 -0.22 [-0.82, 0.38]
Harmer 2009 Current Depression 0.35 [-0.37, 1.06]
Harmer 2009 Healthy 0.40 [-0.33, 1.12]
Chandra 2010 -0.24 [-0.88, 0.39]
Cerit 2015 0.09 [-0.56, 0.75]
Bukala 2019 -0.57 [-1.27, 0.13]
Browning 2007 -0.37 [-1.08, 0.34]
-0.14 [-0.34, 0.05]
Word Recall (Accuracy)
Walsh 2018b 0.07 [-0.55, 0.69]
Pringle 2012 0.02 [-0.68, 0.72]
Pringle 2011 -0.39 [-1.09, 0.31]
Komulainen 2016 -0.78 [-1.55, -0.02]
Harmer 2011 -0.36 [-0.97, 0.24]
Harmer 2009 Current Depression 0.04 [-0.66, 0.73]
Harmer 2009 Healthy 0.05 [-0.67, 0.77]
Harmer 2008 -0.51 [-1.32, 0.31]
Harmer 2003b 0.39 [-0.42, 1.19]
Chandra 2010 -0.76 [-1.42, -0.10]
Cerit 2015 0.22 [-0.43, 0.88]
Bukala 2019 0.55 [-0.14, 1.25]
Browning 2007 -0.07 [-0.78, 0.63]
-0.12 [-0.32, 0.09]
Facial Emotion Recognition (Accuracy)
Walsh 2018b 0.10 [-0.52, 0.72]
Pringle 2012 -0.25 [-0.96, 0.45]
Pringle 2011 -0.07 [-0.76, 0.63]
Murphy 2009a 0.06 [-0.71, 0.83]
Merens 2008 0.13 [-0.50, 0.76]
Harmer 2011 -0.12 [-0.73, 0.48]
Harmer 2009 Healthy -0.40 [-1.11, 0.31]
Harmer 2009 Current Depression 0.08 [-0.61, 0.76]
Harmer 2008 0.73 [-0.09, 1.56]
Harmer 2006 -0.42 [-1.23, 0.39]
Harmer 2003b -0.16 [-0.96, 0.64]
DiSimplicio 2014 -0.32 [-0.96, 0.32]
Chandra 2010 0.03 [-0.60, 0.67]
Cerit 2015 -0.36 [-1.02, 0.30]
Capitao 2015 -0.33 [-1.00, 0.33]
Bukala 2019 0.03 [-0.67, 0.72]
Browning 2007 0.37 [-0.34, 1.08]
Bamford 2015 0.05 [-0.59, 0.68]
Anderson 2011b Remitted Depression -0.37 [-0.90, 0.15]
Anderson 2011b Current Depression 0.23 [-0.49, 0.95]
-0.07 [-0.23, 0.09]
Dot Probe (Explicit Vigilance RT)
Walsh 2018b -0.62 [-1.25, 0.02]
Stein 2012 -0.43 [-1.09, 0.23]
Pringle 2012 0.19 [-0.51, 0.90]
Pringle 2011 -0.04 [-0.73, 0.65]
Merens 2008 -0.26 [-0.89, 0.37]
Harmer 2011 0.37 [-0.24, 0.97]
Chandra 2010 0.16 [-0.47, 0.80]
Cerit 2015 0.50 [-0.16, 1.17]
Capitao 2015 0.38 [-0.29, 1.05]
Bukala 2019 0.74 [0.03, 1.44]
0.09 [-0.13, 0.31]
Word Categorisation (RT)
Walsh 2018b 0.51 [-0.12, 1.14]
Pringle 2012 0.17 [-0.54, 0.87]
Pringle 2011 -0.36 [-1.06, 0.34]
Norbury 2008 1.25 [0.38, 2.13]
Miskowiak 2007 0.64 [-0.22, 1.49]
Lemogne 2010 -0.94 [-1.97, 0.09]
Komulainen 2018 0.62 [-0.11, 1.34]
Komulainen 2016 0.21 [-0.54, 0.95]
Harmer 2009 Healthy -0.07 [-0.79, 0.65]
Harmer 2009 Current Depression 0.07 [-0.63, 0.76]
Harmer 2008 -0.13 [-0.93, 0.67]
DiSimplicio 2012 -0.01 [-0.69, 0.66]
Chandra 2010 -0.23 [-0.87, 0.41]
Cerit 2015 -0.04 [-0.69, 0.62]
Bukala 2019 -0.14 [-0.83, 0.54]
Browning 2007 -0.15 [-0.86, 0.55]
0.09 [-0.10, 0.27]
-0.04 [-0.12, 0.04]
-3
-2
-1
0
1
2
3
Standardized Mean Difference

Figure 4

Standardised mean differences (Hedge's g) effect sizes and 95% CIs for the effect of antidepressants on emotional processing of negative stimuli grouped by cognitive domain. Positive effect sizes represent an increase in emotional processing (increased accuracy, reduced reaction times), whereas negative effect sizes represent a decrease in emotional processing (decreased accuracy, increased reaction times).

Discussion

The cognitive neuropsychological model proposes that antidepressants operate by remediating negative emotional biases, increasing processing of positive emotional stimuli and decreasing processing of negative stimuli [1,2]. In contrast to these predictions, narrative synthesis and meta-analysis indicated that antidepressants were not associated with reliable differences in the processing of positive or negative emotional stimuli across a range of cognitive domains, including categorisation, attention, and memory.

The findings of this systematic review differ to those of previous meta-analyses and narrative reviews within this field. This may be partly due to differences in the focus of previous evidence. In contrast with our findings, earlier meta-analyses focused on brain activity reported effects in line with the cognitive neuropsychological model [23,24]. It is possible that effects of antidepressants occur at a neural level in the absence of behavioural markers. Alternatively, the tasks utilised in research at present may not effectively measure behaviours correlated with observed neural activity. Differences in findings from previous narrative reviews may be due to the methods employed. Of previous reviews examining behavioural evidence none took a systematic approach to identifying literature. Furthermore, in comparison to previous reviews which reported effects at a study level we compared effects within individual cognitive domains. As some studies examined multiple domains with several outcome measures, if summarised at the study level a significant effect for a single individual task or outcome may be taken as supportive evidence despite other null effects. However, when all reported effects are included, and the evidence for each cognitive task and behavioural measure are assessed independently, as in this review, results appear to be more disparate.

When examining effects according to the cognitive task employed, we did not find evidence of group differences in either the narrative summary or meta-analysis. Across all tasks and outcomes included in our meta-analysis effects overlapped with the null. However, the strongest evidence was found for the effect of antidepressants on accurate recognition of positive facial emotions. Of studies reporting significant effects for the recognition of positive emotions, all reported an increase associated with antidepressants. Additionally, in our meta-analysis effects for positive emotions for accuracy in the facial emotion recognition task (FERT) overlapped with the null by only a small margin (SMD = 0.20, 95% CI: -0.03, 0.44). Previous research has reported an association between depression and reduced recognition of happiness [15]. An increase in the recognition of positive emotions from

antidepressant treatment may address this deficit. Our results tentatively suggest that further research examining the effects of antidepressants on emotional processing should focus on accuracy in facial emotion recognition as a potential candidate mechanism.

However, the validity of the estimated effect for positive emotional processing, and the FERT, may be questioned. Whilst the size of the aggregated effect for positive facial emotion recognition is the largest of all cognitive domains, the effect is still small. To detect an effect of this magnitude at conventional levels of power and alpha would require 788 participants. All studies included in this review are underpowered to detect effects of this size, increasing the likelihood of false positive findings [112]. Additionally, we found evidence of publication bias for positive emotional processing, suggesting that some negative or null effects may not be reported. Although, there were several good practices within this field such as use of randomisation and blinding, there was also unclear evidence of selective reporting due to a lack of pre-registration. This again may increase the risk of false-positive findings due to potential undisclosed flexibility in statistical analyses [113]. The use of multiple cognitive tasks and outcome measures per study may further increase this risk. The aggregated effects we observed may therefore be an overestimation. Our results indicate the need for studies within this field to recruit larger samples for adequate power to detect small effects and to pre-register studies to address potential issues with publication bias and undisclosed flexibility of analyses.

Clinical Implications

Antidepressant treatment is characterised by limited response to initial treatment and delayed therapeutic effects [8,10]. Negative cognitive biases are believed to be important in the maintenance of depression [2]. Change in emotional processing has therefore been proposed as an early predictor of treatment response [2]. Most studies identified in this review examined an acute dose of an antidepressant in comparison to a placebo control, limiting direct conclusions about this possibility. However, our findings suggest that antidepressants do not reliably alter emotional processing in either acute or longer-term administration. Change in emotional processing is therefore unlikely to be a clear predictor of response. Additionally, there are suggestions that emotional processing may act as a potential biomarker of efficacy for novel antidepressants treatments [114]. Again, the results of this review suggest limited evidence for this possibility as an association with emotional processing was not reliably observed in antidepressants with demonstrated efficacy. Further research is required to

understand the early therapeutic action of antidepressants and to identify potential biomarkers of treatment response.

Limitations

We took a broad approach in this review integrating evidence across a variety of samples, antidepressants, and study designs. It is possible that specific effects on emotional processing were obscured. However, narrative synthesis of studies did not indicate substantial variations in results. Additionally, meta-analysis indicated low levels of heterogeneity for negative emotional processing.

This review focused on emotional processing; however, the cognitive neuropsychological model also proposes a role for reward processing. Whilst emotion and reward are intrinsically linked, distinct pathways have been identified (121,122). It is possible that antidepressants may operate by remediating reward processing biases.

Only a subset of studies reported or provided data for meta-analysis. The findings of our meta-analysis may therefore be unrepresentative of the wider field. However, results were consistent with narrative synthesis which included all studies identified in the review. Future studies within this field should publish data as open access to allow for aggregation of effects.

Conclusions

We did not find evidence of a reliable association between antidepressant administration and increased processing of positive information or decreased processing of negative information. Our findings do not support theories that antidepressants may operate by remediating negative emotional biases associated with depression. Future research should recruit larger samples to ensure adequate statistical power to detect small effects and publish data as open access to allow for further clarification of the effect of antidepressants on emotional processing.

References

1. Harmer CJ, Hill SA, Taylor MJ, Cowen PJ, Goodwin GM. Toward a neuropsychological theory of antidepressant drug action: Increase in positive emotional bias after potentiation of norepinephrine activity. *Am J Psychiatry*. 2003;160(5):990-992. doi:10.1176/appi.ajp.160.5.990

2. Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*. 2012;37(1):117-136. doi:10.1038/npp.2011.183

3. Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9

4. Scott KM, Bruffaerts R, Tsang A, et al. Depression–anxiety relationships with chronic physical conditions: results from the World Mental Health Surveys. *J Affect Disord*. 2007;103(1-3):113-120.

5. Laursen TM, Musliner KL, Benros ME, Vestergaard M, Munk-Olsen T. Mortality and life expectancy in persons with severe unipolar depression. *J Affect Disord*. 2016;193:203-207. doi:10.1016/j.jad.2015.12.067

6. Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry*. 2015;76(2):155-162. doi:10.4088/JCP.14m09298

7. Cipriani A, Furukawa TA, Salanti G, et al. Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *Focus (Madison)*. 2018;16(4):420-429. doi:10.1176/appi.focus.16407

8. Warden D, Rush AJ, Trivedi MH, Fava M, Wisniewski SR. The STAR*D project results: A comprehensive review of findings. *Curr Psychiatry Rep*. 2007;9(6):449-459. doi:10.1007/s11920-007-0061-3

9. Thase ME, Rush AJ. When at first you don't succeed: Sequential strategies for antidepressant nonresponders. *J Clin Psychiatry*. 1997;58(SUPPL. 13):23-29.

10. The National Institute for Health and Care Excellence. *The NICE Guideline on the*

Treatment and Depression the Treatment and Management of Depression.; 2009.

11. Masand PS. Tolerability and adherence issues in antidepressant therapy. *Clin Ther*. 2003;25(8):2289-2304. doi:10.1016/S0149-2918(03)80220-5

12. Sansone RA, Sansone LA. Antidepressant adherence: Are patients taking their medications? *Innov Clin Neurosci*. 2012;9(5-6):41-46.

13. Breitenstein B, Scheuer S, Holsboer F. Are there meaningful biomarkers of treatment response for depression? *Drug Discov Today*. 2014;19(5):539-561. doi:10.1016/j.drudis.2014.02.002

14. Dalgleish T, Watts FN. Biases of attention and memory in disorders of anxiety and depression. *Clin Psychol Rev*. 1990;10(5):589-604. doi:10.1016/0272-7358(90)90098-U

15. Dalili MN, Penton-Voak IS, Harmer CJ, Munafò MR. Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychol Med*. 2015;45(6):1135-1144. doi:10.1017/S0033291714002591

16. Roiser JP, Sahakian BJ. Hot and cold cognition in depression. *CNS Spectr*. 2013;18(3):139-149. doi:10.1017/S1092852913000072

17. Pringle A, Parsons E, Cowen LG, McTavish SF, Cowen PJ, Harmer CJ. Using an experimental medicine model to understand the antidepressant potential of the N-Methyl-D-aspartic acid (NMDA) receptor antagonist memantine. *J Psychopharmacol*. 2012;26(11):1417-1423. doi:10.1177/0269881112446535

18. Godlewska BR, Harmer CJ. Cognitive neuropsychological theory of antidepressant action: a modern-day approach to depression and its treatment. *Psychopharmacology (Berl)*. 2021;238(5):1265-1278. doi:10.1007/s00213-019-05448-0

19. Merens W, Willem Van der Does AJ, Spinhoven P. The effects of serotonin manipulations on emotional information processing and mood. *J Affect Disord*. 2007;103(1-3):43-62. doi:10.1016/j.jad.2007.01.032

20. Pringle A, Browning M, Cowen PJ, Harmer CJ. A cognitive neuropsychological model of antidepressant drug action. *Prog Neuro-Psychopharmacology Biol Psychiatry*. 2011;35(7):1586-1592. doi:10.1016/j.pnpbp.2010.07.022

21. Warren MB, Pringle A, Harmer CJ. A neurocognitive model for understanding treatment action in depression. *Philos Trans R Soc B Biol Sci*. 2015;370(1677). doi:10.1098/rstb.2014.0213

22. Godlewska BR. Cognitive neuropsychological theory: Reconciliation of psychological and biological approaches for depression. *Pharmacol Ther*. 2019;197:38-51. doi:10.1016/j.pharmthera.2018.12.010

23. Delaveau P, Jabourian M, Lemogne C, Guionnet S, Bergouignan L, Fossati P. Brain effects of antidepressants in major depression: A meta-analysis of emotional processing studies. *J Affect Disord*. 2011;130(1-2):66-74. doi:10.1016/j.jad.2010.09.032

24. Ma Y. Neuropsychological mechanism underlying antidepressant effect: A systematic meta-analysis. *Mol Psychiatry*. 2015;20(3):311-319. doi:10.1038/mp.2014.24

25. Harrer M, Cuijpers P, Furukawa T., Ebert DD. *Doing Meta-Analysis in R: A Hands-on Guide*.; 2019. doi:10.5281/zenodo.2551803

26. Sang HK, Hamann S. Neural correlates of positive and negative emotion regulation. *J Cogn Neurosci*. 2007;19(5):776-798. doi:10.1162/jocn.2007.19.5.776

27. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions*.; 2019. doi:10.1002/9781119536604

28. Viechtbauer W. Package "metafor." Published online 2015.

29. Harrer M, Cuijpers P, Furukawa T, Ebert DD. dmetar: Companion R package for the Guide'Doing Meta-Analysis in R'. Published online 2019.

30. Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics*. 2018;74(3):785-794. doi:10.1111/biom.12817

31. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629

32. Anderson IM, Shippen C, Juhasz G, et al. State-dependent alteration in face emotion recognition in depression. *Br J Psychiatry*. 2011;198(4):302-308. doi:10.1192/bjp.bp.110.078139

33. Arce E, Simmons AN, Lovero KL, Stein MB, Paulus MP. Escitalopram effects on insula and amygdala BOLD activation during emotional processing. *Psychopharmacology (Berl)*. 2008;196(4):661-672. doi:10.1007/s00213-007-1004-8

34. Arnone D, Horder J, Cowen PJ, Harmer CJ. Early effects of mirtazapine on emotional processing. *Psychopharmacology (Berl)*. 2009;203(4):685. doi:10.1007/s00213-008-1410-6

35. Bamford S, Penton-Voak I, Pinkney V, Baldwin DS, Munafò MR, Garner M. Early effects of duloxetine on emotion recognition in healthy volunteers. *J Psychopharmacol*. 2015;29(5):634-641. doi:10.1177/0269881115570085

36. Bershad AK, Ruiz NA, De Wit H. Effects of buprenorphine on responses to emotional stimuli in individuals with a range of mood symptomatology. *Int J Neuropsychopharmacol*. 2018;21(2):120-127. doi:10.1093/ijnp/pyx077

37. Bhagwagar Z, Cowen PJ, Goodwin GM, Harmer CJ. Normalization of Enhanced Fear Recognition by Acute SSRI Treatment in Subjects with a Previous History of Depression. *Am J Psychiatry*. 2004;161(1):166-168. doi:10.1176/appi.ajp.161.1.166

38. Browning M, Reid C, Cowen PJ, Goodwin GM, Harmer CJ. A single dose of citalopram increases fear recognition in healthy subjects. *J Psychopharmacol*. 2007;21(7):684-690. doi:10.1177/0269881106074062

39. Brühl AB, Kaffenberger T, Herwig U. Serotonergic and noradrenergic modulation of emotion processing by single dose antidepressants. *Neuropsychopharmacology*. 2010;35(2):521-533. doi:10.1038/npp.2009.159

40. Bukala BR, Browning M, Cowen PJ, Harmer CJ, Murphy SE. Overnight transdermal scopolamine patch administration has no clear effect on cognition and emotional processing in healthy volunteers. *J Psychopharmacol*. 2019;33(2):255-257. doi:10.1177/0269881118817161

41. Capitão LP, Murphy SE, Browning M, Cowen PJ, Harmer CJ. Acute fluoxetine modulates emotional processing in young adult volunteers. *Psychol Med.* 2015;45(11):2295-2308. doi:10.1017/S0033291715000240

42. Carvalho AF, Köhler CA, Cruz EP, et al. Acute treatment with the antidepressants bupropion and sertraline do not influence memory retrieval in man. *Eur Arch*

Psychiatry Clin Neurosci. 2006;256(5):320-325. doi:10.1007/s00406-006-0640-z

43. Cerit H, Veer IM, Dahan A, et al. Testing the antidepressant properties of the peptide ARA290 in a human neuropsychological model of drug action. *Eur Neuropsychopharmacol*. 2015;25(12):2289-2299. doi:10.1016/j.euroneuro.2015.09.005

44. Chandra P, Hafizi S, Massey-Chase RM, Goodwin GM, Cowen PJ, Harmer CJ. NK1 receptor antagonism and emotional processing in healthy volunteers. *J Psychopharmacol*. 2010;24(4):481-487. doi:10.1177/0269881109103101

45. Constant EL, Adam S, Gillain B, Seron X, Bruyer R, Seghers A. Effects of sertraline on depressive symptoms and attentional and executive functions in major depression. *Depress Anxiety*. 2005;21(2):78-89. doi:10.1002/da.20060

46. Cooper CM, Whiting DA, Cowen PJ, Harmer CJ. Tianeptine in an experimental medicine model of antidepressant action. *J Psychopharmacol*. 2015;29(5):582-590. doi:10.1177/0269881115573810

47. Cremers H, Lee R, Keedy S, Phan KL, Coccaro E. Effects of escitalopram administration on face processing in intermittent explosive disorder: An fMRI Study. *Neuropsychopharmacology*. 2016;41(2):590-597. doi:10.1038/npp.2015.187

48. Dannlowski U, Kersting A, Donges US, Lalee-Mentzel J, Arolt V, Suslow T. Masked facial affect priming is associated with therapy response in clinical depression. *Eur Arch Psychiatry Clin Neurosci*. 2006;256(4):215-221. doi:10.1007/s00406-005-0628-0

49. Di Simplicio M, Norbury R, Harmer CJ. Short-term antidepressant administration reduces negative self-referential processing in the medial prefrontal cortex in subjects at risk for depression. *Mol Psychiatry*. 2012;17(5):503-510. doi:10.1038/mp.2011.16

50. Di Simplicio M, Doallo S, Costoloni G, Rohenkohl G, Nobre AC, Harmer CJ. "Can you look me in the face?" Short-term SSRI administration reverts avoidant ocular face exploration in subjects at risk for psychopathology. *Neuropsychopharmacology*. 2014;39(13):3059-3066. doi:10.1038/npp.2014.159

51. Duschek S, Werner NS, Limbert N, Winkelmann A, Montoya P. Attentional bias toward negative information in patients with fibromyalgia syndrome. *Pain Med*. 2014;15(4):603-612. doi:10.1111/pme.12360

52. Fales CL, Barch DM, Rundle MM, et al. Antidepressant treatment normalizes hypoactivity in dorsolateral prefrontal cortex during emotional interference processing in major depression. *J Affect Disord*. 2009;112(1-3):206-211. doi:10.1016/j.jad.2008.04.027

53. Fu CHY, Costafreda SG, Sankar A, et al. Multimodal functional and structural neuroimaging investigation of major depressive disorder following treatment with duloxetine. *BMC Psychiatry*. 2015;15(1):1-11. doi:10.1186/s12888-015-0457-2

54. Furey ML, Drevets WC, Szczepanik J, Khanna A, Nugent A, Zarate CA. Pretreatment differences in BOLD response to emotional faces correlate with Antidepress ant response to scopolamine. *Int J Neuropsychopharmacol*. 2015;18(8). doi:10.1093/ijnp/pyv028

55. Gibbs AA, Bautista CE, Mowlem FD, Naudts KH, Duka T. Alpha 2B adrenoceptor genotype moderates effect of Reboxetine on negative emotional memory bias in healthy volunteers. *J Neurosci*. 2013;33(43):17023-17028. doi:10.1523/JNEUROSCI.2124-13.2013

56. Gibbs AA, Bautista CE, Mowlem FD, Naudts KH, Duka DT. Catechol-O-methyltransferase val158met genotype determines effect of reboxetine on emotional memory in healthy male volunteers. *J Psychiatry Neurosci*. 2014;39(3):E24. doi:10.1503/jpn.130131

57. Gorka SM, Young CB, Klumpp H, et al. Emotion-based brain mechanisms and predictors for SSRI and CBT treatment of anxiety and depression: a randomized trial. *Neuropsychopharmacology*. 2019;44(9):1639-1648. doi:10.1038/s41386-019-0407-7

58. Greer TL, Sunderajan P, Grannemann BD, Kurian BT, Trivedi MH. Does duloxetine improve cognitive function independently of its antidepressant effect in patients with major depressive disorder and subjective reports of cognitive dysfunction? *Depress Res Treat*. Published online 2014. doi:10.1155/2014/627863

59. Grimm O, Kraehenmann R, Preller KH, Seifritz E, Vollenweider FX. Psilocybin modulates functional connectivity of the amygdala during emotional face discrimination. *Eur Neuropsychopharmacol*. 2018;28(6):691-700. doi:10.1016/j.euroneuro.2018.03.016

60. Harmer CJ, Bhagwagar Z, Perrett DI, Völlm BA, Cowen PJ, Goodwin GM. Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology*. 2003;28(1):148-152. doi:10.1038/sj.npp.1300004

61. Harmer CJ, Shelley NC, Cowen PJ, Goodwin GM. Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *Am J Psychiatry*. 2004;161(7):1256-1263. doi:10.1176/appi.ajp.161.7.1256

62. Harmer CJ, Mackay CE, Reid CB, Cowen PJ, Goodwin GM. Antidepressant Drug Treatment Modifies the Neural Processing of Nonconscious Threat Cues. *Biol Psychiatry*. 2006;59(9):816-820. doi:10.1016/j.biopsych.2005.10.015

63. Harmer CJ, Heinzen J, O'Sullivan U, Ayres RA, Cowen PJ. Dissociable effects of acute antidepressant drug administration on subjective and emotional processing measures in healthy volunteers. *Psychopharmacology (Berl)*. 2008;199(4):495-502. doi:10.1007/s00213-007-1058-7

64. Harmer CJ, O'Sullivan U, Favaron E, et al. Effect of acute antidepressant administration on negative affective bias in depressed patients. *Am J Psychiatry*. 2009;166(10):1178-1184. doi:10.1176/appi.ajp.2009.09020149

65. Harmer CJ, De Bodinat C, Dawson GR, et al. Agomelatine facilitates positive versus negative affective processing in healthy volunteer models. *J Psychopharmacol*. 2011;25(9):1159-1167. doi:10.1177/0269881110376689

66. Harmer CJ, Dawson GR, Dourish CT, et al. Combined NK1 antagonism and serotonin reuptake inhibition: Effects on emotional processing in humans. *J Psychopharmacol*. 2013;27(5):435-443. doi:10.1177/0269881112472558

67. Kamboj SK, Curran H V. Scopolamine induces impairments in the recognition of human facial expressions of anger and disgust. *Psychopharmacology (Berl)*. 2006;185(4):529-535. doi:10.1007/s00213-006-0332-4

68. Kerestes R, Labuschagne I, Croft RJ, et al. Evidence for modulation of facial emotional processing bias during emotional expression decoding by serotonergic and noradrenergic antidepressants: An event-related potential (ERP) study. *Psychopharmacology (Berl)*. 2009;202(4):621-634. doi:10.1007/s00213-008-1340-3

69. Kometer M, Schmidt A, Bachmann R, Studerus E, Seifritz E, Vollenweider FX. Psilocybin biases facial recognition, goal-directed behavior, and mood state toward positive relative to negative emotions through different serotonergic subreceptors. *Biol Psychiatry*. 2012;72(11):898-906. doi:10.1016/j.biopsych.2012.04.005

70. Komulainen E, Heikkilä R, Meskanen K, et al. A single dose of mirtazapine attenuates neural responses to self-referential processing. *J Psychopharmacol*. 2016;30(1):23-32. doi:10.1177/0269881115616384

71. Komulainen E, Heikkilä R, Nummenmaa L, et al. Short-term escitalopram treatment normalizes aberrant self-referential processing in major depressive disorder. *J Affect Disord*. 2018;236:222-229. doi:10.1016/j.jad.2018.04.096

72. Kraehenmann R, Preller KH, Scheidegger M, et al. Psilocybin-induced decrease in amygdala reactivity correlates with enhanced positive mood in healthy volunteers. *Biol Psychiatry*. 2015;78(8):572-581. doi:10.1016/j.biopsych.2014.04.010

73. Labuschagne I, Croft RJ, Phan KL, Nathan PJ. Augmenting serotonin neurotransmission with citalopram modulates emotional expression decoding but not structural encoding of moderate intensity sad facial emotional stimuli: An event-related potential (ERP) investigation. *J Psychopharmacol*. 2010;24(8):1153-1164. doi:10.1177/0269881108097878

74. Lemogne C, Mayberg H, Bergouignan L, et al. Self-referential processing and the prefrontal cortex over the course of depression: A pilot study. *J Affect Disord*. 2010;124(1-2):196-201. doi:10.1016/j.jad.2009.11.003

75. Lochner C, Simmons C, Kidd M, et al. Differential effects of escitalopram challenge on disgust processing in obsessive-compulsive disorder. *Behav Brain Res*. 2012;226(1):274-280. doi:10.1016/j.bbr.2011.09.029

76. Ma Y, Li B, Wang C, Zhang W, Rao Y, Han S. Allelic variation in 5-HTTLPR and the effects of citalopram on the emotional neural network. *Br J Psychiatry*. 2015;206(5):385-392. doi:10.1192/bjp.bp.114.150128

77. Matthews SC, Simmons AN, Strigo IA, Arce E, Stein MB, Paulus MP. Escitalopram attenuates posterior cingulate activity during self-evaluation in healthy volunteers. *Psychiatry Res - Neuroimaging*. 2010;182(2):81-87.

doi:10.1016/j.pscychresns.2010.02.003

78. McCabe C, Cowen PJ, Harmer CJ. NK1 receptor antagonism and the neural processing of emotional information in healthy volunteers. *Int J Neuropsychopharmacol*. 2009;12(9):1261-1274. doi:10.1017/S1461145709990150

79. Merens W, Booij L, Van Der Does AJW. Residual cognitive impairments in remitted depressed patients. Study 1. *Depress Anxiety*. 2008;25(6):E27-E36. doi:10.1002/da.20391

80. Merens W, Booij L, Van Der Does AJW. Residual cognitive impairments in remitted depressed patients. Study 2. *Depress Anxiety*. 2008;25(6):E27-E36. doi:10.1002/da.20391

81. Miskowiak K, Papadatou-Pastou M, Cowen PJ, Goodwin GM, Norbury R, Harmer CJ. Single dose antidepressant administration modulates the neural processing of self-referent personality trait words. *Neuroimage*. 2007;37(3):904-911. doi:10.1016/j.neuroimage.2007.05.036

82. Miskowiak K, O'Sullivan U, Harmer CJ. Erythropoietin Reduces Neural and Cognitive Processing of Fear in Human Models of Antidepressant Drug Action. *Biol Psychiatry*. 2007;62(11):1244-1250. doi:10.1016/j.biopsych.2007.01.011

83. Miskowiak K, Inkster B, Selvaraj S, Wise R, Goodwin GM, Harmer CJ. Erythropoietin improves mood and modulates the cognitive and neural processing of emotion 3 days post administration. *Neuropsychopharmacology*. 2008;33(3):611-618. doi:10.1038/sj.npp.1301439

84. Miskowiak KW, Favaron E, Hafizi S, et al. Effects of erythropoietin on emotional processing biases in patients with major depression: An exploratory fMRI study. *Psychopharmacology (Berl)*. 2009;207(1):133-142. doi:10.1007/s00213-009-1641-1

85. Miskowiak KW, Favaron E, Hafizi S, et al. Erythropoietin modulates neural and cognitive processing of emotional information in biomarker models of antidepressant drug action in depressed patients. *Psychopharmacology (Berl)*. 2010;210(3):419-428. doi:10.1007/s00213-010-1842-7

86. Mogg K, Baldwin DS, Brodrick P, Bradley BP. Effect of short-term SSRI treatment on cognitive bias in generalised anxiety disorder. *Psychopharmacology (Berl)*.

2004;176(3-4):466-470. doi:10.1007/s00213-004-1902-y

87. Murphy SE, Norbury R, O'Sullivan U, Cowen PJ, Harmer CJ. Effect of a single dose of citalopram on amygdala response to emotional faces. *Br J Psychiatry*. 2009;194(6):535-540. doi:10.1192/bjp.bp.108.056093

88. Murphy SE, Yiend J, Lester KJ, Cowen PJ, Harmer CJ. Short-term serotonergic but not noradrenergic antidepressant administration reduces attentional vigilance to threat in healthy volunteers. *Int J Neuropsychopharmacol*. 2009;12(2):169-179. doi:10.1017/S1461145708009164

89. Murrough JW, Collins KA, Fields J, et al. Regulation of neural responses to emotion perception by ketamine in individuals with treatment-resistant major depressive disorder. *Transl Psychiatry*. 2015;5(2):e509-e509. doi:10.1038/tp.2015.10

90. Naudin M, Carl T, Surguladze S, et al. Perceptive biases in major depressive episode. *PLoS One*. 2014;9(2):e86832. doi:10.1371/journal.pone.0086832

91. Norbury R, Mackay CE, Cowen PJ, Goodwin GM, Harmer CJ. The effects of reboxetine on emotional processing in healthy volunteers: An fMRI study. *Mol Psychiatry*. 2008;13(11):1011-1020. doi:10.1038/sj.mp.4002091

92. Norbury R, Taylor MJ, Selvaraj S, Murphy SE, Harmer CJ, Cowen PJ. Short-term antidepressant treatment modulates amygdala response to happy faces. *Psychopharmacology (Berl)*. 2009;206(2):197-204. doi:10.1007/s00213-009-1597-1

93. Outhred T, Das P, Felmingham KL, et al. Impact of acute administration of escitalopram on the processing of emotional and neutral images: A randomized crossover fMRI study of healthy women. *J Psychiatry Neurosci*. 2014;39(4):267. doi:10.1503/jpn.130118

94. Papadatou-Pastou M, Miskowiak KW, Williams JMG, Harmer CJ, Reinecke A. Acute antidepressant drug administration and autobiographical memory recall: A functional magnetic resonance imaging study. *Exp Clin Psychopharmacol*. 2012;20(5):364. doi:10.1037/a0027969

95. Pringle A, McTavish SF, Williams C, Smith R, Cowen PJ, Harmer CJ. Short-term NK1 receptor antagonism and emotional processing in healthy volunteers. *Psychopharmacology (Berl)*. 2011;215(2):239-246. doi:10.1007/s00213-010-2133-z

96. Reed JL, Nugent AC, Furey ML, Szczepanik JE, Evans JW, Zarate CA. Ketamine normalizes brain activity during emotionally valenced attentional processing in depression. *NeuroImage Clin*. 2018;20:92-101. doi:10.1016/j.nicl.2018.07.006

97. Robertson B, Wang L, Diaz MT, et al. Effect of bupropion extended release on negative emotion processing in major depressive disorder: A pilot functional magnetic resonance imaging study. *J Clin Psychiatry*. 2007;68(2):261-267. doi:10.4088/JCP.v68n0212

98. Scheidegger M, Henning A, Walter M, et al. Ketamine administration reduces amygdalo-hippocampal reactivity to emotional stimulation. *Hum Brain Mapp*. Published online 2016. doi:10.1002/hbm.23148

99. Shiroma PR, Thuras P, Johns B, Lim KO. Emotion recognition processing as early predictor of response to 8-week citalopram treatment in late-life depression. *Int J Geriatr Psychiatry*. 2014;29(11):1132-1139. doi:10.1002/gps.4104

100. Simmons AN, Arce E, Lovero KL, Stein MB, Paulus MP. Subchronic SSRI administration reduces insula response during affective anticipation in healthy volunteers. *Int J Neuropsychopharmacol*. 2009;12(8):1009-1020. doi:10.1017/S1461145709990149

101. Stein A, Murphy S, Arteche A, et al. Effects of reboxetine and citalopram on appraisal of infant facial expressions and attentional biases. *J Psychopharmacol*. 2012;26(5):670-676. doi:10.1177/0269881111421970

102. Sterpenich V, Vidal S, Hofmeister J, et al. Increased Reactivity of the Mesolimbic Reward System after Ketamine Injection in Patients with Treatment-resistant Major Depressive Disorder. *Anesthesiology*. 2019;130(6):923-935. doi:10.1097/ALN.0000000000002667

103. Tendolkar I, Van Wingen G, Urner M, Jan Verkes R, Fernández G. Short-term duloxetine administration affects neural correlates of mood-congruent memory. *Neuropsychopharmacology*. 2011;36(11):2266-2275. doi:10.1038/npp.2011.114

104. Tranter R, Bell D, Gutting P, Harmer C, Healy D, Anderson IM. The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *J Affect Disord*. 2009;118(1-3):87-93.

doi:10.1016/j.jad.2009.01.028

105. van Marle HJF, Tendolkar I, Urner M, Verkes RJ, Fernández G, van Wingen G. Subchronic duloxetine administration alters the extended amygdala circuitry in healthy individuals. *Neuroimage*. 2011;55(2):825-831. doi:10.1016/j.neuroimage.2010.12.051

106. Walsh AEL, Browning M, Drevets WC, Furey M, Harmer CJ. Dissociable temporal effects of bupropion on behavioural measures of emotional and reward processing in depression. *Philos Trans R Soc B Biol Sci*. 2018;373(1742):20170030. doi:10.1098/rstb.2017.0030

107. Walsh AEL, Huneke NTM, Brown R, Browning M, Cowen P, Harmer CJ. A dissociation of the acute effects of bupropion on positive emotional processing and reward processing in healthy volunteers. *Front Psychiatry*. 2018;9(482). doi:10.3389/fpsyt.2018.00482

108. Wang Y, Xu C, Cao X, et al. Effects of an antidepressant on neural correlates of emotional processing in patients with major depression. *Neurosci Lett*. 2012;527(1):55-59. doi:10.1016/j.neulet.2012.08.034

109. Warren MB, Cowen PJ, Harmer CJ. Subchronic treatment with St John's wort produces a positive shift in emotional processing in healthy volunteers. *J Psychopharmacol*. 2019;33(2):194-201. doi:10.1177/0269881118812101

110. Wells TT, Clerkin EM, Ellis AJ, Beevers CG. Effect of antidepressant medication use on emotional information processing in major depression. *Am J Psychiatry*. 2014;171(2):195-200. doi:10.1176/appi.ajp.2013.12091243

111. Windischberger C, Lanzenberger R, Holik A, et al. Area-specific modulation of neural activation comparing escitalopram and citalopram revealed by pharmaco-fMRI: A randomized cross-over study. *Neuroimage*. 2010;49(2):1161-1170. doi:10.1016/j.neuroimage.2009.10.013

112. Button KS, Ioannidis JPA, Mokrysz C, et al. Power failure: Why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14(5):365-376. doi:10.1038/nrn3475

113. Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant.

Psychol Sci. 2011;22(11):1359-1366. doi:10.1177/0956797611417632

114. Harmer CJ, Cowen PJ, Goodwin GM. Efficacy markers in depression. *J Psychopharmacol*. 2011;25(9):1148-1158. doi:10.1177/0269881110367722

This declaration concerns the article entitled:			
The effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Hobbs, C., Murphy, S. E., Wright, L., Carson, J., Van Assche, I., O'Brien, J., ... & Button, K. S. (2020). Effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers. *BJPsych open*, *6*(6).		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to the formulation of ideas (70%), design of methodology (70%), and experimental work including creation of study materials (50%), participant recruitment (70%), data collection (70%), data cleaning, analysis, visualisation, and archiving (100%), interpretation of results (70%). The candidate predominantly executed the presentation of the data in journal format (90%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	05/01/2022

Chapter 4: The effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers

Catherine Hobbs[a], Susannah E Murphy[b c], Lucy Wright[b c], James Carson[b c], Indra Van Assche[b c], Jessica O'Brien[b c], Mayowa Oyesanya[b c], Jie Sui[d] Marcus R Munafò[e f g], David Kessler[h], Catherine J Harmer[b c], Katherine S Button[a]

[a] Department of Psychology, University of Bath

[b] Department of Psychiatry, University of Oxford

[c] NHS Foundation Trust, Warneford Hospital, Oxford

[d] School of Psychology, University of Aberdeen

[e] School of Psychological Science, University of Bristol

[f] MRC Integrative Epidemiology Unit at the University of Bristol

[g] National Institute of Health Research Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol

[h] Population Health Sciences, University of Bristol

Chapter Rationale

From systematic review and meta-analysis reported in chapter 3, I found that there is not currently reliable evidence that antidepressants alter sensitivity to positive or negative emotional information, at least amongst cognitive tasks currently used. In chapters 1 and 2 I found evidence that self-referential affective biases were associated with depression in a dose-response manner. Self-referential affective processing may therefore be a sensitive indicator of early antidepressant action. In this chapter I therefore used the cognitive tasks employed in chapter 1 to understand the acute effects of antidepressants on self-referential affective processing. Additionally, I adapted a widely used cognitive task of emotional processing identified in chapter 3 to understand the importance of referential focus on change in affective processing following antidepressant administration. This was a preliminary study to examine the potential mechanistic effects of antidepressants on self-referential affective processing. I therefore conducted a double-blind randomised controlled trial examining differences in task performance between healthy volunteers administered an acute dose of

citalopram versus placebo. Whilst this study design lacks ecological validity, it allowed me to examine the very early mechanistic effects of antidepressants on self-referential processing under controlled conditions.

Open Science Statement

This study was pre-registered on Open Science Framework, where study materials are also openly available (https://osf.io/nhjvs/). The data that support the findings of this study are openly available in the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-00891.

Abstract

Background: Depression is characterised by negative views of the self. Antidepressant treatment may remediate negative self-schema through increasing processing of positive information about the self. Changes in affective processing during social interactions may increase expression of prosocial behaviours, improving interpersonal communications.

Aims: To examine whether acute administration of citalopram is associated with an increase in positive affective learning biases about the self and prosocial behaviour.

Method: Healthy volunteers (n = 41) were randomised to either an acute 20 mg dose of citalopram or matched placebo in a between-subjects double-blind design. Participants completed computer-based cognitive tasks designed to measure referential affective processing, social cognition, and expression of prosocial behaviours.

Results: Participants administered citalopram made more cooperative choices than those administered placebo in a prisoner's dilemma task (β = 20%, 95% CI: 2%, 37%). Exploratory analyses indicated that participants administered citalopram showed a positive bias when learning social evaluations about a friend (β = 4.06, 95% CI: 0.88, 7.24), but not about the self or a stranger. Similarly, exploratory analyses found evidence of increased recall of positive and reduced recall of negative words about others (β = 2.41, 95% CI: 0.89, 3.93), but not the self, in the citalopram group.

Conclusions: Participants administered citalopram showed greater prosocial behaviours, increased positive recall, and increased positive learning of social evaluations towards others. The increase in positive affective bias and prosocial behaviours towards others may, at least partially, be a possible mechanism of antidepressant effect. However, we found no evidence that citalopram influenced self-referential processing.

Introduction

Addressing negative perceptions of the self is believed to be central to the successful treatment of depression. According to cognitive theories, individuals with depression hold negative views and expectations about the self, developed as an internalised response to repeated negative social experiences (e.g., peer victimisation). When activated by external stressors these negative self-schema dominate information processing, increasing automatic processing of negative information about the self. Deliberative cognitive processing is attenuated, preventing reappraisal of these automatic biases (1). Supportive of this theory, peer victimisation in childhood is associated with increased negative and reduced positive perceptions of the self (2). Furthermore, negative affective biases are more likely to be observed in depression if stimuli are encoded in reference to the self (2–5). Change in self-referential affective processing may therefore be a key mechanism of treatments for depression.

The role of antidepressants in addressing negative self-schema

Depression is commonly treated using antidepressant medication, predominantly selective-serotonin reuptake inhibitors (SSRIs) (6). SSRIs operate by blocking the reuptake of 5-HT (5-hydroxytryptpamine). It is believed that the subsequent increase in serotonin may remediate negative affective biases early in treatment (7). Supportive of this, short-term administration of antidepressants is associated with increased processing of positive stimuli in both depressed and healthy volunteer samples (8). Negative self-schemas may be addressed through these changes in automatic affective processing (9). Individuals with depression preferentially process negative information about the self (10). Shifting affective processing through antidepressant treatment may expose individuals to more positive information about the self, remediating negative schema. In keeping with this argument, antidepressants have been found to increase recall of positive characteristics encoded to the self in healthy and depressed individuals (11,12).

Importance of social cognition

In order for changes in affective processing from antidepressant treatment to alter mood, it has been suggested that individuals must engage with their social environment to relearn associations in a more positive context (13). However, this aspect of the model is yet to be fully explored. Greater depression severity is associated with poorer quality social interactions (14), and increased expectations of rejection (15). Raised expectations of rejection may evoke hostile or non-responsive social behaviours, increasing the likelihood of

reciprocal negative behaviours from others, and reinforcing negative expectations of social interactions (16,17). In keeping with previous evidence of increases in positive affective biases and behaviours following antidepressant administration (11,12,18), antidepressant treatment is likely to strengthen positive learning and prosocial behaviours during social interactions. Repeated social interactions with remediated positive affective biases may therefore reinforce engagement in future social interactions, potentially addressing issues of social withdrawal characteristic of depression.

Altering self-schema through changes in social cognition

The self is a social construct, shaped by our perceptions of others' evaluations towards us (19). During social interactions, healthy individuals preferentially incorporate positive evaluations into their self-concept (20). In contrast, individuals with greater depression expect more negative evaluations, (21), selectively engage in negative feedback (22), and show reduced learning of positive evaluations (23), about the self. Preferential engagement with negative social evaluations about the self may reinforce negative self-schema and increase social withdrawal in a vicious cycle. Increasing positive affective biases through antidepressant treatment may increase learning of positive social evaluations from others, altering the affective content of self-schema, and breaking the pattern of maladaptive learning in social interactions. Changes in self-referential affective learning within social contexts may be an important pathway in antidepressant action.

Aims

In this study we examined the influence of an acute administration of citalopram on affective self-referential cognition and social behaviours in healthy volunteers using a double-blind placebo-controlled design. We hypothesised that acute administration of citalopram would be associated with an increase in positive affective biases about the self and increased prosocial behaviour.

Method

This study was pre-registered on Open Science Framework (https://osf.io/nhjvs/). Study materials are also available on the Open Science Framework (https://osf.io/nhjvs/). The data that support the findings of this study are openly available in the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-00891.

Subjects

Participants aged 18 to 45, fluent in English, with normal or corrected-to-normal vision, were recruited through advertisement to the local community. We excluded participants meeting diagnostic criteria for past or current Axis 1 DSM-V Psychiatric Disorder identified using the Structured Clinical Interview for DSM-V Axis I disorders (SCID-V; 24). Other exclusion criteria were current use of psychoactive medication (excluding contraceptive medication), current or past drug or alcohol dependency, current or past significant neurological condition, known hypersensitivity to the study drug, current pregnancy or breast feeding, current significant medical condition, consumption of more than 5 cigarettes per day or more than 6 caffeinated drinks per day, lactose intolerance, previous participation in a study using similar cognitive tasks, previous participation in a study involving medication within the last three months, or recreational psychoactive drug use within the last three months.

Ethical Approval

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. All procedures involving human subjects were approved by the University of Oxford Medical Sciences Interdivisional Research Committee (R64589). Written informed consent was obtained from all participants.

Design

This study used a between-subject, double-blind, placebo controlled design. Participants were randomised to receive a single acute oral dose of 20 mg citalopram or lactose placebo encapsulated in identical white capsules. Blocked randomisation, stratified by sex, was generated using an online randomisation tool (25).

Procedure

Participants first completed self-report questionnaires of mood and personality (baseline) and were administered the study medication. Citalopram is rapidly absorbed, with peak concentrations reached within 2-4 hours (26). Cognitive testing therefore started following a

3-hour rest period after drug administration, in order to maximise drug levels during testing. Participants repeated state measures of mood (post-drug timepoint) and completed the following cognitive tests in a fixed order; Social Evaluation Learning; Associative Learning; Prisoners' Dilemma; Go/No-Go Self-Esteem; Referential categorisation and recall. Participants then repeated the state measures of mood (post-testing timepoint), before completing the Oxford Cognition Stress Task (reported elsewhere). Testing lasted approximately 1.5-2 hours. Prior to participating, participants were asked to eat a light meal and were provided with light refreshments after the rest period. Participants were informed that the study aimed to examine how citalopram alters processing of emotional and social information about the self and others, but were blinded to the specific study hypotheses.

Materials

Questionnaires

Depression was measured using the Patient Health Questionnaire (PHQ-9) (27) and Beck Depression Inventory (BDI-II) (28). Anxiety was measured using the Generalised Anxiety Disorder Scale (GAD-7) (29), the Brief Fear of Negative Evaluation Scale (BFNE) (30) and the Trait Anxiety Inventory (31). Personality traits were measured using the Eysenck Personality Questionnaire Abbreviated (EPQR-A) (32). State mood was measured using the State Anxiety Inventory (31), Positive and Negative Affect Scale (PANAS) (33), and Visual Analogue Scales of sadness, disgust, anger, fear, anxiety, and alertness. State mood measures were completed at baseline, post-drug, and post-testing timepoints.

Blinding

To assess the effectiveness of blinding, at the end of testing participants and the administrating researcher guessed the study drug administered, and their certainty of this guess using a visual analogue scale. Side effects were also monitored using participant self-reports of nausea, dizziness, dry mouth, headaches, alertness, and agitation (absent to severe) at each timepoint.

Cognitive Tasks

For brevity a short description of each task is provided below, full details are available in the supplementary material.

Prisoners' Dilemma: Antidepressants may promote positive social relationships with others by increasing prosocial behaviours. We therefore measured cooperative behaviours using a prisoners' dilemma task. Participants won points based on their decision to cooperate or

defect in combination with the computer-simulated opponent's decision (Figure 1 a). If both chose to cooperate then the points were equally shared, if one defected and the other cooperated then the defector gained all the points, and if both defected neither player gained any points. Participants were unaware of the other player's decision when making their choice. Social context was manipulated so that the other player could initially choose to cooperate (positive) or defect (negative). The proportion of cooperative choices was recorded.

Social Evaluation Learning: It is possible that antidepressants may alter negative self-schema by increasing learning of positive evaluations about the self. We therefore measured learning of social evaluations within a reinforcement learning task (34,35). Participants learnt how much the computer 'liked' the self, a friend and a stranger based on feedback to a forced choice selection between positive and negative social evaluation pairings (Figure 1b). Participants learnt two rules based on the probability of the positive evaluations being 'correct' ('Like' 60-80%, 'Dislike' 20-40%). The number of errors made before reaching the criterion of eight consecutive rule-congruent responses were recorded. Bias scores were calculated by subtracting errors to criterion made when learning the dislike rule from the like rule.

Referential Categorisation and Recall: Previous research has indicated increases in recall of positive characteristics following antidepressant administration. However, the specificity of these effects to the self are unclear. To measure recall of characteristics about the self and others, participants first categorised whether 40 presented positive and negative traits described themselves or a familiar other (yes/no). Participants were then given two minutes to recall those personality traits, using the keyboard to enter their responses (Figure 1c). Separate blocks were completed for each referential condition in a randomised order. The total number of words categorised and correctly recalled were recorded.

Self-Esteem Go/No-Go: To measure self-referential processing occurring in interaction with emotion processing we used a go/no-go task (36). Participants categorised words relating to the self or others, and positive or negative traits, by pressing the spacebar if a presented word belonged to a specified paired referential-emotion category (Figure 1d). Discriminative accuracy (d') was calculated according to referential-emotion condition.

Associative Learning: To measure self, emotion, and reward processing occurring independently we used three simple associative learning tasks (37,38). In each task

participants were presented with a combination of stimuli-shape pairings, relating to the relevant area of processing, and used the keyboard to indicate whether the presented pairings matched with previously learnt associations (Figure 1e). Stimuli varied according to the area of processing examined. Accuracy (% correct) and reaction times were recorded.

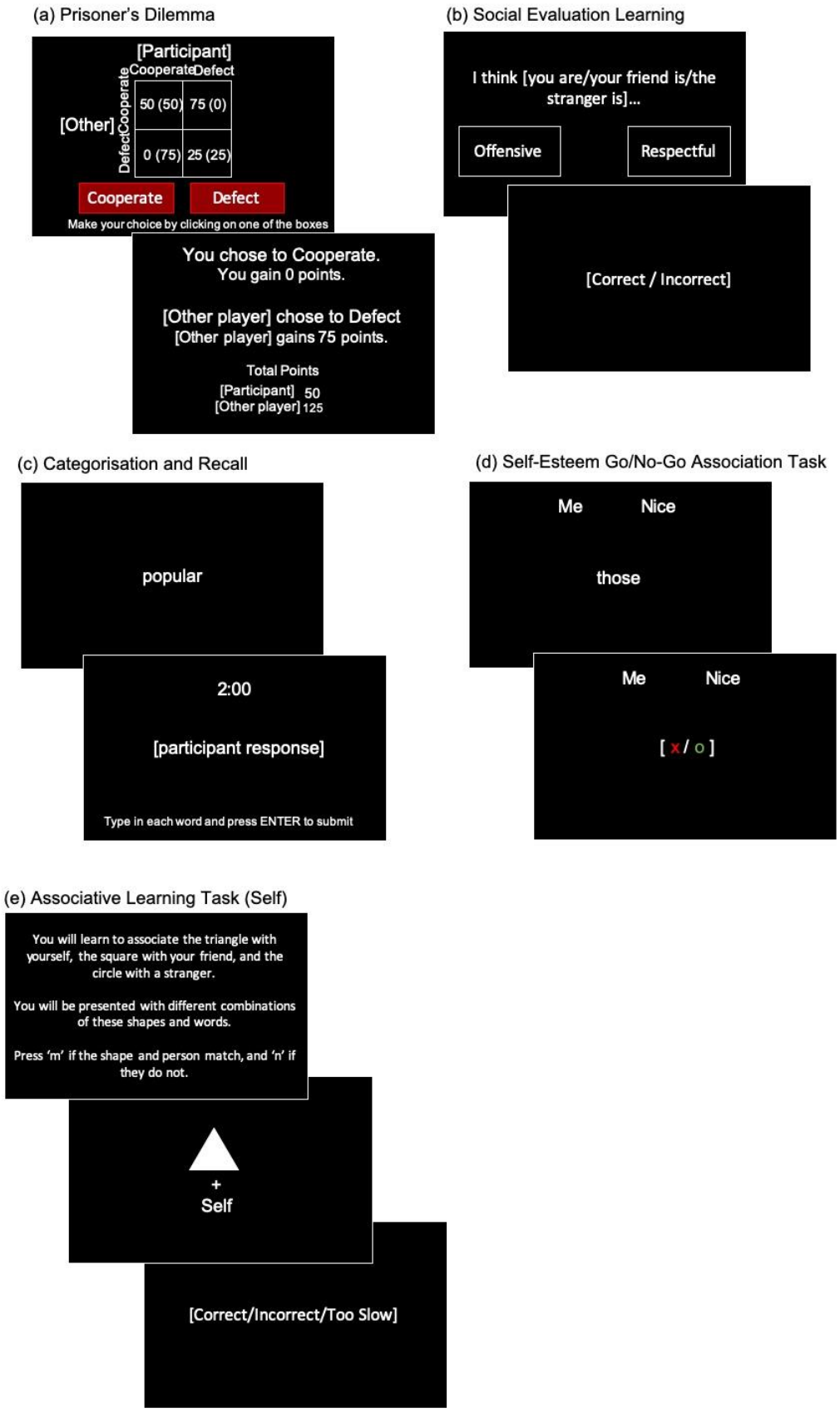


Figure 1

Cognitive Task Procedures

Statistical Analyses

Analyses were conducted in R 3.6.

Sample size calculation

We aimed to recruit 44 participants to provide 90% power to detect previous changes in emotion processing observed in healthy volunteer studies, (drug mean 10.64 (SD 9.77), placebo mean 3.36 (SD 5.96)) (12). However, due to COVID-19, recruitment was terminated at 41 participants in March 2020. With the recruited sample, we were able to detect an effect size of d = 1.04 at 90% power and an alpha of 0.05.

State mood and side effects

The influence of citalopram on state mood and side effects were tested using mixed-effects linear regression models. Separate models were conducted for each measure, with drug group, timepoint, and the interaction between these, as predictors. Subject was entered as a random effect to account for the effect of time.

Cognitive tasks

A series of mixed-effect linear regression models were used to assess the influence of citalopram on task performance. For all models, subject was entered as a random effect to account for the repeated measures elements of tasks, drug group was entered as a predictor, and the task outcome as the outcome. For tasks including a referential and/or valence (e.g., emotion or rule) condition, these were entered into the models as additional categorical predictors, independently and in interaction with drug group. Full model details are available in supplementary materials. For exploratory analyses, p-values are not reported due to undetermined inflation of the alpha rate (39).

Drug group guess and certainty

Differences in group assignment guesses according to drug group were assessed using chi-square tests. Differences in certainty of group assignment according to drug group were evaluated using t-tests.

Results

Sample

Participants (n = 41) were randomly allocated to the citalopram (n = 20) or placebo group (n = 21). Sample characteristics, according to drug group, are presented in Table 1.

Table 1

Sample Demographic Characteristics and Baseline Trait Mood and Personality Self-report Measures

	Citalopram (N = 20)[a]	Placebo (N = 21)
Age, M (SD)	23.90 (3.24)	22.86 (3.58)
Female, N (%)	16 (80)	17 (81)
Ethnicity, N (%)		
Asian	5 (25)	6 (29)
Black	0 (0)	1 (5)
Caucasian	14 (70)	12 (57)
Mixed	1 (5)	2 (9)
Occupation, N (%)		
Employed	5 (25)	4 (19)
Student	15 (75)	17 (81)
Educational Attainment, N (%)		
6th Form College	5 (25)	9 (43)
Undergraduate	7 (33)	7 (33)
Postgraduate	8 (38)	5 (24)
English Spoken as First Language, N (%)	10 (50)	19 (90)
PHQ-9, M (SD)	1.05 (1.43)	1.33 (1.93)
BDI-II, M (SD)	1.58 (2.09)	2.24 (3.65)
GAD-7, M (SD)	0.47 (0.70)	0.71 (1.38)
BFNE, M (SD)	26.89 (7.42)	29.57 (6.61)
STAI-T, M (SD)	31.16 (6.32)	32.0 (7.30)
EPQR-A, M (SD)		
Extraversion	9.21 (2.42)	9.43 (3.06)
Neuroticism	2.63 (2.50)	3.00 (2.53)
Lie	4.42 (1.89)	4.81 (2.54)
Psychoticism	2.37 (1.83)	2.71 (1.93)

[a] Data was missing for one participant for Mood and Personality measures in the Citalopram group (N = 19) due to a technical error

PHQ-9 = Patient Health Questionnaire; BDI-II = Beck Depression Inventory II; GAD-7 = Generalised Anxiety Disorder Questionnaire; BFNE = Brief Fear of Negative Evaluation Scale; STAI-T = Trait Anxiety Inventory; EPQR-A = Eysenck Personality Questionnaire Abbreviated

Self-Reported State Mood

We found no evidence that citalopram altered mood. Participants showed some evidence of a decrease in positive mood (PANAS positive) between baseline and post-drug (β = -1.90, 95% CI: -3.77, -0.02, p = 0.051), and post-testing (β = -2.47, 95% CI: -4.35, -0.60, p = 0.012), but this did not differ between groups (post-drug: β = -0.25, 95% CI: -2.84, 2.34, p = 0.852, post-testing: β = -1.00, 95% CI: -3.59, 1.59, p = 0.450). Participants showed a slight increase in state anxiety (STAI-S) between baseline and post-testing (β = 2.31, 95% CI: 0.00, 4.63, p = 0.054), but this did not differ by drug group (β = 0.64, 95% CI: -2.56, 3.83, p = 0.697). No significant changes over time or differences between groups were observed for the PANAS negative subscale or ratings of disgust, anger, fear, anxiety, or alertness.

Differences were observed between drug groups at baseline in VAS ratings of happiness and sadness, with the placebo group showing higher levels of sadness (β = 8.86, 95% CI: 1.50, 16.22, p = 0.021) and lower levels of happiness (β = -8.22, 95% CI: -15.93, -0.51, p = 0.041). However, exploratory follow-up pairwise comparisons of drug group according to timepoint, found no group differences at future time-points (Supplementary Table 1).

Cognitive Tasks

Prisoner's Dilemma

An effect of drug was observed on the proportion of cooperative choices; participants in the placebo group made 20% fewer cooperative choices compared to the citalopram group (95% CI: -37%, -2%, p = 0.030). There was some suggestion of a main effect of social context in the expected direction, although the confidence intervals included the null; on average, participants made 13% fewer cooperative choices when the other player had initially defected versus cooperated (95% CI: -26%, 1%, p = 0.073). We did not find evidence of an interaction between drug group and social context, indicating that participants in the placebo group made fewer cooperative choices irrespective of whether the other player initially cooperated or defected (β = 5%, 95% CI: -14%, 24%, p = 0.580).

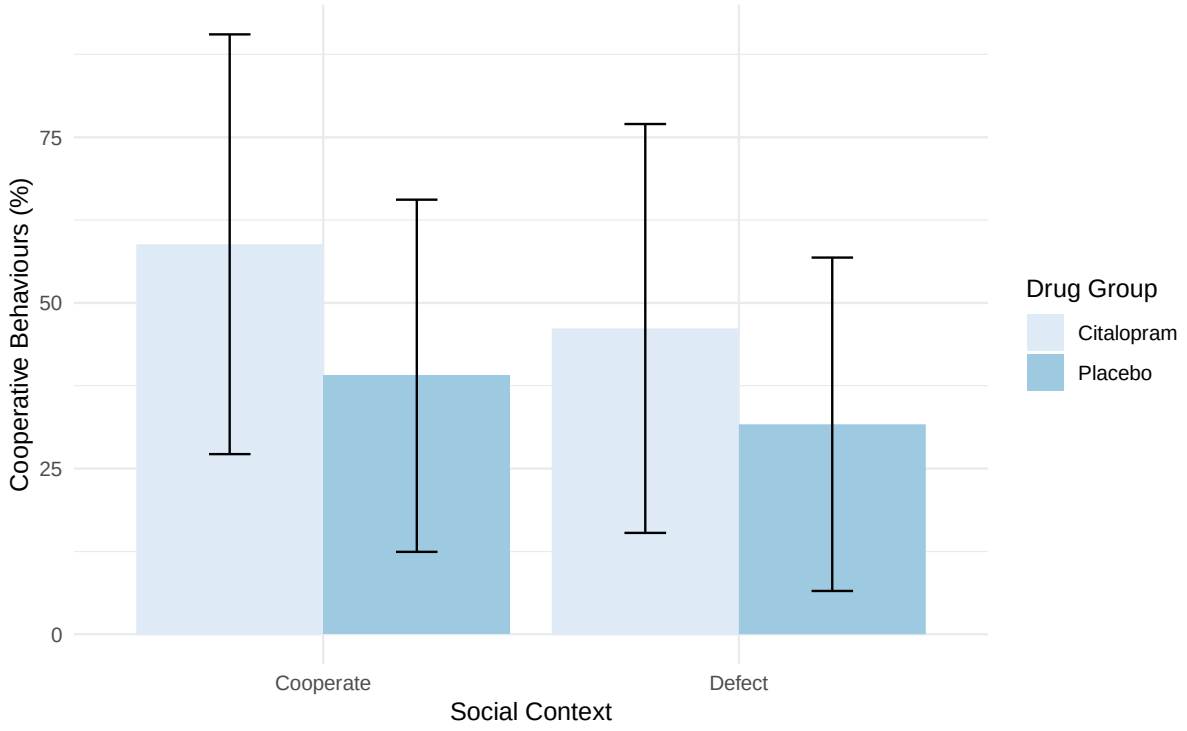


Figure 2

Mean proportion of cooperative behaviours according to drug group and social context. Error bars represent standard deviations.

Social Evaluation Learning

There were some indications that participants in the citalopram group showed a greater positive bias overall as indicated by bias scores. Participants in the citalopram group, on average made 4.8 (SD 6.64) more errors learning negative relative to positive evaluations. In comparison, participants in the placebo group made 2.25 (SD 5.37) more errors on average learning negative relative to positive evaluations. This group effect was particularly heightened for the friend condition (Citalopram: -5.70, SD 5.45, Placebo: -1.64, SD 4.59; Supplementary Figure 1). However, in our initial confirmatory model examining main and interaction effects of referential condition and drug group we found no evidence of group differences (Supplementary Table 2).

To examine whether effects of bias scores were obscured by learning within a particular rule (e.g., better learning of 'dislike' or worse learning of 'like') we examined the effect of drug group on errors to criterion according to referential condition and rule. Participants made a greater number of errors before learning the negative 'dislike' rule ($\beta = 4.53$, 95% CI: 2.21, 6.84, $p < 0.001$). However, this effect did not vary according to drug group or referential condition. No main effects of drug, referential condition, or an interaction between these, were observed (Supplementary Table 2).

Given our previous findings of increased prosocial behaviours towards others in the citalopram group in the prisoners' dilemma task, that the sample size may be small for testing interaction effects, and that the descriptive statistics suggested a drug group effect in the friend condition, we conducted additional exploratory analyses focusing on the effect of drug group in the friend condition. We therefore examined the effect of drug group on bias scores separately for each referential condition. Participants in the citalopram group showed a greater positive bias in the friend condition ($\beta = 4.06$, 95% CI: 0.88, 7.24), but there was little evidence of a difference between drug groups in the self ($\beta = 1.95$, -2.25, 6.16) or stranger conditions ($\beta = 1.63$, -2.46, 5.72).

To further explore the effect of drug group in the friend condition, we conducted additional exploratory analyses examining the effect of rule and drug group on errors to criterion for each referential condition separately. For the friend condition, an interaction between drug group and rule was observed ($\beta = -4.06$, 95% CI: -7.14, -0.98). Participants in the citalopram group on average made 3.55 (SD 2.97) errors before learning the 'like' rule and 9.25 (SD 4.42) errors before learning the 'dislike' rule. A smaller difference was observed between

rules for the placebo group, with participants on average making 5.19 (SD 3.16) errors before learning the ‘like’ rule and 6.38 (SD 3.58) errors before learning the ‘dislike’ rule (Figure 3). The greater positive bias towards the friend, observed with bias scores, was therefore driven by both reduced learning of the friend being disliked and increased learning of the friend being liked. We found no evidence for an interaction between drug group and rule when learning about the self (β = -1.95, 95% CI: -5.73, 1.83), or the stranger (β = -1.63, 95% CI: -5.16, 1.91).

However, these results are exploratory and require further replication to determine the reliability of effects.

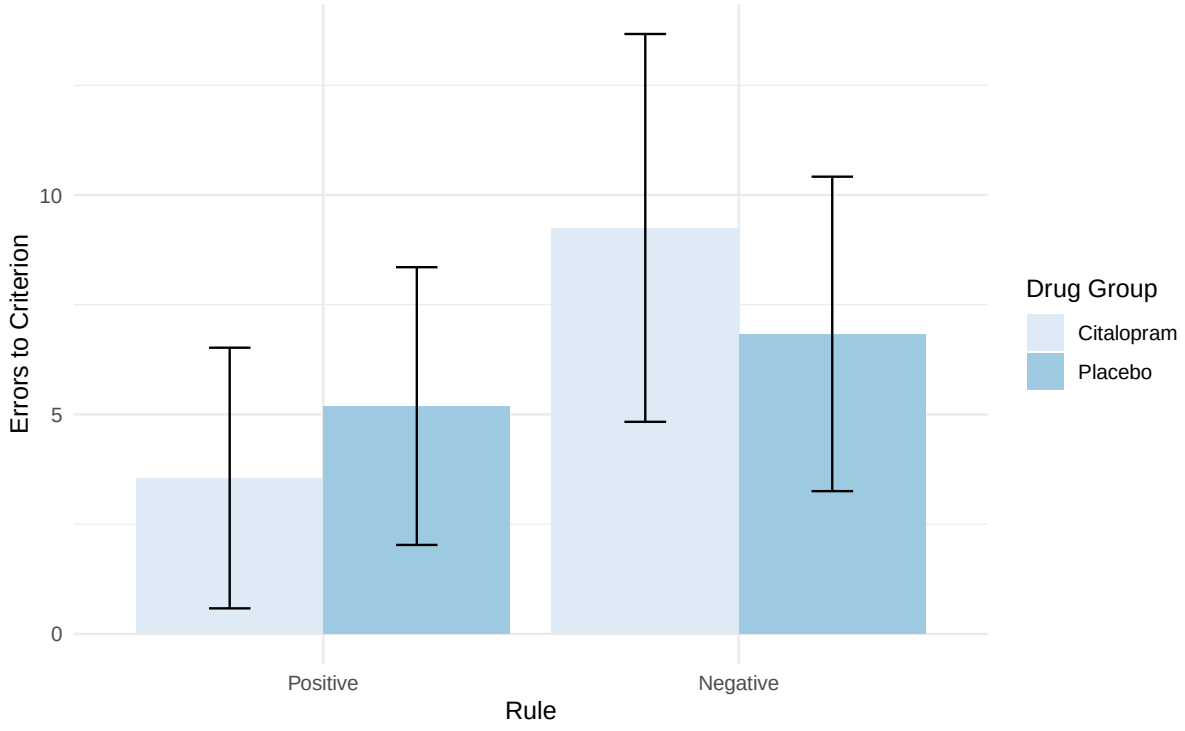


Figure 3

Mean errors to criterion in the friend condition according to drug group and rule. Error bars represent standard deviations.

Referential Emotional Categorisation and Recall

Participants endorsed a greater number of positive words as descriptive (β = -11.70, 95% CI: -13.50, -9.90, p <0.001), for both the self and other. There was no evidence that this effect differed between drug groups (Supplementary Table 3).

When examining the number of correctly recalled words, we found that participants recalled more positive than negative words (β = -1.30, 95% CI: -2.37, -0.23, p = 0.019). There was weak evidence that this differed according to referential condition, with participants recalling fewer dislikeable words in the other versus self condition, although confidence intervals included the null (β = -1.30, 95% CI: -2.82, 0.22, p = 0.096). We found no evidence of an interaction between valence and drug group, or referential condition and drug group (Supplementary Table 3).

There was weak evidence of an interaction between drug group, referential condition, and valence (β = 1.92, 95% CI: -0.20, 4.04, p = 0.079), although confidence intervals included the null. To explore this effect further we conducted additional exploratory analyses examining the interaction between drug group and valence for each referential condition separately. When recalling words about the self, participants recalled fewer dislikeable words (β = -1.30, 95% CI: -2.34, -0.27) but this did not vary by drug group (β = 0.49, 95% CI: -0.96, 1.94). Conversely, when recalling words about others, we observed an interaction between drug group and valence (β = 2.41, 95% CI: 0.89, 3.93). Participants in the citalopram group showed a positive bias towards others, recalling more likeable words and fewer dislikeable words, compared to the placebo group (Figure 4). However, these findings are exploratory and require further replication.

The citalopram group had a higher proportion of participants that did not speak English as a first language (Table 1), which may have impacted recall. Sensitivity analyses indicated that effects were consistent when this was taken into account. English as a first language was not associated with recall (β = -1.13, 95% CI: -2.48, 0.23, p = 0.112).

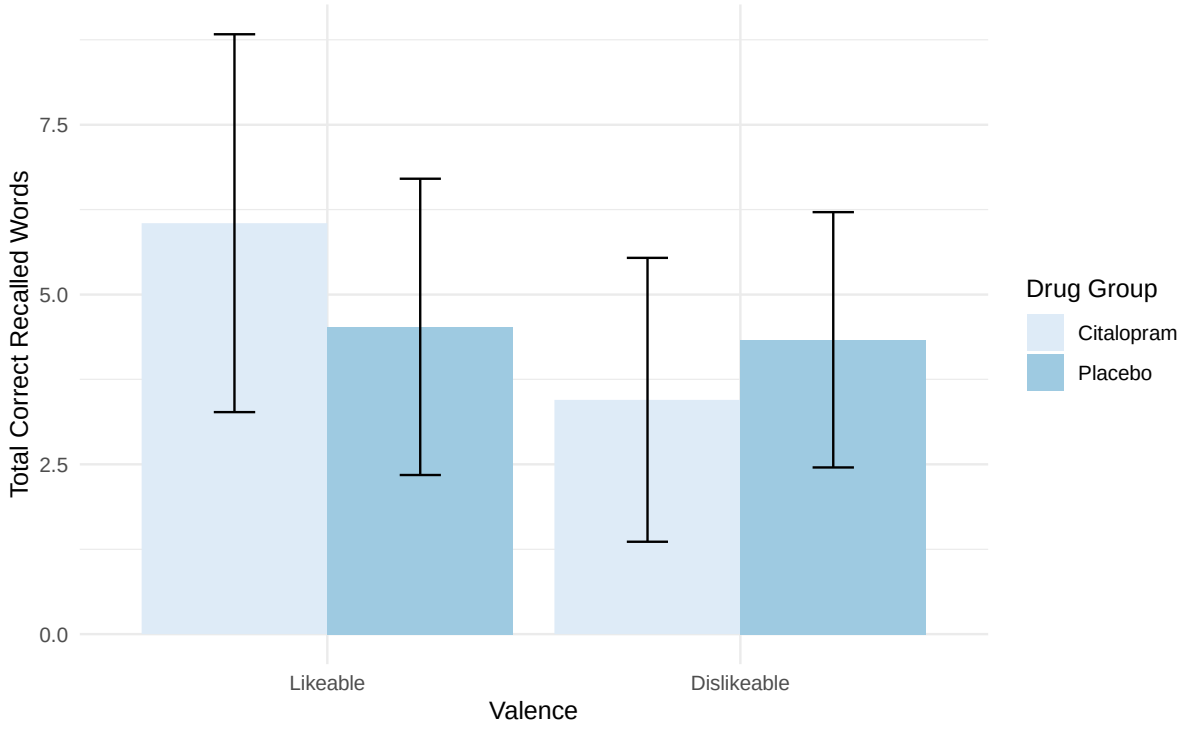


Figure 4

Mean number of correctly recalled words in the other condition according to drug group and valence. Error bars represent standard deviations.

Go/No-Go Self-Esteem

Ten participants were excluded from analyses for the Go/No-Go Association Self-Esteem task as their responses indicated non-compliance according to a priori data exclusion criteria

Participants showed reduced discriminative accuracy for words relating to others versus self (β = -0.47, 95% CI: -0.84, -0.11, p = 0.012), and for negative versus positive words (although confidence intervals included the null; β = -0.36, 95% CI: -0.73, 0.00, p = 0.052). An interaction between referential condition and valence was observed (β = 0.65, 95% CI: 0.14, 1.16, p = 0.015). Participants showed higher discriminative accuracy for positive versus negative words in the Self condition, but the opposite pattern in the Other condition, suggesting a positive self-bias. However, there was no evidence that these effects varied by drug group (Supplementary Table 4).

Additional analyses suggested that these effects were driven by hits rather than false alarms (Supplementary Table 4).

When participants excluded according to a priori data exclusion criteria were included in sensitivity analyses, the main and interactive effects of referential condition and valence were further strengthened, but the lack of an effect of drug group remained.

Associative learning

A main effect of stimuli was observed for the self and emotion tasks, with participants showing greater accuracy and faster reaction times when matching shapes with the self and a happy face (Supplementary Table 5). For the reward task, there was no evidence of differences in accuracy according to level of reward, although participants were slightly slower to match shapes with the medium (£3) than high (£9) level of reward. There was no evidence that performance on any of the associative learning tasks varied by drug group, either as a main effect or in interaction with stimuli (Supplementary Table 5).

Blinding

A greater proportion of participants in the citalopram group (74%) believed that they had taken citalopram compared to the placebo group (14%), at a greater level of certainty ($t(30.09)=3.47$, $p = 0.002$). Conversely, researchers administering the drug did not significantly differ in their group guesses and certainty according to drug group (Supplementary Table 6).

Failure of participant blinding was possibly attributable to the side effects of citalopram, with the citalopram group only showing increases in nausea ($p < 0.001$) and dizziness ($p = 0.012$) over time. An increase in agitation ($p = 0.011$), and some evidence of an increase in headaches ($p = 0.078$), was also observed over time, but this varied little by group ($p = 0.100$ and $p = 0.474$ respectively). There was no evidence of changes for dry mouth or alertness in either group (Supplementary Table 7).

Discussion

Our results tentatively support the theory that antidepressants increase positive affective biases and prosocial behaviours in healthy volunteers, in the absence of change in mood. Participants administered citalopram cooperated more in a prisoners' dilemma game. There was also some evidence, based on exploratory analyses, that participants administered citalopram showed a greater positive bias when recalling words about others, and when learning social evaluations about a friend.

In support of our pre-registered hypothesis, participants administered citalopram cooperated on a greater proportion of trials in a prisoners' dilemma task. Our findings add to literature indicating that serotonin is instrumental in modulating social behaviour. Previous research has indicated that temporary reduction of serotonin, through tryptophan depletion, reduces cooperative behaviours in a prisoners' dilemma game (40). In line with our findings, increasing serotonin through administration of citalopram over a two-week period has previously been found to increase cooperative communication and behaviour towards others (18). One week administration of an SSRI has also been found to reduce self-reported hostile behaviours (41). Increases in prosocial behaviour through antidepressant treatment may increase quality of social interactions, positively reinforcing engagement in interpersonal communications and increasing social support.

However, our findings of an increase in cooperative behaviours following a *single* dose of citalopram differs from previous research. In a previous study, a single dose of citalopram was not associated with greater cooperation in healthy volunteers (42). Disagreement between these findings may be attributable to variations in the Prisoners' Dilemma task. In Tse and Bond's (42) study, participants controlled the amount of points allocated to the other player. In this study, allocation of points was dependent on both the participant and the other players' decisions. Prosocial behaviours are believed to be motivated by the aim of eliciting reciprocal altruistic behaviours from others (43). In this study participants may have therefore been more motivated to engage in cooperative actions, providing a more sensitive marker of change following serotonin modulation.

Additionally, we found some evidence that antidepressants increase prosocial affective biases towards familiar others. Participants administered citalopram showed better learning of friends being liked and reduced learning of friends being disliked by a computer persona. Affective recall was also altered. Participants administered citalopram recalled more likeable

characteristics, and less dislikeable characteristics of others. However, these were exploratory analyses and require further replication. One potential mechanism of antidepressants may be blunting perceptions of negative characteristics and increasing sensitivity to positive characteristics in others. Although this requires further exploration in clinical samples. Increasing positive perceptions of others through antidepressant treatment may increase engagement in social interactions, addressing issues of social withdrawal and anhedonia associated with depression.

In contrast to our expectations, there was no evidence to support our hypothesis that the effects of antidepressants on affective processing would be strongest for self-referential stimuli. This contrasts with previous research indicating change in positive self-referential biases following antidepressant administration (11,12). However, in these studies participants only encoded information in reference to the self. The specificity of this effect to self-related information is therefore unclear. In a study including both a self and friend condition, subjects administered escitalopram endorsed fewer negative characteristics about themselves, but also more positive characteristics about others (44). In this study, we found no evidence to suggest that citalopram selectively affected positive learning about the self. If anything, our exploratory analyses indicated that citalopram produced the largest group effects in the friend condition with increased positive learning of social evaluations of friends following antidepressant treatment. However, our confirmatory test did not support a referential condition by drug group interaction.

Self-schemas are pervasive, dominate information processing and are resistant to disconfirmatory evidence (1). A single dose of an antidepressant may not be sufficient to address entrenched self-referential negative biases. Affective processing of information related to others may be more flexible, and therefore more sensitive to change from acute administration of antidepressants. In support of this theory, pharmacological induction of anxiety was found to influence other-referential processing, whilst self-referential processing was preserved (45). We may therefore only see changes in self-referential affective biases from longer periods of antidepressant treatment. Alternatively, addressing negative self-schema may also require remediation of top-down, deliberative biases through treatments such as cognitive behavioural therapy (CBT) (9). Further longitudinal studies are required examining changes in self-referential affective biases during long-term use of antidepressants, to assess their effectiveness in remediating negative self-schema.

We found no evidence that acute citalopram influenced inhibitory control, measured using an affective go/no-go association task, or simple associative learning of emotional, self and reward stimuli. Previous research has reported inconsistent findings regarding the relationship between depressive symptoms and performance on affective inhibitory control tasks (46,47). Similarly, we have previously found no association between simple associative learning and depression severity (48). Serotonin therefore appears to have little influence over these particular cognitive processes.

Clinical Implications

Variation in individual treatment response and delays in therapeutic action currently expose patients to considerable periods of potentially ineffective antidepressant treatment (49). Our findings suggest that change in prosocial behaviours, and more tentatively positive affective biases towards others, may be a sensitive marker of early changes in antidepressant treatment. Extension of our findings to clinical samples may offer a promising marker of treatment response that could allow clinicians to identify effective treatments for individuals experiencing depression at earlier timepoints.

Limited change in negative self-schema from antidepressant treatment, as we observed in this study, may provide an explanation for the relatively high relapse rates following treatment discontinuation (50). At present, it is unclear whether change in affective processing is sustained after antidepressant treatment is discontinued. If increased positive affective biases do not translate to a sustained change in self-schema, depressive symptoms are likely to return following antidepressant discontinuation. Sustained changes in self-schema through treatments such as CBT may explain the substantially lower relapse rates compared to antidepressant treatment (51).

Limitations

This study had lower statistical power than planned as recruitment was terminated prematurely due to COVID-19. We were therefore powered to detect large effects and may have been underpowered to detect smaller effects, such as the hypothesised interaction effects for social evaluation learning. To address this limitation, we have made all materials and data publicly available for further replication and extension of our work.

Additionally, despite strict blinding procedures there was evidence that blinding was not effective at a participant level. It is possible that our results may be partially driven by response biases. Although mitigating this possibility, we predominantly used implicit

affective tasks, where the main purpose of the tasks was not explicitly stated, and participants were blinded to the study hypotheses.

This study used a placebo drug as the comparator condition. However, high placebo response rates have been observed in antidepressant trials (52), suggesting placebo expectations may influence similar cognitive processes. Future studies including a no-treatment-control arm condition, or a placebo lead-in phase, would allow more precise evaluation of the effects of citalopram on social cognition.

Finally, this study examined changes in social cognition and behaviour following antidepressant administration in healthy volunteers. The use of healthy volunteers offers the opportunity to study affective biases unconfounded by 'cold' cognitive biases (53). However, it limits our insight into the mechanisms of antidepressants in improving mood. It is also possible that our findings may not directly translate to clinically depressed patients. Further research would benefit from extending our findings to longer term changes in affective processing in depressed individuals.

Summary

Overall, our findings suggest that acute administration of citalopram in healthy volunteers is associated with increased prosocial behaviour towards others. In contrast to our expectations, we found little evidence that the effect of acute citalopram on affective processing was heightened for information related to the self. Exploratory analyses instead suggested increased positive affective biases towards others. Changes in affective processing and prosocial behaviours towards others may, at least partially, be a possible mechanism of antidepressant effect. Further research in clinical samples is required to examine this possibility.

References

1. Beck AT. The evolution of the cognitive model of depression and its neurobiological correlates. Am J Psychiatry. 2008;165(8):969–77.

2. Cole DA, Dukewich TL, Roeder K, Sinclair KR, McMillan J, Will E, et al. Linking peer victimization to the development of depressive self-schemas in children and adolescents. J Abnorm Child Psychol. 2014;42(1):149–60.

3. Gaddy MA, Ingram RE. A meta-analytic review of mood-congruent implicit memory in depressed mood. Clin Psychol Rev. 2014;34(5):402–16.

4. Ji JL, Grafton B, MacLeod C. Referential focus moderates depression-linked attentional avoidance of positive information. Behav Res Ther. 2017;93:47–54.

5. Hertel PT, El-Messidi L. Am I Blue? Depressed Mood and the Consequences of Self-Focus for the Interpretation and Recall of Ambiguous Words. Behav Ther. 2006;37(3):259–68.

6. Kendrick T, Dowrick C, McBride A, Howe A, Clarke P, Maisey S, et al. Management of depression in UK general practice in relation to scores on depression severity questionnaires: analysis of medical record data. BMJ. 2009 Mar;338:b750.

7. Harmer CJ, Goodwin GM, Cowen PJ. Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. Br J Psychiatry. 2009;195(2):102–8.

8. Walsh AEL, Harmer CJ. The cognitive neuropsychological model of antidepressant response. Curr Opin Psychol. 2015;4:124–30.

9. Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. Neuropsychopharmacology. 2012;37(1):117–36.

10. Phillips WJ, Hine DW, Thorsteinsson EB. Implicit cognition and depression: A meta-analysis. Clin Psychol Rev. 2010;30(6):691–709.

11. Harmer CJ, O'Sullivan U, Favaron E, Massey-Chase R, Ayres R, Reinecke A, et al. Effect of acute antidepressant administration on negative affective bias in depressed patients. Am J Psychiatry. 2009;166(10):1178–84.

12. Harmer CJ, Shelley NC, Cowen PJ, Goodwin GM. Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. Am J Psychiatry. 2004;161(7):1256–63.

13. Godlewska BR, Harmer CJ. Cognitive neuropsychological theory of antidepressant action: a modern-day approach to depression and its treatment. Psychopharmacology (Berl). 2020;1–14.

14. Teo AR, Choi HJ, Valenstein M. Social Relationships and Depression: Ten-Year Follow-Up from a Nationally Representative Study. PLoS One. 2013;8(4):e62396.

15. Caouette JD, Guyer AE. Cognitive distortions mediate depression and affective response to social acceptance and rejection. J Affect Disord. 2016;190:792–9.

16. Lewinsohn PM. The behavioral study and treatment of depression. In: Progress in behavior modification. Elsevier; 1975. p. 19–64.

17. Tse WS, Bond AJ. Consequences of displaying abnormal social behaviour: Avoidance and reduction of social reinforcement. J Affect Disord. 2003;75(1):49–58.

18. Tse WS, Bond AJ. Serotonergic intervention affects both social dominance and affiliative behaviour. Psychopharmacology (Berl). 2002;161(3):324–30.

19. Shrauger JS, Schoeneman TJ. Symbolic interactionist view of self-concept: Through the looking glass darkly. Psychol Bull. 1979;86(3):549.

20. Korn CW, Prehn K, Park SQ, Walter H, Heekeren HR. Positively biased processing of self-relevant social feedback. J Neurosci. 2012;32(47):16832–44.

21. Moritz D, Roberts JE. Self-Other Agreement and Metaperception Accuracy Across the Big Five: Examining the Roles of Depression and Self-Esteem. J Pers. 2018;86(2):296–307.

22. Giesler RB, Josephs RA, Swann WB. Self-verification in clinical depression: The desire for negative evaluation. J Abnorm Psychol. 1996;105(3):358.

23. Hobbs C, Faraway J, Kessler D, Munafò MR, Sui J, Button KS. Social evaluation learning in social anxiety disorder and depression: a mega-analysis presented at the Summer Meeting of the Bristish Association for Psychopharmacology. J Psychopharmacol. 2019;Supplement(8):A33.

24. First M, Williams J, Karg R, Spitzer R. Structured Clinical Interview for DSM-5. Arlington, VA: American Psychiatric Association; 2015.

25. Sealed Envelope Ltd. Simple randomisation service [Internet]. 2019 [cited 2020 May 21]. Available from: https://www.sealedenvelope.com/simple-randomiser/v1/

26. Kragh-Sørensen P, Overø KF, Petersen OL, Jensen K, Parnas W. The Kinetics of Citalopram: Single and Multiple Dose Studies in Man. Acta Pharmacol Toxicol (Copenh). 1981;48(1):53–60.

27. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire (PHQ-9) - Overview. J Gen Intern Med. 2001;16:606–16.

28. Beck AT, Steer RA, Brown GK. Manual for the Beck depression inventory-II. San Antonio, TX Psychol Corp. 1996;

29. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med. 2006;

30. Leary MR. A Brief Version of the Fear of Negative Evaluation Scale. Personal Soc Psychol Bull. 1983;9(3):371–5.

31. Spielberger CD. Assessment of state and trait anxiety: Conceptual and methodological issues. South Psychol. 1985;2(4):6–16.

32. Francis LJ, Brown LB, Philipchalk R. The development of an abbreviated form of the revised Eysenck personality questionnaire (EPQR-A): Its use among students in England, Canada, the U.S.A. and Australia. Pers Individ Dif. 1992;13(4):443–9.

33. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. J Pers Soc Psychol. 1988;54(6):1063.

34. Button KS, Browning M, Munafò MR, Lewis G. Social inference and social anxiety: Evidence of a fear-congruent self-referential learning bias. J Behav Ther Exp Psychiatry. 2012;43(4):1082–7.

35. Button KS, Kounali D, Stapinski L, Rapee RM, Lewis G, Munafò MR. Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. PLoS One. 2015;10(4):e0119456.

36. Gregg AP, Sedikides C. Narcissistic fragility: Rethinking its links to explicit and implicit self-esteem. Self Identity. 2010;9(2):142–61.

37. Sui J, Ohrling E, Humphreys GW. Negative mood disrupts self- and reward-biases in perceptual matching. Q J Exp Psychol. 2016;69(7):1438–48.

38. Stolte M, Humphreys G, Yankouskaya A, Sui J. Dissociating biases towards the self and positive emotion. Q J Exp Psychol. 2017;70(6):1011–22.

39. Nosek BA, Lakens D. Registered reports: A method to increase the credibility of published results. Soc Psychol (Gott). 2014;45(3):137–41.

40. Wood RM, Rilling JK, Sanfey AG, Bhagwagar Z, Rogers RD. Effects of tryptophan depletion on the performance of an iterated Prisoner's Dilemma game in healthy adults. Neuropsychopharmacology. 2006;31(5):1075–84.

41. Knutson B, Wolkowitz OM, Cole SW, Chan T, Moore EA, Johnson RC, et al. Selective alteration of personality and social behavior by serotonergic intervention. Am J Psychiatry. 1998;155(3):373–9.

42. Tse WS, Bond AJ. Difference in serotonergic and noradrenergic regulation of human social behaviours. Psychopharmacology (Berl). 2002;159(2):216–21.

43. Trivers RL. The Evolution of Reciprocal Altruism. Q Rev Biol. 1971;46(1):35–57.

44. Matthews SC, Simmons AN, Strigo IA, Arce E, Stein MB, Paulus MP. Escitalopram attenuates posterior cingulate activity during self-evaluation in healthy volunteers. Psychiatry Res - Neuroimaging. 2010;182(2):81–7.

45. Button KS, Karwatowska L, Kounali D, Munafò MR, Attwood AS. Acute anxiety and social inference: An experimental manipulation with 7.5% carbon dioxide inhalation. J Psychopharmacol. 2016;30(10):1036–46.

46. Lewis G, Button KS, Pearson RM, Munafò MR, Lewis G. Inhibitory control of positive and negative information and adolescent depressive symptoms: a population-based cohort study. Manuscr Submitt Publ. 2020;

47. Kilford EJ, Foulkes L, Potter R, Collishaw S, Thapar A, Rice F. Affective bias and current, past and future adolescent depression: A familial high risk study. J Affect Disord. 2015;174:265–71.

48. Hobbs C, Sui J, Munafò MR, Kessler D, Button KS. Biased Self-Referential Processing in Depression. Manuscr Prep. 2020;

49. Warden D, Rush AJ, Trivedi MH, Fava M, Wisniewski SR. The STAR*D project results: A comprehensive review of findings. Curr Psychiatry Rep. 2007;9(6):449–59.

50. Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E, et al. Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. Lancet. 2003;361(9358):653–61.

51. Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: A review of meta-analyses. Clin Psychol Rev. 2006;26(1):17–31.

52. Timothy Walsh B, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: Variable, substantial, and growing. J Am Med Assoc. 2002;287(14):1840–7.

53. Roiser JP, Sahakian BJ. Hot and cold cognition in depression. CNS Spectr. 2013;18(3):139–49.

This declaration concerns the article entitled:			
The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [X] **In review** [ ] **Accepted** [ ] **Published** [ ]			
Publication details (reference)	Hobbs, C., Beck, M., Denham, F., Pettit, L., Faraway, J., Munafò, M.R., Sui, J., Kessler, D., & Button, K. S. (2022). The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care. *Submitted to Journal of Psychopharmacology*.		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to the formulation of ideas (60%), design of methodology (70%), and experimental work including creation of study materials (50%), participant recruitment (100%), data collection (80%), data cleaning, analysis, visualisation, and archiving (100%), interpretation of results (90%). The candidate predominantly executed the presentation of the data in journal format (90%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	05/01/2022

Chapter 5: The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care

Catherine Hobbs[a], Milly Beck[a], Faye Denham[a], Laura Pettit[a], Julian Faraway [b], Marcus R Munafò [c d e], Jie Sui [f], David Kessler [g], Katherine S Button[s]

[a] Department of Psychology, University of Bath

[b] Department of Mathematical Sciences, University of Bath

[c] School of Psychological Science, University of Bristol

[d] MRC Integrative Epidemiology Unit at the University of Bristol

[e] National Institute of Health Research Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol

[f] School of Psychology, University of Aberdeen

[g] Population Health Sciences, University of Bristol

Chapter Rationale

Most research examining the effect of antidepressants on emotional processing has been conducted using healthy volunteers and short-term administration of antidepressants (see chapter 3). In chapter 4, I examined the early mechanisms of antidepressants on self-referential affective processing using an acute dose of citalopram in healthy volunteers. In chapter 5 I build upon this, using a naturalistic design to examine change in self-referential affective learning and depression over the first eight weeks of antidepressant treatment in patients recruited from primary care. I initially used a range of cognitive tasks that I tested in previous chapters of this PhD (chapters 1, 2, and 4) and identified from systematic review (chapter 3) as having the strongest antidepressant effects. However, due to the COVID-19 pandemic I amended our data collection procedure to allow participants to complete the study remotely from March 2020 onwards. To reduce potential fatigue effects associated with remote data collection I reduced the cognitive battery to one task that provided the strongest and most reliable effects across chapters; the social evaluation learning task which measures self-referential affective learning within social contexts. In chapter 6, I therefore examine social evaluation learning as a potential mechanism of early antidepressant effects by examining the association between change in learning of social evaluations and change in depression and anxiety symptoms.

Open Science Statement

This study was pre-registered on Open Science Framework (https://osf.io/z9p8a), where study materials are also openly available (https://osf.io/8a95j/). The data that support the findings of this study are openly available in the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-01107. A preprint is available on https://psyarxiv.com/4sn9y/.

Abstract

Background: It has been proposed that antidepressants operate by increasing sensitivity to positive versus negative information. Increasing positive affective learning within social contexts may play a key role in remediating negative self-schema. We investigated the association between change in biased learning of social evaluations about the self and others, and mood during early antidepressant treatment.

Method: Prospective cohort assessing patients recruited from primary care sites in South West England at four timepoints over the first eight weeks of antidepressant treatment. At each timepoint participants completed self-report measures of depression (BDI-II, PHQ-9), anxiety (GAD-7), and a computerised task measuring learning of social evaluations about the self, a friend, and a stranger.

Results: We did not find evidence that learning about the self was associated with a reduction in PHQ-9 (*b*=0.08, 95%CI:-0.05,0.20, p=0.239) or BDI-II scores (*b*=0.10, 95%CI:-0.18,0.38, p=0.469). We found some weak evidence that increased positive learning about the friend was associated with a reduction in BDI-II scores (*b*=0.30, 95%CI:-0.02, 0.62, p=0.069). However, exploratory analyses indicated stronger evidence that increased positive learning about the self (*b*=0.18, 95% CI:0.07,0.28, p=0.002) and a friend (*b*=0.22, 95%CI:0.10,0.35, p=0.001) was associated with a reduction in anxiety.

Conclusions: Change in social evaluation learning was more reliably associated with a reduction in anxiety rather than depression. Antidepressants may treat anxiety symptoms by remediating negative affective biases towards socially threatening information directed towards the self and close others. However, our findings are based on exploratory analyses within a small sample and require further replication.

Keywords

Depression, Antidepressants, Self-Schema, Social Cognition, Emotional Processing

Introduction

Depression is the leading mental health problem worldwide (Vos et al., 2015), affecting 4.4% of the global population at a given point (World Health Organization, 2017). Most patients are treated within primary care settings using antidepressants (McManus et al., 2016). Whilst antidepressants are effective at reducing depression symptoms (Cipriani et al., 2018), half of patients do not respond to the first antidepressant prescribed (Warden et al., 2007). Additionally, there is a delay in therapeutic action. NICE guidelines recommend altering treatment if there is inadequate response only after three to four weeks (NICE, 2009). Patients are therefore currently exposed to long periods of potentially ineffective treatment.

To address these issues, research has focused on identifying the early mechanisms of antidepressant action. By understanding how antidepressants operate it may be possible to identify patients that are not likely to respond and alter treatment at an earlier timepoint. One potential mechanism that has generated considerable research is change in affective processing. Depression is associated with increased sensitivity to negative relative to positive information across cognitive domains (Dalgleish and Watts, 1990; Dalili et al., 2015; Roiser and Sahakian, 2013). According to the cognitive neuropsychological theory, antidepressants work by remediating these negative affective biases early in treatment, increasing sensitivity to positive relative to negative affective information. An improvement in mood is produced when individuals have interacted with their social environment with remediated affective biases, allowing them to relearn associations from a more positive perspective (Godlewska and Harmer, 2020; Harmer et al., 2003; Roiser et al., 2012).

A considerable number of laboratory-based studies have been conducted on the effect of antidepressants on affective processing. However, results are somewhat inconsistent. Whilst some studies have reported greater recall of positive words in participants administered an antidepressant versus placebo (Arnone et al., 2009; Harmer et al., 2009, 2011), others have reported no differences (Browning et al., 2007; Komulainen et al., 2018; Walsh et al., 2018), and others have reported a decrease (Cooper et al., 2015; Komulainen et al., 2016). Similarly, whilst some evidence has been found of increased processing of positive (Di Simplicio et al., 2014; Harmer et al., 2009; Murphy et al., 2009) and decreased processing of negative (Harmer et al., 2006, 2011; Kamboj and Curran, 2006) facial emotions following antidepressant administration, other studies have reported no association (Cremers et al., 2016; Labuschagne et al., 2010). Additionally, inclusion criteria for depression tends to be unrepresentative of patients within primary care, who experience a wider range of severity of

depression than represented in these studies. (Berardi et al., 2005). Furthermore, most research has been conducted under controlled laboratory conditions with short-term administration of antidepressants. It is therefore unclear whether current findings translate to clinical settings.

A small number of studies have been conducted in primary care, with mixed findings. In a large scale RCT there was no evidence of a difference in recall of emotional words in primary care patients administered sertraline versus placebo over 6 weeks (Ahmed et al., 2021). However, supportive of the cognitive neuropsychological theory, in a prospective cohort study of patients treated with citalopram or reboxetine, an increase in the recognition of positive facial emotions between baseline and two-week follow-up was found to correlate with a reduction in mood by six weeks (Tranter et al., 2009). Additionally, a machine learning algorithm based on change in facial emotion recognition and baseline depression severity has been found to predict patients' response to citalopram with 77% accuracy (Browning et al., 2019). However, attempts to tailor treatment based on this algorithm within primary care had mixed findings. No differences were found between patients randomised to the algorithm arm versus treatment as usual in reduction of depressive symptoms at eight weeks. However, patients in the algorithm arm showed a greater reduction in anxiety (Browning et al., 2021). At present evidence for change in affective processing associated with antidepressant treatment in primary care settings is therefore inconsistent.

It may be possible to improve the current evidence base for the role of affective processing in early antidepressant action by identifying a more precise psychological mechanism. One such mechanism may be affective processing that maintains depressive self-schema. According to the cognitive model of depression, adverse social experiences in early life lead individuals to develop core sets of negative beliefs about the self, termed negative self-schema. When negative self-schema are activated by stressors in later life, they promote automatic processing of negative and punishing information about the self. Deliberative cognitive processes are attenuated, preventing reappraisal of these automatic biases. Negative self-schema are therefore reinforced by affective biases in a vicious cycle (Beck, 1987; Beck and Dozois, 2011). Supportive of this theory, affective biases associated with depression are heightened when information is processed about the self (Ji et al., 2017; Phillips et al., 2010). Focusing on the role of antidepressants in remediating negative affective biases occurring in reference to self-schema may therefore provide a more sensitive measure of antidepressant action.

Social evaluation learning is believed to be a key mechanism linking change in affective processing to change in self-schema. Perceptions of the self are informed by how we believe others view us (Shrauger and Schoeneman, 1979). Biased processing of social feedback is likely to inform self-schema. Within social interactions, healthy individuals demonstrate better learning of positive evaluations (Button et al., 2015) and update perceptions of the self to a greater degree in response to positive feedback (Korn et al., 2012). Conversely, individuals experiencing depression expect that others will perceive them negatively (Moritz and Roberts, 2018), preferentially engage with unfavourable social evaluations (Giesler et al., 1996), respond more negatively to social situations (Setterfield et al., 2016), and show poorer learning of positive evaluations about the self (Hobbs et al., 2021). Repeated exposure to negative social feedback about the self is likely to reinforce existing negative self-schema, perpetuating depressive symptoms. Antidepressants may operate by increasing sensitivity to positive social feedback, subsequently remediating negative self-schema by exposing individuals to increased positive evaluations about the self.

We have previously investigated the acute effect of antidepressants on social evaluation learning in a placebo-controlled randomised study of citalopram (Hobbs, Murphy, et al., 2020). In contrast to our expectations, citalopram appeared to increase positive affective biases towards familiar others, rather than the self. However, this study examined a single dose of citalopram in a healthy volunteer sample. Self-schemas are believed to be heavily entrenched and resistant to disconfirmatory evidence (Beck, 2008). Longer-term antidepressant treatment may be required to address pervasive self-referential affective biases. In keeping with this possibility, the cognitive neuropsychological model theorises that depressive symptoms are only addressed when individuals have interacted with their social environment with increased positive affective biases (Godlewska and Harmer, 2020). As change in depressive symptoms is commonly only observed from between three to eight weeks of antidepressant treatment (Uher et al., 2011), we may only expect to see change in social evaluation learning around this period.

We investigated whether antidepressants increase positive learning of social evaluations about the self and if change in social evaluation learning was associated with a change in depression. We took a naturalistic approach, observing change in social evaluation learning in primary care patients prescribed an antidepressant under the care of their general practitioner (GP) over the first eight weeks of treatment. We hypothesised that patients would become

better at learning positive social evaluations about the self and that this would be associated with a reduction in depression.

Method

The study protocol was pre-registered on Open Science Framework (https://osf.io/z9p8a), where study materials are also available (https://osf.io/8a95j/). The data that support the findings of this study are openly available in the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-01107 (Hobbs et al., 2022).

Participants

Participants were recruited through GP referrals from primary care sites in South West England. Eligible patients were aged 18 to 65 years, fluent in English, with normal or corrected-to-normal vision. Patients were eligible if they were considering antidepressant treatment but had not yet started treatment or were within the first two weeks of antidepressant treatment. A wash-out period of 8 weeks was required between previous and current courses of antidepressants.

Exclusion criteria included experience of a mental health disorder or developmental difficulties other than depression and anxiety (e.g., bipolar disorder, psychosis, autism, personality disorder and/or eating disorders), current treatment for substance misuse, and receiving care or being referred to secondary mental health care services. Due to the potential influence on affective processing, we excluded participants receiving Cognitive Behavioural Therapy at baseline or patients prescribed amitriptyline, pregabalin, benzodiazepines or any major tranquiliser currently or in the last 8 weeks.

Data were collected at 4 main time-points: baseline, two-, six-, and eight-week follow-up. An additional long-term follow-up timepoint was completed at six months by a small proportion of participants. Prior to the COVID-19 pandemic, data was collected face to face by researchers. To account for social distancing measures, from April 2020 onwards all data was collected remotely using online survey (Qualtrics, 2020) and cognitive task (Inquisit, 2020) software.

Ethical Approval

All participants provided written or digital informed consent. Ethical approval was obtained from the South West Frenchay NHS Research Ethics Committee (18/SW/0287).

Measures

At each timepoint participants completed self-report measures of mood and a computerised cognitive task measuring social evaluation learning.

Self-Report Measures of Mood

We used the Patient Health Questionnaire (PHQ-9) as a primary measure of depression, and the Beck Depression Inventory (BDI-II) as a secondary measure. Both questionnaires measure depression symptoms in the preceding two weeks with greater scores indicating greater severity. The PHQ-9 consists of nine items measuring DSM-IV diagnostic criteria symptoms with total scores ranging from 0 to 27. The BDI-II consists of 21 items with scores ranging from 0 to 63 and has a greater focus on cognitive symptoms. Both measures have good psychometric properties (Cameron et al., 2008; Wang and Gorenstein, 2013), and are used in primary care (Kendrick et al., 2009).

We used the Generalised Anxiety Disorder Questionnaire (GAD-7) to measure generalised anxiety symptoms. The GAD-7 has been validated as a screening tool for generalised anxiety disorder within primary care (Spitzer et al., 2006). Participants are asked to reflect on their experiences of seven anxiety symptoms within the previous two weeks. Scores range from 0 to 21, with higher scores indicating greater severity.

We used a single-item global rating of change (GRC) scale ('How have your moods and feelings changed?') to measure participants' perceptions of change in mood. At baseline participants were asked to respond based on change in the previous 2 weeks, at follow-up participants were asked to respond based on change since the previous timepoint. Following previous research (Hobbs, Lewis, et al., 2020), we collated GRC responses into a binary outcome of feeling better versus the feeling the same or worse to reflect that neither feeling the same nor worse is a positive therapeutic outcome.

At baseline only, participants completed the Clinical Interview Schedule-Revised (CIS-R) (Lewis et al., 1992), a self-administered computerised assessment that determines ICD-10 diagnoses of common mental health disorders. The CIS-R was completed in face-to-face sessions only.

Social Evaluation Learning Task

We used a computerised social evaluation learning task to measure affective learning within social contexts (Button et al., 2015). Participants learnt how much the computer 'liked' the self, a friend and a stranger based on feedback to a forced choice selection between positive and negative social evaluation pairings (Figure 1). Participants learnt two rules based on the probability of the positive evaluations being 'correct' (positive 'like' 60–80%, negative 'dislike' 20–40%). No time limit was imposed on selection of words. Individual blocks were

completed for each referential condition and rule. Order of referential condition, and rule nested within referential condition, was randomised. Participants completed 24 trials per referential condition-rule block.

To measure learning we calculated the number of errors made before reaching the criterion of eight consecutive rule-congruent responses. We then calculated bias scores to reflect learning of the positive relative to the negative rule, by subtracting errors to criterion made when learning the negative rule from the positive rule. Lower scores indicate a more positive bias as more errors have been made learning the negative relative to the positive rule.

Additional Measures

Additional measures were completed by participants prior to COVID-19. Full details are reported in supplementary materials. To allow for remote data collection and to reduce potential fatigue effects, only the measures outlined above were completed by participants following the pandemic. We chose to focus on the social evaluation learning task as we have previously found a reliable relationship between task outcomes and depression symptoms (Hobbs et al., 2019, 2021), as well as evidence of modulation by antidepressant administration (Hobbs, Murphy, et al., 2020). Due to low statistical power associated with a reduced sample size, aside from the CIS-R which has been included for descriptive purposes, this data has not been analysed and is not reported within this paper. However, we have made all outcome data from these measures openly available.

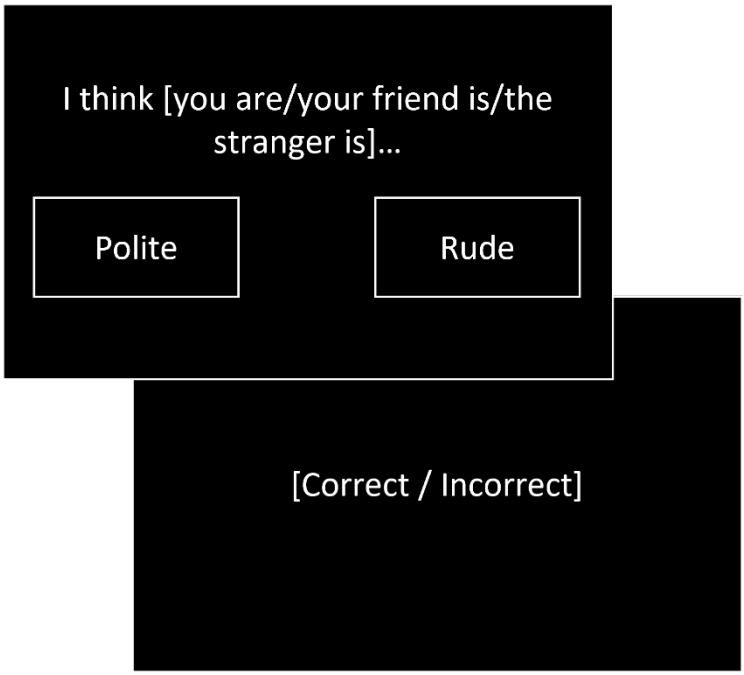


Figure 1

Example of a trial in the social evaluation learning task. Participants learnt how much the computer 'liked' or 'disliked' the self, a friend and a stranger in separate blocks based on feedback to selection of positive or negative social evaluation words. Participants learnt two rules, a positive 'like' rule where 'correct' feedback was given upon selection of the positive evaluation on 60-80% of trials, and a negative 'dislike' rule where 'correct' feedback was given upon selection of positive evaluations on 20-40% of trials.

Statistical Analyses

Analyses were conducted in R version 4.0.5.

Sample Size

We aimed to recruit 52 participants to provide 80% power to detect small to medium effects (d_z= 0.40), equivalent to an approximately 2-point change in social evaluation learning bias scores with an assumed standard deviation of 5. However, the COVID-19 pandemic occurred midway through the study impacting recruitment. Based on the number of participants that provided data at baseline and follow-up timepoints we are powered to detect medium to large within-subject effects (d_z = 0.61), equivalent to an approximately 3-point change in bias scores.

Hypothesis 1: Social evaluation learning regarding the self will become more positively biased during antidepressant treatment, as measured by a better learning of positive evaluations relative to negative evaluations towards the self.

We used mixed-effects linear regression models with bias scores (positive errors to criterion – negative errors to criterion) as a continuous outcome, participant as a random effect, and timepoint as categorical predictors. Additionally, to investigate the specificity of effects to the self we included referential condition and an interaction between referential condition and timepoint as categorical predictors.

Hypothesis 2: Change in social evaluation learning will be associated with a reduction in depressive symptoms, as indicated by a decrease in PHQ-9 scores.

We calculated change in PHQ-9 scores and bias scores in each referential condition separately by subtracting the current timepoint from the previous timepoint. We then used a mixed-effect linear regression model with change in PHQ-9 scores as a continuous outcome and change in bias scores in each referential condition as predictors. Time-point was entered as a fixed effect and participant as a random effect to account for the repeated measures design. As change in depression is influenced by baseline severity (Bauer-Staeb et al., 2021), we entered baseline PHQ-9 scores as a fixed effect. To assess the reliability of these findings we repeated this analysis with another measure of depression, the BDI-II.

Exploratory Analyses

We explored whether change in social evaluation learning was associated with a change in anxiety symptoms by repeating the analysis for hypothesis 2, substituting GAD-7 scores for PHQ-9 scores. To assess whether these findings persisted when co-morbid depression was taken into account, we repeated this analysis adjusting for depression by including change in PHQ-9 and BDI-II scores as additional predictors.

Results

Sample

Of 170 patients referred to the study, 33 patients took part in the baseline data collection session (Figure 2). Four participants with baseline data only were excluded due to extensive missing data, leaving 29 participants for analysis. At two-, six- and eight-week follow-up; 23 (79%), 21 (72%), and 22 (76%) participants provided data respectively. Eleven participants (38%) provided data at 6-month long-term follow-up.

Baseline sample characteristics are presented in Table 1. Patients were aged 18-60 years (mean 38, SD 11), predominantly female (62%), and all identified as white. At baseline all but one participant was taking an antidepressant for an average of 9 days. The most common antidepressants were sertraline (54%) and citalopram (32%). Approximately half of patients reported previous antidepressant treatment.

Details of treatment characteristics by timepoint are reported in Supplementary Table S1. Two participants discontinued antidepressant treatment (n = 1 two-weeks, n = 1 six-weeks). There was high treatment adherence across timepoints. A small proportion of participants reported also receiving psychological therapy (7-19% across timepoints).

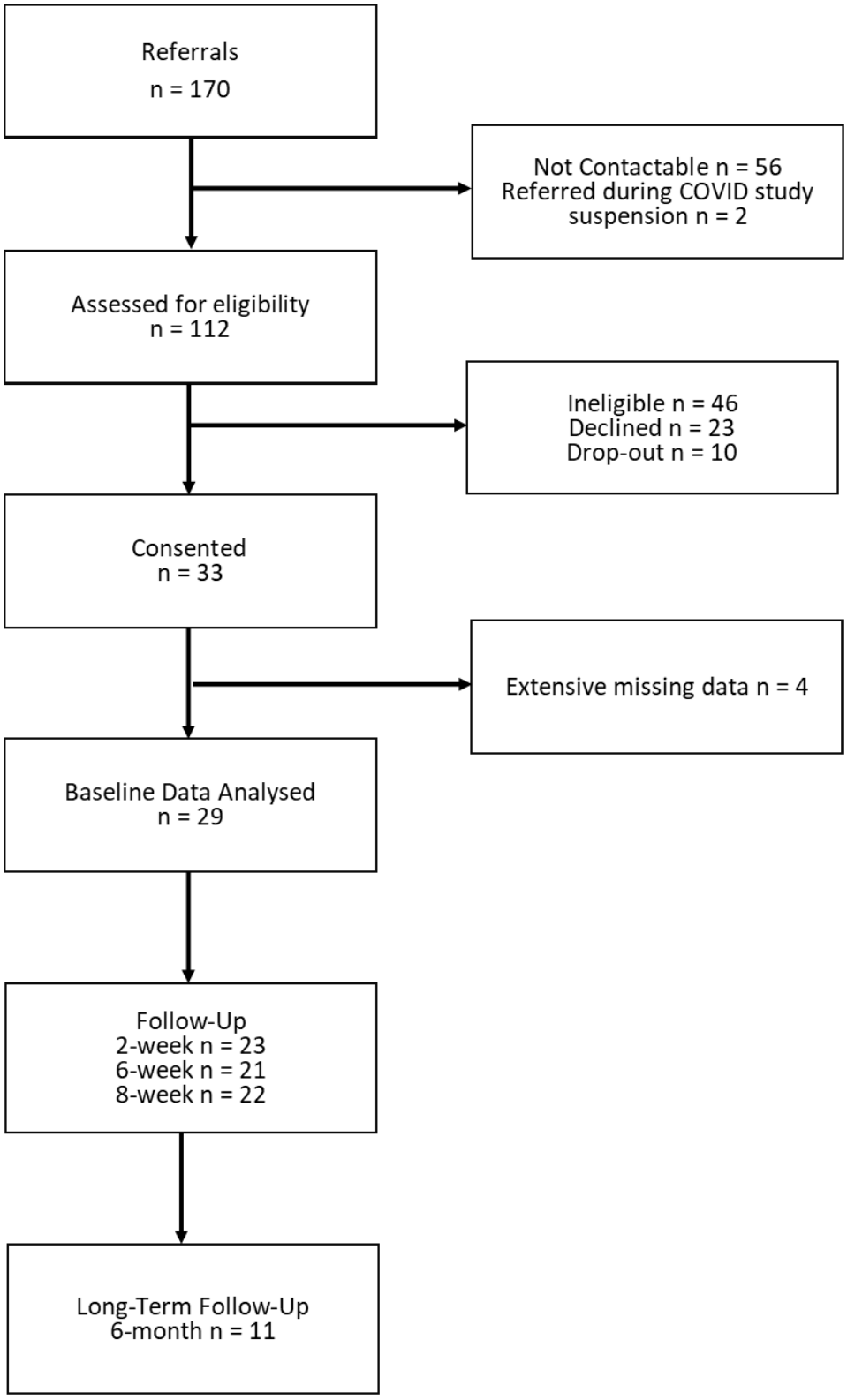


Figure 2

Flow diagram illustrating participant recruitment

Table 1

Baseline Sample Characteristics

N	29
Age, M (SD)	38.3 (11.3)
Gender, N (%)	
Male	11 (37.9)
Female	18 (62.1)
Ethnicity, N (%)	
White	29 (100)
Occupation, N (%)	
Employed	22 (75.9)
Student	2 (6.9)
Unemployed	3 (10.3)
Other	2 (6.9)
Educational attainment [a], N (%)	
GCSE or equivalent	5 (17.2)
A-Level or equivalent	8 (27.6)
Diploma or equivalent	6 (20.7)
Degree or equivalent	10 (34.5)
Relationship, N (%)	
Married/Living as Married	17 (58.6)
Single	8 (27.6)
Divorced / Separated	4 (13.8)
Living situation, N (%)	
Homeowner	17 (58.6)
Renting	4 (13.8)
Living with a relative/friend	8 (27.6)
Depression duration, N (%)	
< 2 weeks	0 (0)
2 weeks - 6 months	9 (32.1)
6 months – 1 year	6 (21.4)
1 – 2 years	2 (7.1)

≥ 2 years	11 (39.3)
CIS-R primary diagnosis [b]	
Depressive Episode	17 (89.5%)
Generalised Anxiety Disorder	2 (10.5%)
CIS-R secondary diagnosis [b]	
Generalised Anxiety Disorder	8 (42.1%)
Mixed Anxiety and Depression	8 (42.1%)
Specific Phobia	3 (15.8%)
Antidepressant medication, N (%) [c]	
Sertraline	15 (53.6)
Citalopram	9 (32.1)
Fluoxetine	2 (7.1)
Mirtazapine	2 (7.1)
Length of current antidepressant treatment (days), M (SD)	9.3 (3.7)
Other medications [d], N (%)	7 (24.1)
Previous antidepressant treatment, N (%)	14 (48.3)
Previous psychological therapy, N (%)	12 (41.4)
Data collection, N (%)	
Face to face	19 (65.5)
Remote	10 (34.5)

CIS-R = Clinical Interview Schedule Revised, GCSE = General Certificate of Secondary Education, A-Level = Advanced Level

[a] GCSEs are an entry-level qualification taken by UK students typically at the end of compulsory education at 16 years of age, equivalent to O-Levels or CSEs; A-Levels are a post-16 pre-university subject-specific qualification, equivalent to the International Baccalaureate; Diploma or equivalent are practical worked-based qualifications in specific industries/careers, includes Regulated Qualification Framework (RQF), National Vocation Qualification (NVQ) or Business and Technology Education Council (BTEC) diplomas; Degree or equivalent refers to undergraduate or postgraduate degree or higher qualification completed at a university or other higher education institution.

[b] The CIS-R was completed in face to face baseline testing sessions only. Data is therefore only available for 19 participants.

[c] One participant was not taking antidepressant medication at baseline

[d] Medication reported included treatment for diabetes and high blood pressure, hormonal treatment, treatment for an underactive thyroid gland, an asthma inhaler, medication for heartburn, antibiotics, painkillers, and the contraceptive pill

Change in Mood

Mean scores for measures of depression (PHQ-9, BDI-II) and anxiety (GAD-7) are reported in table 2. At baseline, participants on average experienced moderate depression and anxiety symptoms. Both depression and anxiety declined over time; by 8-week follow-up average scores reflected mild symptoms. Most participants reported feeling better at follow-up timepoints on the GRC.

Table 2

Descriptive statistics for self-report measures of depression (PHQ-9, BDI-II), anxiety (GAD-7), change in mood (GRC) and bias scores in the Social Evaluation Learning task by timepoint. Greater bias scores indicate a more negative bias (relatively better learning of the negative relative to the positive rule).

	Baseline	**2-weeks**	**6-weeks**	**8-weeks**	**6-months**
Self-Reported Mood					
N	29	23	21	21	11
PHQ-9, M (SD)	14.31 (4.81)	9.09 (5.01)	7.62 (4.84)	6.90 (5.39)	8.73 (6.39)
BDI-II, M (SD)	26.50 (8.74)	19.78 (9.71)	13.53 (8.55)	13.62 (10.70)	12.10 (9.17)
GAD-7, M (SD)	12.83 (4.72)	8.26 (4.74)	6.38 (5.45)	5.29 (4.74)	7.45 (6.12)
GRC, N (%) [a]					
Worse/Same	15 (53.57)	4 (17.39)	6 (28.57)	3 (14.29)	1 (9.09)
Better	13 (46.43)	19 (82.61)	15 (71.43)	18 (85.71)	10 (90.91)
Social Evaluation Learning Bias Scores					
N	29	22	20	21	10
Self, M (SD)	-0.84 (8.30)	-2.34 (7.22)	-2.85 (7.67)	-2.48 (5.69)	-6.05 (7.77)
Friend, M (SD)	-5.28 (8.18)	-4.23 (7.20)	-4.00 (6.57)	-4.07 (6.59)	-2.94 (8.57)
Stranger, M (SD)	-1.36 (5.60)	-3.16 (6.05)	-2.33 (6.16)	-1.45 (6.71)	-7.56 (8.58)

Bias Scores = positive errors to criterion – negative errors to criterion, PHQ-9 = Patient Health Questionnaire, BDI-II = Beck Depression Inventory II, GAD-7 = Generalised Anxiety Disorder Questionnaire, GRC = Global Rating of Change Scale, SEL = Social Evaluation Learning

[a] At baseline, participants responded to the question "Compared to two weeks ago, how have your moods and feelings changed?", at follow-up time points participants responded to the question "Compared to when you last answered these questions, how have your moods and feelings changed?"

Note: BDI-II scores were missing for one participant at baseline, two participants at 6-weeks, and one participant at 6-months. Bias scores were missing for the friend and stranger conditions for one participant at 6-month follow-up.

Hypothesis 1: Social evaluation learning regarding the self will become more positively biased during antidepressant treatment, as measured by a better learning of positive evaluations relative to negative evaluations towards the self.

Participants were more positively biased when learning about the friend versus the self at baseline (*b* = -4.43, 95% CI: -7.76, -1.11, p = 0.010). However, there was no evidence of a change in bias scores over time, or that this differed by referential condition (Table 3; session: p = 0.934, session x referential condition p = 0.834). We therefore did not find evidence to support our hypothesis. Full results are reported in Table 3 and mean bias scores are presented in Table 2 and Figure 3.

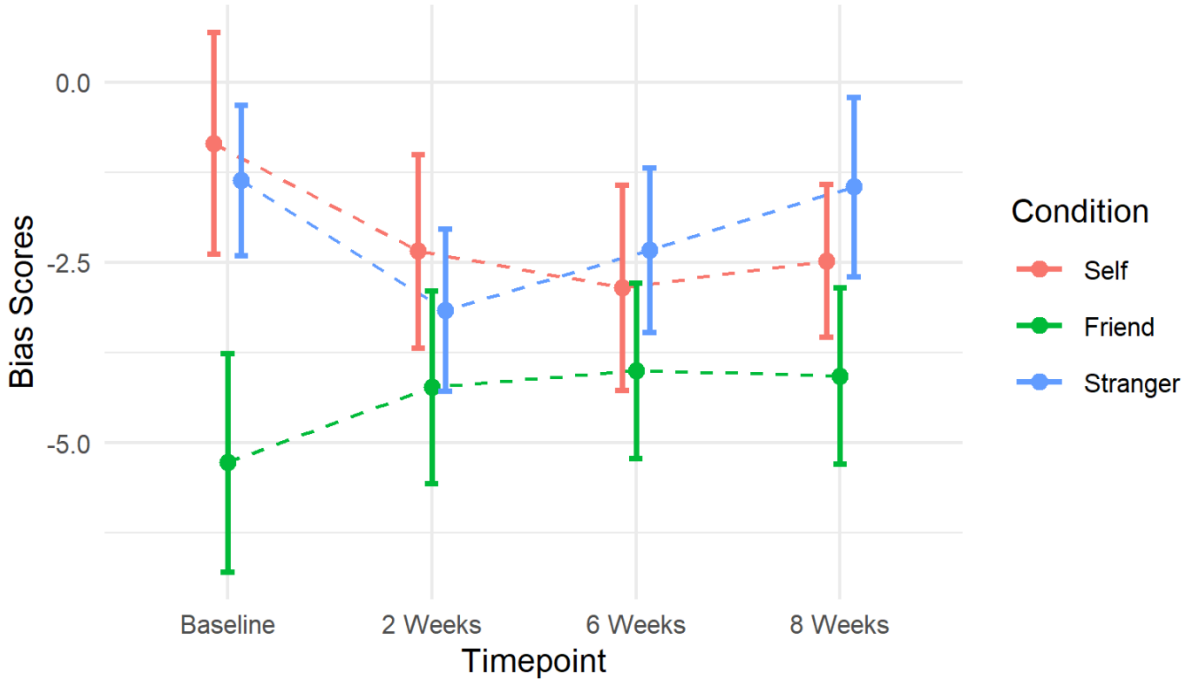


Figure 3

Change in bias scores according to referential condition over 8 weeks of antidepressant treatment. Lower bias scores indicate a more positive bias as participants have made a greater number of errors learning the negative 'dislike' rule versus the positive 'like' rule. Error bars represent standard errors.

Table 3

Mixed-effects linear regression models examining differences in SEL bias scores (outcome) by timepoint and referential condition

	b	***b* 95% CI**	**β**	**β 95% CI**	**p**
Intercept	-0.85	-3.38, 1.69	0.29	-0.08, 0.65	0.514
Session					0.934
Baseline	Reference				
2-week	-1.57	-5.18, 2.03	-0.23	-0.75, 0.29	0.393
6-week	-2.26	-5.97, 1.45	-0.33	-0.86, 0.21	0.234
8-week	-1.83	-5.48, 1.83	-0.26	-0.79, 0.26	0.328
Referential Condition					0.160
Self	Reference				
Friend	-4.43	-7.76, -1.11	-0.64	-1.12, -0.16	0.010
Stranger	-0.52	-3.84, 2.81	-0.07	-0.55, 0.40	0.761
Session * Referential Condition					0.834
2-week * Friend	2.55	-2.52, 7.61	0.37	-0.36, 1.10	0.326
6-week * Friend	3.28	-1.92, 8.49	0.47	-0.28, 1.22	0.218
8-week * Friend	2.84	-2.29, 7.97	0.41	-0.33, 1.15	0.280
2-week * Stranger	-0.30	-5.36, 4.76	-0.04	-0.77, 0.69	0.907
6-week * Stranger	1.04	-4.16, 6.25	0.15	-0.60, 0.90	0.695
8-week * Stranger	1.54	-3.59, 6.67	0.22	-0.52, 0.96	0.557

b = unstandardised regression coefficients, β = standardised regression coefficient.

Hypothesis 2: Change in social evaluation learning will be associated with a reduction in depressive symptoms, as indicated by a decrease in PHQ-9 scores.

We did not find evidence in support of our hypothesis. Change in depression severity, measured by PHQ-9 scores, was not associated with change in biased learning about the self (*b* = 0.08, 95% CI: -0.05, 0.20, p = 0.239), the friend (*b* = 0.09, 95% CI: -0.05, 0.23, p = 0.229), or the stranger (*b* = -0.01, 95% CI: -0.15, 0.13, p = 0.871; Table 4, Figure 4).

Table 4

Mixed-effects linear regression models examining the association between change in PHQ-9 scores (outcome) and change in SEL bias scores by referential condition

	b	***b* 95% CI**	**β**	**β 95% CI**	**p**
Intercept	-2.02	-5.58. 1.55	-0.52	-0.90, -0.14	0.272
Bias Scores Change					
Self	0.08	-0.05, 0.20	0.15	-0.10, 0.39	0.239
Friend	0.09	-0.05, 0.23	0.15	-0.09, 0.38	0.229
Stranger	-0.01	-0.15, 0.13	-0.02	-0.27, 0.23	0.871
Baseline PHQ-9	-0.22	-0.44, -0.01	-0.24	-0.47, -0.01	0.050
Session					0.003
Baseline to 2-weeks	Reference				
2 to 6-weeks	3.97	1.41, 6.53	0.86	0.31, 1.42	0.004
6 to 8-weeks	3.62	1.03, 6.22	0.79	0.22, 1.35	0.008

b = unstandardised regression coefficients, β = standardised regression coefficient.

When we examined the association between change in learning and BDI-II scores, our secondary measure of depression, we found weak evidence of an association in the friend condition. An increase in learning of positive relative to negative evaluations about the friend was associated with a reduction in BDI-II scores, although statistical evidence for this was weak (*b* = 0.30, 95% CI: -0.02, 0.62, p = 0.069). We did not find evidence of an association with biased learning about the self or the stranger (Table 5).

Table 5

Mixed-effects linear regression models examining the association between change in BDI-II scores (outcome) and change in SEL bias scores by referential condition

	b	***b* 95% CI**	**β**	**β 95% CI**	**p**
Intercept	1.52	-6.32, 9.37	-0.25	-0.65, 0.15	0.705
Bias Scores Change					
Self	0.10	-0.18, 0.38	0.10	-0.17, 0.36	0.469
Friend	0.30	-0.02, 0.62	0.24	-0.01, 0.49	0.069
Stranger	-0.03	-0.37, 0.31	-0.02	-0.29, 0.25	0.874
Baseline BDI-II	-0.32	-0.58, -0.07	-0.31	-0.56, -0.07	0.016
Session					0.039
Baseline to 2-weeks	Reference				
2 to 6-weeks	0.90	-4.64, 6.45	0.10	-0.50, 0.70	0.751
6 to 8-weeks	6.37	0.89, 11.85	0.69	0.10, 1.29	0.027

b = unstandardised regression coefficients, β = standardised regression coefficient.

Exploratory Analysis

In exploratory analyses we found evidence of an association between change in anxiety and change in bias scores (Table 6, Figure 4). Increased positive learning about the self and the friend, indicated by a reduction in bias scores, was associated with a small reduction in GAD-7 scores (Self: *b* = 0.18, 95% CI: 0.07, 0.28, p = 0.002, Friend: *b* = 0.22, 95% CI: 0.10, 0.35, p = 0.001). We did not find evidence of an association between change in anxiety and learning about a stranger (*b* = 0.01, 95% CI: -0.11, 0.13, p = 0.835). When we adjusted for change in PHQ-9 and BDI-II scores these findings were unchanged (Supplementary Table S2).

Table 6

Mixed-effects linear regression models for an exploratory analysis examining the association between change in GAD-7 scores (outcome) and change in SEL bias scores by referential condition

	b	***b* 95% CI**	**β**	**β 95% CI**	**p**
Intercept	-0.17	-3.38, 3.03	-0.34	-0.66, -0.02	0.917
Bias Scores Change					
Self	0.18	0.07, 0.28	0.34	0.14, 0.55	0.002
Friend	0.22	0.10, 0.35	0.36	0.16, 0.56	0.001
Stranger	0.01	-0.11, 0.13	0.02	-0.19, 0.23	0.835
Baseline GAD-7	-0.33	-0.55, -0.11	-0.30	-0.50, -0.10	0.004
Session					0.020
Baseline to 2-weeks	Reference				
2 to 6-weeks	2.98	0.79, 5.16	0.65	0.17, 1.12	0.010
6 to 8-weeks	1.97	-0.25, 4.19	0.43	-0.05, 0.91	0.088

b = unstandardised regression coefficients, β = standardised regression coefficient.

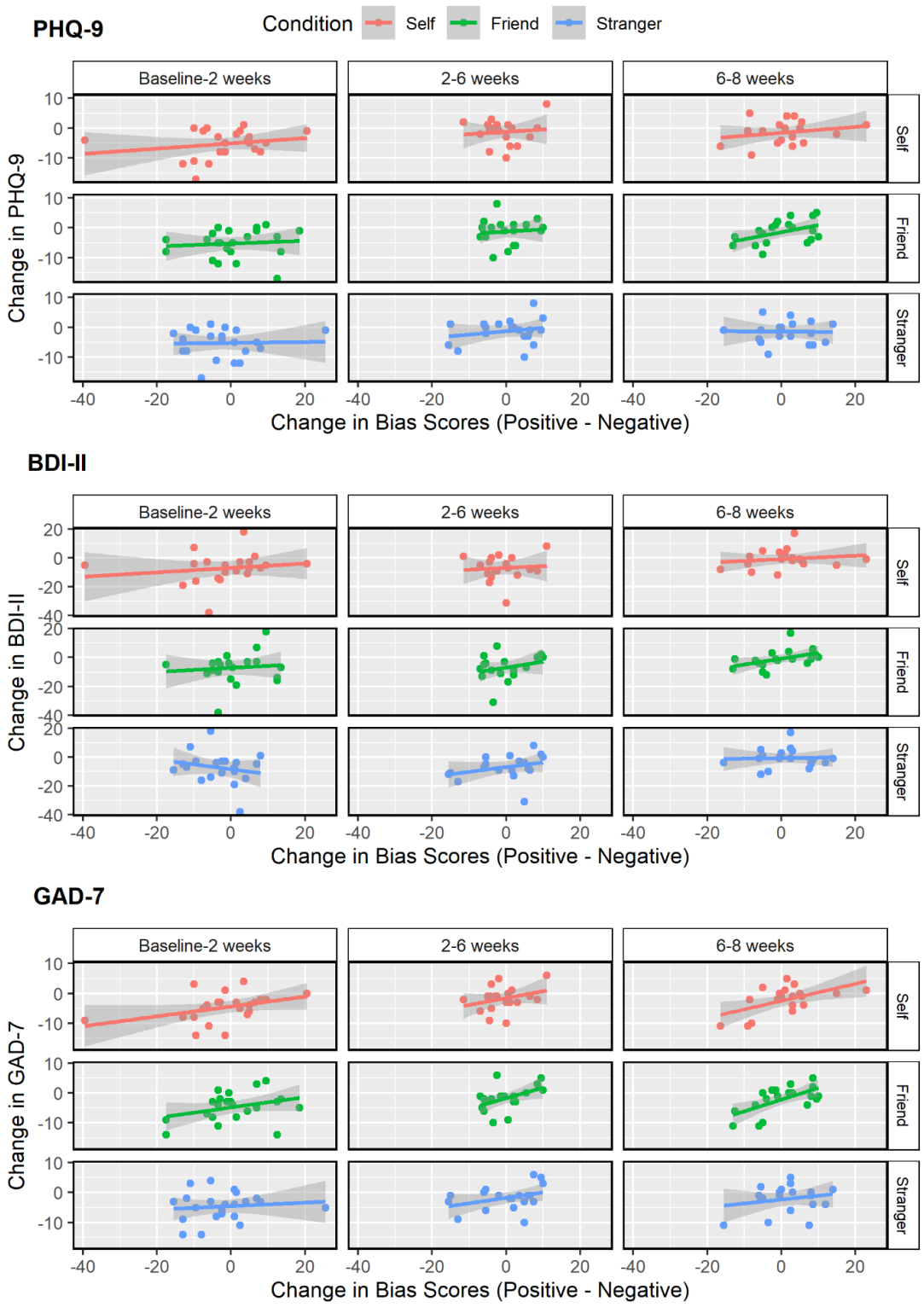


Figure 4

Association between change in biased learning about the self, a friend, and a stranger and self-report measures of depression (PHQ-9 and BDI-II) and anxiety (GAD-7) by timepoint. Errors bars represent 95% confidence intervals.

Long-Term Follow-Up

At six-month follow-up participants on average showed a small increase in PHQ-9 and GAD-7 scores from eight-week follow-up, although scores remained substantially lower than baseline. Bias scores declined for the self and the stranger but remained relatively stable for the friend (Table 2). However, as only a small proportion of participants provided data at this timepoint (n = 11, 37.9%) the majority of whom reported feeling better (90.91%), this pattern may only represent a particular subset of participants.

Discussion

We investigated the association between change in learning of social evaluations and depression symptoms over the first eight weeks of antidepressant treatment in primary care patients. In contrast to our expectations, we did not find evidence of an association between learning of social evaluations about the self and reductions in depression. In our secondary measure of depression, we instead found weak evidence of an association with learning about a friend. Exploratory analyses suggested that change in social evaluation learning about both the self and a friend was instead better associated with a reduction in anxiety rather than depression symptoms.

We speculated that antidepressants may remediate negative self-schema by increasing learning of positive social evaluations about the self. However, we did not find evidence of an association between change in learning about the self and a reduction in PHQ-9 or BDI-II scores. Despite most patients showing an improvement in depression, learning about the self was relatively stable throughout treatment. These findings are in line with our previous research, where we did not find evidence that acute citalopram influenced learning about the self in healthy volunteers (Hobbs, Murphy, et al., 2020). It is possible that self-schemas may be more effectively targeted using a combination of antidepressants and cognitive behavioural therapy to address both top-down and bottom-up affective biases (Dozois et al., 2009; Roiser et al., 2012). However, our findings do not support the theory that change in self-referential learning of social evaluations plays a central role in addressing depression symptoms in antidepressant treatment alone.

Using a secondary measure of depression, the BDI-II, we found weak evidence of an association between increased positive learning about a friend and a reduction in depression. These findings are consistent with our previous research, where citalopram increased positive affective biases towards familiar others but not the self (Hobbs, Murphy, et al., 2020). It is possible that antidepressants may partly operate by increasing sensitivity to positive characteristics in close others such as friends. However, as we did not find evidence of this with the PHQ-9 it is possible that effects are specific to symptoms measured by the BDI-II. One possible symptom may be social anhedonia, measured by the BDI-II item ‘interest in other people’. Social anhedonia is believed to be important in the development and relapse of depression (Barkus and Badcock, 2019), as well as expression of social behaviours (Setterfield et al., 2016). It is possible that antidepressants may increase learning of positive characteristics in close others subsequently increasing reward and engagement in social

interactions. However, evidence for this effect was weak and these findings should therefore be treated with caution.

In contrast to our hypotheses, which focused on the role of social evaluation learning in depression, exploratory analyses indicated that change in affective learning was associated with a reduction in anxiety. On average, patients that became better at learning positive versus negative evaluations about both the self and the friend showed a reduction in anxiety symptoms. Cognitive models propose that individuals with generalised anxiety hold self-schemas focused on personal threat (Beck, 1976). When activated these schemas produce affective biases that heighten sensitivity towards threatening environmental stimuli. Supportive of this theory, anxious individuals show greater attention towards and implicit memory for threatening information and are more likely to interpret neutral information as threatening (Bar-Haim et al., 2007; Dalgleish and Watts, 1990; Eysenck et al., 1991; Mathews et al., 1997). We have previously found that individuals with greater social anxiety show better learning of negative relative to positive social evaluations (Button et al., 2012, 2015). Negative social evaluations, such as criticism, may represent a source of social threat that individuals with generalised anxiety implicitly interpret as potential acts of social exclusion or aggression. Antidepressants may reduce sensitivity to negative information, helping to remediate these threat-related biases and subsequently reducing anxiety symptoms.

Our findings are similar to those of a recent RCT within primary care. Patients experiencing depression were randomly assigned to receive treatment as usual or to have their treatment guided by an algorithm based predominantly on change in facial emotion recognition. There was no evidence that patients in the algorithm arm showed a greater decrease in the primary outcome of depression symptoms at eight weeks. However, patients in the algorithm arm showed a greater decline in the secondary outcome of anxiety (Browning et al., 2021). Altering treatment based on affective biases was therefore more beneficial for anxiety rather than depression. Additionally, previous research has found that short-term antidepressant treatment of generalised anxiety reduced threat-related interpretative biases (Mogg et al., 2004). In combination with previous evidence, our findings suggest that change in affective processing may be a potential antidepressant mechanism for anxiety.

Furthermore, recent evidence suggests that change in anxiety rather than depression, may be a more sensitive measure of therapeutic outcome of antidepressant treatment. Antidepressants are used to manage anxiety symptoms where patients have declined or not responded to

psychological interventions (NICE, 2011, 2013), and are generally effective (Cassano et al., 2002). Recent research has suggested that antidepressants may benefit primary care patients through improvements in anxiety as opposed to depression. In a large-scale RCT of primary care patients, there was evidence of a clinically meaningful reduction in anxiety but not depression symptoms at 6-weeks in patients treated with sertraline versus placebo (Lewis et al., 2019). Notably, most participants in our sample were receiving sertraline suggesting potentially similar effects may have occurred on anxiety. Researchers have subsequently proposed that focusing on change in affective processing as a predictor of anxiety may be more useful than depression (Browning et al., 2021). Our findings suggest that change in social evaluation learning about the self and familiar others may be a sensitive marker of antidepressant response based on change in anxiety symptoms.

Strengths

Whereas most previous research examining affective processing in antidepressant action has been conducted under controlled laboratory settings, we took a naturalistic approach observing primary care patients receiving antidepressant treatment. Our findings are therefore more representative of antidepressant treatment for depression and anxiety in the United Kingdom. We also employed a prospective cohort design, allowing us to investigate the relationship between change in mood and social evaluation learning occurring over time.

Additionally, we employed a novel task that was able to integrate self, emotion, and reward processing. We have previously validated the use of this task in individuals experiencing varying levels of depression (Hobbs et al., 2019, 2021) and anxiety (Button et al., 2012, 2015; Hopkins et al., 2021).

Limitations

Recruitment was impacted by the COVID-19 pandemic, and we were unable to reach our target sample. Due to our small sample, we were underpowered to detect small effects potentially contributing to false-negative findings. Additionally, our sample was limited in its demographic diversity with all participants identifying as white. Although antidepressant response has not been found to differ according to ethnicity (Lesser et al., 2010), our findings may not be reflective of all individuals within primary care. Further research evaluating our findings in a larger and more demographically diverse sample is required.

This study lacked a control group limiting our ability to understand the causal role of change in social evaluation learning on anxiety or depression. It is possible that change in social

evaluation learning may be an epiphenomenon of mood disorders rather than playing a mechanistic role. Recruiting larger samples would allow for more complex statistical analyses, such as cross-lagged regression models (Hecht and Zitzmann, 2021), to investigate this possibility.

Conclusions

In contrast to our hypotheses, we did not find evidence of an association between change in social evaluation learning and depression symptoms. Change in social evaluation learning was instead more reliably associated with a reduction in anxiety. Patients that became more positively biased when learning about both the self and the friend on average showed a reduction in anxiety symptoms. Antidepressants may treat anxiety symptoms by remediating negative affective biases towards socially threatening information. However, these findings are based on exploratory analyses and require further replication.

References

Ahmed N, Bone JK, Lewis Gemma, et al. (2021) The effect of sertraline on emotional processing: Secondary analyses of the PANDA randomised controlled trial. *Psychological Medicine*: 1–8. DOI: 10.1017/S0033291720004985.

Arnone D, Horder J, Cowen PJ, et al. (2009) Early effects of mirtazapine on emotional processing. *Psychopharmacology* 203(4): 685. DOI: 10.1007/s00213-008-1410-6.

Bar-Haim Y, Lamy D, Pergamin L, et al. (2007) Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*. DOI: 10.1037/0033-2909.133.1.1.

Barkus E and Badcock JC (2019) A transdiagnostic perspective on social anhedonia. *Frontiers in Psychiatry* 10: 216. DOI: 10.3389/FPSYT.2019.00216/BIBTEX.

Bauer-Staeb C, Kounali D-Z, Welton NJ, et al. (2021) Effective dose 50 method as the minimal clinically important difference: Evidence from depression trials. *Journal of Clinical Epidemiology* 137: 200–208. DOI: 10.1016/j.jclinepi.2021.04.002.

Beck AT (1976) *Cognitive Therapy and the Emotional Disorders*. New York: International University Press.

Beck AT (1987) Cognitive models of depression. *Journal of Cognitive Psychotherapy* 1: 5–37.

Beck AT (2008) The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry* 165(8): 969–977. DOI: 10.1176/appi.ajp.2008.08050721.

Beck AT and Dozois DJA (2011) Cognitive therapy: Current status and future directions. *Annual Review of Medicine* 62: 397–409. DOI: 10.1146/annurev-med-052209-100032.

Berardi D, Menchetti M, Cevenini N, et al. (2005) Increased recognition of depression in primary care: Comparison between primary-care physician and ICD-10 diagnosis of depression. *Psychotherapy and Psychosomatics* 74(4): 225–230. DOI: 10.1159/000085146.

Browning M, Reid C, Cowen PJ, et al. (2007) A single dose of citalopram increases fear recognition in healthy subjects. *Journal of Psychopharmacology* 21(7): 684–690. DOI: 10.1177/0269881106074062.

Browning M, Kingslake J, Dourish CT, et al. (2019) Predicting treatment response to antidepressant medication using early changes in emotional processing. *European Neuropsychopharmacology* 29(1): 66–75. DOI: 10.1016/j.euroneuro.2018.11.1102.

Browning M, Bilderbeck AC, Dias R, et al. (2021) The clinical effectiveness of using a predictive algorithm to guide antidepressant treatment in primary care (PReDicT): an open-label, randomised controlled trial. *Neuropsychopharmacology* 46(7): 1307–1314. DOI: 10.1038/s41386-021-00981-z.

Button KS, Browning M, Munafò MR, et al. (2012) Social inference and social anxiety: Evidence of a fear-congruent self-referential learning bias. *Journal of Behavior Therapy and Experimental Psychiatry* 43(4): 1082–1087. DOI: 10.1016/j.jbtep.2012.05.004.

Button KS, Kounali D, Stapinski L, et al. (2015) Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. *PLoS ONE* 10(4): e0119456. DOI: 10.1371/journal.pone.0119456.

Cameron IM, Crawford JR, Lawton K, et al. (2008) Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *British Journal of General Practice* 58(546): 32–36. DOI: 10.3399/bjgp08X263794.

Cassano GB, Rossi NB and Pini S (2002) Psychopharmacology of anxiety disorders. *Dialogues in Clinical Neuroscience* 4(3): 271. DOI: 10.31887/DCNS.2002.4.3/GCASSANO.

Cipriani A, Furukawa TA, Salanti G, et al. (2018) Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *FOCUS* 16(4): 420–429. DOI: 10.1176/appi.focus.16407.

Cooper CM, Whiting DA, Cowen PJ, et al. (2015) Tianeptine in an experimental medicine model of antidepressant action. *Journal of Psychopharmacology* 29(5): 582–590. DOI: 10.1177/0269881115573810.

Cremers H, Lee R, Keedy S, et al. (2016) Effects of escitalopram administration on face processing in intermittent explosive disorder: An fMRI Study. *Neuropsychopharmacology* 41(2): 590–597. DOI: 10.1038/npp.2015.187.

Dalgleish T and Watts FN (1990) Biases of attention and memory in disorders of anxiety and

depression. *Clinical Psychology Review* 10(5): 589–604. DOI: 10.1016/0272-7358(90)90098-U.

Dalili MN, Penton-Voak IS, Harmer CJ, et al. (2015) Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine* 45(6): 1135–1144. DOI: 10.1017/S0033291714002591.

Di Simplicio M, Doallo S, Costoloni G, et al. (2014) 'Can you look me in the face?' Short-term SSRI administration reverts avoidant ocular face exploration in subjects at risk for psychopathology. *Neuropsychopharmacology* 39(13): 3059–3066. DOI: 10.1038/npp.2014.159.

Dozois DJA, Bieling PJ, Patelis-Siotis I, et al. (2009) Changes in Self-Schema Structure in Cognitive Therapy for Major Depressive Disorder: A Randomized Clinical Trial. *Journal of Consulting and Clinical Psychology* 77(6): 1078–1088. DOI: 10.1037/a0016886.

Eysenck MW, Mogg K, May J, et al. (1991) Bias in Interpretation of Ambiguous Sentences Related to Threat in Anxiety. *Journal of Abnormal Psychology* 100(2): 144–150. DOI: 10.1037/0021-843X.100.2.144.

Giesler RB, Josephs RA and Swann WB (1996) Self-verification in clinical depression: The desire for negative evaluation. *Journal of Abnormal Psychology* 105(3): 358. DOI: 10.1037/0021-843X.105.3.358.

Godlewska BR and Harmer CJ (2020) Cognitive neuropsychological theory of antidepressant action: a modern-day approach to depression and its treatment. *Psychopharmacology*: 1–14. DOI: 10.1007/s00213-019-05448-0.

Harmer CJ, Hill SA, Taylor MJ, et al. (2003) Toward a neuropsychological theory of antidepressant drug action: Increase in positive emotional bias after potentiation of norepinephrine activity. *American Journal of Psychiatry*. DOI: 10.1176/appi.ajp.160.5.990.

Harmer CJ, Mackay CE, Reid CB, et al. (2006) Antidepressant Drug Treatment Modifies the Neural Processing of Nonconscious Threat Cues. *Biological Psychiatry* 59(9): 816–820. DOI: 10.1016/j.biopsych.2005.10.015.

Harmer CJ, O'Sullivan U, Favaron E, et al. (2009) Effect of acute antidepressant

administration on negative affective bias in depressed patients. *American Journal of Psychiatry* 166(10): 1178–1184. DOI: 10.1176/appi.ajp.2009.09020149.

Harmer CJ, De Bodinat C, Dawson GR, et al. (2011) Agomelatine facilitates positive versus negative affective processing in healthy volunteer models. *Journal of Psychopharmacology* 25(9): 1159–1167. DOI: 10.1177/0269881110376689.

Hecht M and Zitzmann S (2021) Sample Size Recommendations for Continuous-Time Models: Compensating Shorter Time Series with Larger Numbers of Persons and Vice Versa. *Structural Equation Modeling* 28(2): 229–236. DOI: 10.1080/10705511.2020.1779069.

Hobbs C, Faraway J, Kessler D, et al. (2019) Social evaluation learning in social anxiety disorder and depression: a mega-analysis presented at the Summer Meeting of the Bristish Association for Psychopharmacology. *Journal of Psychopharmacology* Supplement(8): A33.

Hobbs C, Lewis G, Dowrick C, et al. (2020) Comparison between self-administered depression questionnaires and patients' own views of changes in their mood: A prospective cohort study in primary care. *Psychological Medicine*. DOI: 10.1017/S0033291719003878.

Hobbs C, Murphy SE, Wright L, et al. (2020) Effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers. *BJPsych Open* 6(6). DOI: 10.1192/bjo.2020.107.

Hobbs C, Sui J, Kessler D, et al. (2021) Self-processing in relation to emotion and reward processing in depression. *Psychological Medicine*: 1–13. DOI: 10.1017/s0033291721003597.

Hobbs C, Beck M, Denham F, et al. (2022) Dataset for 'The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care'. University of Bath Research Data Archive. DOI: 10.15125/BATH-01107.

Hopkins AK, Dolan R, Button KS, et al. (2021) A Reduced Self-Positive Belief Underpins Greater Sensitivity to Negative Evaluation in Socially Anxious Individuals. *Computational Psychiatry* 5(1): 21. DOI: 10.5334/cpsy.57.

Inquisit (2020) Inquisit 6. Available at: www.millisecond.com.

Ji JL, Grafton B and MacLeod C (2017) Referential focus moderates depression-linked attentional avoidance of positive information. *Behaviour Research and Therapy* 93: 47–54. DOI: 10.1016/j.brat.2017.03.004.

Kamboj SK and Curran H V. (2006) Scopolamine induces impairments in the recognition of human facial expressions of anger and disgust. *Psychopharmacology* 185(4): 529–535. DOI: 10.1007/s00213-006-0332-4.

Kendrick T, Dowrick C, McBride A, et al. (2009) Management of depression in UK general practice in relation to scores on depression severity questionnaires: Analysis of medical record data. *BMJ (Online)* 338(7697): b750. DOI: 10.1136/bmj.b750.

Komulainen E, Heikkilä R, Meskanen K, et al. (2016) A single dose of mirtazapine attenuates neural responses to self-referential processing. *Journal of Psychopharmacology* 30(1): 23–32. DOI: 10.1177/0269881115616384.

Komulainen E, Heikkilä R, Nummenmaa L, et al. (2018) Short-term escitalopram treatment normalizes aberrant self-referential processing in major depressive disorder. *Journal of Affective Disorders* 236: 222–229. DOI: 10.1016/j.jad.2018.04.096.

Korn CW, Prehn K, Park SQ, et al. (2012) Positively biased processing of self-relevant social feedback. *Journal of Neuroscience* 32(47): 16832–16844. DOI: 10.1523/JNEUROSCI.3016-12.2012.

Labuschagne I, Croft RJ, Phan KL, et al. (2010) Augmenting serotonin neurotransmission with citalopram modulates emotional expression decoding but not structural encoding of moderate intensity sad facial emotional stimuli: An event-related potential (ERP) investigation. *Journal of Psychopharmacology* 24(8): 1153–1164. DOI: 10.1177/0269881108097878.

Lesser IM, Myers HF, Lin KM, et al. (2010) Ethnic differences in antidepressant response: a prospective multi-site clinical trial. *Depression and Anxiety* 27(1): 56–62. DOI: 10.1002/DA.20619.

Lewis G, Pelosi A, Araya R, et al. (1992) Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine* 22(2): 465–486. DOI: 10.1017/S0033291700030415.

Lewis Gemma, Duffy L, Ades A, et al. (2019) The clinical effectiveness of sertraline in primary care and the role of depression severity and duration (PANDA): a pragmatic, double-blind, placebo-controlled randomised trial. *The Lancet Psychiatry* 6(11): 903–914. DOI: 10.1016/S2215-0366(19)30366-9.

Mathews A, Mackintosh B and Fulcher EP (1997) Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences* 1(9). Elsevier Current Trends: 340–345. DOI: 10.1016/S1364-6613(97)01092-9.

McManus S, Bebbington P, Jenkins R, et al. (2016) Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. A survey carried out for NHS Digital by NatCen Social Research and the Department of Health Sciences, University of Leicester. NHS digital.

Mogg K, Baldwin DS, Brodrick P, et al. (2004) Effect of short-term SSRI treatment on cognitive bias in generalised anxiety disorder. *Psychopharmacology* 176(3–4): 466–470. DOI: 10.1007/s00213-004-1902-y.

Moritz D and Roberts JE (2018) Self-Other Agreement and Metaperception Accuracy Across the Big Five: Examining the Roles of Depression and Self-Esteem. *Journal of Personality* 86(2): 296–307. DOI: 10.1111/jopy.12313.

Murphy SE, Norbury R, O'Sullivan U, et al. (2009) Effect of a single dose of citalopram on amygdala response to emotional faces. *British Journal of Psychiatry* 194: 535–540. DOI: 10.1192/bjp.bp.108.056093.

Phillips WJ, Hine DW and Thorsteinsson EB (2010) Implicit cognition and depression: A meta-analysis. *Clinical Psychology Review* 30(6): 691–709. DOI: 10.1016/j.cpr.2010.05.002.

Qualtrics (2020) Qualtrics. Provo, Utah, USA: Qualtrics. Available at: https://www.qualtrics.com.

Roiser JP and Sahakian BJ (2013) Hot and cold cognition in depression. *CNS spectrums* 18(3): 139–149. DOI: 10.1017/S1092852913000072.

Roiser JP, Elliott R and Sahakian BJ (2012) Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology* 37(1): 117–136. DOI: 10.1038/npp.2011.183.

Setterfield M, Walsh M, Frey AL, et al. (2016) Increased social anhedonia and reduced

helping behaviour in young people with high depressive symptomatology. *Journal of Affective Disorders* 205: 372–377. DOI: 10.1016/j.jad.2016.08.020.

Shrauger JS and Schoeneman TJ (1979) Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin* 86(3): 549. DOI: 10.1037/0033-2909.86.3.549.

Spitzer RL, Kroenke K, Williams JBW, et al. (2006) A Brief Measure for Assessing Generalized Anxiety Disorder. *Archives of Internal Medicine* 166(10): 1092–1097. DOI: 10.1001/archinte.166.10.1092.

The National Institute for Health and Care Excellence (2009) *The NICE Guideline on the Treatment and Depression the Treatment and Management of Depression*.

The National Institute for Health and Care Excellence (2011) *Generalised anxiety disorder and panic disorder in adults: management. Clinical guideline [CG113]*.

The National Institute for Health and Care Excellence (2013) *Social anxiety disorder: recognition, assessment and treatment. Clinical guideline [CG159]*.

Tranter R, Bell D, Gutting P, et al. (2009) The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *Journal of Affective Disorders* 118(1–3): 87–93. DOI: 10.1016/j.jad.2009.01.028.

Uher R, Mors O, Rietschel M, et al. (2011) Early and Delayed Onset of Response to Antidepressants in Individual Trajectories of Change During Treatment of Major Depression. *The Journal of Clinical Psychiatry* 72(11): 1586–1592. DOI: 10.4088/jcp.10m06419.

Vos T, Barber RM, Bell B, et al. (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 386(9995): 743–800.

Walsh AEL, Huneke NTM, Brown R, et al. (2018) A dissociation of the acute effects of bupropion on positive emotional processing and reward processing in healthy volunteers. *Frontiers in Psychiatry* 9(482). DOI: 10.3389/fpsyt.2018.00482.

Wang YP and Gorenstein C (2013) Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Brazilian Journal of Psychiatry* 35(4): 416–431.

DOI: 10.1590/1516-4446-2012-1048.

Warden D, Rush AJ, Trivedi MH, et al. (2007) The STAR*D project results: A comprehensive review of findings. *Current Psychiatry Reports* 9(6): 449–459. DOI: 10.1007/s11920-007-0061-3.

World Health Organization (2017) Depression and Other Common Mental Disorders: Global Health Estimates. *No. WHO/MSD/MER/2017.2*. Genva. Available at: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf.

This declaration concerns the article entitled:			
Is depression associated with reduced optimistic belief updating?			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Hobbs, C., Vozarova, P., Sabharwal, A., Shah, P., Button, K.S. (2022). Is depression associated with reduced optimistic belief updating? *Royal Society Open Science, 9*. doi:10.1098/rsos.190814.		
Copyright status (tick the appropriate statement) I hold the copyright for this material [X] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to the formulation of ideas (70%), design of methodology (70%), and experimental work including creation of study materials (80%), participant recruitment (100%), data collection (100%), data cleaning, analysis, visualisation, and archiving (100%), interpretation of results (75%). The candidate predominantly executed the presentation of the data in journal format (90%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	05/01/2022

Chapter 6: Is depression associated with reduced optimistic belief updating?

Catherine Hobbs[a], Petra Vozarova[a], Aarushi Sabharwal[a], Punit Shah[a], Katherine S Button[a]

[a] Department of Psychology, University of Bath

Chapter Rationale

In this chapter I report an additional paper that moves beyond the main focus of this PhD to examine another potential cognitive mechanism underlying depression; optimistic belief updating. In chapters 1 and 2 I found evidence of a reliable association between self-referential affective biases and depression severity. However, in chapters 4 and 5 I found mixed evidence for the importance of self-referential change in affective learning in antidepressant treatment. Systematic review of literature reported in chapter 3 identified the need for novel measures of psychopharmacological antidepressant effects. In this chapter I therefore investigated a different area of cognition in this thesis – optimistic belief updating. Previous research has indicated that healthy individuals change their beliefs about negative life events more following good versus bad news [1]. However, studies suggest that individuals with depression lack this optimistic belief updating bias [2,3]. In chapter 6 I aimed to replicate this effect and to expand upon previous findings to evaluate belief updating for positive life events.

I chose to focus on optimistic belief updating as (1) it incorporates self-referential affective processing with participants asked to reflect on the likelihood of themselves experiencing positive and negative life events, (2) it has similar mechanisms to social evaluation learning focused on in previous chapters, namely how people learn from novel environmental information, and (3) it may be a potentially more sensitive measure of biases in depression as it incorporates another element of the cognitive theory of depression; Beck's cognitive triad [4]. This refers to theory that individuals experiencing depression tend to have negative views of the self, the world, and the future when negative self-schema are activated. Optimistic belief updating incorporates these concepts by asking individuals to estimate their likelihood (the self) experiencing life events (the world) in their lifetime (the future). Additionally, preliminary research has indicated change in optimistic belief updating following

administration of ketamine [5], which has documented antidepressant effects [6]. Optimistic belief updating may therefore provide a potentially more sensitive measure of cognitive biases in depression and could therefore be a possible mechanism through which antidepressants operate.

Open Science Statement

I conducted this paper in the format of a registered report. The stage 1 protocol (https://osf.io/f2t9p), and study materials (https://osf.io/aqsrb/) are openly available on the Open Science Framework. The data and code that support the findings of stage 2 are openly available in the University of Bath Research Data Archive (https://doi.org/10.15125/BATH-01078).

Abstract

When asked to evaluate their probability of experiencing a negative life event, healthy individuals update their beliefs more following good news than bad. This is referred to as optimistic belief updating. In contrast, individuals with depression update their beliefs by a similar amount, showing reduced optimism. We conducted the first independent replication of this effect and extended this work to examine whether reduced optimistic belief updating in depression also occurs for positive life events. Replicating previous research, healthy and depression groups differed in belief updating for negative events (β=0.71,95%CI:0.24,1.18). Whereas healthy participants updated their beliefs more following good news than bad, individuals experiencing depression lacked this bias. However, our findings for positive events were inconclusive. Whilst we did not find statistical evidence that patterns of belief updating between groups varied by valence (β =-0.51,95% CI:-1.16,0.15), mean update scores suggested that both groups showed largely similar updating for positive life events. Our results add confidence to previous findings that depression is characterised by negative future expectations maintained by reduced updating in response to good news. However, further research is required to understand the specificity of this to negative events, and into refining methods for quantifying belief updating in clinical and non-clinical research.

Introduction

Cognitive neuropsychological models of depression emphasise the role of maladaptive negative beliefs in contributing to the development and maintenance of depression [7]. Individuals experiencing depression are believed to hold pessimistic views of themselves, their future, and the world around them. Information is processed in a manner consistent with these views, resulting in negative cognitive biases [8]. Symptom improvement can be observed when dysfunctional schema and biased cognition are addressed in a therapy, suggesting a potential causal role [9].

However, while depressed individuals' self-beliefs and world views may be more pessimistic relative to healthy controls, evidence is emerging that this may be because healthy controls hold overly optimistic beliefs about themselves and their world views. Thus, rather than being pessimistic per se, depression may be better characterised as a loss of the optimistic outlook seen in healthy individuals. For example, people in the general population are overly optimistic in their judgements of self-performance. In comparison, depression is associated with less optimistic, but more accurate estimates [10]. Positive processing biases may be protective for mental health in increasing self-esteem, confidence, and life satisfaction [11,12]. A loss of these positive biases, sometimes referred to as 'depressive realism', may perpetuate depression symptoms through reinforcing negative views of the self.

Recent work has extended this theory to beliefs about the probability of experiencing future life events. Individuals experiencing high levels of depressive symptoms showed a reduced optimism bias when estimating the probability of experiencing a life event. Participants with high levels of depression were less likely to predict experiencing a positive event and more likely to predict experiencing a negative event [13]. However, this does not explain how these biases are maintained when individuals are presented with information that challenges existing perceptions. To examine this a belief updating task has been developed. In this task participants are asked to estimate their chances of experiencing a negative life event and are then presented with the average probability for someone in a similar sociocultural environment. When asked to re-estimate their chances, healthy individuals have been found to preferentially update their initial beliefs after receiving desirable information (i.e., the probability is lower than their initial estimate), compared to undesirable information (i.e., the probability is greater than their initial estimate). This has been deemed as evidence of

optimistic belief updating in healthy individuals [1,14,15], and posited as a mechanism which serve to maintain positive self-esteem.

In contrast, individuals with depression have not been found to display this asymmetry in belief updating. Korn et al (2014) [3] reported that individuals with major depressive disorder showed no differences in the probability of changing their beliefs following desirable or undesirable information. In contrast, healthy controls continued to preferentially update their beliefs following desirable information. These results were later replicated within the same research group, using functional MRI to examine the brain activity underlying belief updating. Whereas healthy individuals displayed diminished neural coding of undesirable information, individuals with depression showed strong neural coding in response to undesirable information in the right inferior parietal lobule and right inferior frontal gyrus [2]. However, despite these findings being upheld as evidence of reduced optimistic belief updating in depression and being widely cited, they have not yet been independently replicated. Replication will increase our certainty regarding the role of belief updating biases in depression.

Additionally, the validity of findings of optimistic belief updating have been questioned [16–21]. At the time of writing, studies investigating loss of optimistic belief updating in depression have used negative life events only. In this case, optimistic updating would be apparent where beliefs are updated more when the event is *less likely* than initially estimated, than when the event is *more likely* than initially estimated. However, for positive life events, optimistic updating would be apparent where beliefs are updated more when the event is *more likely* than initially estimated, than when the event is *less likely* than initially estimated. If optimistic belief updating biases are indeed present they should be greater in response to desirable information for both positive and negative life events for healthy individuals, although this would result in updates in numerically opposite directions [19].

Although updating of beliefs for positive outcomes have not been investigated in depressed samples, they have been investigated in healthy samples with mixed results. In one study, while participants continued to show biased updates towards desirable information for negative life events, the opposite effect occurred for positive life events. Participants were more likely to update their beliefs following undesirable information for positive life events, therefore displaying a pessimism bias [19]. However, this work has been criticised on the

basis of including life events with a high probability of occurring (e.g., 'Going to your favourite restaurant'), and using simulated probabilities. Participants may therefore have been aware that the presented average probabilities were inaccurate, influencing their responses [22]. When using these events with actual probabilities of occurring within the next month, participants were found to preferentially update beliefs in response to desirable information for both positive and negative events [22]. However, another study observed optimistic biased updating only for negative events, no bias was observed for positive life events [23]. It is therefore unclear whether optimistic belief updating is present in healthy individuals for both positive and negative life events.

Optimistic belief updating for positive events has not yet been explored in relation to depression symptoms. It has been argued that depression is characterised by a blunting of emotional response to both positive and negative stimuli [24]. However, others have suggested information processing is relatively intact for negative information and only aberrant for positive information [25–27]. Research examining future belief updating to date suggests an emotional blunting in depression, where participants were not influenced by the desirability of estimates (i.e., whether it was good or bad news for them). However, it is possible that this is specific to negative life events, and effects may differ according to desirability for positive life events. Incorporating positive life events into this task would allow us to examine whether this is a general bias apparent across events of different valences, or whether this may be dependent on the information presented. This would improve our understanding of the mechanisms of reduced optimistic belief updating in depression and may have implications for clinical strategies to remediate reduced optimism.

In this study we aim to replicate findings of Korn et al (2014) [3] indicating reduced optimistic belief updating in individuals with depression. Specifically, we aim to replicate their finding of preferential updates in beliefs of experiencing negative life events following desirable versus undesirable information in healthy controls, but no differences in individuals experiencing depression. This would be the first independent replication of this effect. Furthermore, we aim to build upon previous results by incorporating positive life events into this paradigm. This will provide us with a greater understanding of the precise mechanisms underlying reduced optimistically biased belief updating previously observed in depression.

Hypotheses

H1 In line with Korn et al (2014) [3] our primary hypothesis was that healthy controls would show optimistically biased belief updating for negative life events, as indicated by a greater change in beliefs following desirable versus undesirable information for negative life events. In contrast, we hypothesised that individuals with depression would show a reduced optimistic belief updating for negative life events, as indicated by smaller differences in change in beliefs following desirable versus undesirable information for negative life events.

H2 We hypothesised that this effect would be consistent across both positive and negative life events. That is, healthy controls would show greater changes in beliefs following desirable versus undesirable information irrespective of the valence of the life event. Individuals with depression would show smaller differences in change in beliefs following desirable versus undesirable information for both positive and negative life events. However, given the previous mixed findings regarding biased belief updating for positive life events, and no current literature for individuals with depression, we acknowledged that our belief in this hypothesis is weak.

H3 In line with Korn et al (2014) [3] we predicted that depressed individuals would on average initially rate their chances of experiencing a negative life event as higher compared to healthy controls.

H4 We extended this hypothesis to initial estimates of positive life events, such that depressed individuals would on average initially rate their chances of experiencing a positive life event as lower compared to healthy controls.

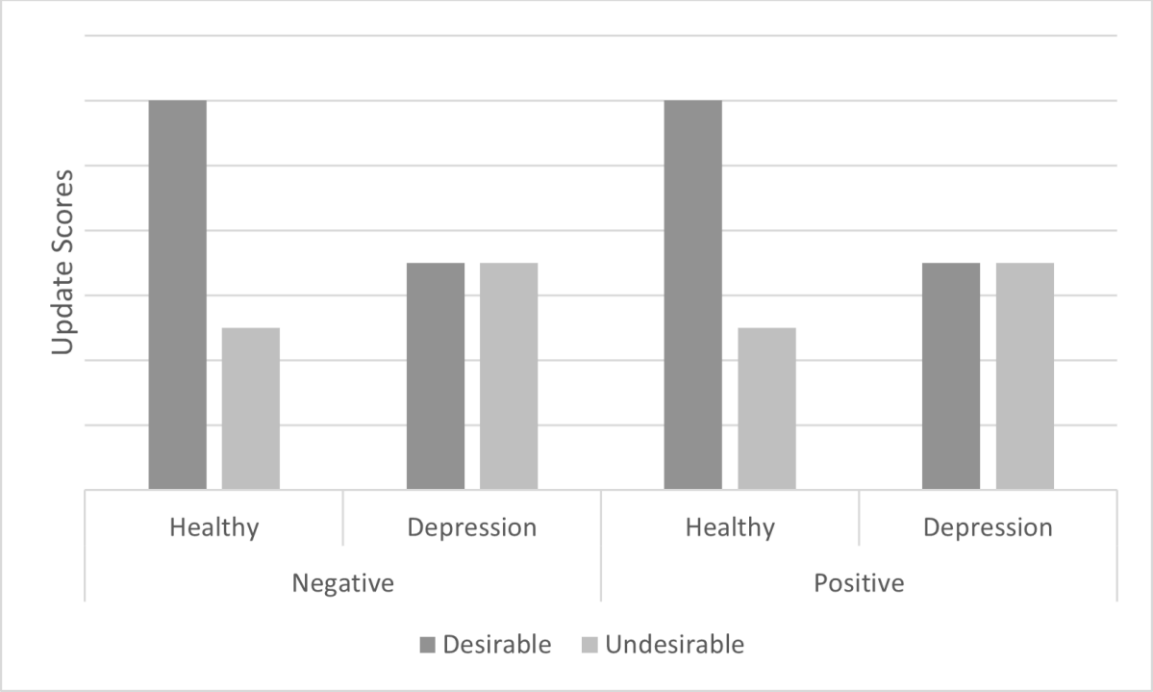


Figure 1

Hypothesised differences between absolute mean updates in the likelihood of life events according to desirability and valence of life event, for healthy controls and participants with moderate to severe levels of depression

Methods

Study Design

A mixed-model design with group as a between-subject factor (depressed versus healthy controls), and desirability of life events (undesirable versus desirable) and valence (positive versus negative) as within-subject effects, was used.

Sampling Plan

Recruitment

Due to the COVID-19 pandemic, recruitment took place online using the participant recruitment platform 'Prolific' [28]. Screening surveys were completed online using 'Qualtrics' [29]. Two groups of participants were screened and recruited on the basis of Patient Health Questionnaire (PHQ-9) scores according to recommended clinical guidelines [30]. A moderately severe to severely depressed sample was recruited with PHQ-9 scores ≥ 15. A healthy control sample was recruited with PHQ-9 scores ≤ 4. Participants were only invited to take part in the test phase of the study if their PHQ-9 scores were within these ranges. To optimise recruitment of participants with high levels of depression we restricted advertisement of this initial survey to participants that have reported experiencing a current diagnosed mental health condition, using custom pre-screening on Prolific. A separate survey with the same procedure was used to recruit healthy controls without this restriction.

To reduce participant burden during online data collection, we did not conduct diagnostic interviews. However, based on previous research within this group [31], we expect 93% of individuals with PHQ-9 scores ≥ 15 to meet diagnostic criteria for a major depressive episode. In combination with targeted recruitment of patients reporting a current mental health condition on Prolific, we anticipated that our sample would have similar clinical profiles to those of previous research within this field [2,3].

Inclusion Criteria

Participants in both groups were aged 18 and over, fluent in written and spoken English, with normal or corrected to normal vision, and were current residents of the United Kingdom. Participants were included irrespective of any comorbid psychiatric disorders, or current treatment for depression. To ensure high quality of data, custom pre-screening was used on

Prolific to identify participants that had previously completed ≥ 5 studies with a 98% acceptance rate.

Exclusion Criteria

As outlined above, only participants with PHQ-9 scores ≤ 4 or ≥ 15 were invited to participate. In keeping with Korn et al (2014) [3], participants who reported a history of, or current substance abuse, were excluded.

Sample Size and Rationale

Korn et al (2014) [3] reported an interaction between desirability (desirable/undesirable) and group (MDD/healthy) of $F(1,35) = 6.9$, $p = 0.013$, $\eta_p^2 = 0.17$. To account for potential publication bias we reduced this effect size by one-third, $\eta_p^2 = 0.113$. As the correlation between within-subject effects has not been reported in previous research we took a conservative approach of assuming no correlation. Sample size calculation performed in G*Power [32] for a within-between interaction using ANOVA indicated that 108 participants would be required to detect an effect of $\eta_p^2 = 0.113$ with 95% power at an alpha level of 0.005. We therefore aimed to recruit two groups of 54 participants.

Blinding

Blinding was not applied as testing took take place online, eliminating the possibility of experimenter effects.

Materials

Testing took place online; self-report measures were completed using 'Qualtrics' [29] and the Belief Updating Task was completed using 'Inquisit' [33].

Self-Report Measures

Prior to completing the task participants completed a number of self-report measures as outlined below.

Depression

The BDI-II [8] was used to measure depression symptoms. The BDI-II is a 21 item self-report questionnaire of depression symptoms. Possible responses for each item range from (0), indicating low levels, to (3) indicating high levels. Higher scores indicate higher levels of depression symptoms.

The PHQ-9 [30] was used as an additional measure of depression symptoms, as this is widely used in UK clinical practice and is a more appropriate depression screening measure due to its brevity [34]. The PHQ-9 is a 9-item self-report questionnaire of depression symptoms. Possible responses range from (0) 'Not at all' to (3) 'Nearly every day'. Again, higher scores indicate higher levels of depression symptoms. The PHQ-9 and BDI-II demonstrate strong correlation in general and clinical populations [35,36].

Anxiety

General anxiety was measured using the Generalised Anxiety Disorder Questionnaire (GAD-7) [37]. The GAD-7 is a 7 item questionnaire of anxiety symptoms, with responses ranging from (0) 'Not at all' to (3) 'Nearly every day'. Higher scores indicate higher levels of anxiety symptoms.

Social anxiety was measured using the Brief Fear of Negative Evaluation Scale (BFNE) [38]. This is a 12-item self-report measure of fear of negative evaluation by others.

Trait Optimism

Trait levels of optimism and pessimism were measured using the Life Orientation Test-Revised (LOT-R) [39]. Participants were asked to indicate their agreement with 10 statements, ranging from (1) 'I agree a lot', to (5) 'I disagree a lot'.

Positive and Negative Mood

The Positive and Negative Affect Scale [40] (PANAS) was used to measure state mood before and after completion of the task. Participants were asked to indicate to what extent they currently experienced ten positive and ten negative emotions, ranging from (1) 'very slightly or not at all' to (5) 'extremely'.

Demographics and Clinical Characteristics

Information was collected on age, gender, ethnicity, relationship status, employment status, highest qualification and living situation. Participants were asked if they have previously experienced depression, whether they have received any treatment for depression (both past and current), and whether a family member has experienced depression.

Belief Updating Task

Following completion of the self-report measures, participants were asked to complete a computerised task measuring beliefs of experiencing a number of life events. A task was created on Inquisit [33], replicating the methods outlined in Korn et al (2014) [3] but with the addition of positive life events. This task has been made publicly available on the Open Science Framework (https://osf.io/aqsrb/).

Stimuli

Task stimuli consisted of short descriptions of 40 negative and 40 positive life events and the average probability of someone in the same environment as the participant experiencing these events in their lifetime. Participants completed two training trials for positive and negative life events, resulting in 42 negative and 42 positive life events overall. In keeping with previous research[11] very common (greater than 70%) and very rare (less than 10%) events were not included. This allowed for change between initial and re-estimates. To allow for variability in ratings for high and low probability life events, participants were told that event probabilities fall within 3% and 77%. Of the 72 negative life events used by Korn et al [3], we used 42 of the same negative life events, although we calculated new estimates to account for possible changes in available statistics for probabilities following the publication of Korn et al (2014) [3] (Supplementary Table 1). We were unable to use 30 of the life events used previously, as the event probabilities were very rare (less than 10%), very common (greater than 70%), or data for lifetime probability was not available (Supplementary Table 2). In addition, we identified estimates for 42 positive life events (Supplementary Table 3). Estimates were derived from reputable resources including peer reviewed academic journal articles, governmental resources (e.g., the Office for National Statistics), non-governmental organisation resources, and marketing and consumer data companies.

Task Procedure

For each life event, participants were presented with a short description of the life event for four seconds and asked to imagine this event happening/not happening to them. Following this, participants were asked to estimate the probability of this event happening/not happening to them in their lifetime. A response threshold of 10 seconds was imposed. A fixation cross was then displayed for 1.2 seconds, before participants were presented with the average probability of this event happening/not happening to someone like them for three seconds. A fixation point was again displayed for 1.2 seconds before this sequence was repeated for the remaining life events (figure 2A).

Participants were then asked to re-estimate the probability of experiencing/not experiencing the events that were presented previously. A short description of the event was displayed for 4 seconds, before participants were again asked to estimate their probability of experiencing/not experiencing the event in their lifetime. Again, a response threshold of 10 seconds was imposed. A fixation cross was then displayed for 1.2 seconds, before this sequence was repeated for the remaining life events (figure 2B). Participants also completed two training trials to ensure that they understand the task instructions.

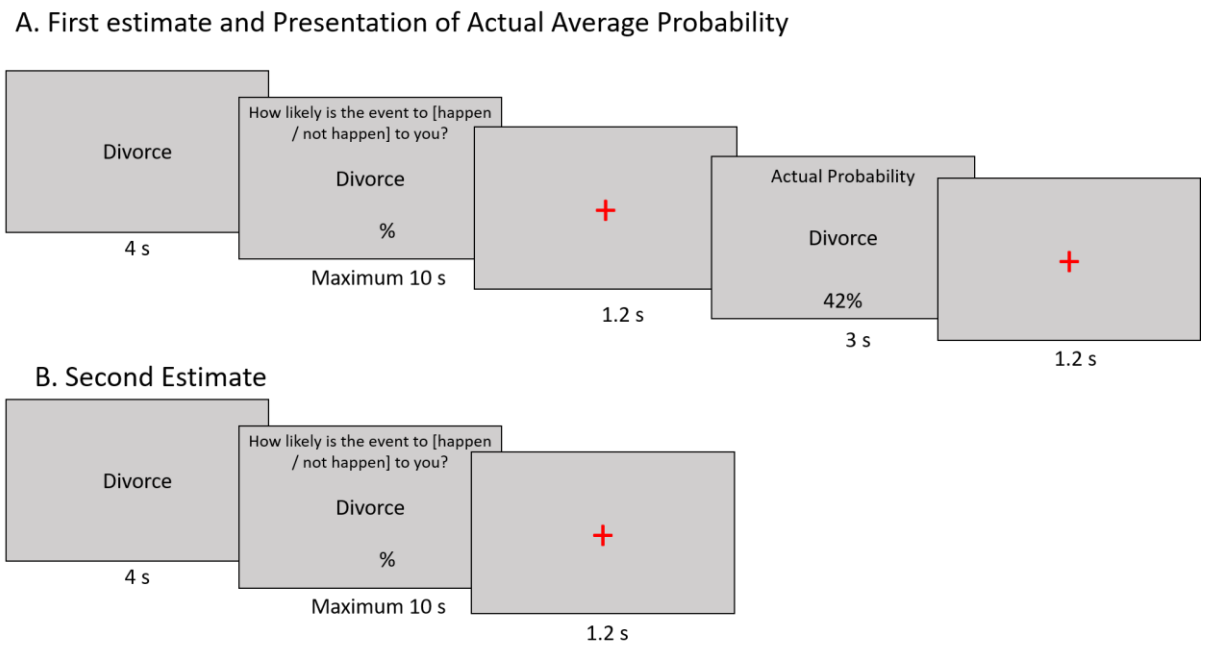


Figure 2

Main Task Procedure

Participant's memory of the average probability of life events was tested following the completion of probability estimations. Participants were asked to recall the average probability of each event happening for someone like them as presented previously (i.e., the actual probability). Subjective ratings for each life event were then collected. Participants were asked to rate each event on seven scales: vividness, familiarity, prior experience, emotional arousal, negativity, positivity, and controllability. Each scale was rated from 1, indicating low levels, to 6, indicating high levels.

Separate blocks were completed for negative and positive life events, with order of completion counterbalanced (i.e., participants completed estimations, recall and ratings for all events of one valence, before repeating each procedure for all events of the remaining valence). This ensures that a pure replication of the procedure used by Korn et al (2014) [3] was available for the subgroup of participants who completed the negative block first. Within each of the positive and negative life events estimation blocks, half of the events were framed as 'happening' whereas the other half were framed as 'not happening'. More specifically, in half of the blocks participants were asked to estimate the probability of the event happening to them in their life time and were presented with the average probability of the event occurring. In the other half of the blocks, participants were asked to estimate the probability of the event not happening to them in their lifetime and were presented with the average probability of the event not occurring. Life events were therefore split into four lists of 20 events each (2 lists of negative events, and 2 list of positive events). Average probabilities did not significantly differ between lists. The order of happening versus not happening, and list completion order, was counterbalanced between participants (Supplementary Tables 4 and 5).

Debriefing

Although presented probabilities originate from validated sources, participants' own perceptions may differ, contributing to scepticism regarding the credibility of probabilities. To account for this, participant's perceptions of the credibility of the presented probabilities for life events were measured through self-report following completion of the belief updating task. Participants were asked to indicate how strongly they agree with the item 'The average probabilities presented in the task were accurate', on a five point scale ranging from 'Strongly Agree' to 'Strongly Disagree'.

Task Outcomes

To allow us to combine estimates between 'not happening' and 'happening', participants' estimates in the 'not happening' blocks were subtracted from 100 to calculate the equivalent probability of the event happening.

Estimation errors were calculated by subtracting the actual probability of events from the participants' initial estimate of the events (i.e. estimation error difference = initial estimate – actual probability). Overestimations are therefore represented by positive values, whereas underestimations are represented by negative values.

Participant trials for each life event were categorised according to whether the actual probability presented is desirable or undesirable. For negative life events, trials were deemed as desirable if the actual probability was lower than the participant's initial estimate (i.e., the participant initially overestimated their probability of experiencing the life event, a positive estimation error). For negative life events, trials were deemed as undesirable if the actual probability was greater than the participant's initial estimate (i.e., the participant underestimated their probability of experiencing the life event, a negative estimation error).

For positive life events, trials were deemed as desirable if the actual probability was greater than the participant's initial estimate (i.e., the participant initially underestimated their probability, a negative estimation error). Trials were deemed as undesirable if the actual probability was lower than the participants initial estimate (i.e., the participant initially overestimated their probability, a positive estimation error).

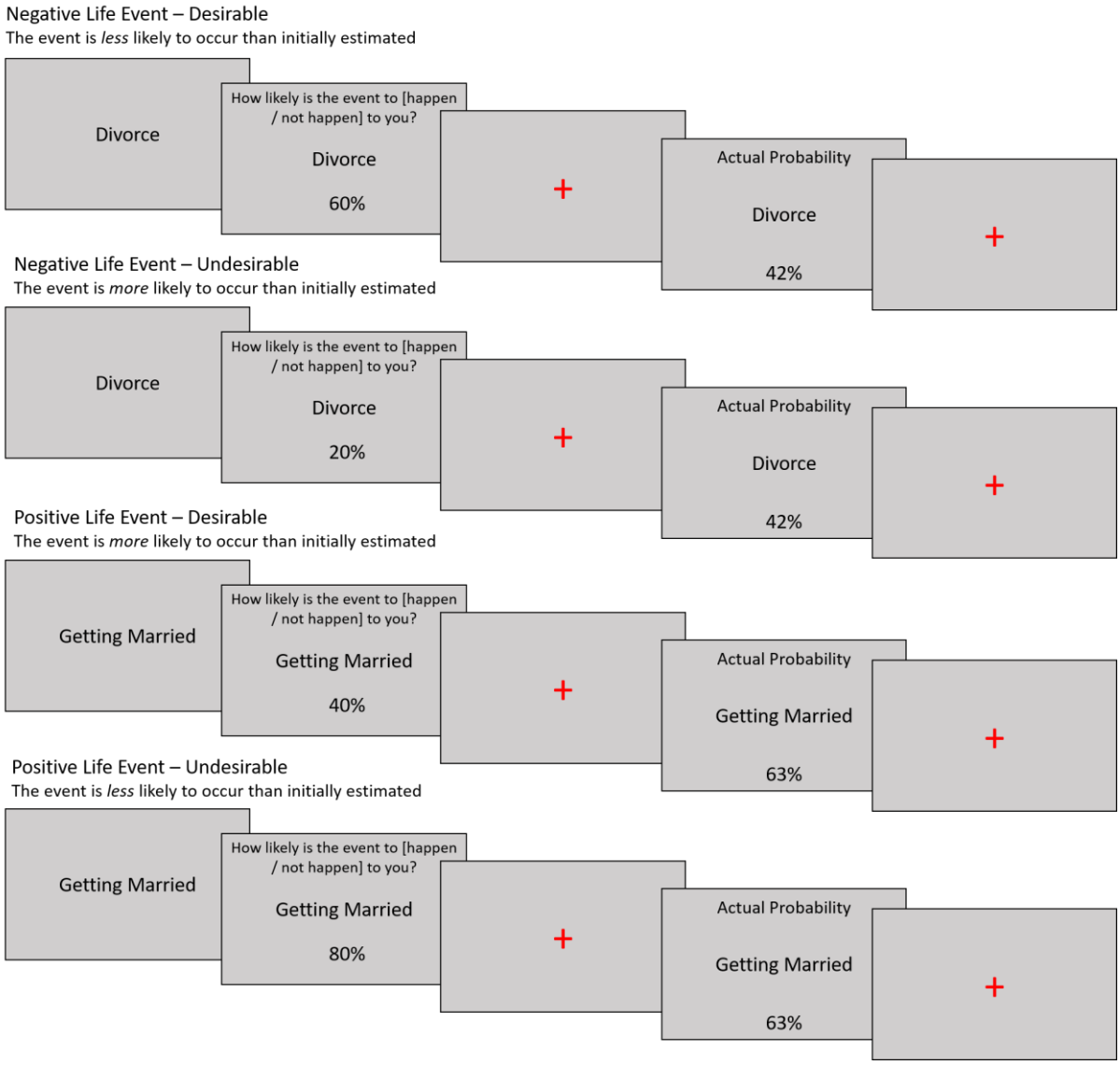


Figure 3

Example of desirable versus undesirable categorisations in positive and negative life events. For negative life events initial overestimations are desirable, whereas initial underestimations are undesirable. For positive life events initial underestimations are desirable, whereas initial overestimations are undesirable.

The primary outcome of this task is the extent to which participants change their initial estimates, following information regarding the average probability of experiencing the life event (continuous variable). Optimistic biased updating is indicated by greater changes towards desirable information relative to undesirable information. To calculate update scores, re-estimates (i.e., estimates made *after* presentation of the actual probability) were subtracted from initial estimates (i.e., update = initial estimate – re-estimate). These values were then coded as positive when the update (the re-estimate) moved towards the actual probability, and negative when the update moved away from the actual probability. Positive scores therefore indicate movement towards the actual probability, whereas negative scores indicate movement away from the actual probability (irrespective of valence and desirability).

Memory errors were calculated by subtracting the participants' recalled average probability from the actual probability of the event (i.e., memory error = actual probability – participants' recollection of probability).

Quality Check and Exclusion Criteria

Participants were not exposed to experimental manipulation within this study and were presented with the average probability of life events based on validated statistics. A positive control does therefore not apply to this study. However, to ensure data quality, following the procedure of Korn et al (2014) [3] trials were excluded if participants did not respond within the maximum response limit of 10 seconds. Responses were restricted to within 3% and 77% for the probability of the event happening, and within 23% and 97% for the event not happening, to reflect the range of probabilities specified to participants at the beginning of the task. Trials where the estimation error is 0 (i.e., the initial estimate is the same as the actual probability) were excluded from analyses including desirability, as this cannot be categorised. However, these trials were retained in analyses where calculation of desirability was not required.

The inclusion of 'happening' and 'not happening' trials controlled for any directional effects of any participant not appropriately engaged in the task. For example, selecting responses at random or selecting the same response will add noise (making our analyses more conservative), but is unlikely to introduce directional bias into the estimates. Therefore, all data from participants completing the task were included in the analyses.

As outlined in the inclusion criteria, to ensure data quality participants were restricted to those that have previously completed ≥ 5 studies with a 98% acceptance rate on Prolific. Attention checks were also placed throughout the self-report measures and the Belief Updating task, in the form of 8 items directing participants to provide a specified response (e.g. 'Please select 'Disagree''). Participants that failed any of these attention checks were considered non-compliant and excluded (Figure 4).

Analysis Plan

The data and code that support the findings of this study are openly available in the University of Bath Research Data Archive (https://doi.org/10.15125/BATH-01078) [41].

Inference Criteria

A frequentist approach was taken using a stringent alpha level of 0.005 [42].

Primary Analyses

Hypothesis 1: Healthy controls will show a greater change in beliefs following desirable versus undesirable information for negative life events, whereas individuals with depression will show a smaller difference in change in beliefs following desirable versus undesirable information for negative life events.

We used a mixed-effects linear regression model using only trials for negative life events. Absolute mean update scores were entered as the outcome, and group (healthy versus depressed), desirability (desirable versus undesirable), and an interaction term between group and desirability entered as predictors. Subject was entered as a random effect to account for the repeated measures design.

ANOVA and linear regression are equivalent statistical models, both following general linear modelling [43]. Our mixed-effects linear regression model therefore corresponds to the mixed-design ANOVA used by Korn et al (2014) [3], but with the additional interpretational advantages provided by the regression coefficients.

Hypothesis 2: The effect of optimism bias will be consistent across both positive and negative life events. Healthy controls will show a greater change in beliefs following desirable versus undesirable information irrespective of the valence of the life event. Individuals with depression will show smaller differences in change in beliefs following desirable versus undesirable information for both positive and negative life events.

We conducted a mixed-effects linear regression model using trials for both positive and negative life events. Absolute mean update scores were entered as the outcome, and group (healthy versus depressed), desirability (desirable versus undesirable), and valence (positive versus negative) entered as predictors. Additionally, interaction terms between group and desirability, group and valence, desirability and valence, and group, desirability and valence were entered as additional predictors. Subject was entered as a random effect to account for the repeated measures design.

Hypothesis 3: Initial estimates of negative life events will be associated with depression, such that depressed individuals will on average initially rate their chances of experiencing a negative life event as higher compared to healthy controls.

This hypothesis was tested using a linear regression using only trials for negative life events. Initial scores were entered as the outcome, and group (healthy versus depressed) was the predictor.

Hypothesis 4: Initial estimates of positive life events will be associated with depression, such that depressed individuals will on average initially rate their chances of experiencing a positive life event as lower compared to healthy controls.

This hypothesis was tested using a linear regression model using only trials for positive life events. Initial scores were entered as the outcome, and group (healthy versus depressed) was the predictor.

Additional Analyses

Following the procedure of Korn et al (2014) [3], we investigated whether any observed differences in update scores between depressed individuals and healthy controls are

attributable to estimation errors (e.g., the difference between initial estimates and presented average probability). Scaled absolute mean update scores will be calculated by dividing mean update scores by absolute mean estimation errors. A mixed-effects linear regression model was created using only trials for negative life events. Scaled absolute mean update scores was entered as the outcome, with group (healthy versus depressed), desirability (desirable versus undesirable), and an interaction term between group and desirability entered as predictors. Subject was entered as a random effect to account for the repeated measures design.

This model was also repeated using both trials for positive and negative events, with valence, and an interaction term between valence and desirability, valence and group, and valence, desirability and group added as additional predictors.

To examine whether there was a dose-response relationship between depression symptoms and optimism bias, the linear regression models outlined in hypothesis 1 and 2 were repeated with continuous BDI-II scores entered as a predictor individually and in interaction with desirability. Group was entered as an additional predictor as a design variable. These models were repeated for PHQ-9 scores.

To assess whether any observed differences in updates according to group are attributable to trait levels of optimism the regression models outlined for hypothesis 1 and 2 were repeated with LOT-R scores as an additional predictor.

Differences in absolute mean memory errors were examined using a mixed-effect linear regression model. Absolute mean memory errors were entered as the outcome, with group (healthy versus depressed), desirability (desirable versus undesirable) and valence (positive versus negative) entered as predictors. Additionally, interaction terms between group and desirability, group and valence, desirability and valence, and group, desirability and valence were entered as additional predictors. Subject was entered as a random effect. The linear regression models outlined in hypotheses 1 and 2 were repeated with memory estimates as a predictor.

To assess whether observed results in the primary analyses were influenced by the framing of life events (happening versus not happening), these models were repeated with framing entered as an additional predictor. Likewise, to assess whether the observed results were

influenced by order effects of the valence of life events, order of completion for positive and negative blocks was entered as an additional predictor to these models.

To examine whether the observed results in the primary analyses were influenced by participants' ratings of vividness, familiarity, prior experience, emotional arousal, negativity, positivity, and controllability, the models specified in the primary analyses section were repeated with ratings added as additional predictors in separate models. If the results of the primary analyses are similar following adjustment for participants' ratings, this suggests that variations in perceptions of life events between participants did not significantly influence our findings regarding belief updating. We also conducted sensitivity analyses, repeating the primary analyses but with positive and negative events classified according to each participant's negativity and positivity ratings, to see how robust our findings were to any discrepancies in classification of positive/negative events. To classify life events according to participants' ratings, we aggregated positivity and negativity ratings by calculating the mean of negativity ratings and reverse coded positivity ratings. We then categorised the life event as being perceived as positive if the value was in the lower half of the scale (< 3.5), neutral if the event was at the midpoint of the scale (3.5), and negative if the value was in the upper half of the scale (> 3.5).

Dependent on evidence from our regression models that the order in which participants complete negative and positive events blocks has a strong influence on the results, we planned to repeat hypotheses 1 and 3 in the half of participants who completed the negative events first, to provide an uncontaminated replication of Korn et al's (2014) [3] finding. This sample (n = 27 per group) would have provided 86% power to detect an effect of $\eta_p^2 = 0.113$ at an alpha level of 0.05. However, as we found no evidence of order effects, we did not conduct this analysis.

To account for potential variations in perceptions of the credibility of probabilities presented for life events, we conducted a sensitivity analysis replicating the primary analyses with data for participant's who indicated scepticism excluded.

Exploratory Analyses

As participants' age varied by group, we conducted exploratory analyses to examine whether our findings were influenced by age by including it as an additional predictor in the models for our primary analyses outlined above.

Results

Sample

The study sample consisted of 56 participants in the healthy group and 54 participants in the depression group. Full recruitment information is available in Figure 4, sample demographics are outlined in Table 1, and clinical characteristics are provided in Table 2. There was some variation in demographics according to group. Participants experiencing depression were on average younger, a greater proportion were female, they were less likely to be employed or a homeowner and were more likely to not be in a relationship.

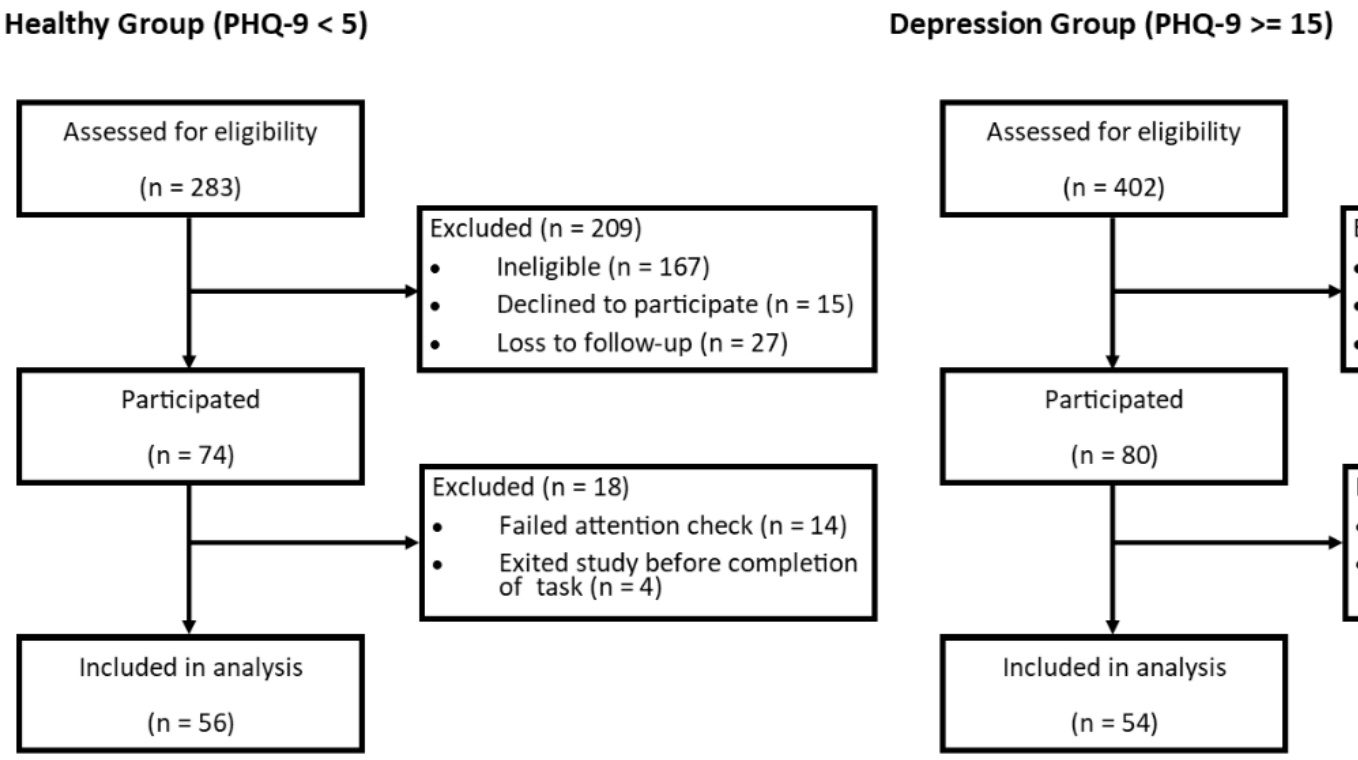


Figure 4

Consort flow diagram for the Healthy and Depression groups

Table 1

Sample demographics by group

	Healthy	**Depression**
N	56	54
Age, M (SD)	37.36 (14.56)	32.56 (12.17)
Gender, N (%)		
Male	19 (33.9)	8 (14.8)
Female	36 (64.3)	43 (79.6)
Non-Binary	1 (1.8)	3 (5.6)
Ethnicity, N (col%)		
White	47 (83.9)	46 (85.2)
Black	0 (0.0)	2 (3.7)
Asian	7 (12.5)	2 (3.7)
Mixed	2 (3.6)	4 (7.4)
Occupation, N (%)		
Employed	36 (64.3)	22 (40.7)
Student	8 (14.3)	10 (18.5)
Unemployed	6 (10.7)	14 (25.9)
Other	6 (10.7)	8 (14.8)
Educational Attainment, N (%)		
Primary Education	0 (0.0)	3 (5.6)
Secondary Education	18 (32.1)	24 (44.4)
Degree	18 (32.1)	18 (33.3)
Higher Degree	10 (35.7)	9 (16.7)
Relationship, N (%)		
Single	13 (23.2)	24 (44.4)
In a relationship	42 (75.0)	29 (53.7)
Other	1 (1.8)	1 (1.9)
Living Situation, N (%)		
Homeowner	30 (53.6)	13 (24.1)
Renting	20 (35.7)	23 (42.6)
Living with a relative/friend	6 (10.7)	18 (33.3)

Note: Percentages represent the proportion of participants within each group

Table 2

Sample Psychiatric Characteristics

	Healthy	**Depression**
N	56	54
PHQ-9, M (SD)		
Screening	2.27 (1.34)	18.69 (2.87)
Testing	2.48 (2.13)	17.41 (3.76)
BDI-II, M (SD)	6.36 (5.65)	38.11 (9.21)
GAD-7, M (SD)	1.84 (2.19)	13.35 (4.08)
BFNE, M (SD)	33.70 (9.88)	50.02 (9.25)
LOT-R, M (SD)	16.09 (5.40)	4.56 (4.22)
PANAS Change, M (SD)		
Positive	-1.78 (5.40)	-0.96 (3.71)
Negative	-1.64 (2.68)	-1.94 (4.18)
Self-reported depression, N (%)		
Current	0 (0%)	46 (85.2%)
Previous	21 (38%)	8 (14.8%)
Current Treatment, N (%)		
Psychological Therapy	0 (0%)	12 (22.2%)
Antidepressants	3 (5.4%)	22 (40.7%)
Family history of depression, N (%)	30 (53.6%)	41 (75.9%)

Note: Percentages represent the proportion of participants within each group. PANAS Change scores represent change in scores before and after completion of the belief updating task.

PHQ-9 = Patient Health Questionnaire, BDI-II = Beck Depression Inventory II, GAD-7 = Generalised Anxiety Disorder Questionnaire, BFNE = Brief Fear of Negative Events Scale, LOT-R = Life Orientation Test- Revised, PANAS = Positive and Negative Affect Scale

Table 3

Mean (standard deviation) belief updating task outcomes by group, valence of life events and desirability.

	Negative				Positive			
	Healthy (n = 56)		Depression (n = 54)		Healthy (n = 56)		Depression (n = 54)	
	Desirable	Undesirable	Desirable	Undesirable	Desirable	Undesirable	Desirable	Undesirable
Total Eligible Trials	22.12 (4.30)	16.52 (4.39)	26.20 (3.72)	12.06 (3.57)	12.29 (5.38)	26.02 (5.49)	17.43 (6.10)	20.78 (6.07)
Ineligible Trials								
Estimation Error of Zero	1.44 (0.69)		1.47 (0.70)		1.47 (0.73)		1.76 (0.91)	
Missing Estimate	1.50 (0.97)		2.14 (4.00)		1.53 (1.17)		1.20 (0.41)	
Outside Specified Range	1.29 (0.76)		1.00 (0.00)		1.57 (1.40)		1.20 (0.42)	
Initial Estimates	35.84 (6.68)		44.63 (7.44)		40.50 (9.30)		33.42 (9.45)	
Estimation Errors	24.13 (5.00)	-22.42 (3.88)	29.37 (5.51)	-20.64 (4.71)	-13.78 (3.61)	30.74 (7.93)	-15.31 (3.04)	29.50 (7.47)
Absolute Estimation Errors	24.13 (5.00)	22.42 (3.88)	29.37 (5.51)	20.64 (4.71)	13.78 (3.61)	30.74 (7.93)	15.31 (3.04)	29.50 (7.47)
Update Scores	12.49 (7.76)	7.98 (6.72)	10.50 (7.30)	11.34 (7.80)	7.82 (5.50)	7.82 (7.70)	8.75 (9.38)	10.26 (6.76)
Scaled Update Scores	0.52 (0.30)	0.36 (0.30)	0.38 (0.28)	0.54 (0.31)	0.58 (0.42)	0.30 (0.31)	0.60 (0.60)	0.38 (0.26)
Absolute Memory Errors	14.91 (6.44)	14.96 (5.33)	18.41 (6.62)	16.52 (6.05)	13.90 (5.96)	16.92 (8.04)	14.48 (5.79)	17.44 (8.98)
Ratings								
Controllability	2.49 (0.64)	2.58 (0.73)	2.35 (0.53)	2.33 (0.71)	3.53 (0.90)	4.11 (0.74)	3.11 (0.82)	3.69 (0.76)
Emotional Arousal	3.06 (1.09)	3.06 (1.08)	3.64 (0.92)	3.37 (1.05)	3.18 (0.94)	3.78 (1.04)	3.34 (1.03)	3.78 (0.89)
Familiarity	2.90 (0.85)	2.60 (0.93)	3.24 (0.82)	2.72 (1.00)	2.69 (0.94)	3.78 (0.77)	2.41 (0.96)	3.44 (0.97)

Negativity	4.97 (0.82)	4.98 (0.83)	4.89 (0.65)	4.87 (0.90)	1.94 (0.69)	1.51 (0.41)	2.19 (0.94)	1.80 (0.57)
Positivity	1.29 (0.30)	1.24 (0.38)	1.37 (0.45)	1.33 (0.55)	4.25 (0.91)	4.91 (0.64)	3.82 (1.04)	4.57 (0.70)
Prior Experience	2.02 (0.61)	1.77 (0.65)	2.50 (0.73)	2.08 (0.84)	2.06 (0.86)	3.26 (0.86)	1.67 (0.74)	2.84 (0.95)
Vividness	3.20 (0.93)	2.95 (1.07)	3.77 (0.88)	3.51 (0.85)	3.19 (1.00)	4.14 (0.78)	3.03 (0.92)	3.96 (0.85)

Hypothesis 1: *Healthy controls will show a greater change in beliefs following desirable versus undesirable information for negative life events, whereas individuals with depression will show a smaller difference in change in beliefs following desirable versus undesirable information for negative life events.*

In support of our hypothesis, we found evidence that the relationship between update scores and group varied according to the desirability of presented information (β = 0.71, β 95% CI: 0.24, 1.18, p = 0.004). Participants in the healthy comparison group displayed optimistic belief updating, changing their beliefs to a greater extent following desirable (M 12.49, SD 7.76) versus undesirable information (M 7.98 SD 6.72; t(108) = 3.54, p < .001). In comparison, individuals experiencing depression on average updated their scores to a largely similar extent following desirable or undesirable information (Desirable: M 10.50, SD 7.30, Undesirable: M 11.34 SD 7.80, t(108) = -0.65, p = 0.518, Figure 5).

Hypothesis 2: *The effect of optimism bias will be consistent across both positive and negative life events. Healthy controls will show a greater change in beliefs following desirable versus undesirable information irrespective of the valence of the life event. Individuals with depression will show smaller differences in change in beliefs following desirable versus undesirable information for both positive and negative life events.*

When analysing both positive and negative life events, we again found evidence of optimistic belief updating in healthy individuals (Group x Desirability: β = 0.71, 95% CI = 0.24, 1.17, p = 0.003). Participants in the healthy group updated their beliefs more towards desirable (M 10.16 SD 7.09) versus undesirable information (M 7.90 SD 7.20; t(324) = 2.55, p = 0.011), whereas no difference was observed in the depression group (Desirable: M 9.62 SD 8.41, Undesirable M 10.80 SD 7.28; t(324) = -1.31, p = 0.192). We did not find statistical evidence that this pattern differed between positive and negative life events (Group x Desirability x Valence: β = - 0.51, β 95% CI = -1.16, 0.15, p = 0.130). However, whilst the depression group showed largely similar updating behaviour for positive (Desirable: M 8.75 SD 9.38, Undesirable M 10.26 SD 6.76; t(324) = -1.19, p = 0.235) and negative (Desirable: M 10.50 SD 7.30, Undesirable M 11.34 SD 7.80; t(324) = -0.66, p = 0.510) life events, a differential pattern was observed in the healthy group. Participants in the healthy group showed optimistic belief updating for negative life events (Desirable M 12.49 SD 7.76, Undesirable

M 7.98 SD 6.72, $t(324) = 3.60$, $p < .001$), but for positive life events showed similar updating towards desirable (M 7.82 SD 5.50) and undesirable (M 7.82 SD 7.70) information ($t(324) = 0.00$, $p = 0.998$). In combination with the wide confidence intervals, we cannot rule out the presence of a Group x Valence x Desirability interaction smaller than we have statistical power to detect.

Full descriptive statistics are provided in Table 3 and analytical results are available in Table 4.

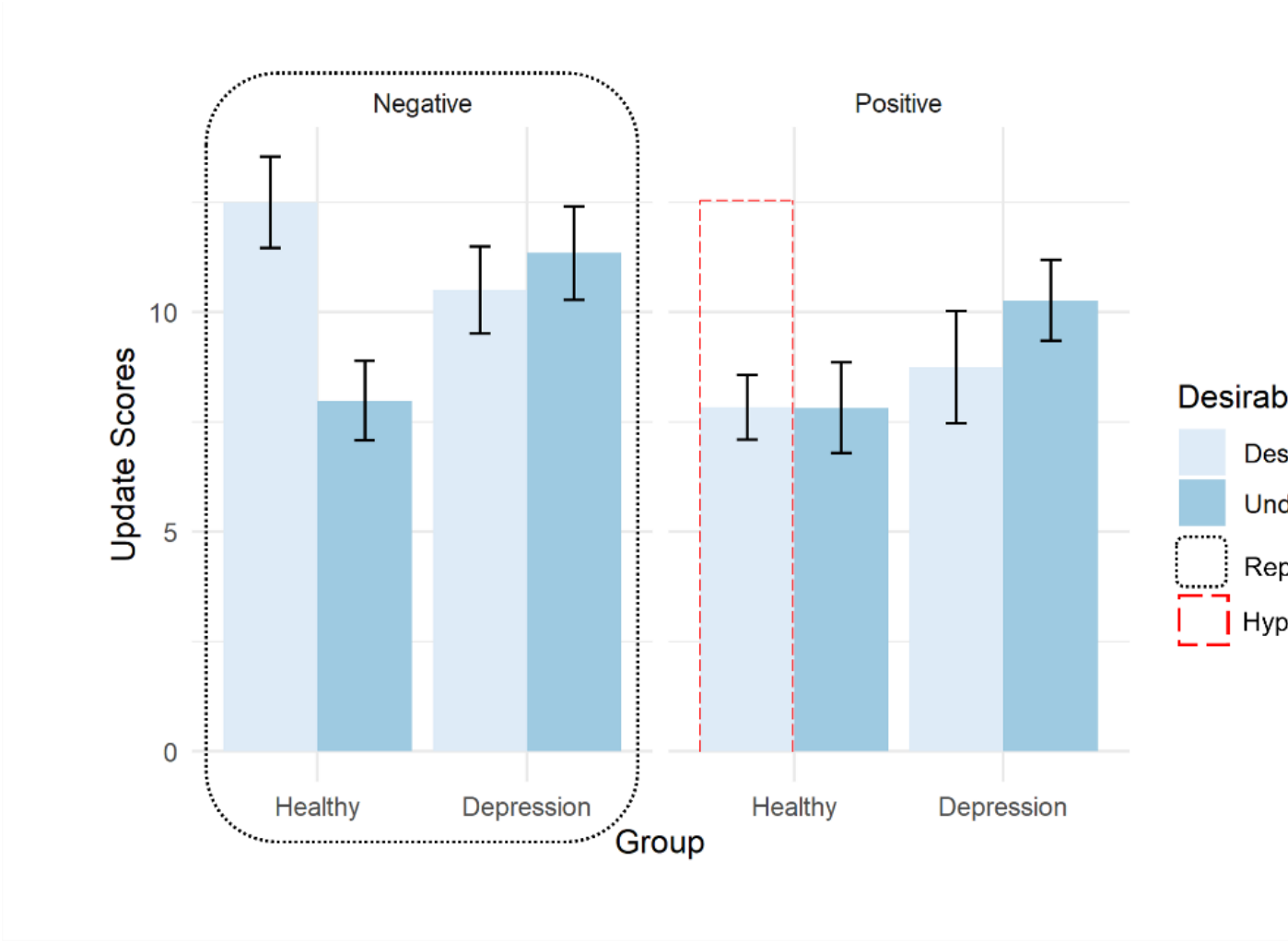


Figure 5

Mean update scores by group, desirability and valence. Error bars represent standard errors. For negative life events (left panel) we replicated Korn et al's (2014) findings of reduced optimistic updating in individuals experiencing depression; whereas participants in the healthy group showed greater updates for desirable versus undesirable information, participants in the depression group showed similar levels of updating for desirable and undesirable information. For positive life events, we hypothesised that the effect of optimism bias would be consistent to those observed for negative life events. As illustrated by the red dashed bar, mean update scores did not support our hypothesis; whilst participants in the healthy group showed greater updating for desirable versus undesirable information for negative life events, they showed a similar change in beliefs for desirable and undesirable information for positive life events, that is, no optimistic belief updating for positive events.

Table 4

Results from mixed-effects linear regression models for hypotheses one and two examining changes in belief updating

	β	95% CI	p
Hypothesis 1 – Negative Life Events			
Intercept	0.25	0.00, 0.51	< .001
Group	-0.26	-0.63, 0.10	0.160
Desirability	-0.60	-0.93, -0.27	0.001
Group x Desirability	0.71	0.24, 1.18	0.004
Hypothesis 2 – Positive and Negative Life Events			
Intercept	0.38	0.12, 0.64	< .001
Group	-0.26	-0.63, 0.10	0.161
Valence	-0.62	-0.94, -0.29	< .001
Desirability	-0.60	-0.92, -0.27	< .001
Group x Valence	0.39	-0.08, 0.85	0.104
Group x Desirability	0.71	0.24, 1.17	0.003
Valence x Desirability	0.60	0.14, 1.06	0.011
Group x Desirability x Valence	-0.51	-1.16, 0.15	0.130

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable

Hypothesis 3: *Initial estimates of negative life events will be associated with depression, such that depressed individuals will on average initially rate their chances of experiencing a negative life event as higher compared to healthy controls.*

As hypothesised, individuals with depression initially estimated their chance of experiencing a negative life event as higher compared to healthy individuals ($\beta = 1.06$, 95% CI: 0.74, 1.38, $p < .001$; Figure 6).

Hypothesis 4: *Initial estimates of positive life events will be associated with depression, such that depressed individuals will on average initially rate their chances of experiencing a positive life event as lower compared to healthy controls.*

In support of our hypothesis, individuals in the depression group initially estimated their chance of experiencing a positive life event as lower compared to the healthy group ($\beta = -0.71$, β 95% CI: -1.06, -0.35, $p < .001$; Figure 6).

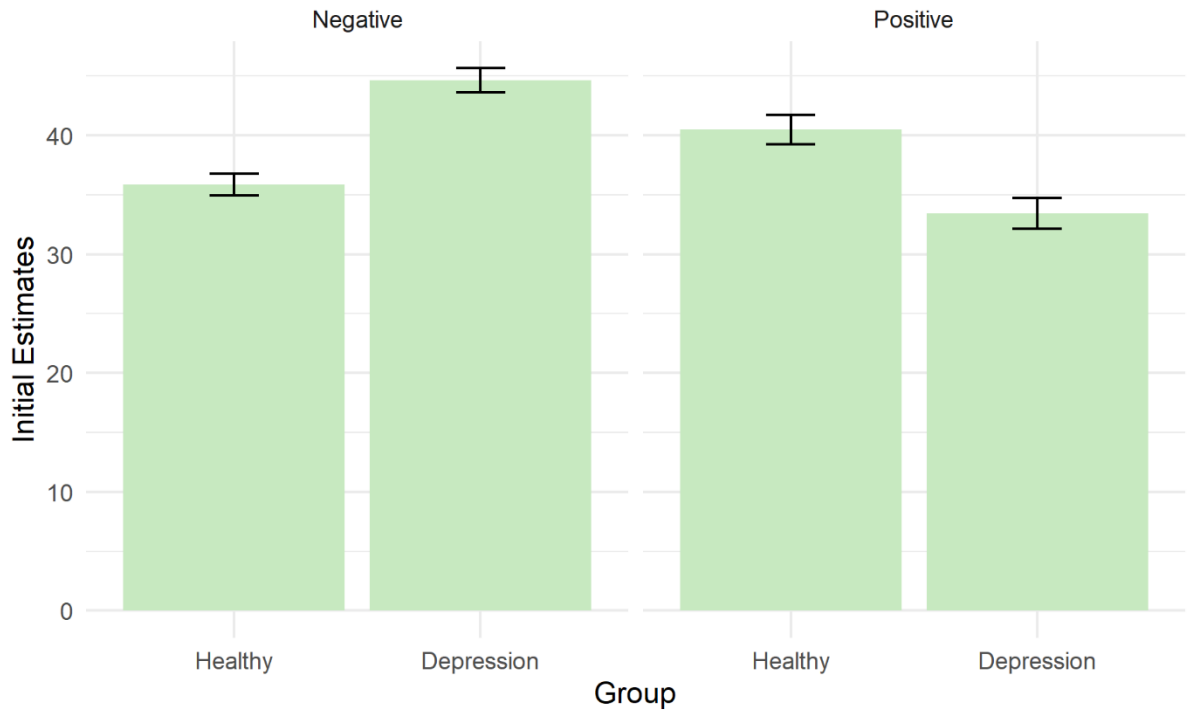


Figure 6

Mean Initial Estimates for Negative and Positive Life Events by Group. Error bars represent standard errors. Participants in the depression group initially estimated their chances of experience a negative life event as greater, and chances of experiencing a positive life event as lower, than participants in the healthy group.

Additional Analyses

Scaled Update Scores

To evaluate whether differences in update scores between groups may be driven by differences in estimation errors, we calculated scaled update scores by dividing update scores by absolute estimation errors. Greater scaled update values indicate greater change in re-estimates relative to the magnitude of estimation errors. When analysing only negative life events we again found evidence of an interaction between group and desirability ($\beta = 1.04$, 95% CI: 0.59, 1.48, $p < .001$). Consistent with hypothesis 1, participants in the healthy group displayed optimistic belief updating ($t(108) = 3.34$, $p = 0.001$), whereas participants in the depression group showed the opposite pattern ($t(108) = -3.11$, $p = 0.002$; Figure 7). Hypothesis 2 was not supported; when analysing both positive and negative life events, we found weak evidence that this pattern differed by valence (Group x Desirability x Valence: $\beta = -0.68$, 95% CI: -1.33, -0.04, $p = 0.039$). In contrast to hypothesis 2, both the healthy ($t(324) = 4.62$, $p < .001$) and depression group ($t(324) = 3.54$, $p = 0.001$) showed greater scaled updating towards desirable versus undesirable information for positive life events (Figure 7). Full results are reported in Table 5.

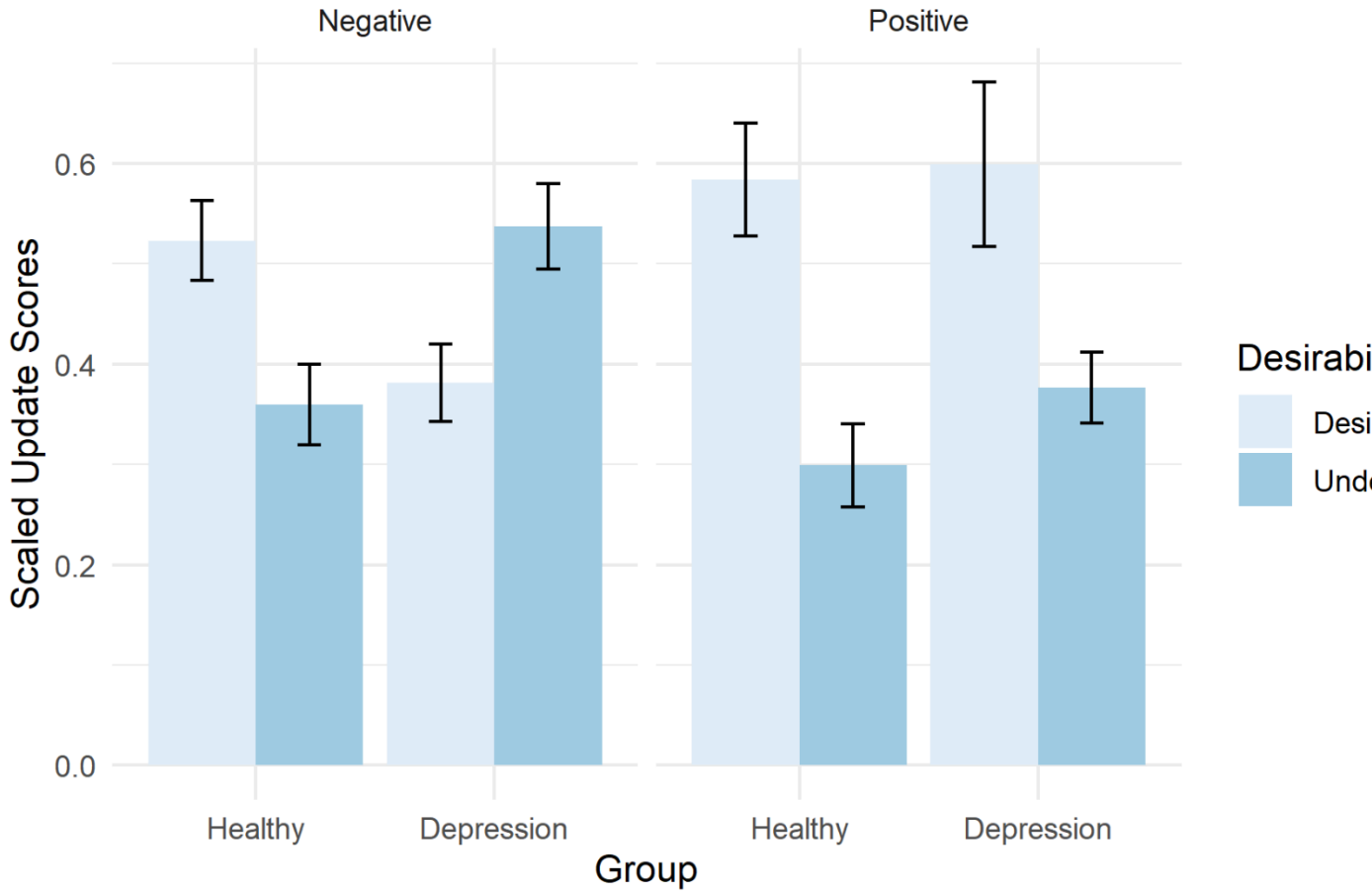


Figure 7

Mean scaled update scores (updates divided by estimation errors) by group, valence, and desirability. Error bars represent standard errors. For negative life events (left panel), whereas participants in the healthy group showed greater updates for desirable versus undesirable information, participants in the depression group showed greater updates for undesirable versus desirable information. For positive life events (right panel), both the healthy and depression groups showed greater updates for desirable versus undesirable information

Table 5

Results from mixed effects linear regression models replicating our primary analyses for update scores scaled for estimation errors.

	β	β 95% CI	p
Hypothesis 1 – Negative Life Events			
Intercept	0.24	-0.02, 0.49	< .001
Group	-0.46	-0.82, -0.10	0.014
Desirability	-0.53	-0.84, -0.22	0.001
Group x Desirability	1.04	0.59, 1.48	< .001
Hypothesis 2 – Positive and Negative Life Events			
Intercept	0.18	-0.08, 0.43	< .001
Group	-0.38	-0.74, -0.02	0.041
Valence	0.16	-0.16, 0.48	0.327
Desirability	-0.43	-0.76, -0.11	0.008
Group x Valence	0.42	-0.04, 0.88	0.075
Group x Desirability	0.85	0.39, 1.31	< .001
Valence x Desirability	-0.32	-0.78, 0.13	0.166
Group x Desirability x Valence	-0.68	-1.33, -0.04	0.039

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Dose-Response Relationship

When measuring depression symptoms using the BDI-II we found little evidence of a relationship with update scores for either positive or negative life events. When measuring depression symptoms using the PHQ-9 we found only weak evidence of a dose-response relationship for negative life events (β = 0.25, 95% CI: 0.01, 0.49, p = 0.046). As depression severity increased there was an increasing overlap in confidence intervals between desirable and undesirable trials, suggesting that greater depression was associated with reduced differences in updating behaviour following desirable or undesirable information (Supplementary Figure 1). We found no evidence that this effect differed by valence when analysing both positive and negative life events (β = -0.12, 95% CI: -0.45, 0.21, p = 0.480). Full results are reported in Supplementary Table 6.

Memory Errors

Whilst there was some evidence that participants in the depression group had worse memory overall (Group: β = 0.51, 95% CI: 0.14, 0.88, p = 0.007), there was only weak evidence that this interacted with valence (β = -0.43, 95% CI: -0.85, -0.01, p = 0.047) and no evidence of an interaction with desirability (β = -0.28, 95% CI: -0.70, 0.14, p = 0.187). There was no evidence of an interaction between group, valence, and desirability (β = 0.28, 95% CI: -0.32, 0.87, p = 0.365). When adjusting for memory errors in our primary analyses, results relating to our hypotheses were unchanged (Supplementary Tables 7-8).

Trait Optimism and Order Effects

The effects outlined above relating to our hypotheses were unchanged when adjusting for trait optimism, framing of life events, and valence order (Supplementary Tables 9-11).

Life Event Ratings

Mean ratings of life events (Table 3, Supplementary Figure 2) suggested some differences in perceptions of events between groups. Most notably, healthy participants tended to rate both positive and negative life events as less controllable than the depression group. Additionally, participants in the healthy group rated negative life events, on average, as less emotionally arousing, familiar and vivid, and reported less prior experience of these events, compared to the depression group. There were relatively small differences in mean ratings of negativity

and positivity. Both groups rated negative events as more negative and less positive, and vice versa for positive events. However, healthy participants tended to rate positive events as slightly less negative and slightly more positive compared to the depression group. Although, when we adjusted for life event ratings our findings within our primary analyses were unchanged (Supplementary Table 12).

To evaluate whether differences in our categorisation of events as positive or negative versus participants' perceptions of valence may have influenced our findings, we conducted sensitivity analyses repeating our primary analyses with events classified according to each participant's negativity and positivity ratings. Results relating to our hypotheses were again unchanged suggesting that this did not account for our findings (Supplementary Table 13).

We have provided the mean ratings for each life event to aid the selection of life events in future research in Supplementary Table 14.

Adjusting for age

We did not find evidence that age influenced our findings (Supplementary Table 15).

Scepticism Over Validity of Life Event Estimates

Most participants in the healthy (n = 32, 58.18%) and depression groups (n = 30, 57.69%) disagreed with the statement "The average probabilities in the task were accurate". When we repeated our primary analyses excluding these participants who indicated scepticism, our results for hypotheses one, three and four were consistent with our previous findings but effects were strengthened (Supplementary Figure 3). For hypothesis two, we did not find statistical evidence of an interaction between valence, group, and desirability (β = -0.69, 95% CI: -1.57, 0.20, p = 0.129), but visual inspection of data indicated differences in a largely similar pattern to those previously observed. For negative life events, healthy individuals updated more towards desirable versus undesirable information whereas individuals with depression showed the opposite pattern (Supplementary Figure 3A). However, for positive life events both groups showed approximately similar levels of updating for undesirable and desirable information. Full results for all models are available in Supplementary Table 16.

Discussion

When estimating the likelihood of experiencing negative life events in the future, healthy individuals are more likely to update their beliefs following desirable information ("good news") versus undesirable information ("bad news") [14,22,44,45] (although see [17–19,46] for critiques of this evidence and [22,47,48] for rebuttals). In contrast, individuals experiencing depression have been found to lack this optimistic belief updating bias [2,3]. We conducted the first independent replication of the effect of depression on optimistic belief updating to negative life events. Additionally, we extended previous research to examine the effect of depression on belief updating for positive life events.

Negative Life Events

This was the first attempt to independently replicate findings of reduced optimistic belief updating for negative life events in depression. In support of hypothesis 1, and replicating previously reported effects [3], we observed a pattern of belief updating in the healthy comparison group consistent with optimism bias; healthy participants changed their beliefs more following desirable versus undesirable information. In contrast, individuals experiencing depression showed a lack of biased updating, changing their beliefs to a similar extent following desirable or undesirable information. When we accounted for estimation errors, the effects of depression were even greater; healthy individuals continued to show optimistic updating whereas individuals with depression showed a pessimistic belief updating bias, changing their beliefs to a greater extent following undesirable versus desirable information. Individuals experiencing depression were also more likely to believe that they would experience a negative life event in the future. As in previous research [3], depression was therefore characterised by both initial pessimistic beliefs of experiencing a negative life event, as well as a reduced ability to update these beliefs following disconfirmatory positive evidence. However, we did not replicate a dose-response association between BDI-II scores and optimistic belief updating. Although we did find weak evidence using another measure of depression severity, the PHQ-9. Rather than a dose-response association, our findings instead suggest that reduced optimistic belief updating may only be apparent at higher severities of depression.

Positive Life Events

To our knowledge, this is the first study to examine the association between depression and belief updating for positive life events. Previous research regarding optimistic belief updating for positive life events has been restricted to healthy samples, producing mixed results [19,22,23]. Additionally, research was limited by the use of simulated probabilities and shorter timeframes of reference than typically used. We attempted to address these methodological limitations by using validated probabilities for the likelihood of experiencing positive life events across one's lifetime.

We hypothesised that we would observe similar patterns of optimistic belief updating in healthy controls versus individuals experiencing depression across both positive and negative life events. However, evidence for this hypothesis was ambiguous. In support of hypothesis 2, we did not find evidence within linear mixed effect models that differences in patterns of belief updating between groups varied according to the valence of events (i.e., there was little support for the Group x Desirability x Valence interaction). However, contrary to hypothesis 2, mean update scores suggested that whilst the depression group showed similar updating, the healthy group displayed optimistic updating for negative events but did not show optimistic belief updating for positive events. Furthermore, when we accounted for estimation errors, we did find evidence of the Group x Desirability x Valence interaction; the healthy and depression groups showed different patterns of belief updating for negative events (consistent with hypothesis 1), but both groups displayed similar optimistic updating for positive life events (contrary to hypothesis 2). Despite the mixed patterns of belief updating according to desirability, group, and valence across our analyses, it is clear that the patterns of belief updating by group were not the same for positive and negative life events. On balance, therefore, we find little support for hypothesis 2. Although further research is required, our findings suggest that reduced optimistic belief updating related to depression predominantly occurred in response to negative life events.

We also hypothesised that depressed individuals would initially rate their chances of experiencing a positive event as lower compared to healthy controls. In keeping with our hypothesis, the depression group showed reduced optimism for positive life events; providing lower estimates of experiencing positive events in their lifetime compared to the healthy group. Individuals experiencing depression therefore have initial reduced optimistic beliefs

regarding their likelihood of experience positive events in the future. However, it is unclear at present from our findings how these beliefs are changed following disconfirmatory evidence.

Our results are in line with previous research in healthy volunteers, where optimistic belief updating was more clear-cut for negative and not positive life events [23]. Additionally, our findings of reduced optimistic belief updating in depression for negative but not positive life events are broadly in keeping with previous research. In a study examining interpretation biases, whilst individuals with greater levels of depression showed a reduced ability to change initially negative interpretations, no differences were observed for revision of initially positive interpretations [49]. Similarly, in another study that induced positive expectations of task performance, whilst individuals with depression showed lower initial positive expectations they did not differ from healthy controls in subsequent updating of their beliefs [50]. Differences in belief updating associated with depression therefore appear to be specific to initial increased negative expectations. Whilst depression appears to be associated with cognitive inflexibility when processing emotional stimuli [51], our findings suggest that, when anticipating future life events, this is heightened for negative events. It is possible that this is driven by differences in the perceptions of positive and negative life events beyond those measured in this study. Alternatively, as processing of negative information elicits distinct brain regions to positive, rewarding information, differences associated with depression may be specific to particular neural pathways [52–54]. Future research is required to disentangle the potential mechanisms underlying the specificity of belief updating effects in depression to negative life events.

Clinical Implications

According to cognitive theories, depression is characterised by core negative beliefs about the self, the environment, and the future [8]. In keeping with this theory, we found that individuals experiencing depression were more likely to believe that negative life events would happen and were less likely to believe that positive life events would happen. When examining how these expectations were updated following novel information, in keeping with theories of depressive realism [10], we found that the healthy controls, rather than individuals experiencing depression, displayed biased processing. Whereas healthy individuals changed their beliefs more after receiving good versus bad news about negative life events, individuals experiencing depression showed approximately equal change

irrespective of the desirability of new information. Within the context of models of belief updating, our findings suggest that depressive core beliefs are maintained by reduced accommodation of positive information, preventing updating of negative beliefs even when presented with good news [55–57]. Treatments for depression could target not only core negative beliefs about the future but also how these beliefs are updated following novel information. In particular, attempting to increase accommodation of positive information into existing beliefs may be beneficial. However, our findings in combination with other research in this field suggests that reduced optimistic belief updating is currently only reliably observed for negative life events. Treatments should currently focus on maintenance of negative beliefs about the future.

Additionally, there is research suggesting that change in optimistic belief updating may be a potential predictor of change in mood disorder symptoms. In a study examining bipolar patients over a five year period, reduced optimistic belief updating predicted earlier relapse [58]. This has yet to be examined in patients experiencing unipolar depression but raises the possibility that measuring belief updating may allow us to identify individuals at risk of developing depression or patients at risk of relapse. Additionally, belief updating has been found to predict change in depression symptoms during treatment. In an observational study, treatment resistant depression patients were initially found to update their beliefs following desirable or undesirable information to a similar extent. However, following a single infusion of ketamine patients displayed an optimistic belief updating bias. Notably this change mediated an improvement in depressive symptoms after one week of treatment [5]. Change in belief updating may therefore be a potential biomarker of early treatment response for depression. Further research examining this possibility within conventional antidepressant treatments for depression may be beneficial in improving our understanding of treatment mechanisms.

Future Research

Whilst our study adds to evidence that depression is associated with reduced optimistic belief updating for negative life events the cognitive mechanisms underlying this effect remain unclear. Our findings were not explained by trait optimism, recall of presented probabilities, or perceptions of the life events such as controllability, positivity, or negativity. One possibility is that initial levels of optimism influenced participants' attention towards

desirable information. In a recent study, inducing optimistic expectancies regarding task performance increased attention towards rewards versus punishments [59]. Another possibility is that depression may be associated with differences in the use of cognitive immunization strategies. Cognitive immunization refers to strategies used to challenge evidence contradictory to existing beliefs, such as questioning the credibility of the source. In line with previous research [56], it is possible that individuals experiencing depression used cognitive immunization strategies to a greater extent when presented with desirable news that contradicted negative expectations. Finally, it is possible that transdiagnostic symptoms underlying a range of mental health disorders, rather than depression symptoms per se, may be at last partially responsible for reduced optimistic belief updating. One possible candidate may be stress. Both induced acute stress and naturalistic stress (firefighters on call) have been associated with an absence of optimistic belief updating [60]. Further exploration of potential cognitive mechanisms underlying reduced optimistic belief updating in depression would be useful in identifying sensitive targets for therapeutic intervention.

Limitations

Due to the COVID-19 pandemic and accompanying social restrictions we collected all data online. Data quality may therefore have been lower than in-person research previously conducted in this field. However, we included attention checks and had a strict exclusion criterion, removing any participant that failed a single attention check. Research has also indicated no evidence of performance differences on cognitive tasks requiring high levels of concentration between participants completing the task online versus lab conditions [61].

Additionally, it is possible that our recruitment methods may have led to differences in sample characteristics to those of previous studies, potentially influencing our findings. Whereas previous research recruited depression groups from clinical settings, we recruited our sample online. Psychiatric characteristics may therefore be expected to differ. However, levels of depression and reported treatment in our depression group were largely similar to that of previous research limiting the possibility that this impacted our findings.

Despite our life event estimates being derived from reputable sources a large proportion of our participants expressed scepticism as to their credibility. Participant's willingness to update their beliefs may have been influenced by perceptions of credibility. Indeed, we found

our effects were strengthened when only participants that did not report scepticism were included. However, it is difficult to tell how much this may have contributed to differences in our findings to that of previous research as to our knowledge we are the first to report data for a measure of scepticism for this task. Additionally, as described above, attempts to discount the credibility of evidence may be a potential strategy within belief updating. Whilst participant scepticism over presented stimuli may traditionally be viewed as a limitation within cognitive studies, within this context it may be a potential mechanism associated with observed effects.

Finally, in replicating Korn et al [3], we employed a widely used version of the belief updating paradigm. However, this task has recently come under increasing scrutiny. In the classic version of the task that we employed, participants are only asked to provide an estimate of their personal likelihood of experiencing life events. It has been suggested that a more appropriate way of administering the task necessitates asking participants about both their personal as well as the average person's likelihood of experiencing life events, that is, 'the base rate' [19]. This task has also been criticised for issues relating to scale attention and base rate regression [16,17] (although see [22,47] for rebuttals).

That we did not find evidence of optimistic belief updating for positive events in neither people with depression nor healthy controls, could arguably be taken as evidence against the existence of a universal optimistic updating bias. Similarly, our pattern of results could be taken as another sign of an underlying problem with the validity of the optimistic updating paradigm, which is currently being debated [16–19,46]. As such, reduced optimistic belief updating effects associated with depression may require further validation after a more universally accepted measure of optimistic updating is developed. Further research examining whether effects relating to depression are replicated in future adaptations of this task (e.g., using neutral life events [46]) may help us to understand the role of belief updating in depression. Given the contentious debate over this paradigm we would encourage researchers within this field to publish materials and data as open access to facilitate further constructive discussion

Summary

In this study we independently replicated findings of reduced optimistic belief updating for negative life events in individuals experiencing depression. This study was conducted within the format of a registered report and recruited a substantially larger sample than that of previous research. Our results therefore add confidence to previous findings that depression is characterised by negative future expectations maintained by a reduced ability to update these expectations in response to good news. Treatments targeting belief updating by increasing the accommodation of positive information may be beneficial in remediating depressive symptoms. However, our findings for positive life events suggest that optimistic belief updating effects are not a universal feature of human cognition. Whilst individuals with depression were less optimistic in their belief of experiencing positive life events in the future, there was little difference in updating of these beliefs compared to healthy controls. Further research understanding the cognitive mechanisms underlying the specificity of reduced optimistic belief updating associated with depression to negative life events is required.

References

1. Sharot T, Korn CW, Dolan RJ. 2011 How unrealistic optimism is maintained in the face of reality. *Nat. Neurosci.* **14**, 1475–1479. (doi:10.1038/nn.2949)

2. Garrett N, Sharot T, Faulkner P, Korn CW, Roiser JP, Dolan RJ. 2014 Losing the rose tinted glasses: Neural substrates of unbiased belief updating in depression. *Front. Hum. Neurosci.* **8**, 639. (doi:10.3389/fnhum.2014.00639)

3. Korn CW, Sharot T, Walter H, Heekeren HR, Dolan RJ. 2014 Depression is related to an absence of optimistically biased belief updating about future life events. *Psychol. Med.* **44**, 579–592. (doi:10.1017/S0033291713001074)

4. Beck AT. 1987 Cognitive models of depression. *J. Cogn. Psychother.* **1**, 5–37.

5. Bottemanne H, Morlaàs O, Claret A, Fossati P, Schmidt L. 2021 Can we change depressive beliefs? Modulation of belief updating mechanisms by ketamine in treatment resistant depression. , 1–38. (doi:10.31234/OSF.IO/ZX5VR)

6. Romeo B, Choucha W, Fossati P, Rotge JY. 2015 Meta-analysis of short- and mid-term efficacy of ketamine in unipolar and bipolar depression. *Psychiatry Res.* **230**, 682–688. (doi:10.1016/j.psychres.2015.10.032)

7. Clark L, Chamberlain SR, Sahakian BJ. 2009 Neurocognitive mechanisms in depression: Implications for treatment. *Annu. Rev. Neurosci.* **32**, 57–74. (doi:10.1146/annurev.neuro.31.060407.125618)

8. Beck AT. 2002 Cognitive Models of Depression. In *Clinical advances in cognitive psychotherapy: Theory and application* (eds R Leahy, ET Dowd), pp. 29–61. Springer Publishing Company.

9. Beck AT, Dozois DJA. 2011 Cognitive therapy: Current status and future directions. *Annu. Rev. Med.* **62**, 397–409. (doi:10.1146/annurev-med-052209-100032)

10. Moore MT, Fresco DM. 2012 Depressive realism: A meta-analytic review. *Clin. Psychol. Rev.* **32**, 496–509. (doi:10.1016/j.cpr.2012.05.004)

11. Cummins R, Nistico H. 2002 Maintaining Life Satisfaction: The Role of Positive Cognitive Bias. *J. Happiness Stud.* **3**, 37–69. (doi:10.1023/A:1015678915305)

12. Blaine B, Crocker J. 1993 Self-Esteem and Self-Serving Biases in Reactions to

Positive and Negative Events: An Integrative Review. In *Self-Esteem* (ed RF Baumeister), pp. 55–85. Springer, Boston, MA. (doi:10.1007/978-1-4684-8956-9_4)

13. Strunk DR, Lopez H, DeRubeis RJ. 2006 Depressive symptoms are associated with unrealistic negative predictions of future life events. *Behav. Res. Ther.* **44**, 861–882. (doi:10.1016/j.brat.2005.07.001)

14. Sharot T. 2011 The optimism bias. *Curr. Biol.* **21**, R941–R945. (doi:10.1016/j.cub.2011.10.030)

15. Sharot T, Garrett N. 2016 Forming Beliefs: Why Valence Matters. *Trends Cogn. Sci.* **20**, 25–33. (doi:10.1016/j.tics.2015.11.002)

16. Harris AJL, Hahn U. 2011 Unrealistic Optimism About Future Life Events: A Cautionary Note. *Psychol. Rev.* **118**, 135–154. (doi:10.1037/a0020997)

17. Harris AJL, De Molière L, Soh M, Hahn U. 2017 Unrealistic comparative optimism: An unsuccessful search for evidence of a genuinely motivational bias. *PLoS One* **12**, e0173136. (doi:10.1371/journal.pone.0173136)

18. Harris A, Shah P, Catmur C, Bird G, Hahn U. 2013 Autism, optimism and positive events: Evidence against a general optimistic bias. *Proc. Annu. Meet. Cogn. Sci. Soc.* **35**, 555–560.

19. Shah P, Harris AJL, Bird G, Catmur C, Hahn U. 2016 A pessimistic view of optimistic belief updating. *Cogn. Psychol.* **90**, 71–127. (doi:10.1016/j.cogpsych.2016.05.004)

20. Coutts A. 2018 Good news and bad news are still news: experimental evidence on belief updating. *Exp. Econ. 2018 222* **22**, 369–395. (doi:10.1007/S10683-018-9572-5)

21. Barron K. 2020 Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains? *Exp. Econ. 2020 241* **24**, 31–58. (doi:10.1007/S10683-020-09653-Z)

22. Garrett N, Sharot T. 2017 Optimistic update bias holds firm: Three tests of robustness following Shah et al. *Conscious. Cogn.* **50**, 12–22. (doi:10.1016/J.CONCOG.2016.10.013)

23. Marks J, Baines S. 2017 Optimistic belief updating despite inclusion of positive events. *Learn. Motiv.* **58**, 88–101. (doi:10.1016/J.LMOT.2017.05.001)

24. Eshel N, Roiser JP. 2010 Reward and punishment processing in depression. *Biol. Psychiatry*. **68**, 118–124. (doi:10.1016/j.biopsych.2010.01.027)

25. Kunisato Y, Okamoto Y, Ueda K, Onoda K, Okada G, Yoshimura S, Suzuki SI, Samejima K, Yamawaki S. 2012 Effects of depression on reward-based decision making and variability of action in probabilistic learning. *J. Behav. Ther. Exp. Psychiatry* **43**, 1088–1094. (doi:10.1016/j.jbtep.2012.05.007)

26. Lewis G, Kounali D-Z, Button KS, Duffy L, Wiles NJ, Munafò MR, Harmer CJ, Lewis G. 2017 Variation in the recall of socially rewarding information and depressive symptom severity: a prospective cohort study. *Acta Psychiatr. Scand.* **135**, 489–498. (doi:10.1111/ACPS.12729)

27. Ji JL, Grafton B, MacLeod C. 2017 Referential focus moderates depression-linked attentional avoidance of positive information. *Behav. Res. Ther.* **93**, 47–54. (doi:10.1016/j.brat.2017.03.004)

28. Prolific. 2021 Prolific. See https://www.prolific.co.

29. Qualtrics. 2020 Qualtrics.

30. Kroenke K, Spitzer RL, Williams JBW. 2001 The PHQ-9: Validity of a brief depression severity measure. *J. Gen. Intern. Med.* **16**, 606–613. (doi:10.1046/j.1525-1497.2001.016009606.x)

31. Hobbs C, Sui J, Kessler D, Munafò MR, Button KS. 2021 Self-processing in relation to emotion and reward processing in depression. *Psychol. Med.* , 1–13. (doi:10.1017/s0033291721003597)

32. Erdfelder E, Faul F, Buchner A. 1996 GPOWER: A general power analysis program. *Behav. Res. Methods, Instruments, Comput.* **28**, 1–11. (doi:10.3758/BF03203630)

33. Inquisit. 2020 Inquisit 6. See www.millisecond.com.

34. Clark DM. 2011 Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *Int. Rev. Psychiatry* **23**, 318–327. (doi:10.3109/09540261.2011.606803)

35. Martin A, Rief W, Klaiberg A, Braehler E. 2006 Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *Gen. Hosp. Psychiatry*

28, 71–77. (doi:10.1016/j.genhosppsych.2005.07.003)

36. Kung S, Alarcon RD, Williams MD, Poppe KA, Jo Moore M, Frye MA. 2013 Comparing the beck depression inventory-II (BDI-II) and Patient Health Questionnaire (PHQ-9) depression measures in an integrated mood disorders practice. *J. Affect. Disord.* **145**, 341–343. (doi:10.1016/j.jad.2012.08.017)

37. Spitzer RL, Kroenke K, Williams JBW, Löwe B. 2006 A Brief Measure for Assessing Generalized Anxiety Disorder. *Arch. Intern. Med.* **166**, 1092–1097. (doi:10.1001/archinte.166.10.1092)

38. Leary MR. 1983 A Brief Version of the Fear of Negative Evaluation Scale. *Personal. Soc. Psychol. Bull.* **9**, 371–375. (doi:10.1177/0146167283093007)

39. Scheier MF, Carver CS, Bridges MW. 1994 Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Personal. Soc. Psychol. Bull.* **67**, 1063–1078.

40. Watson D, Clark LA, Tellegen A. 1988 Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Pers. Soc. Psychol.* **54**, 1063. (doi:10.1037//0022-3514.54.6.1063)

41. Hobbs C, Shah P, Button KS. In press. Dataset for 'Is depression associated with reduced optimistic belief updating?' (doi:https://doi.org/10.15125/BATH-01078)

42. Benjamin DJ *et al.* 2018 Redefine statistical significance. *Nat. Hum. Behav.* **2**, 6–10. (doi:10.1038/s41562-017-0189-z)

43. Rutherford A. 2001 Introducing Anova and Ancova: A GLM Approach. SAGE.

44. Kuzmanovic B, Rigoux L. 2017 Valence-dependent belief updating: Computational validation. *Front. Psychol.* **8**, 1087. (doi:10.3389/fpsyg.2017.01087)

45. Kuzmanovic B, Jefferson A, Vogeley K. 2016 The role of the neural reward circuitry in self-referential optimistic belief updates. *Neuroimage* **133**, 151–162. (doi:10.1016/j.neuroimage.2016.02.014)

46. Burton J, Harris A, Shah P, Hahn U. 2021 Optimism where there is none: Asymmetric belief updating observed with valence-neutral life events. [under revision]. *Cognition*

47. Sharot T, Garrett N. 2021 A Guideline and Cautionary Note: How to Use the Belief

Update Task Correctly. *PsyArXiv Prepr.* (doi:10.31234/osf.io/st4vu)

48. Garrett N, Sharot T. 2021 Failure to Replicate Burton, Harris, Shah & Hahn (2021): There is No Belief Update Bias for Neutral Events. *PsyArXiv Prepr.*

49. Everaert J, Bronstein M V., Cannon TD, Joormann J. 2018 Looking Through Tinted Glasses: Depression and Social Anxiety Are Related to Both Interpretation Biases and Inflexible Negative Interpretations. *Clin. Psychol. Sci.* **6**, 517–528. (doi:10.1177/2167702617747968)

50. Kube T, Kirchner L, Rief W, Gärtner T, Glombiewski JA. 2019 Belief updating in depression is not related to increased sensitivity to unexpectedly negative information. *Behav. Res. Ther.* **123**, 103509. (doi:10.1016/j.brat.2019.103509)

51. Murphy FC, Michael A, Sahakian BJ. 2012 Emotion modulates cognitive flexibility in patients with major depression. *Psychol. Med.* **42**, 1373–1382. (doi:10.1017/S0033291711002418)

52. Wrase J, Kahnt T, Schlagenhauf F, Beck A, Cohen MX, Knutson B, Heinz A. 2007 Different neural systems adjust motor behavior in response to reward and punishment. *Neuroimage* **36**, 1253–1262. (doi:10.1016/j.neuroimage.2007.04.001)

53. Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C. 2006 Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* **26**, 9530–9537. (doi:10.1523/JNEUROSCI.2915-06.2006)

54. Camara E, Rodriguez-Fornells A, Münte TF. 2009 Functional connectivity of reward processing in the brain. *Front. Hum. Neurosci.* **2**, 19. (doi:10.3389/neuro.09.019.2008)

55. Pinquart M, Endres D, Teige-Mocigemba S, Panitz C, Schütz AC. 2021 Why expectations do or do not change after expectation violation: A comparison of seven models. *Conscious. Cogn.* **89**, 103086. (doi:10.1016/j.concog.2021.103086)

56. Kube T, Rief W, Gollwitzer M, Gärtner T, Glombiewski JA. 2019 Why dysfunctional expectations in depression persist - Results from two experimental studies investigating cognitive immunization. *Psychol. Med.* **49**, 1532–1544. (doi:10.1017/S0033291718002106)

57. Kube T, Rozenkrantz L. 2021 When Beliefs Face Reality: An Integrative Review of

Belief Updating in Mental Health and Illness. *Perspect. Psychol. Sci.* **16**, 247–274. (doi:10.1177/1745691620931496)

58. Ossola P, Garrett N, Sharot T, Marchesi C. 2020 Belief updating in bipolar disorder predicts time of recurrence. *Elife* **9**, 1–17. (doi:10.7554/eLife.58891)

59. Kress L, Bristle M, Aue T. 2018 Seeing through rose-colored glasses: How optimistic expectancies guide visual attention. *PLoS One* **13**, e0193311. (doi:10.1371/journal.pone.0193311)

60. Garrett N, González-Garzón AM, Foulkes L, Levita L, Sharot T. 2018 Updating beliefs under perceived threat. *J. Neurosci.* **38**, 7901–7911. (doi:10.1523/JNEUROSCI.0716-18.2018)

61. Gould SJJ, Cox AL, Brumby DP, Wiseman S. 2015 Home is Where the Lab is: A Comparison of Online and Lab Data From a Time-sensitive Study of Interruption. *Hum. Comput.* **2**, 45–67. (doi:10.15346/hc.v2i1.4)

General Discussion

Depression is characterised by a heightened internal focus on the self (Northoff, 2007), enhanced sensitivity to negative information (Everaert, Duyck, & Koster, 2014), and reduced sensitivity to reward (Huys, Pizzagalli, Bogdan, & Dayan, 2013; Keren et al., 2018). Differences in processing of self-related information are believed to be linked to emotion and reward cognition (Beck, 1987; Northoff & Hayes, 2011). The interaction between self, emotion, and reward processing may therefore be a key cognitive vulnerability in depression and may be a sensitive target for therapeutic intervention. In this thesis I aimed to evaluate whether depression is associated with differences in processing of self-related information and whether this varies depending on the emotional valence or rewarding nature of information (chapters 1 and 2). Secondly, I aimed to understand whether administration of antidepressants is associated with change in self-referential affective processing (chapters 3, 4 and 5).

Summary and discussion of main findings

Chapter 1: Self processing in relation to emotion and reward processing in depression

In chapter 1 I investigated the relationship between self, emotion, and reward processing occurring independently and in interaction with depression using three types of cognitive tasks. Individuals from the community (n = 144) with none, mild, and moderate to severe levels of depression completed cognitive tasks measuring self, emotion, and reward processing, and self-report measures of depression in two testing sessions.

In relation to the first aim of my thesis I found that depression was associated with differences in processing of self-related information, but that this was dependent on both the emotional and rewarding nature of information. When learning positive and negative social evaluations as part of a computerised reinforcement learning task, individuals with depression were impaired in their ability to learn a positive 'like' rule about the self. Depression was therefore characterised by reduced learning of positive and socially rewarding information about the self.

However, when self and emotion occurred in interaction, I found evidence of increased positive biases towards others but not reliable evidence of biases towards the self. Within a go/no-go task measuring inhibitory control of emotional information in relation to the self and others, there was not reliable evidence that individuals experiencing depression differed

in their ability to associate emotional information with the self. However, I found evidence that participants with greater depression demonstrated both a greater sensitivity when associating positive words and a reduced sensitivity when associating negative words with others.

When the self was processed independently from emotional or rewarding information using an associative learning task, I did not find differences in relation to depression severity. In keeping with previous research, participants showed greater accuracy and faster reaction times when learning to pair the neutral shapes with the self versus others, positive versus negative faces, and high versus low amounts of monetary reward (Stolte, Humphreys, Yankouskaya, & Sui, 2017; Sui & Humphreys, 2015; Sui, He, & Humphreys, 2012). However, I did not find reliable evidence that these effects varied by depression.

Depression was therefore associated with differences in self, emotion, and reward processing when these areas of cognition occurred in interaction within a social context, but not when only self and emotion were examined, or when these areas were measured independently.

Chapter 2: Individual differences in social evaluation learning and their association with psychopathology

In chapter 2, I aimed to validate my previous findings of an association between reduced positive social evaluation learning biases about the self and depression severity. I collated all previous research conducted using the social evaluation learning task within this research group to create a relatively large individual-level participant dataset (n = 552). Additionally, I recruited a general population sample via the online recruitment platform prolific (n = 807) to further assess the reliability of findings outside of supervised lab settings. Both datasets included the social evaluation learning task, and self-report measures of depression and social anxiety. As previous research has focused on the role of social evaluation learning in social anxiety (Button, Browning, Munafò, & Lewis, 2012; Button et al., 2015), I evaluated potential transdiagnostic biases underlying both disorders as well as potential differential patterns of learning.

Overall participants demonstrated a positive bias in learning of social evaluations, making fewer errors learning the positive versus negative rule. Depression and social anxiety were independently associated with a loss of this positive bias. However, different patterns of learning underpinned this. In both datasets I found strong evidence that social anxiety was associated with better learning of the negative ‘dislike’ rule but not the positive ‘like’ rule. In

contrast, depression was associated with differences in both learning of being liked and disliked. Participants with greater depression made fewer errors learning the negative 'dislike' rule and greater errors learning the positive 'like' rule, although evidence for this effect was weaker in the second dataset of individuals recruited online. Notably, all effects were specific to learning about the self; I did not find evidence of an association between learning about others with either social anxiety or depression.

My findings in chapter 2 support the findings of chapter 1 in establishing further evidence of reduced positive biases when learning social evaluations in depression, as well as the specificity of these effects to the self. Chapter 2 therefore addresses the first aim of this thesis in suggesting that depression is associated with differences in processing of self-related information when this occurs in interaction with emotion and reward within a social context. However, my findings in chapter 2 did differ from chapter 1 slightly in that whilst I replicated findings of reduced learning of positive social evaluations, I also found evidence of increased learning of negative evaluations. This may be attributable to the larger sample size of datasets within this chapter providing greater sensitivity to detect small effects. However, mitigating this possibility, effects were attenuated in the second larger dataset of participants who completed the social evaluation learning task remotely online. Collectively however both chapters 1 and 2 were consistent in suggesting that depression is characterised by a reduced positive social evaluation learning style. However, it may vary across samples as to whether this is predominantly driven by increased sensitivity to negative social evaluations or decreased sensitivity to positive social evaluations.

Chapter 3: The cognitive neuropsychological theory of antidepressants: a systematic review and meta-analysis of behavioural evidence

Antidepressants are proposed to operate by remediating negative affective biases associated with depression, by increasing sensitivity to positive relative to negative information (Harmer, Hill, Taylor, Cowen, & Goodwin, 2003; Roiser, Elliott, & Sahakian, 2012). Understanding the early psychopharmacological mechanisms of antidepressants may be helpful in refining the effectiveness of current antidepressant treatment. In preparation for the second aim of this thesis, I aimed to establish the current evidence base for change in emotional processing following antidepressant administration by conducting a comprehensive systematic review and meta-analysis.

I did not find reliable evidence from 82 studies included in the narrative synthesis portion of the review or 28 studies included in meta-analyses that antidepressants were associated with reliable differences in the processing of positive or negative emotional stimuli. Effects overall did not vary by cognitive domain, although I found some weak evidence that antidepressants may be associated with enhanced recognition of positive facial emotions. Most research was conducted using healthy volunteer samples (k = 50), and most studies examined acute doses (k = 43) or short-term administration of antidepressants (≤ 2 weeks; k = 18). There is therefore currently limited evidence on change in behavioural measures of emotional processing during antidepressant treatment in clinical settings.

Overall, my findings suggest that, at least amongst the measures currently used, there is not presently reliable evidence of a change in emotional processing following antidepressant administration. It is possible that identifying a more sensitive measure of differences in emotional processing associated with depression, such as the social evaluation learning task identified in chapters 1 and 2, may provide a more sensitive measure of antidepressant action. Additionally, my findings highlight the need for more ecologically valid research to understand whether findings translate to clinical practice.

Chapter 4: The effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers

In chapter 4 I addressed the second aim of this PhD in investigating whether administration of antidepressants is associated with a change in self-referential affective processing. Using a double-blind between-groups randomised design, healthy volunteers received an acute dose of citalopram (n = 20) or matched placebo (n = 21) and completed cognitive tasks measuring self-referential affective processing and social cognition. I hypothesised that acute administration of citalopram would be associated with an increase in positive affective biases about the self and increased prosocial behaviour.

In contrast to my hypothesis, I did not find evidence that change in affective biases following acute citalopram was heightened for information processed in reference to the self. Instead, exploratory analyses indicated that citalopram seemed to increase positive affective biases towards others. On the social evaluation learning task participants administered citalopram showed better learning of the positive ‘like’ rule and worse learning of the negative ‘dislike’ rule about a friend, but not the self. Additionally, participants administered citalopram recalled more likeable characteristics and less dislikeable characteristics about others, but no

differences were found for the self. Whilst these analyses were exploratory and based on a small sample, the consistency across these tasks supports the conclusion that change in affective biases was heightened for others rather than the self.

Whilst I previously found evidence in chapter 1 that depression was associated with increased positive biases towards others on a self-esteem go/no-go task, I did not find that citalopram impacted performance on this task. Previous research has reported inconsistent findings for the relationship between depression and response inhibition (Kilford et al., 2015; Lewis, Button, Pearson, Munafò, & Lewis, 2020). Additionally, through systematic review in chapter 3 I found that of the eight studies using a response inhibition task, only one found evidence in support of the cognitive neuropsychological model (Kometer et al., 2012). Therefore, measuring self-emotional processing using a go/no-go task does not appear to be a reliable measure of depression or antidepressant action. Additionally, I did not find evidence that citalopram altered performance on self-referential processing when measured independently of emotion and reward in a simple associative learning task. Participants prioritised learning to associate neutral shapes with the self and positive information, but task performance did not vary between participants administered citalopram versus placebo. I previously found no evidence of an association between accuracy or reaction times on this task with depression severity. Associative learning tasks where self, emotion and reward are processed independently do therefore not appear to be related to depression or antidepressant mechanisms.

In this chapter I also extended my previous work focusing on cognition to examine social behaviours. In support of my hypothesis, I found evidence that citalopram increased expressions of prosocial behaviours. Participants administered citalopram on average chose to cooperate on a greater number of trials in a prisoner's dilemma task compared to those administered placebo. Social context did not moderate this effect; participants in the citalopram group made greater cooperative choices irrespective of whether their opponent initially chose to cooperate or defect. A single dose of an antidepressant therefore appeared to increase prosocial behaviours regardless of the negativity or positivity of others' behaviours. In combination with our findings of increased positive social evaluation learning and recall biases towards others, my findings suggest that citalopram decreased sensitivity to negative characteristics and increased sensitivity to positive characteristics in others, which may have contributed to the greater expression of positive social behaviours in the citalopram group.

Chapter 5: The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care

In chapter 4 I found that in contrast to my expectations, antidepressants altered affective biases towards others rather than the self. However, self-schemas are argued to be heavily entrenched, to monopolise information processing and to be highly resistant to contradictory information (Beck, 2008). Acute pharmacological manipulation does not appear to be sufficient to alter learning of social evaluation biases about the self (Button, Karwatowska, Kounali, Munafò, & Attwood, 2016). Change in self-referential affective biases may therefore require longer-term antidepressant treatment than the single dose administered in chapter 4. Additionally, chapter 4 used a healthy volunteer sample. Although this allowed me to examine change in affective biases whilst controlling for 'cold' non-emotional cognitive impairments associated with depression, such as reduced working memory (Roiser & Sahakian, 2013), antidepressants are likely to alter emotional processing differently in healthy volunteers versus individuals experiencing depression (Ma, 2015). Moreover, chapter 3 highlighted the need for more ecologically valid research investigating antidepressant effects on affective processing to understand whether findings translate to clinical practice.

In chapter 5, I therefore investigated whether change in self-referential affective biases was associated with change in mood during the first eight-weeks of antidepressant treatment. Patients from primary care in the southwest of England completed self-report measures of depression and anxiety, and the social evaluation learning task at four timepoints over the first eight weeks of antidepressant treatment (n = 29). I hypothesised that patients would become better at learning positive social evaluations about the self and that this would be associated with a reduction in depression.

In contrast to my hypothesis, I did not find evidence that change in affective biases was heightened for the self. There was no evidence that biased learning of social evaluations changed over time, or that this differed by referential condition. Additionally, I did not find evidence that change in social evaluation learning over eight weeks of antidepressant treatment was associated with change in depression symptoms. I previously speculated that change in self-referential affective biases may only be observed from longer-term treatment. However, I did not find evidence to support this possibility. Despite most patients showing an improvement in depression, learning about the self was relatively stable throughout treatment. In relation to the second aim of this PhD my findings suggest that administration of antidepressants is not associated with change in self-referential affective processing.

However, in keeping with my findings in chapter 4, I found weak evidence of an association between my secondary measure of depression, the BDI-II, and change in social evaluation learning biases about a friend. On average, patients that showed better learning of positive versus negative evaluations about a friend showed a reduction in BDI-II scores. As we did not find evidence of this when depression was measured with the PHQ-9 it is possible that effects are related to specific symptoms measured by the BDI-II. One possibility is that social evaluation learning biases towards familiar others contribute to social anhedonia, measured by the BDI-II item 'interest in other people'. An increase in learning of positive characteristics in familiar others may help address social anhedonia by increasing positive perceptions of social interactions. However, evidence for this effect was weak, and these findings should therefore be treated with caution.

In contrast to my expectations, we found stronger evidence that change in social evaluation learning biases was more reliably associated with a change in anxiety symptoms rather than depression. On average, patients that became better at learning positive versus negative evaluations about both the self and the friend showed a reduction in anxiety symptoms. Cognitive models emphasise the role of a heightened sensitivity to threat in maintaining anxiety disorders (Mathews, Mackintosh, & Fulcher, 1997). It is possible that antidepressants may reduce sensitivity to negative information, such as critical social evaluations, helping to remediate these threat-related biases and subsequently reducing anxiety symptoms. Whilst these findings were based on exploratory analyses and must therefore be treated with caution, they are in line with recent research suggesting that change in affective biases may be more strongly linked to anxiety (Browning et al., 2021). Additionally, there is some evidence that change in anxiety may be a more important therapeutic indicator of response to the antidepressant sertraline than depression (Lewis et al., 2019). Focusing on change in affective processing as an indicator of a reduction in anxiety, rather than depression, may therefore be more beneficial (Browning et al., 2021).

However, these findings must be contextualised within the sample size for this study. Due to the COVID-19 pandemic our sample was substantially smaller than planned. We were therefore underpowered to detect small effects potentially contributing to false-negative findings (Vadillo, Konstantinidis, & Shanks, 2016). It is possible that there were some small changes in self-referential social evaluation learning that I could not detect. Subsequently whilst this research attempted to address the second aim of this PhD in understanding whether administration of antidepressant is associated with change in self-

referential affective processing, further research is required to draw more definitive conclusions. Similarly, whilst I found some evidence that change in social evaluation learning predicted change in anxiety (GAD-7) these findings are based on exploratory analyses with undetermined inflation of the alpha rate, potentially contributing to false-positive findings (Button et al., 2013).

Chapter 6: Is depression associated with reduced optimistic belief updating?

In my final chapter I moved beyond the original aims of this PhD to examine an alternative cognitive bias in depression, optimistic belief updating. Previous research has indicated that healthy individuals change their beliefs about negative life events more following good versus bad news (Sharot, Korn, & Dolan, 2011). However, studies suggest that individuals with depression lack this optimistic belief updating bias (Garrett et al., 2014; Korn, Sharot, Walter, Heekeren, & Dolan, 2014). In chapter 6 I aimed to replicate this effect and to expand upon previous findings to evaluate belief updating for positive life events. Due to COVID-19 I recruited individuals experiencing moderate to severe levels of depression (n = 54) and healthy controls (n = 56) online. Participants completed self-report measures of mood and a computerised belief-updating task for future positive and negative life events. I hypothesised that individuals with depression would show reduced optimistic belief updating relative to healthy controls for both positive and negative life events.

In keeping with my hypotheses, I replicated previously observed effects of reduced optimistic belief updating for negative life events in depression. Whereas healthy participants changed their beliefs more following good versus bad news, individuals experiencing depression changed their beliefs to a similar extent. Additionally, individuals experiencing depression had greater initial expectations of experiencing a negative life event in the future. Depression was therefore characterised by both initial pessimistic beliefs of experiencing a negative life event and reduced updating of these beliefs following novel positive information. Results for positive life events were less conclusive. In keeping with our expectations, the depression group showed reduced optimism for positive life events; providing lower estimates of experiencing positive events in their lifetime compared to the healthy group. However, we did not observe the same pattern of belief updating between groups as we found for negative life events. Scaled mean update scores suggested that both the healthy and depression groups showed greater updating towards good versus bad news for positive life events.

Our findings suggest that pessimistic beliefs about negative life events in depression may be maintained through a lack of biased updating. Whereas healthy individuals change their beliefs more when they receive good versus bad news, individuals with depression change their beliefs equally in response to good and bad news. It may be beneficial for depression treatments to attempt to increase accommodation of positive information to help remediate existing pessimistic beliefs. However, our findings suggest that this may only be effective in addressing beliefs regarding negative rather than positive life events.

Overall

Chapters 1 and 2 addressed the first aim of this PhD in evaluating whether depression is associated with differences in processing of self-related information, and whether this varies depending on the emotional valence or rewarding nature of information. Findings from chapter 1 indicated that depression was most reliably associated with change in processing of self-related information when this occurred in reference to emotional and rewarding information. Participants with greater depression were worse at learning positive versus negative social evaluations about the self in a reinforcement learning task. Chapter 2 validated these findings in two independent datasets; depression was again associated with reduced learning of positive versus negative social evaluations, although effects were weaker in participants that completed the study online independently. Treatments attempting to increase learning of positive social evaluations about the self may be helpful in addressing depressive symptoms.

Chapters 3 to 5 addressed the second aim of this PhD in understanding whether administration of antidepressants is associated with change in self-referential affective processing. Systematic review and meta-analysis in chapter 3 found that there is not currently reliable evidence that antidepressants alter positive and negative emotional processing using behavioural measures of various cognitive domains. In chapters 4 and 5 I therefore investigated whether affective biases occurring in reference to the self may provide a more sensitive indicator of early antidepressant action. I did not find evidence that antidepressants selectively altered self-referential affective biases in either an RCT of acute citalopram in healthy volunteers (chapter 4) or a prospective cohort of primary care patients over eight weeks of antidepressant treatment (chapter 5).

Contrarily, antidepressants seemed to increase positive affective biases towards others (chapter 4) which was associated with a reduction in the secondary measure of depression,

the BDI-II (chapter 5). Increased positive biases towards friends may address social anhedonia by increasing positive associations with social interactions, leading to greater future engagement, and exposing patients to greater environmental positive feedback that may over time remediate negative schema.

Additionally, exploratory analyses in chapter 5 also indicated strong evidence that change in social evaluation learning about both the self and a friend was associated with a reduction in anxiety. My findings were similar to recent research suggesting that change in affective biases may be more strongly linked to anxiety rather than depression (Browning et al., 2021). Antidepressants may treat anxiety symptoms by remediating negative affective biases towards socially threatening information directed towards the self and familiar others.

I also found some evidence that antidepressants increased the expression of positive social behaviours. In chapter 4, healthy volunteers administered a single dose of citalopram choose to cooperate on a greater number of trials on a prisoners' dilemma task irrespective of the actions of their opponent. An increase in positive social behaviours may be beneficial in remediating depression by increasing the quality of social relationships, subsequently providing greater social support.

Finally, in chapter 6, I examined a novel area of cognitive biases, optimistic belief updating. I successfully replicated previous evidence of reduced optimistic belief updating for negative life events in individuals experiencing depression. Whereas healthy individuals updated their beliefs more following good versus bad news, individuals experiencing depression updated their beliefs equally to good and bad news. These findings were in line with biases in social evaluation learning observed in chapter 1. In both the optimistic belief updating task and social evaluation learning task, individuals with greater depression showed a reduced ability to accommodate positive information.

Clinical Implications

Across studies, individuals experiencing depression consistently demonstrated a reduced ability to learn positive, socially rewarding information versus negative, socially critical information about the self. Reduced positive self-referential social evaluation learning biases may play a role in maintaining depressive self-schema by reinforcing negative perceptions of the self and reducing exposure to alternative, more positive information. Raised expectations of negative social evaluations may also evoke hostile or withdrawn social behaviours, increasing the likelihood of reciprocal negative behaviours from others. This may decrease

the quality of current relationships as well as preventing individuals from developing new social connections, reducing the availability of social support (Lewinsohn, 1975; Tse & Bond, 2003). Social evaluation learning may therefore provide an important and potentially reversible target for therapeutic intervention that can address impairments in social functioning, negative self-schema, and depressive symptoms.

Whilst I did not find reliable evidence that antidepressants alone altered learning of social evaluations about the self, it is possible that this may be more effectively targeted using a combination of antidepressants and cognitive behavioural therapy (CBT). Previous research has reported change in self-schema following treatment combining antidepressants and CBT but not antidepressants alone (Dozois et al., 2009). CBT is likely to be particularly beneficial in remediating social evaluation learning biases as it aims to modify negative cognitive styles maintaining negative self-schema and depressive symptoms (Power, 2013). Additionally, combining antidepressants and CBT allows remediation of both bottom-up, automatic, and top-down, deliberative processes believed to drive negative affective biases (Roiser et al., 2012). However, my findings also highlight the importance of explicitly addressing the referential perspective of social evaluation learning biases. Whilst acknowledging that CBT techniques already typically aim to address self-referential biases, my findings emphasise the importance of explicitly manipulating the referential focus towards the self when attempting to remediate negative cognitive styles within social contexts.

Despite research indicating that antidepressants are effective in treating depression (Cipriani et al., 2018), a large proportion of individuals do not respond to the first antidepressant prescribed (Warden, Rush, Trivedi, Fava, & Wisniewski, 2007). It is not currently possible to predict early in treatment whether an antidepressant will be effective. It has been suggested that change in affective processing may be used as a potential biomarker of antidepressant efficacy (Browning et al., 2019). By identifying patients with limited change in affective processing shortly after starting antidepressant treatment it may be possible to alter their treatment at an earlier timepoint. However, I did not find evidence from systematic review and meta-analysis that antidepressants were reliably associated with a change in processing of emotional information. Whilst most research was conducted within healthy volunteers using short-term administration of antidepressants, this suggests that use of affective processing as a biomarker of response may have limited effectiveness. Attempts to personalise treatment through change in affective processing does not at present seem feasible, at least amongst the behavioural measures currently used.

Future research

In this PhD I focused on the role of affective biases occurring in relation to the self. In contrast to my expectations, I found that affective biases towards others was more reliably altered following antidepressant administration. However, the importance of such change is unclear. Whilst I found weak evidence that an increased positive bias when learning about a friend was associated with change in self-reported depression during antidepressant treatment, in previous chapters I did not find evidence that individuals experiencing depression differed in bias towards a friend. It is therefore unclear as to how altering affective social biases towards friends has therapeutic benefits. One possibility is that antidepressants produce indirect changes in social cognition that are not a reversal of a cause of depression but have beneficial implications for treatment (Healy & McMonagle, 1997). Whilst an increased sensitivity to negative evaluations about a friend does not appear to be associated with maintenance of depression, increasing sensitivity to positive characteristics in friends may be one pathway to address wider depressive symptoms. For example, increased positive biases towards friends may address social anhedonia by increasing positive perceptions within social contexts, leading to greater engagement in future social interactions, and exposing patients to greater environmental positive feedback (Lewinsohn, 1975). Further research is required to clarify how change in social cognition towards friends may provide therapeutic benefits.

Another possibility is that an increase in positive affective biases towards others increases expressions of positive social behaviours. It has been suggested that antidepressants may operate in part by reducing expression of hostile behaviours and increasing affiliative behaviours (Young, Moskowitz, & aan het Rot, 2014). In support of this theory and previous research (Knutson et al., 1998; Tse & Bond, 2002), I found that participants administered a single dose of citalopram chose to cooperate on a greater number of trials on a prisoners' dilemma task regardless of the actions of their opponent. Additionally, in this study I also found that citalopram increased learning and recall of positive characteristics about a friend. It is possible that increased positive affective biases towards others may have contributed to the increased expression of cooperative behaviours. However, there has been relatively little research investigating the effect of antidepressants on social behaviours, particularly increased positive behaviours (Young et al., 2014). Furthermore, to the best of my knowledge there has been no research examining change in affective biases in relation to change in social behaviour during antidepressant treatment. Future research understanding the temporal

relationship between affective cognition, social behaviours, and depression symptoms would be beneficial in refining our understanding of antidepressant mechanisms.

Furthermore, based on exploratory analyses I found stronger evidence that change in social evaluation learning was associated with anxiety rather than depression during antidepressant treatment. Primary care patients that showed an increase in learning of positive relative to negative evaluations about both the self and friend on average showed a reduction in anxiety symptoms over eight weeks of antidepressant treatment. My findings were in line with a recent RCT in primary care where patients whose treatment was guided by change in affective processing showed a greater decline in anxiety but not depression (Browning et al., 2021). As with depression, individuals with anxiety are believed to hold information processing biases, but these tend to be towards threatening rather than emotional stimuli (Beck, 1976). It is possible that some antidepressants are more effective at addressing threat-related biases. Negative social evaluations, such as criticism, may represent a source of social threat that anxious individuals implicitly interpret as potential acts of aggression. Antidepressants may reduce sensitivity to negative information, helping to remediate these threat-related biases within social contexts, and subsequently reducing anxiety symptoms. However, these findings are based on exploratory analyses. Further research investigating change in affective processing in relation to anxiety symptoms in antidepressant treatment is required to validate my findings.

Moreover, it is also possible that optimistic belief updating may be a sensitive psychopharmacological measure for antidepressant action in relation to a reduction in depression. Replicating previous research, I found that individuals experiencing depression show reduced optimistic belief updating for negative life events. Whereas healthy individuals preferentially update their expectations of experiencing a negative life event in response to good news, individuals experiencing depression lack this bias, altering their expectations equally in response to good and bad news. Reduced optimistic belief updating may play a key role in reinforcing negative beliefs regarding the self, the world, and the future, associated with depression (Beck, 1987). In line with the predictions of the cognitive neuropsychological model we would expect antidepressants to increase sensitivity to positive environmental information, potentially remediating reduced optimism by increasing updating of expectations in response to positive news. Providing preliminary support of this possibility, recent research in treatment resistant patients reported that ketamine increased positive belief updating which mediated a reduction in depression (Bottemanne et al., 2021).

Whilst research within this area is still at a very early stage and has not examined conventional antidepressant treatments, it supports the possibility that optimistic belief updating may be associated with antidepressant response.

My findings were restricted to adult samples. However, it is possible that the relationship between social evaluation learning and depression varies across developmental periods. Adolescence may be a sensitive period for development of reduced positive social biases as it is a period in which perceptions of the self are developing (Sebastian, Burnett, & Blakemore, 2008) and the opinion of peers is highly important (Rubin, Bukowski, & Parker, 2006). In line with my findings, recent research has indicated that adolescents with greater depression show reduced positive social evaluation learning biases about the self (Bone et al., 2021). However, in contrast to my findings, adolescents also showed reduced positive social evaluation learning biases towards a stranger (Bone et al., 2021). The nature of social evaluation learning in relation to depression therefore appears to differ by age. However, this study lacked an adult comparison group. It is possible that differences in findings may reflect differences in study methodologies. Bone and colleagues recruited and collected data in school classrooms whereas, participants in my studies completed the tasks independently. It is likely that completing the task in the presence of peers alters the social context of data collection, potentially altering task performance (Breiner et al., 2018). Further research examining the relationship between social evaluation learning and depression by age would be helpful in refining appropriate targets for potential interventions specific to each developmental period.

Strengths

I took a rigorous scientific approach by incorporating open science principles throughout my thesis. I pre-registered all primary analyses and hypotheses for each study to ensure the integrity of my findings. Additionally, I published chapter 6 in the format of a registered report with peer review conducted before and after data collection. I have made all data collected during this PhD, and code for statistical analyses using non-proprietary software ('R'), available as open access to allow others to replicate my findings. I have also published all materials for which myself or my co-authors hold copyright to allow replication of my methods.

Furthermore, I used a wide range of methodologies to triangulate my findings. This included cross-sectional studies examining a range of depression severity (chapters 1 and 2),

systematic review and meta-analysis (chapter 3), a double-blind RCT (chapter 4), a prospective cohort of patients recruited from primary care (chapter 5), and a case-control study examining differences in cognitive biases between depression versus healthy controls (chapter 6). Additionally, I examined the influence of antidepressants in both controlled conditions, administering an acute dose in healthy volunteers, as well as naturalistically, observing primary care patients receiving antidepressant treatment. By combining methods to address the aims of thesis I have been able to provide a comprehensive overview of the role of self-referential affective processing in depression and antidepressant treatment.

Limitations

Although I found reliable evidence that depression was associated with social evaluation learning and optimistic belief updating biases, I was unable to comment on the causal importance of these associations. It is possible that change in these biases is symptomatic of depression rather than playing a mechanistic role. Examining change in social evaluation learning and optimistic belief updating in relation to mood within a longitudinal cohort would help clarify their causal importance. Alternatively, attempting to directly manipulate these biases through techniques such as cognitive bias modification and observing subsequent change in mood would be helpful in clarifying the direction of their relationship with depression.

Additionally, my sample sizes were smaller than planned in chapters 4 and 5 due to the COVID-19 pandemic occurring during recruitment. A smaller sample size than anticipated is likely to have limited my ability to detect hypothesised effects, potentially producing false-negative findings (Vadillo et al., 2016). It is possible that antidepressants influence self-referential affective processing but that effects were smaller than I was able to detect. Alternatively, small sample sizes are believed to inflate false discovery rates (Button et al., 2013). It is therefore possible that some of my findings are false positives. Additionally, due to my small sample size in chapter 5 I was limited in my ability to run more complex exploratory analyses that may have allowed me to comment on causal associations (Hecht & Zitzmann, 2021).

Furthermore, I made several adaptations to study methodologies to comply with COVID-19 social distancing measures that may have influenced the scientific rigour of my research. Most data for chapters 5 and 6 was collected remotely. Data quality may therefore have been lower than in-person research previously conducted in this field. Additionally, for chapter 6, I

originally planned to recruit individuals meeting diagnostic criteria for depression to replicate previous research (Korn et al., 2014). To run the study remotely within the timeframe available I altered our inclusion criteria to participants reporting moderate to severe symptoms of depression on the PHQ-9. It is possible that my recruitment methods may have led to differences in sample characteristics to those of previous studies, potentially influencing our findings.

Finally, I used computerised cognitive tasks to measure affective processing. Whilst this allowed me to obtain standardised quantitative measures of implicit affective processing under controlled conditions across relatively large samples, it is unclear whether my findings translate to real life settings. Future research employing a more naturalistic approach to understanding the relationship between affective processing, social cognition, and social behaviours in relation to depression using techniques such as ecological momentary assessment would help to validate my findings.

Conclusion

Overall, depression was reliably associated with a reduced ability to learn positive versus negative evaluations about the self. Although I did not find evidence that social evaluation learning biases towards the self were altered by antidepressants, psychological therapies aiming to increase sensitivity to positive social evaluations may be beneficial in remediating negative self-schema believed to maintain depression. I also found reliable evidence that individuals experiencing depression show a reduced ability to update beliefs regarding negative life events in response to positive news. Again, psychological therapies may be useful in remediating these biases, although preliminary research suggests that antidepressants may also be beneficial (Bottemanne et al., 2021). Additionally, although I did not find evidence that self-referential affective biases were important in remediating depression in early antidepressant treatment, I identified several potential psychopharmacological mechanisms of antidepressants for future research. Firstly, antidepressants may operate by increasing sensitivity to positive information about familiar others. This may treat depression by increasing enjoyment and engagement in social interactions. Secondly, increased positive affective biases towards others may also contribute to increased expressions of positive social behaviours. This is likely to increase the likelihood of reciprocal positive behaviours from others, further reinforcing reward from social interactions. Finally, exploratory analyses suggested that change in affective biases may be important in remediating anxiety. It is possible that antidepressants operate in part by

remediating threat-related biases, such as an increased sensitivity to socially critical information, that maintain anxiety symptoms.

References

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International University Press.

Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy, 1*, 5–37.

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry, 165*(8), 969–977. https://doi.org/10.1176/appi.ajp.2008.08050721

Bone, J. K., Pike, A. C., Lewis, G., Lewis, G., Blakemore, S.-J., & Roiser, J. P. (2021). Computational mechanisms underlying social evaluation learning and associations with depressive symptoms during adolescence [Preprint]. *PsyArXiv*, 1–21. https://doi.org/10.31234/OSF.IO/9M7VR

Bottemanne, H., Morlaàs, O., Claret, A., Fossati, P., & Schmidt, L. (2021). Can we change depressive beliefs? Modulation of belief updating mechanisms by ketamine in treatment resistant depression [Preprint]. *PsyArXiv*, 1–38. https://doi.org/10.31234/OSF.IO/ZX5VR

Breiner, K., Li, A., Cohen, A. O., Steinberg, L., Bonnie, R. J., Scott, E. S., … Galván, A. (2018). Combined effects of peer presence, social cues, and rewards on cognitive control in adolescents. *Developmental Psychobiology, 60*(3), 292–302. https://doi.org/10.1002/dev.21599

Browning, M., Bilderbeck, A. C., Dias, R., Dourish, C. T., Kingslake, J., Deckert, J., … Dawson, G. R. (2021). The clinical effectiveness of using a predictive algorithm to guide antidepressant treatment in primary care (PReDicT): an open-label, randomised controlled trial. *Neuropsychopharmacology, 46*(7), 1307–1314. https://doi.org/10.1038/s41386-021-00981-z

Browning, M., Kingslake, J., Dourish, C. T., Goodwin, G. M., Harmer, C. J., & Dawson, G. R. (2019). Predicting treatment response to antidepressant medication using early changes in emotional processing. *European Neuropsychopharmacology, 29*(1), 66–75. https://doi.org/10.1016/j.euroneuro.2018.11.1102

Button, K. S., Browning, M., Munafò, M. R., & Lewis, G. (2012). Social inference and social

anxiety: Evidence of a fear-congruent self-referential learning bias. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(4), 1082–1087. https://doi.org/10.1016/j.jbtep.2012.05.004

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Button, K. S., Karwatowska, L., Kounali, D., Munafò, M. R., & Attwood, A. S. (2016). Acute anxiety and social inference: An experimental manipulation with 7.5% carbon dioxide inhalation. *Journal of Psychopharmacology*, *30*(10), 1036–1046. https://doi.org/10.1177/0269881116653105

Button, K. S., Kounali, D., Stapinski, L., Rapee, R. M., Lewis, G., & Munafò, M. R. (2015). Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. *PLoS ONE*, *10*(4), e0119456. https://doi.org/10.1371/journal.pone.0119456

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., … Geddes, J. R. (2018). Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *FOCUS*, *16*(4), 420–429. https://doi.org/10.1176/appi.focus.16407

Dozois, D. J. A., Bieling, P. J., Patelis-Siotis, I., Hoar, L., Chudzik, S., McCabe, K., & Westra, H. A. (2009). Changes in Self-Schema Structure in Cognitive Therapy for Major Depressive Disorder: A Randomized Clinical Trial. *Journal of Consulting and Clinical Psychology*, *77*(6), 1078–1088. https://doi.org/10.1037/a0016886

Everaert, J., Duyck, W., & Koster, E. H. W. (2014). Attention, interpretation, and memory biases in subclinical depression: A proof-of-principle test of the combined cognitive biases hypothesis. *Emotion*, *14*(2), 331–340.

Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the rose tinted glasses: Neural substrates of unbiased belief updating in depression. *Frontiers in Human Neuroscience*, *8*, 1–9. https://doi.org/10.3389/fnhum.2014.00639

Harmer, C. J., Hill, S. A., Taylor, M. J., Cowen, P. J., & Goodwin, G. M. (2003). Toward a neuropsychological theory of antidepressant drug action: Increase in positive emotional bias after potentiation of norepinephrine activity. *American Journal of Psychiatry*, *160*(5), 990–992. https://doi.org/10.1176/appi.ajp.160.5.990

Healy, D., & McMonagle, T. (1997). The enhancement of social functioning as a therapeutic principle in the management of depression. *Journal of Psychopharmacology*, *11*(4), S25–S31.

Hecht, M., & Zitzmann, S. (2021). Sample Size Recommendations for Continuous-Time Models: Compensating Shorter Time Series with Larger Numbers of Persons and Vice Versa. *Structural Equation Modeling*, *28*(2), 229–236. https://doi.org/10.1080/10705511.2020.1779069

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 1–16. https://doi.org/10.1186/2045-5380-3-12

Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., … Stringaris, A. (2018). Reward processing in depression: A conceptual and meta-analytic review across fMRI and EEG studies. *American Journal of Psychiatry*, *175*(11), 1111–1120. https://doi.org/10.1176/appi.ajp.2018.17101124

Kilford, E. J., Foulkes, L., Potter, R., Collishaw, S., Thapar, A., & Rice, F. (2015). Affective bias and current, past and future adolescent depression: A familial high risk study. *Journal of Affective Disorders*, *174*, 265–271. https://doi.org/10.1016/j.jad.2014.11.046

Knutson, B., Wolkowitz, O. M., Cole, S. W., Chan, T., Moore, E. A., Johnson, R. C., … Reus, V. I. (1998). Selective alteration of personality and social behavior by serotonergic intervention. *American Journal of Psychiatry*, *155*(3), 373–379. https://doi.org/10.1176/ajp.155.3.373

Kometer, M., Schmidt, A., Bachmann, R., Studerus, E., Seifritz, E., & Vollenweider, F. X. (2012). Psilocybin biases facial recognition, goal-directed behavior, and mood state toward positive relative to negative emotions through different serotonergic subreceptors. *Biological Psychiatry*, *72*(11), 898–906. https://doi.org/10.1016/j.biopsych.2012.04.005

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*(3), 579–592. https://doi.org/10.1017/S0033291713001074

Lewinsohn, P. M. (1975). The behavioral study and treatment of depression. In *Progress in behavior modification* (Vol. 1, pp. 19–64). Elsevier.

Lewis, G., Button, K. S., Pearson, R. M., Munafò, M. R., & Lewis, G. (2020). Inhibitory control of positive and negative information and adolescent depressive symptoms: a population-based cohort study. *Psychological Medicine*, 1–11. https://doi.org/https://doi.org/10.1017/S0033291720002469

Lewis, G., Duffy, L., Ades, A., Amos, R., Araya, R., Brabyn, S., … Lewis, G. (2019). The clinical effectiveness of sertraline in primary care and the role of depression severity and duration (PANDA): a pragmatic, double-blind, placebo-controlled randomised trial. *The Lancet Psychiatry*, *6*(11), 903–914. https://doi.org/10.1016/S2215-0366(19)30366-9

Ma, Y. (2015). Neuropsychological mechanism underlying antidepressant effect: A systematic meta-analysis. *Molecular Psychiatry*, *20*(3), 311–319. https://doi.org/10.1038/mp.2014.24

Mathews, A., Mackintosh, B., & Fulcher, E. P. (1997). Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences*, *1*(9), 340–345. https://doi.org/10.1016/S1364-6613(97)01092-9

Northoff, G. (2007). Psychopathology and pathophysiology of the self in depression - Neuropsychiatric hypothesis. *Journal of Affective Disorders*, *104*(1–3), 1–14. https://doi.org/10.1016/j.jad.2007.02.012

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*(11), 1019–1025. https://doi.org/10.1016/j.biopsych.2010.12.014

Ossola, P., Garrett, N., Sharot, T., & Marchesi, C. (2020). Belief updating in bipolar disorder predicts time of recurrence. *ELife*, *9*, 1–17. https://doi.org/10.7554/eLife.58891

Power, M. (2013). CBT for Depression. In *The Wiley-Blackwell Handbook of Mood Disorders* (pp. 173–191). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118316153.ch7

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive mechanisms of treatment in

depression. *Neuropsychopharmacology*, *37*(1), 117–136. https://doi.org/10.1038/npp.2011.183

Roiser, J. P., & Sahakian, B. J. (2013). Hot and cold cognition in depression. *CNS Spectrums*, *18*(3), 139–149. https://doi.org/10.1017/S1092852913000072

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer Interactions, Relationships, and Groups. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 571–645). John Wiley & Sons, Ltd.

Sebastian, C., Burnett, S., & Blakemore, S. J. (2008). Development of the self-concept during adolescence. *Trends in Cognitive Sciences*, *12*(11), 441–446. https://doi.org/10.1016/j.tics.2008.07.008

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. https://doi.org/10.1038/nn.2949

Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, *70*(6), 1011–1022. https://doi.org/10.1080/17470218.2015.1101477

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1105. https://doi.org/10.1037/a0029792

Sui, J., & Humphreys, G. W. (2015). The interaction between self-bias and reward: Evidence for common and distinct processes. *Quarterly Journal of Experimental Psychology*, *68*(10), 1952–1964. https://doi.org/10.1080/17470218.2015.1023207

Tse, W. S., & Bond, A. J. (2002). Serotonergic intervention affects both social dominance and affiliative behaviour. *Psychopharmacology*, *161*(3), 324–330. https://doi.org/10.1007/s00213-002-1049-7

Tse, W. S., & Bond, A. J. (2003). Consequences of displaying abnormal social behaviour: Avoidance and reduction of social reinforcement. *Journal of Affective Disorders*, *75*(1), 49–58. https://doi.org/10.1016/S0165-0327(02)00036-8

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin and Review*, *23*(1), 87–102. https://doi.org/10.3758/s13423-015-0892-6

Warden, D., Rush, A. J., Trivedi, M. H., Fava, M., & Wisniewski, S. R. (2007). The STAR*D project results: A comprehensive review of findings. *Current Psychiatry Reports*, *9*(6), 449–459. https://doi.org/10.1007/s11920-007-0061-3

Young, S. N., Moskowitz, D. S., & aan het Rot, M. (2014). Possible role of more positive social behaviour in the clinical effect of antidepressant drugs. *Journal of Psychiatry and Neuroscience*, *39*(1), 60–65. https://doi.org/10.1503/jpn.130165

Appendices

Chapter 1 Supplementary Materials

Supplementary materials for Hobbs, C., Sui, J., Kessler, D., Munafò, M. R., & Button, K. S. (2021). Self-processing in relation to emotion and reward processing in depression. *Psychological Medicine*, 1-13.

Statistical Models

Associative Learning Task

In all analyses, tasks (e.g., self, reward and emotion) were modelled separately.

We first examined whether task performance differed according to condition using mixed-effects linear regression models. Accuracy (%) and reaction times (ms) were entered as outcomes in separate models, and condition entered as a categorical predictor. Subject was included as a random effect to account for the within-subject effect of condition.

We then examined whether task performance (accuracy and reaction times) was associated with depression severity using linear regression models. In each model task outcomes were entered as predictors in wide format according to each condition (e.g., for the self task average accuracy/reaction times in the self, friend and stranger condition were entered as separate predictors). PHQ-9 or BDI-II total scores were used as continuous outcomes in separate models.

Go/No-Go Self-Esteem Task

We first examined whether task performance differed according to condition using mixed-effects linear regression models. Discriminative accuracy was entered as the outcome, and referential-emotion condition was entered as the predictor. Subject was included as a random effect to account for the within-subject effect of condition.

To explore whether depression was associated with discriminative accuracy when categorising positive and negative words with the self and others we used linear regression models with PHQ-9 and BDI-II scores as a continuous outcomes in individual models. Discriminative accuracy in each referential-valence combination (self-positive, self-negative, other-positive, other-negative) were entered as predictors.

As we excluded a high proportion of participants (25%) due to non-compliance in our main analysis, we repeated this model including all participants as a sensitivity analysis.

Social Evaluation Learning Task

We first examined whether task performance differed according to condition using mixed-effects linear regression models. As this task was completed over two sessions we included session as a categorical predictor in all models. We accounted for within-subject effects by including subject as a random effect. Our first model used bias scores as a continuous outcome and referential condition as a categorical predictor. Our second model used errors to criterion as a continuous outcome, and referential condition, rule and an interaction between referential condition and rule as predictors. Our third model used the same predictors, with global ratings as a continuous outcome.

We then used a mixed-effects linear regression model to estimate the relationship between biased learning and depression severity. PHQ-9 or BDI-II scores were entered as a continuous outcome and bias scores (like-dislike) in each referential condition were entered as predictors in wide format. To account for the within-subject design across testing sessions, session was entered as a fixed effect and subject was entered as a random effect.

To examine whether the relationship between self bias scores and depression symptoms was consistent across sessions we conducted an exploratory analysis using mixed-effects linear regression. Again PHQ-9 and BDI-II scores were used as outcomes in separate models. Bias scores in the self condition, session and an interaction between these terms were used as predictors.

To understand the relative contribution of learning the ‘like’ versus ‘dislike’ rules on overall biased learning we conducted another mixed-effects linear regression model. PHQ-9 and BDI-II scores were entered as a continuous outcomes in separate models, and errors to criterion in each referential condition – rule combination were entered as predictors in wide format. To account for the multiple testing sessions, session was entered as a fixed effect and subject was entered as a random effect. We also repeated this model with global ratings for each referential-condition-rule combination as predictors, to assess the consistency of deliberative versus automatic learning.

Finally, we assessed whether these our findings regarding the association between depression and performance on this task were maintained when social anxiety was taken into account by repeating these models with BFNE scores entered as an additional predictor.

Primary Diagnosis of Major Depressive Episode

To examine whether our findings were valid for participants meeting clinical diagnostic criteria for depression, we repeated the primary analyses examining the association between task performance and depression for each task using logistic regression models with a binary variable of meeting diagnostic criteria as a primary diagnosis of Major Depressive Episode on the CIS-R as the outcome. Not meeting diagnostic criteria for a primary MDE was used as the reference category in all analyses. As the CIS-R was only completed at session 1, for tasks with multiple sessions only data from session 1 was used.

Adjusting for Age and Gender

We repeated the analyses for each task examining the association between task performance and depression with age (continuous) and gender (categorical) included as additional predictors.

Table S1

Mean (SD) accuracy (%) and reaction times for the Associative Learning Tasks

	Self			Reward			Emotion		
	Self	Friend	Stranger	£9	£3	£1	Happy	Neutral	Sad
Accuracy (%)	88.01 (10.34)	83.55 (11.67)	80.77 (13.09)	79.67 (15.39)	75.23 (15.61)	74.98 (17.04)	80.06 (15.49)	69.28 (16.07)	64.64 (16.26)
Reaction Times, (ms)	676 (70)	706 (80)	712 (77)	681 (91)	712 (98)	700 (101)	710 (114)	750 (121)	748 (121)

Table S2

Results from a mixed-effects linear regression model examining differences in task performance (accuracy and reaction times; outcomes) according to stimuli in self, reward and emotion associative learning tasks (predictors)

		Accuracy (%)				**Reaction Times (ms)**				
		b	***b* 95% CI**	**β**	**β 95 % CI**	**p**	***b***	***b* 95% CI**	**β**	**β 95 % CI**	**p**
Self											
	Self (reference)	88.01	86.09, 89.94	0.32	0.16, 0.48	< .001	676.20	663.79, 688.60	-0.28	-0.44, -0.12	< .001
	Friend	-4.46	-7.19, -1.74	-0.37	-0.51, -0.23	0.001	29.29	11.75, 46.83	0.38	0.29, 0.47	0.001
	Stranger	-7.24	-9.97, -4.52	-0.60	-0.74, -0.46	< .001	36.21	18.67, 53.75	0.47	0.38, 0.56	< .001
Reward											
	High (£9) (reference)	79.67	77.05, 82.30	0.19	0.03, 0.35	< .001	680.72	664.91, 696.52	-0.17	-0.33, -0.01	< .001
	Medium (£3)	-4.45	-8.16, -0.73	-0.28	-0.41, -0.14	0.019	31.25	53.60, 2.75	0.32	0.24, 0.40	0.006

	Low (£1)	-4.69	-8.41, --0.98	-0.29	-0.43, -0.15	0.013	19.07	-3.28, 41.42	0.20	0.11, 0.28	0.094
Emotion											
	Happy (reference)	80.06	77.45, 82.67	0.51	0.36, 0.66	< .001	709.46	690.05, 728.87	-0.22	-0.38, -0.06	< .001
	Neutral	-10.78	-14.47, -7.08	-0.63	-0.76, -0.49	< .001	40.03	12.58, 67.48	0.33	0.26, 0.41	0.004
	Sad	-15.41	-19.11, -11.72	-0.90	-1.03, -0.76	< .001	38.01	10.56, 65.46	0.32	0.25, 0.39	0.007

b = unstandardised regression coefficients, β = standardised regression coefficients

Note: Accuracy and reaction times were entered as outcomes in separate models, and tasks were analysed separately.

Table S3

Results from a mixed-effects linear regression model examining differences in discriminative accuracy (outcome) according to referential condition (self/other) and emotion (positive/negative) (predictors)

	b	*b* 95% CI	β	β 95% CI	p
Intercept	1.40	1.30, 1.50	0.56	0.38, 0.73	< .001
Referential Condition	-0.69	-0.81, -0.56	-1.14	-1.36, -0.93	< .001
Emotion	-0.38	-0.50, -0.25	-0.63	-0.84, -0.41	< .001
Condition*Emotion	0.79	0.61, 0.97	1.31	1.01, 1.61	< .001

b = unstandardised regression coefficients, β = standardised regression coefficients

Note: Self was the reference category for the condition variable, positive was the reference category for the emotion variable

Table S4

Results from linear regression models examining the association between discriminative accuracy in each referential-emotion condition (predictors) with PHQ-9 and BDI-II Scores (outcomes) in participants demonstrating task compliance according to a priori criteria (n = 108) and all participants (n = 144)

	PHQ-9					BDI-II				
	b	*b* 95% CI	β	β 95% CI	p	β	95% CI	*b*	*b* 95% CI	p
Task Compliant (n = 108)										
Intercept	12.08	8.81, 15.35	0.00	-0.17, 0.17	< .001	20.84	13.87, 27.82	0.00	-0.18, 0.18	<.001
Self-Positive	-2.47	-4.54, -0.39	-0.24	-0.44, -0.04	0.020	-3.20	-7.62, 1.23	-0.15	-0.36, 0.06	0.155
Self-Negative	-0.59	-2.57, 1.39	-0.05	-0.24, 0.13	0.553	-0.81	-5.03, 3.41	-0.04	-0.22, 0.15	0.704
Other-Positive	3.51	1.24, 5.79	0.30	0.10, 0.49	0.003	6.78	1.93, 11.64	0.28	0.08, 0.47	0.007
Other-Negative	-2.46	-4.24, -0.68	-0.27	-0.46, -0.07	0.007	-5.13	-8.92, -1.34	-0.27	-0.46, -0.07	0.008
All participants (n = 144)										
Intercept	9.84	7.83, 11.85	0.00	-0.16, 0.16	< .001	16.92	12.72, 21.11	0.00	-0.16, 0.16	< .001

Self-Positive	-2.10	-3.78, -0.42	-0.24	-0.44, -0.05	0.015	-2.39	-5.89, 1.12	-0.13	-0.33, 0.06	0.181
Self-Negative	0.05	-1.56, 1.66	0.00	-0.17, 0.18	0.949	0.32	-3.04, 3.68	0.02	-0.16, 0.20	0.850
Other-Positive	2.81	0.97, 4.66	0.28	0.10, 0.46	0.003	5.17	1.31, 9.02	0.25	0.006, 0.44	0.009
Other-Negative	-1.35	-2.82, 0.11	-0.17	-0.35, 0.01	0.070	-3.24	-6.30, -0.18	-0.20	-0.38, -0.01	0.038

b = unstandardised regression coefficients, β = standardised regression coefficients

Note: A priori criteria for non-compliance was discrimination scores lower than 5 and/or bias scores less than 12 or greater than 36. Separate analyses were conducted including task compliant participants (n = 108) and all participants (n = 144).

Table S5

Results from mixed-effects linear regression models examining differences in measures of performance on the Social Evaluation Learning task (outcome) according to referential condition and rule (predictors)

	b	***b* 95% CI**	**β**	**β 95% CI**	**p**
Bias Scores					
Intercept	-1.26	-2.55, 0.04	0.14	0.02, 0.27	0.057
Referential Condition					< .001
Self (reference)					
Friend	-2.07	-2.93, -1.21	-0.35	-0.49, -0.20	< .001
Stranger	-0.44	-1.31, 0.42	-0.07	-0.22, 0.07	0.318
Session	-0.13	-0.83, 0.58	-0.01	-0.07, 0.05	0.726
Errors to Criterion					
Intercept	6.19	5.46, 6.92	-0.17	-0.29, -0.05	< .001
Referential Condition					< .001
Self (reference)					
Friend	-1.12	-1.75, -0.50	-0.27	-0.41, -0.12	< .001
Stranger	-0.16	-0.78, 0.46	-0.04	-0.19, 0.11	0.615

Rule					< .001
Positive (reference)					
Negative	1.45	0.82, 2.07	0.34	0.19. 0.49	< .001
Session	0.21	-0.15, 0.57	0.02	-0.02, 0.07	0.254
Condition*Rule					< .001
Self (reference)					
Friend	2.07	1.19, 2.95	0.49	0.28, 0.70	< .001
Stranger	0.44	-0.44, 1.32	0.10	-0.10, 0.31	0.328
Global Ratings					
Intercept	5.61	5.39, 5.83	0.66	0.58, 0.75	< .001
Referential					< .001
Condition					
Self (reference)					
Friend	0.32	0.14, 0.50	0.17	0.07, 0.27	0.001
Stranger	0.09	-0.10, 0.27	0.05	-0.05, 0.15	0.354
Rule					< .001
Positive (reference)					
Negative	-2.67	-2.85, -2.49	-1.47	-1.57, -1.37	< .001
Session	0.22	0.12, 0.33	0.06	0.03, 0.09	< .001
Condition*Rule					0.985
Self (reference)					
Friend	0.01	-0.25, 0.26	0.00	-0.14, 0.15	0.947
Stranger	-0.01	-0.27, 0.24	-0.01	-0.15, 0.13	0.915

b = unstandardised regression coefficients, β = standardised regression coefficients

Table S6

Results from mixed-effects linear regression models examining the effect of Social Evaluation Learning task outcomes according to referential condition and rule on depression severity after adjusting for social anxiety (BFNE scores)

	PHQ-9					BDI-II				
	b	*b* 95% CI	β	β 95% CI	p	β	95% CI	*b*	*b* 95% CI	p
Bias Scores										
Intercept	3.44	0.91, 5.97	0.00	-0.14, 0.14	0.008	4.88	0.00, 9.77	0.00	-0.14, 0.14	0.051
Self	0.10	0.04, 0.16	0.11	0.04, 0.19	0.002	0.21	0.10, 0.32	0.12	0.06, 0.18	< .001
Friend	-0.03	-0.09, 0.03	-0.03	-0.10, 0.03	0.292	0.01	-0.10, 0.11	0.00	-0.05, 0.06	0.895
Stranger	0.01	-0.06, 0.07	0.01	-0.06, 0.07	0.859	0.03	-0.09, 0.14	0.01	-0.04, 0.07	0.678
Session	-0.75	-1.18, -0.32	-0.07	-0.11, -0.03	0.001	-0.46	-1.23, 0.31	-0.02	-0.05, 0.01	0.244
BFNE	0.13	0.07, 0.19	0.24	0.13, 0.35	< .001	0.26	0.15, 0.37	0.23	0.13, 0.33	< .001
Errors to Criterion										
Intercept	2.55	-0.11, 5.20	0.00	-0.14, 0.14	0.061	4.06	-1.10, 9.23	0.00	-0.14, 0.14	0.124
Self-Positive	0.15	0.06, 0.24	0.12	0.05, 0.19	0.001	0.28	0.11, 0.44	0.10	0.04, 0.16	0.001
Self-Negative	-0.04	-0.13, 0.04	-0.03	-0.10, 0.03	0.317	-0.16	-0.32, -0.01	-0.06	-0.12, 0.00	0.038
Friend-Positive	0.04	-0.05, 0.12	0.03	-0.03, 0.08	0.387	0.02	-0.13, 0.18	0.01	-0.04, 0.06	0.794
Friend-Negative	0.09	0.01, 0.16	0.07	0.01, 0.13	0.031	0.00	-0.14, 0.14	0.00	-0.05, 0.05	0.999

Stranger-Positive	-0.04	-0.13, 0.05	-0.03	-0.09, 0.03	0.380	0.02	-0.14, 0.18	0.01	-0.05, 0.06	0.788
Stranger-Negative	-0.06	-0.16, 0.03	-0.05	-0.11, 0.02	0.179	-0.02	-0.19, 0.15	-0.01	-0.07, 0.05	0.804
Session	-0.74	-1.17, -0.31	-0.07	-0.11, -0.03	0.001	-0.45	-1.24, 0.33	-0.02	-0.05, 0.01	0.260
BFNE	0.13	0.07, 0.19	0.24	0.13, 0.35	< .001	0.25	0.14, 0.37	0.23	0.13, 0.33	< .001
Global Ratings										
Intercept	4.37	0.81, 7.94	0.00	-0.14, 0.14	0.017	7.61	0.77, 14.45	0.00	-0.14, 0.14	0.030
Self-Positive	-0.52	-0.82, -0.21	-0.12	-0.19, -0.05	0.001	-0.75	-1.32, -0.18	-0.08	-0.15, -0.02	0.010
Self-Negative	0.11	-0.20, 0.42	0.03	-0.05, 0.10	0.479	-0.01	-0.59, 0.57	0.00	-0.07, 0.06	0.976
Friend-Positive	-0.09	-0.40, 0.24	-0.02	-0.08, 0.05	0.632	-0.01	-0.60, 0.59	0.00	-0.06, 0.06	0.987
Friend-Negative	0.23	-0.05, 0.50	0.05	-0.01, 0.12	0.103	0.34	-0.17, 0.86	0.04	-0.02, 0.09	0.196
Stranger-Positive	0.31	0.01, 0.60	0.07	0.00, 0.13	0.046	-0.17	-0.74, 0.39	-0.02	-0.08, 0.04	0.543
Stranger-Negative	-0.17	-0.46, 0.12	-0.04	-0.10, 0.03	0.251	0.18	-0.36, 0.73	0.02	-0.04, 0.08	0.502
Session	-0.76	-1.19, = 0.32	-0.07	-0.11, -0.03	0.001	-0.46	-1.27, 0.35	-0.02	-0.06, 0.02	0.267
BFNE	0.13	0.07, 0.19	0.24	0.14, 0.35	< .001	0.27	0.16, 0.38	0.24	0.14, 0.34	< .001

b = unstandardised regression coefficients, β = standardised regression coefficients, BFNE = Brief Fear of Negative Evaluation

Reliability of findings with clinical diagnosis of a Major Depressive Episode

To examine whether our findings applied to participants meeting diagnostic criteria for depression we repeated our analyses, examining the odds of meeting diagnostic criteria for a primary Major Depressive Episode (MDE) associated with task outcomes.

For the associative learning tasks, we again observed little evidence of a relationship between task performance and depression (supplementary table 7). There was weak evidence of increased reaction times when pairing shapes with the stranger being associated with an increased odds of a MDE (OR: 1.02, 95% CI: 1.00, 1.03, p = 0.034). However, this was not observed for PHQ-9 or BDI-II scores, suggesting an unreliable association.

For the Go/No-Go task, as we observed for the PHQ-9 and BDI-II, depression was predominantly associated with discriminative accuracy in the other condition. Increased discriminative accuracy when associating positive words with others was associated with an increased odds of experiencing an MDE (OR: 4.82, 95% CI: 1.51, 15.35, p = 0.008). Increased discriminative accuracy when associating negative words with others was associated with a reduced odds of experiencing an MDE, although effects overlapped with the null (OR: 0.43, 95% CI: 0.18, 1.01, p = 0.052). There was little evidence of an association between discriminative accuracy when associating words with the self with MDE (supplementary table 8).

Finally, we observed highly similar findings for the association between performance on the social reinforcement learning task with odds of experiencing a primary MDE to those we observed for the PHQ-9 and BDI-II (supplementary table 9). Reduced positive biases when learning about the self, driven by a greater number of errors when learning the positive rule, was associated with a greater odds of experiencing a primary MDE (self bias scores; OR: 1.16, 95% CI: 1.08, 1.25, p < .001, self-positive errors to criterion; OR: 1.23, 95% CI: 1.10, 1.38, p < .001). Increased global perceptions of being liked by the computer were associated with a reduced odds of experiencing depression (OR: 0.49, 95% CI: 0.32, 0.74, p < .001).

Table S7

Results from logistic regression models examining the association between accuracy and reaction times in the Associative Learning tasks and odds of meeting diagnostic criteria for a primary major depressive episode

Task	Stimuli	Accuracy (% correct)			Reaction Times (ms)		
		OR	95% CI	p	OR	95% CI	p
Self	Self	0.99	0.94, 1.05	0.794	0.99	0.98, 1.00	0.122
	Friend	1.02	0.97, 1.07	0.464	0.99	0.98, 1.00	0.197
	Stranger	1.02	0.98, 1.07	0.383	1.02	1.00, 1.03	0.034
Reward	High (£9)	0.99	0.95, 1.03	0.558	1.00	0.99, 1.01	0.669
	Medium (£3)	1.04	1.00, 1.09	0.066	1.00	0.99, 1.01	0.904
	Low (£1)	1.01	0.98, 1.04	0.498	1.00	0.99, 1.01	0.930
Emotion	Happy	0.99	0.96, 1.03	0.753	1.00	0.99, 1.01	0.719
	Neutral	1.00	0.97, 1.04	0.857	1.01	1.00, 1.01	0.293
	Sad	1.03	0.99, 1.06	0.138	0.99	0.98, 1.00	0.234

OR = Odds Ratio

Table S8

Results from a logistic regression model examining the association between discriminative accuracy in the Go/No-Go Task and odds of meeting diagnostic criteria for a primary major depressive episode

	OR	95% CI	p
Self Positive	0.71	0.28, 1.84	0.482
Self Negative	0.71	0.29, 1.78	0.471
Other Positive	4.82	1.51, 15.35	0.008
Other Negative	0.43	0.18, 1.01	0.052

OR = Odds Ratio

Table S9

Results from logistic regression models examining the association between social reinforcement learning task outcomes and odds of meeting diagnostic criteria for a primary major depressive episode

	OR	95% CI	p
Bias Scores			
Self	1.16	1.08, 1.25	< .001
Friend	0.99	0.92, 1.06	0.725
Stranger	0.95	0.88, 1.03	0.253
Errors to Criterion			
Self-Positive	1.23	1.10, 1.38	< .001
Self-Negative	0.90	0.80, 1.01	0.069
Friend-Positive	1.00	0.88, 1.14	0.992
Friend-Negative	1.03	0.92, 1.16	0.594
Stranger-Positive	0.98	0.86, 1.11	0.720
Stranger-Negative	1.07	0.95, 1.21	0.236
Global Ratings			
Self-Positive	0.49	0.32, 0.74	< .001
Self-Negative	1.16	0.76, 1.76	0.496
Friend-Positive	1.12	0.76, 1.66	0.567
Friend-Negative	0.68	0.43, 1.07	0.097
Stranger-Positive	1.08	0.73, 1.59	0.691
Stranger-Negative	1.23	0.80, 1.90	0.340

OR = Odds Ratio

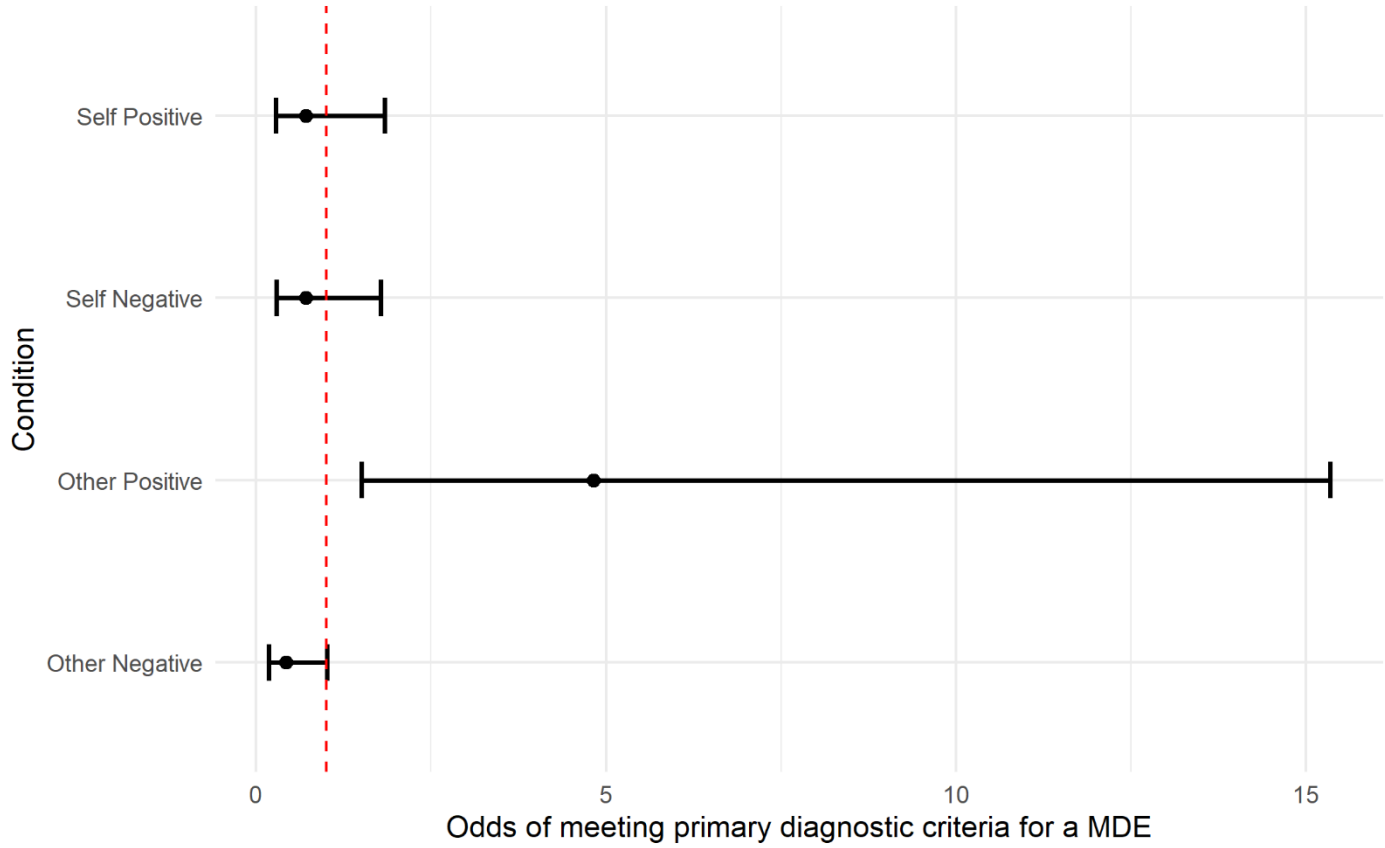


Figure S1

Results from a logistic regression model examining the association between discriminative accuracy in the self-esteem go/no-go task according to referential-emotion condition and odds of meeting criteria for a major depressive episode as a primary diagnosis.

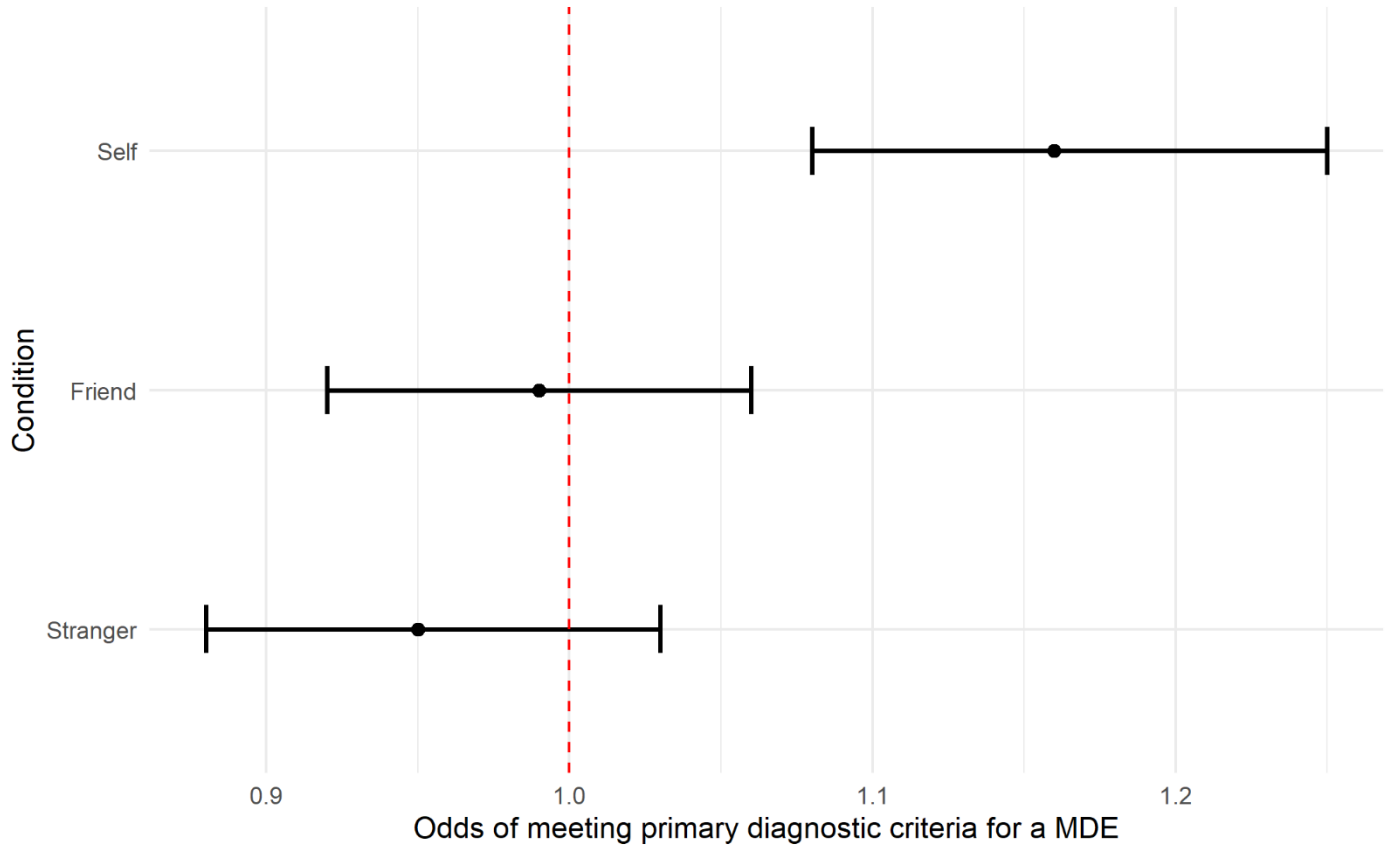


Figure S2

Results from a logistic regression model examining the association between bias scores in the social evaluation learning task according to referential condition and odds of meeting criteria for a major depressive episode as a primary diagnosis.

Table S10

Results from linear regression models examining the association between task performance outcomes with depression (PHQ-9/BDI-II) adjusting for gender and age.

Associative Learning										
	PHQ-9					BDI-II				
	b	*b* 95% CI	β	β 95% CI	p	*b*	*b* 95% CI	β	β 95% CI	p
Accuracy (% correct)										
Self										
Intercept	15.10	6.00, 24.21	-0.14	-0.51, 0.22	0.001	19.50	-0.09, 39.09	-0.03	-0.63, 0.11	0.051
Self	-0.06	-0.17, 0.06	-0.11	-0.33, 0.11	0.334	-0.14	-0.39, 0.11	-0.12	-0.34, 0.10	0.270
Friend	-0.06	-0.17, 0.05	-0.12	-0.36, 0.11	0.288	-0.01	-0.24, 0.22	-0.01	-0.24, 0.22	0.922
Stranger	0.05	-0.05, 0.15	0.12	-0.12, 0.35	0.337	0.12	-0.10, 0.33	0.13	-0.11, 0.37	0.281
Gender										
Male (reference)										

Female	1.02	-1.22, 3.26	0.19	-0.22, 0.60	0.369	3.80	-1.01, 8.62	0.33	-0.09, 0.74	0.121
Other	-2.32	-13.20, 8.56	-0.43	-2.44, 1.58	0.674	4.88	-18.51, 28.27	0.42	-1.60, 2.44	0.680
Age (years)	-0.15	-0.28-0.01	-0.18	-0.35, -0.01	0.034	-0.25	-0.54, 0.05	-0.14	-0.31, 0.03	0.099
Reward										
Intercept	8.88	1.51, 16.25	-0.08	-0.44, 0.28	0.019	9.77	-5.90, 25.43	-0.25	-0.61, 0.11	0.220
High (£9)	-0.06	-0.15, 0.03	-0.18	-0.43, 0.08	0.171	-0.17	-0.36, 0.02	-0.23	-0.48, 0.03	0.078
Medium (£3)	0.09	0.00, 0.18	0.26	0.00, 0.53	0.053	0.21	0.02, 0.41	0.28	0.02, 0.55	0.034
Low (£1)	-0.03	-0.09, 0.04	-0.08	-0.30, 0.13	0.432	0.03	-0.11, 0.17	0.04	-0.17, 0.25	0.686
Gender										
Male (reference)										

	Female	0.60	-1.64, 2.83	0.11	-0.30, 0.52	0.600	3.70	-1.05, 8.45	0.32	-0.09, 0.73	0.125
	Other	-4.49	-15.36, 6.38	-0.83	-2.84, 1.18	0.415	0.32	-22.77, 23.42	0.03	-1.96, 2.02	0.978
Age (years)		-0.11	-0.25, 0.03	-0.14	-0.32, 0.04	0.117	-0.18	-0.48, 0.12	-0.10	-0.28, 0.07	0.244
Emotion											
Intercept		9.80	2.95, 16.65	-0.12	-0.49, 0.24	0.005	15.98	1.34, 30.61	-0.27	-0.63, 0.09	0.033
Happy		-0.03	-0.11, 0.04	-0.10	-0.31, 0.12	0.372	-0.08	-0.23, 0.08	-0.11	-0.32, 0.11	0.324
Neutral		0.03	-0.05, 0.11	0.08	-0.16, 0.31	0.521	0.06	-0.11, 0.22	0.08	-0.16, 0.31	0.514
Sad		0.01	-0.07, 0.08	0.02	-0.22, 0.25	0.879	0.04	-0.13, 0.21	0.06	-0.18, 0.29	0.638
Gender											
	Male (reference)										

Female	0.88	-1.34, 3.10	0.16	-0.25, 0.57	0.435	3.99	-0.75, 8.72	0.34	-0.06, 0.75	0.098
Other	-3.44	-14.39, 7.51	-0.63	-2.66, 1.39	0.536	2.60	-20.79, 25.99	0.22	-1.79, 2.24	0.826
Age (years)	-0.14	-0.28, 0.00	-0.18	-0.35, 0.00	0.045	-0.26	-0.55, 0.04	-0.15	-0.32, 0.02	0.091
Reaction Times (ms)										
Self										
Intercept	11.00	1.67, 20.32	-0.11	-0.47, 0.26	0.021	21.27	1.37, 41.17	-0.25	-0.62, 0.11	0.036
Self	0.00	-0.02, 0.03	0.03	-0.29, 0.36	0.848	0.00	-0.06, 0.05	-0.01	-0.33, 0.32	0.962
Friend	-0.01	-0.04, 0.01	-0.18	-0.58, 0.21	0.357	0.00	-0.09, 0.02	-0.23	-0.63, 0.16	0.242
Stranger	0.01	-0.02, 0.04	0.10	-0.34, 0.54	0.648	0.03	-0.04, 0.09	0.18	-0.26, 0.62	0.423
Gender										

	Male (reference)									
	Female	0.765	-1.46, 2.99	0.14	-0.27, 0.55	0.498	3.71	-1.04, 8.45	0.32	-0.09, 0.73	0.125
	Other	-3.10	-14.05, 7.84	-0.57	-2.59, 1.45	0.576	4.07	-19.29, 27.43	0.35	-1.66, 2.37	0.731
Age (years)		-0.13	-0.27, 0.02	-0.15	-0.33, 0.02	0.087	-0.23	-0.53, 0.08	-0.13	-0.31, 0.04	0.143
Reward											
Intercept		6.40	-1.23, 14.02	-0.10	-0.47, 0.26	0.099	8.71	-7.56, 24.98	-0.26	-0.62, 0.10	0.292
High (£9)		0.01	-0.01, 0.04	0.22	-0.15, 0.59	0.234	0.03	-0.02, 0.08	0.25	-0.12, 0.61	0.189
Medium (£3)		-0.01	-0.04, 0.02	-0.18	-0.64, 0.28	0.446	-0.01	-0.06, 0.05	-0.05	-0.51, 0.40	0.813
Low (£1)		0.00	-0.02, 0.02	0.02	-0.36, 0.40	0.911	-0.01	-0.06, 0.03	-0.12	-0.49, 0.26	0.545
Gender											

	Male (reference)									
	Female	0.76	-1.47, 2.98	0.14	-0.27, 0.55	0.504	3.79	-1.00, 8.54	0.33	-0.08, 0.74	0.117
	Other	-3.54	-14.43, 7.34	-0.65	-2.66, 1.36	0.521	2.89	-20.34, 26.12	0.25	-1.75, 2.25	0.806
Age (years)		-0.14	-0.28, -0.01	-0.18	-0.34, -0.01	0.039	-0.26	-0.55, 0.02	-0.15	-0.32, 0.01	0.072
Emotion											
Intercept		10.05	3.22, 16.88	-0.11	-0.48, 0.25	0.004	17.40	2,77, 32.03	-0.26	-0.62, 0.10	0.020
Happy		0.00	-0.01, 0.02	0.08	-0.30, 0.46	0.667	0.00	-0.04, 0.04	0.02	-0.36, 0.40	0.914
Neutral		0.00	-0.03, 0.02	-0.10	-0.60, 0.40	0.699	-0.01	-0.05, 0.04	-0.06	-0.56, 0.44	0.821
Sad		0.00	-0.02, 0.02	-0.01	-0.51, 0.49	0.967	0.00	-0.05, 0.05	0.02	-0.49, 0.52	0.943
Gender											

	Male (reference)									
	Female	0.80	-1.42, 3.02	0.15	-0.26, 0.56	0.477	3.79	-0.96, 8.54	0.33	-0.08, 0.74	0.117
	Other	-2.80	-13.85, 8.25	-0.52	-2.56, 1.52	0.617	3.80	-19.86, 27.46	0.33	-1.71, 2.37	0.751
Age (years)		-0.14	-0.28, 0.00	-0.17	-0.34, 0.00	0.044	-0.26	-0.55, 0.03	-0.15	-0.32, 0.02	0.082
Self-Esteem											
GNAT											
		PHQ-9					BDI-II				
		b	*b* 95% CI	β	β 95% CI	p	*b*	*b* 95% CI	β	β 95% CI	p
d'											
Intercept		14.60	9.83, 1.937	-0.08	-0.55, 0.39	< .001	24.07	13.87, 34.27	-0.20	-0.68, 0.28	< .001
Self Positive		-2.52	-4.61, -0.43	-0.25	-0.45, -0.04	0.019	-3.52	-7.98, 0.95	-0.17	-0.38, 0.04	0.122
Self Negative		-0.53	-2.55, 1.49	-0.05	-0.23	0.605	-0.95	-5.26, 3.37	-0.04	-0.23	0.664

				, 0.14					, 0.15	
Other Positive	3.85	1.53, 6.18	0.32	0.13, 0.52	0.001	7.55	2.59, 12.51	0.31	0.10, 0.51	0.003
				-					-	
Other Negative	-2.49	-4.28, -0.69	-0.27	0.46, -0.07	0.007	-5.09	-8.93, -1.25	-0.27	0.47, -0.07	0.010
Gender										
Male (reference)										
				-					-	
Female	0.54	-2.39, 3.46	0.09	0.42, 0.60	0.718	2.70	-3.55, 8.96	0.23	0.30, 0.75	0.393
				-			15.80,		-	
Other	1.52	-9.34, 12.37	0.26	1.63, 2.16	0.782	7.41	30.62	0.62	1.33, 2.58	0.528
				-					-	
Age (years)	-0.14	-0.28, 0.00	-0.18	0.36, 0.00	0.044	-0.25	-0.54, 0.05	-0.15	0.34, 0.03	0.099
Social Evaluation Learning										
	PHQ-9					BDI-II				
	b	*b* 95% CI	β	β 95% CI	p	*b*	*b* 95% CI	β	β 95% CI	p
Bias Scores										

Intercept	10.46	7.15, 13.77	-0.07	-0.41, 0.26	< .001	17.09	10.18, 23.99	-0.22	-0.56, 0.12	< .001
Self	0.10	0.04, 0.19	0.12	0.05, 0.19	0.001	0.23	0.12, 0.34	0.12	0.06, 0.19	< .001
Friend	-0.03	-0.09, 0.02	-0.03	-0.09, 0.03	0.272	0.01	-0.09, 0.12	0.01	-0.05, 0.06	0.792
Stranger	-0.01	-0.07, 0.06	-0.01	-0.07, 0.06	0.814	0.00	-0.12, 0.11	0.00	-0.06, 0.05	0.947
Session	-0.88	-1.30, -0.46	-0.08	-0.12, -0.04	< .001	-0.72	-1.46, 0.03	-0.03	-0.06, 0.00	0.062
Gender										
Male (reference)										
Female	0.55	-1.52, 2.61	0.10	-0.28, 0.48	0.605	3.17	-1.15, 7.51	0.28	-0.10, 0.66	0.153
Other	-2.19	-12.33, 7.95	-0.40	-2.26, 1.46	0.673	3.43	-17.86, 24.71	0.30	-1.57, 2.17	0.753

Age (years)	-0.10	-0.23, 0.02	-0.12	-0.27, 0.03	0.108	-0.20	-0.46, 0.07	-0.11	-0.26, 0.04	0.147

b = unstandardised regression coefficients, β = standardised regression coefficients

Chapter 3 Supplementary Materials

Supplementary materials for Hobbs, C., Bone, J., Ruzickova, T., Sui, J., Kessler, D., Munafò, M.R., & Button, K. S. (2022). The cognitive neuropsychological theory of antidepressants: a systematic review and meta-analysis of behavioural evidence.

Study Characteristics

The majority of studies compared a single experimental condition against a single comparator condition (k = 64). Of those that included multiple conditions, 9 studies tested multiple types of antidepressant treatments, 1 study administered multiple doses of the same drug, 7 studies administered the same drug but in different samples, and 1 study tested multiple comparator conditions against a single experimental condition.

Of studies administering antidepressants to a single sample, a healthy volunteer sample was most used (k = 50). Of 19 studies recruiting samples based on depression, 15 studies used a sample of individuals with current depression, two studies examined individuals with remitted depression, and two studies examined individuals with treatment resistant depression. Single studies recruited individuals with current generalised anxiety disorder (GAD), current depression or anxiety, high neuroticism, fibromyalgia, and varying levels of anxiety and depression.

Of studies administering antidepressants to multiple samples, three studies compared individuals with current depression to healthy controls, one study compared individuals with current and remitted depression, one study compared individuals with remitted depression to healthy controls, one study compared individuals with intermittent explosive disorder to healthy controls, and one study compared individuals with OCD that were taking chronic antidepressants, individuals with OCD not taking chronic antidepressants, and healthy controls. One study compared individuals with current depression in the antidepressant group to both a healthy control group and individuals with current depression in the control group.

The majority of antidepressants tested were selective serotonin reuptake inhibitors (SSRIs) (n =33; citalopram = 20, escitalopram = 8, fluoxetine = 2, paroxetine = 1, sertraline = 2). Thirteen studies examined the noradrenaline reuptake inhibitor reboxetine, 6 studies examined the selective noradrenaline reuptake inhibitor duloxetine, and 9 studies examined atypical antidepressants (bupropion n = 4, agomelatine n = 1, tianeptine n = 1, mirtazapine n

= 2, buprenorphine n = 1). One study examined a herbal remedy (St John's Wort), nine studies examined mixed antidepressant treatment. Twenty studies examined experimental antidepressant treatments (ketamine n = 4, aprepitant n = 3, erythropoietin (epo) n = 4, ARA290 n = 1, memantine n = 1, psilocybin n = 3, scopolamine n = 3, GSK424887 = 1). One observational study did not report details of the antidepressant treatment examined.

Most studies examined the influence an acute dose of an antidepressant (k = 43). Of studies examining longer durations drug administration length was as follows; 1 week (k = 12), 7-10 days (k = 1), 10 days (k = 1), 2 weeks (k = 3), 3 weeks (k = 3), 4 weeks (k = 1), 7 weeks (k =1), 8 weeks (k = 3), 12 weeks (k = 3) and for infusion of Scopolamine, 3 sessions of 15 minutes each (k = 1). Two studies used multiple timepoints of two and six weeks. Eight observational studies did not have a fix period of administration.

70 studies were experimental, in that they controlled the drug type and dose, and 12 studies were observational.

The majority of studies compared the experimental condition(s) against placebo (k = 59). Twenty-two studies compared the experimental condition(s) against participants not taking an antidepressant, 14 of which employed before-after within-subject comparisons. One observational study examined the association between emotional processing and antidepressant treatment over time in participants on mixed lengths and doses of antidepressants.

48 studies compared experimental and comparator conditions using a between-subject comparison, 19 studies employed a crossover design, and 15 studies used a within-subject comparison.

On average, studies used 2 tasks to measure emotional processing (SD = 1.6), although the maximum of 6 tasks was used in 6 studies.

Narrative Summaries

Facial Emotion Processing

Facial Emotion Recognition

Forty-one studies examined the influence of antidepressants on recognition of facial emotions. Narrative synthesis of results indicated largely mixed effects of antidepressants on emotion recognition (Table S1). Consistent with the predictions of the cognitive neuropsychological theory, 18 studies reported an increase in the recognition of positive emotions [1,2,11–18,3–10], and no studies reported a decrease. Additionally, 16 studies reported a reduction in the recognition of negative emotions associated with antidepressant administration [1,6,25–30,11,16,19–24].

However, in contrast with the predictions of this theory, eight studies reported an increase in the processing of negative emotions [7,9,10,12,14,20,22,31]. Additionally, five studies reported main effects of antidepressants across both positive and negative emotions, two in increasing recognition [10,32] and three in decreasing recognition [2,19,33]. Finally, ten studies reported no effects of antidepressants [9,28,34–41].

Implicit Facial Emotion Processing

One study examined implicit processing of facial emotions using a masked affective priming task [42]. Task performance was not associated with antidepressant treatment.

Table S1

Studies assessing the influence of antidepressants on the recognition of facial emotions. ↑ indicates enhanced emotional processing (e.g. increased accuracy, faster reaction times), ↓ indicates diminished emotional processing (e.g. reduced accuracy, slower reaction times) , - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**							**Drug Effect**	**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Type**	**Emotions**	**Presentation**	**Stimuli**	**Morphed**	**Response**	**Outcomes**		
Anderson (2011b)	Mixed	16	Current MDD	None	14	Current MDD [a]	Between-Subject, Observational	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	1000 – 2000 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, *d'*, *β*	↑ All emotions (Accuracy, *d'*, *β)*	+
	Mixed	17	Remitted MDD	None	82	Remitted MDD [a]										
Arnone (2009)	Mirtazapine, 15 mg, Acute	15	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT	↓ Fear (Accuracy)	

Bamford (2015)	Duloxetine, 30-60 mg, 2 weeks	19	Healthy	Placebo	19	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	150 ms	Faces (Other)	yes	Categorical	Accuracy Misclassifications, *d'*	↓ Sadness (Accuracy, *d'*)	+
Bhagwagar (2004)	Citalopram, 10 mg, Acute	10	Healthy	Placebo	10	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy	↑ Fear (Accuracy)	
	Citalopram, 10 mg, Acute	10	Remitted MDD	Placebo	10	Remitted MDD									↓ Fear (Accuracy)	
Browning (2007)	Citalopram, 20 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, Misclassifications, *d'*, RT	↑ Neutral (Accuracy, Misclassifications) ↑ Fear (*d'*)	+
Bukala (2019)	Scopolamine, 1.5	16	Healthy	Placebo	17	Healthy	Between-Subject,	Facial Emotion	Anger, Disgust,	Unclear	Faces	Unclear	Categorical	Accuracy, RT	-	+

	mg, Acute						Experimental	Recognition	Fear, Happiness, Sadness, Surprise, Neutral							
Capitao (2015)	Fluoxetine, 20 mg, Acute	17	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT (correct responses), Misclassifications	↓ Anger (Accuracy, RT) ↓ Sadness (Accuracy) ↑ Happiness [b] (Accuracy)	+
Cerit (2015)	ARA 290, 2mg, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT, *d'*, β	↓ All emotions (Accuracy) ↓ Disgust [b] (Accuracy) ↓ Sadness (β)	+
Chandra (2010)	Aprepitant, 125 mg, Acute	20	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happine	500 ms	Faces (Ekman & Fri	yes	Categorical	Accuracy, RT	↓ All emotions (RT) ↑ Happiness (Accuracy)	+

									ss, Sadness, Surprise, Neutral		esen)					
Cooper (2015)	Tianeptine, 12.5 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT (correct responses), Misclassifications	↓ All emotions (Accuracy)	
Cremers (2016)	Escitalopram, 0.375 mg/kg, Acute	14	Intermittent Explosive Disorder	Placebo	14	Intermittent Explosive Disorder	Crossover, Experimental	Facial Emotion Recognition	Anger, Fear, Neutral	4000 ms	Faces (Ekman & Friesen)	no	Categorical	Accuracy, RT	-	
	Escitalopram, 0.375 mg/kg, Acute	14	Healthy	Placebo	14	Healthy									-	
Di Simplicio (2014)	Citalopram, 20 mg, 1 week	16	Healthy	Placebo	17	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness,	500 ms	Faces (Ekman & Friesen)	yes	Categorical	*d'*	↑ Positive [c] (*d'*)	+

									Surprise, Neutral							
Grimm (2018)	Psilocybin, 0.16 mg/kg, Acute	18	Healthy	Placebo	18	Healthy	Crossover, Experimental	Facial Emotion Recognition	Happiness, Sad, Anger, Fear, Neutral[d]	2000 ms	Faces (Other)	no	Categorical	Accuracy, RT	-	
Harmer (2003a)	Citalopram, 10 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT (correct responses), Misclassifications	↑ Happiness (Accuracy, RT) ↑ Fear (Accuracy, RT)	
Harmer (2003b)	Reboxetine, 4 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, *d'*	↑ Happiness (Accuracy, *d'*)	+
Harmer (2004)	Reboxetine, 4 mg b.i.d., 1 week	14	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Hap	500 ms	Faces (Ekman & Fri	yes	Categorical	Accuracy, RT (correct responses), Misclas	↓ Fear (Accuracy, Misclassifications)	

									pineSS, Sadness, Surprise, Neutral		esen)			sificatiOns	↓ Anger (Accuracy) ↑ Surprise (Misclassifications)	
	Citalopram, 20 mg, 1 week	14	Healthy												↓ Fear (Accuracy, Misclassifications) ↓ Anger (Accuracy) ↓ Disgust (Accuracy) ↓ Surprise (Accuracy) ↑ Surprise (Misclassifications) ↑ Happiness (Accuracy, Misclassifications)	
Harmer (2006)	Citalopram, 20 mg, 1 week	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy	↓ Fear (Accuracy)	+

									, Surprise, Neutral							
Harmer (2008)	Duloxetine, 60 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	NR	yes	Categorical	Accuracy, RT, *d'*	↑ Happiness (Accuracy, *d'*) ↑ Disgust (Accuracy, *d'*)	+
Harmer (2009)	Reboxetine, 4 mg, Acute	15	Current MDD	Placebo	18	Current MDD	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT (correct responses)	↑ Happiness (Accuracy)	+
	Reboxetine, 4 mg, Acute	16	Healthy	Placebo	15	Healthy									↑ Happiness (Accuracy)	
Harmer (2011)	Agomelatine, 25 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surp	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, Misclassifications, RT	↓ Sad (Accuracy, Misclassifications) ↑ Neutral (Accuracy, Misclassifications)	+

	Agomelatine, 50mg, 1 week	16	Healthy						rise, Neutral						-	+
Harmer (2013)	Citalopram, dose unclear, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, Misclassifications, RT	↑ Anger (Accuracy) ↑ Happiness (Accuracy, Misclassifications)	
	GSK424887, dose unclear, Acute	18	Healthy												-	
Kamboj (2006)	Scopolamine, 0.6 mg, Acute	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	3000 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, Misclassifications	↓ Disgust (Accuracy) ↓ Anger (Accuracy, Misclassifications)	
Kerestes (2009)	Citalopram, 20 mg, Acute	12	Healthy	Placebo	12	Healthy	Crossover, Experimental	Facial Emotion Recognition	Happiness, Sad, Neutral	507 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy	-	
	Reboxetine, 4 mg b.i.d, Acute	12	Healthy												-	

Kometer (2012)	Psilocybin, 215 ug/kg, Acute	17	Healthy	Placebo + Placebo	17	Healthy	Crossover, Experimental	Reading the Mind in the Eyes	Positive, Negative, Neutral	NR	Eye-region of faces	no	Categorical	Error rates, valence of chosen word	↑ Positive (Error Rate) ↓ Negative (Error Rate) [e]	
Labuschagne (2010)	Citalopram, 20 mg, Acute	14	Healthy	Placebo	14	Healthy	Crossover, Experimental	Facial Emotion Recognition	Happiness, Sad, Neutral	507 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy	-	
Lochner (2012)	Escitalopram, 20 mg, Acute	19	Healthy	Placebo	18	Healthy	Crossover, Experimental	Facial Emotion Recognition	Disgust [f]	NR	Faces (Other)	yes	Categorical	Accuracy, *d'*	↓ Disgust (Accuracy)	
	Escitalopram, 20 mg, Acute	19	OCD chronic antidepressant use	Placebo	18	OCD chronic antidepressant use									↑ Disgust (Accuracy)	
	Escitalopram, 20 mg, Acute	19	OCD no chronic antidepressant use	Placebo	18	OCD no chronic antidepressant use									↑ Disgust (Accuracy)	
Merens (2008) Study 1	Mixed, minimum 4 weeks	19	Remitted MDD	None	20	Healthy	Between-Subject, Observational	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness,	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT	-	+

									Neutral							
Miskowiak (2007b)	Epo, 40,000 IU/ml, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy	↓ Fear (Accuracy)	
Miskowiak (2008)	Epo, 40,000 IU/ml, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT (correct responses), *d'*, Misclassifications	↑ All Emotions (*d'*) ↑ Happiness (Accuracy) ↑ Fear (Accuracy)	
Miskowiak (2010)	Epo, 40,000 IU/ml, Acute	10	Current MDD	Placebo	9	Current MDD	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy	↓ Fear (Accuracy)	

Murrough (2015)	Ketamine hydrochloride, 0.5 mg/kg, Acute dose	18	Treatment Resistant MDD	None	18	Treatment Resistant MDD [a]	Within-Subject, Experimental	Facial Emotion Perception	Happiness, Sadness, Neutral	2000 ms	Faces (Ekman & Friesen)	yes	Likert	RT	↑ Happiness (RT)	
Murphy (2009a)	Citalopram, 20 mg, Acute	13	Healthy	Placebo	13	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT	↑ Happiness (Accuracy) ↓ Disgust [d] (Accuracy)	+
Naudin (2014)	Mixed	22	Current MDD	None	22	Current MDD [a]	Within-Subject, FU at point of clinical improvement, Observational	Facial Emotion Recognition	Happiness, Sad, Anger, Fear, Neutral [f]	500 / 2000 ms	Faces (Other)	no	Categorical	*d'*, *β*	↑ Anger (*d'*) ↑ Fear (*d'*) ↑ Happiness (*d'*) ↑ Sad (*d'*) ↓ Sadness *(β)*	
Pringle (2011)	Aprepitant, 125 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness,	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT, Misclassifications	-	+

									Sadness, Surprise, Neutral							
Pringle (2012)	Memantine, 10 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT, Misclassifications	-	+
Shiroma (2014)	Citalopram, 10 mg, 1 week	27	Veterans aged ≥ 55 current MDD	None	30	Veterans aged ≥ 55 current MDD	Within-Subject, FU after 1 week drug administration, Experimental	Facial Emotion Recognition	Happiness, Neutral	500 / 5000 ms	Faces (Other)	yes	Categorical	Accuracy, RTs	↑ Happiness (Accuracy, RT)	
Stein (2012)	Reboxetine, 4 mg b.i.d, 7 days	12	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Infant Face Ratings Task	Positive, Muted Positive, Neutral, Muted Negative, Neg	100 / 2000 ms	Infant Faces	NR	Likert	Ratings	↑ Positive (Ratings)	
	Citalopram, 20 mg, 7 days	13	Healthy												↑ Positive (Ratings)	

									ative							
Tranter (2009)	Citalopram, 20 mg, 2 & 6 weeks	2 week: 32, 6 week: 39	Current MDD	None	57	Current MDD	Within-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, *d'*, *β*	↑ Happiness (Accuracy) ↑ Disgust (Accuracy) ↑ Surprise (Accuracy) ↑ Anger (*β*)	
	Reboxetine, 4 mg b.i.d, 2 & 6 weeks	2 weeks: 27, 6 weeks: 30	Current MDD	None	51	Current MDD									↑ Happiness (Accuracy) ↑ Disgust (Accuracy) ↑ Surprise (Accuracy) ↑ Anger (*β*)	
Walsh (2018)	Bupropion, 150 mg 7-10 days then 150 mg 2/day, 6 weeks total	2 weeks: 43, 6 weeks: 60	Current MDD	None	46	Current MDD[a]	Within-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surp	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, *d'*, Misclassifications, *β*	↓ Sadness (Misclassifications, Bias)	

									rise, Neutral							
Walsh (2018b)	Bupropion, 150 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, *d'*, RT (correct responses)	↑ Happiness (Accuracy, *d'*)	+
Warren (2019)	St John's Wort, 200 mg 3/day, 1 week	23	Healthy	Placebo	24	Healthy	Between-Subject, Experimental	Facial Emotion Recognition	Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral	500 ms	Faces (Ekman & Friesen)	yes	Categorical	Accuracy, RT	↓ Disgust (Accuracy)	

[a] Study also included healthy volunteer group who did not receive any medication. Results are not included as they are outside the focus of this review.

[b] Trend effect

[c] Data aggregated for happiness and surprise

[d] Separate blocks completed per emotion with comparisons to neutral

[e] The effects of Psilocybin were further modulated pre-treatment with placebo versus Ketanserin (50 mg); Psilocybin increased error rates for negative faces only after placebo but not Ketanserin pre-treatment.

[f] Results only reported for disgust, but anger, fear, happiness, sadness included in the task.

Word Categorisation

The influence of antidepressants on categorisation of positive and negative characteristics was examined in 24 studies (Table S2). In this task participants were most commonly instructed to categorise words according to whether they would like or dislike overhearing someone describing them using these characteristics.

Mixed results were reported for the effect of antidepressants on word categorisation. In keeping with the predictions of the cognitive neuropsychological theory, antidepressants reduced reaction times for categorisation of positive words relative to negative words in four studies [4,6,8,19]. Additionally, one study reported an increase in reaction times that was specific to positive words [43], and another study reported greater endorsement of positive words about others and less endorsement of negative words about the self [44]. However, 17 studies reported no effects of antidepressants on categorisation of words [2,7,40,41,43,45–49,9,15,21,28,29,31,33,34]. Additionally, one study reported only a main effect of mirtazapine in increasing reaction times for both positive and negative words [26].

Table S2

Studies assessing the influence of antidepressants on categorisation of emotional words. ↑ indicates enhanced emotional processing (e.g. increased accuracy, faster reaction times), ↓ indicates diminished emotional processing (e.g. reduced accuracy, slower reaction times), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**					**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Referential Condition**	**Emotions**	**No. Words**	**Presentation**	**Outcomes**	**Positive**	**Negative**	
Arnone (2009)	Mirtazapine, 15 mg, Acute	15	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT (correct classifications)	↓ (RT)	↓ (RT)	
Browning (2007)	Citalopram, 20 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT (correct classifications)	-	-	+
Bukala (2019)	Scopolamine, 1.5 mg, Acute	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	Accuracy, RT	-	-	+
Cerit (2015)	ARA290, 2 mg, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	Accuracy, RT	↑ (RT)	↓ (RT)	+
Chandra (2010)	Aprepitant, 125 mg, Acute	20	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT (correct classifications	-	-	+
Cooper (2015)	Tianeptine, 12.5 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject,	Self	Positive, Negative	60	500 ms	RT (correct classifications	-	-	

							Experimental								
Di Simplicio (2012)	Citalopram, 20 mg, 1 week	18	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self	Positive, Negative	90	500 ms	Accuracy, RT	-	-	+
Harmer (2003b)	Reboxetine, 4 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Unclear	Positive, Negative	60	500 ms	RT	↑ (RT)	↓ (RT)	
Harmer (2004)	Citalopram, 20 mg, 1 week	14	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT (correct classifications	↑[a] (RT)	↓[a] (RT)	
	Reboxetine, 4 mg bid, 1 week	14	Healthy										↑ (RT)	↓ (RT)	
Harmer (2008)	Duloxetine, 60 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	Accuracy, RT (correct classifications)	-	-	+
Harmer (2009)	Reboxetine, 4 mg, Acute	15	Current MDD	Placebo	18	Current MDD	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT (correct classifications	↑ (RT)	↓ (RT)	+
	Reboxetine, 4 mg, Acute	16	Healthy	Placebo	15	Healthy							↑ (RT)	↓ (RT)	+
Harmer (2011)	Agomelatine, 25 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self; Non-Self (Animal)	Positive, Negative	120	500 ms	RT (correct classifications	-	-	
	Agomelatine, 50 mg, 1 week	16	Healthy										-	-	
Harmer (2013)	Citalopram, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject	Self	Positive,	60	500 ms	RT (correct	-	-	

	GSK424887, Acute	18	Healthy				, Experimental		Negative			classifications	-	-	
Komulainen (2016)	Mirtazapine, 15 mg, Acute	15	Healthy	No Drug	15	Healthy	Between-Subject, Experimental	Self; Neutral Control (directions)	Positive, Negative, Neutral	80	1000 ms	Accuracy, RT	-	-	+
Komulainen (2018)	Escitalopram, 10 mg, 1 week	17	Current MDD	Placebo	15	Current MDD	Between-Subject, Experimental	Self; Neutral Control (directions)	Positive, Negative, Neutral	80	1000 ms	RT	↑ (RT)	-	
Lemogne (2010)	Mixed [b], 6-14 weeks of treatment	8	Current MDD	Mixed [b], ≤ 1 week of treatment	8	Current MDD [c]	Within-Subject, Observational	Self; General; Neutral [d]	Positive, Negative, Neutral	190	1000 ms	RT	-	-	+
Matthews (2010)	Escitalopram, 5 mg 3 days then 10 mg 18 days, 21 days total	15	Healthy	Placebo	15	Healthy	Crossover, Experimental	Self; Other; Neutral (‘This word is’)	Positive, Negative	96	1500 ms	Accuracy	↑ Other (Accuracy)	↓ Self (Accuracy)	
Miskowiak (2007a)	Reboxetine, 4 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	Accuracy, RT	-	-	+
Norbury (2008)	Reboxetine, 4 mg, 1 week	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Self [b]	Positive, Negative	120	500 ms	Accuracy, RT	-	-	+
Pringle (2011)	Aprepitant, 125	16	Healthy	Placebo	16	Healthy	Between-Subject,	Self	Positive, Negative	60	500 ms	Accuracy	-	-	+

	mg, 1 week						Experimental								
Pringle (2012)	Memantine, 10 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT	-	-	+
Walsh (2018)	Bupropion, 150 mg 7-10 days then 150 mg 2/day, 6 weeks total	2 weeks: 43, 6 weeks: 60	Current MDD	None	46	Current MDD [c]	Within-Subject, Experimental	Self	Positive, Negative	60	500 ms	Accuracy	-	-	
Walsh (2018 b)	Bupropion, 150 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT	-	-	+
Warren (2019)	St John's Wort, 200 mg 3/day, 1 week	23	Healthy	Placebo	24	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	500 ms	RT	-	-	

[a] Trend Effect

[b] At session 1 (≤ 1 week of treatment) 4 patients were taking a selective serotoninreuptake inhibitor (SSRI), 2 a serotonin–norepinephrine reuptake inhibitor (SNRI), and 2 a tricyclic antidepressant. At session 2 (6-14 weeks of treatment), SSRI had been switched for a SNRI for 1 patient.

and the treatment had been augmented with mirtazapine for 2

patients taking either a SSRI or a SNRI. N

[c] Study also included healthy volunteer group who did not receive any medication. As we are interested in the effect of drug in this study we have restricted analyses to groups of participants with drug comparisons

[d] Neutral control condition not analysed or reported

Perceptions of Emotional Images

Eight studies examined the influence of antidepressants on perceptions of emotional images in relation to ratings of valence and/or emotional arousal (Table S3). Results were generally unsupportive of the cognitive neuropsychological theory. Only one study reported a change in emotional processing; Buprenorphine versus placebo increased the positivity of ratings, with this effect enhanced for social images [50]. Two studies, both examining drugs related to Ketamine, reported main effects in increasing reaction times for ratings for both positive and negative images [51,52]. The remaining studies reported no effect of antidepressants [53–57].

Table S3

Studies assessing the influence of antidepressants on perceptions of emotional images. ↑ indicates enhanced emotional processing (e.g. increased ratings), ↓ indicates diminished emotional processing (e.g. decreased ratings), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**							**Drug Effect**	
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Stimuli**	**Presentation**	**Trials**	**Matched**	**Ratings**	**Ratings Format**	**Outcomes**	**Positive**	**Negative**
Bershad (2018)	Buprenorphine, 0.2mg, Acute	38	Healthy	Placebo	38	Healthy	Crossover, Experimental	IAP	6000 ms	54	yes	Valence	Evaluative Space Grid	Ratings	↑[a] (Valence Rating)	-
Bruhl (2010)[b]	Citalopram, 40 mg, Acute	16	Healthy	Placebo	15	Healthy	Crossover, Experimental	IAP	7920 ms	56	yes	Valence	Likert	Ratings	-	-
	Reboxetine, 8 mg, Acute	14	Healthy												-	-
Gibbs (2013)	Reboxetine, 4 mg, Acute	57	Healthy[c]	Placebo	62	Healthy[c]	Between-Subject, Experimental	IAP	2500 ms	72	yes	Valence & Arousal	Likert	Ratings	-	-
Gibbs (2014)	Reboxetine, 4 mg, Acute	38	Healthy[d]	Placebo	34	Healthy[d]	Between-Subject, Experimental	IAP	2500 ms	72	yes	Valence & Arousal	Likert	Ratings	-	-
Outhred (2014)	Escitalopram, 20 mg, Acute	36	Healthy	Placebo	36	Healthy	Crossover, Experimental	IAP	4000 ms	45	NR	Valence[e]	Categorical	Ratings	-	-
Scheidegger (2016)	S-Ketamine, 0.12-0.25 mg/kg/h, Acute[f]	23	Healthy	None (baseline)	23	Healthy	Within-subject, Experimental	IAP	NR	90	yes	Valence	Categorical	Reaction Times	↑ (RT)	↑ (RT)

Sterpenich (2019)	Ketamine Hydrochloride, 0.5 mg/kg, Acute	10	Treatment Resistant, Current MDD	None (baseline)	10	Treatment Resistant, Current MDD	Within-subject, Experimental	NR	3000 ms	90	yes	Valence	Likert	Ratings, Reaction Times	↑ (RT)	↑ (RT)
Wang (2012)	Fluoxetine, 20-80 mg, 8 weeks	18	Current MDD	None (baseline)	18	Current MDD[g]	Within-subject, Experimental	IAP	2000 ms	54	NR	Valence	Categorical	Ratings, Reaction Times	-	-

IAP = International Affective Picture System, NR = Not Reported

[a] Stimuli could be either social or non-social. Buprenorphine was found to selective enhance positivity ratings for images with social content.

[b] This task also included a priming condition; in half of trials participants were presented with a symbol indicating the upcoming emotion, in the other half participants were presented with a symbole indicating that the upcoming emotion was unknown.

[c] Varied according to Alpha 2B Adrenoceptor Genotype, in this review we have aggregated groups according to drug administration.

[d] Varied according to COMT Genotype, in this review we have aggregated groups according to drug administration.

[e] Participants also rated brightness of backgrounds as a control condition, we have not included these results as they do not reflect emotional processing.

[f] 0.12 mg/kg 15 mins prior fMRI, continuous infusion 0.25 mg/kg/h during scanning period

[g] Study also included a healthy control group that did not receive a drug intervention. As this review focuses on the effects of antidepressant administration we have focused on the before and after comparisons of the group receiving an antidepressant.

Memory

Word Recall

Following completion of the word categorisation task, 19 studies asked participants to recall the words previously categorised (Table S4).

Two studies reported an increase in accuracy of positive recall and a decrease in accuracy of negative recall [26,28]. Four studies reported an increase in positive emotional processing; two an increased accuracy for positive words [6,8] and two a greater number of false positive responses [7,21]. Two studies reported a decrease in negative emotional processing; one a decreased accuracy for negative words [4], and one a decrease in false negative responses [29].

However, 11 studies reported results inconsistent with the cognitive neuropsychological theory. One study reported weak evidence of a relative decrease in positive versus negative processing [34]. Another study reported similar effects for an experimental compound GSK424887 and no effects for citalopram [9]. One study reported a decrease in the proportion of positive words recalled [33], and two studies reported main effects of decreased accuracy across both positive and negative words [2,46]. No effect of antidepressants on recall was observed in six studies [15,19,31,40,41,43].

Table S4

Studies assessing the influence of antidepressants on recall of emotional words. ↑ indicates enhanced emotional processing (e.g. increased accuracy, faster reaction times), ↓ indicates diminished emotional processing (e.g. decreased accuracy, slower reaction times) , - indicates no effect.

Study	Experimental Condition			Comparator Condition			Design	Task						Drug Effect		MA
	Drug	N	Sample	Drug	N	Sample		Referential [illegible]	Emotions	No. Words	Time after	Time for Recall	Outcomes	Positive	Negative	
Arnone (2009)	Mirtazapine, 15 mg, Acute	15	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	NR	2	Accuracy	↑ (Accuracy)	↓ (Accuracy)	
Browning (2007)	Citalopram, 20 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	15 mins	NR	Accuracy	-	-	+
Bukala (2019)	Scopolamine, 1.5 mg, Acute	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	NR	NR	Accuracy	↓[a] (Accuracy)	↑[a] (Accuracy)	+
Cerit (2015)	ARA290, 2 mg, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	15 mins	2	Accuracy, Misclassifications	-	-	+
Chandra (2010)	Aprepitant, 125 mg, Acute	20	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	15 mins	NR	Accuracy, Misclassifications	↓ (Accuracy)	↓ (Accuracy)	+
Cooper (2015)	Tianeptine, 12.5 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	NR	2	Accuracy, Misclassifications, Proportion positive recalled	↓ (Proportion positive)	-	
Harmer	Reboxetine, 4	12	Healthy	Placebo	12	Healthy	Between-Subject,	Unclear	Positive,	60	0 m	NR	Accuracy	-	↓ (Accuracy)	+

(2003 b)	mg, Acute						Experimental		Negative		ins					
Harmer (2004)	Citalopram, 20 mg, 1 week	14	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	0 mins	NR	Accuracy, Proportion positive	↑ (Proportion positive)	-	
	Reboxetine, 4 mg bid, 1 week	14	Healthy											↑ (Proportion positive)	-	
Harmer (2008)	Duloxetine, 60 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	15 mins	NR	Proportion positive, Misclassifications	↑ (Misclassifications)	-	+
Harmer (2009)	Reboxetine, 4 mg, Acute	15	Current MDD	Placebo	18	Current MDD	Between-Subject, Experimental	Self	Positive, Negative	60	15 mins	NR	Accuracy	↑ (Accuracy)	-	+
	Reboxetine, 4 mg, Acute	16	Healthy	Placebo	15	Healthy								-	-	+
Harmer (2011)	Agomelatine, 25 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self; Non-Self (Animal)	Positive, Negative	120	5 mins	2	Accuracy, Misclassifications	↑ (Accuracy)	↓ (Accuracy)	+
	Agomelatine, 50 mg, 1 week	16	Healthy											-	-	+
Harmer (2013)	Citalopram, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	5 mins	2	Accuracy	-	-	
	GSK424887, Acute	18	Healthy											↓[a] (Accuracy)	↑[a] (Accuracy)	
Komulainen (2016)	Mirtazapine, 15 mg, Acute	15	Healthy	No Drug	15	Healthy	Between-Subject, Experimental	Self; Neutral Control (directi	Positive, Negative, Neutral	80	1.5 hrs	NR	Accuracy	↓ (Accuracy)	↓ (Accuracy)	+

								ons)								
Komulainen (2018)	Escitalopram, 10 mg, 1 week	17	Current MDD	Placebo	15	Current MDD	Between-Subject, Experimental	Self; Neutral Control (directions)	Positive, Negative, Neutral	80	NR	NR	Accuracy	-	-	
Pringle (2011)	Aprepitant, 125 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	0 mins	2	Accuracy, Misclassifications	-	-	+
Pringle (2012)	Memantine, 10 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	0 mins	2	Accuracy	-	-	+
Walsh (2018)	Bupropion, 150 mg 7-10 days then 150 mg 2/day, 6 weeks total	2 weeks: 43, 6 weeks: 60	Current MDD	None	46	Current MDD [b]	Within-Subject (2 and 6 week FU), Experimental	Self	Positive, Negative	60	NR	2	Accuracy, Misclassifications	-	↓ (Misclassifications)	
Walsh (2018b)	Bupropion, 150 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60	NR	2	Accuracy, Misclassifications	-	-	+
Warren (2019)	St John's Wort, 200	23	Healthy	Placebo	24	Healthy	Between-Subject,	Self	Positive, Negative	60	NR	2	Accuracy, Misclassifications	↑ (Misclassifications)	-	

	mg 3/day, 1 week						Experimental									

Note: Times of 0 mins for 'Time after categorisation' indicate that recall occurred immediately after categorisation.

NR = Not Reported

[a] Trend Effect

[b] Study also included healthy volunteer group who did not receive any medication. As we are interested in the effect of drug in this study we have restricted analyses to groups of participants with drug comparisons

Image Recall

Three studies asked participants to recall emotional images that they had previously been presented (Table S5). One study reported results supportive of the cognitive neuropsychological theory; healthy participants recalled fewer negative images following acute administration of reboxetine versus placebo [54]. The remaining two studies reported only a main effect of antidepressants [55,58].

Table S5

Studies assessing the influence of antidepressants on recall of emotional images. ↑ indicates enhanced emotional processing (e.g. better memory), ↓ indicates diminished emotional processing (e.g. worse memory), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Other Therapy**	**Task Design**					**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**			**Stimuli**	**Number**	**Time after**	**Time for**	**Outcomes**	**Positive**	**Negative**	
Gibbs (2013)	Reboxetine, 4 mg, Acute	57	Healthy [a]	Placebo	62	Healthy [a]	Between-Subject, Experimental	None	IAP	72	30 mins	10 mins	EEMI	-	↓[c] (EEMI)	
Gibbs (2014)	Reboxetine, 4 mg, Acute	38	Healthy [b]	Placebo	34	Healthy [b]	Between-Subject, Experimental	None	IAP	72	30 mins	10 mins	EEMI, Accuracy	↑[d] (EEMI)	↑[d] (EEMI)	
Miskowiak (2009)	Epo, 40,000 IU/ml, Acute	9	Current MDD	Placebo	8	Current MDD	Between-Subject, Experimental	Yes[e]	IAP	60	0 mins	NR	Accuracy, Misclassifications	↑ (Misclassifications)	↑ (Misclassifications)	

IAP = International Affective Picture System, EEMI = Emotional Enhancement Memory Index (percentage increase in recall of emotional, relative to neutral pictures), NR = Not Reported

Note: All studies used positive, negative, neutral pictures. Times of 0 mins for 'Time after categorisation' indicate that recall occurred immediately after categorisation.

[a] Varied according to Alpha 2B Adrenoceptor Genotype, in this review we have grouped together

[b] Varied according to COMT Genotype, in this review we have grouped together

[c] Additional effect of Alpha 2B Adrenoceptor Deletion noncarriers versus carriers; recall of negative words decreased in deletion noncarriers compared with carriers only for participants administered reboxetine.

[d] Additional effect of COMT Genotype; recall of emotional words was enhanced for COMT Met/Met group for participants administered reboxetine

[e] All patients but two were on antidepressant medication. No change in medication was made in the 2 weeks prior to or during the study. Epo group: any antidepressant n = 7 (SSRI n = 4, Dual Action = 2, MAOI n = 1), Antipsychotic n = 2, Benzodiazepine n = 3, No meds n = 1. Placebo group: any antidepressant n = 5 (SSRI: n = 3, Dual Action = 1, mAOI n = 2), Antipsychotic n = 1, Lithium n = 1, Benzodiazepine n = 1, Thyroxin n = 1, no meds n = 1

Word Recognition

Fifteen studies examined the influence of antidepressants on recognition of emotional words (Table S6). In this task, participants are asked to identify familiar words from a list of words previously used in the categorisation task and novel words. Five studies reported effects consistent with the cognitive neuropsychological theory. Three studies reported increased processing of positive versus negative words, although outcomes varied; one study reported an effect on accuracy [40]; another on misclassifications [15], and one on reaction times [48]. Additionally, one study reported an increase in accuracy specific to positive words [21], and one study reported a decrease in misclassifications of negative words [41]. Inconsistent with the cognitive neuropsychological theory, two studies reported main effects of antidepressants in opposing directions [33,46], and eight studies reported no effects [2,9,19,28,31,34,43,49].

Table S6

Studies assessing the influence of antidepressants on recognition of emotional words. ↑ indicates enhanced emotional processing (e.g. better memory), ↓ indicates diminished emotional processing (e.g. worse memory), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**					**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Referential Condition**	**Emotions**	**No. Words**	**Presentation**	**Outcomes**	**Positive**	**Negative**	
Browning (2007)	Citalopram, 20 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative			Accuracy	-	-	+
Bukala (2019)	Scopolamine, 1.5 mg, Acute	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self	Positive, Negative	NR	NR	Accuracy, RT	-	-	+
Cerit (2015)	ARA290, 2 mg, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	NR	Accuracy, RT, *d'*, *β*	-	-	+
Chandra (2010)	Aprepitant, 125 mg, Acute	20	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	NR	Accuracy, RT	-	-	+
Cooper (2015)	Tianeptine, 12.5 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	NR	*d'*, RT, Misclassifications	↑ (*d'*)	↑ (*d'*)	

Harmer (2011)	Agomelatine, 25 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Self; Non-Self (Animal)	Positive, Negative	60 familiar, 60 unfamiliar	NR	Accuracy, Misclassifications, RTs	-	-	+
	Agomelatine, 50 mg, 1 week	16	Healthy										-	-	+
Harmer (2013)	Citalopram, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	NR	*d'*	-	-	
	GSK424887, Acute	18	Healthy										-	-	
Komulainen (2016)	Mirtazapine, 15 mg, Acute	15	Healthy	No Drug	15	Healthy	Between-Subject, Experimental	Self; Neutral Control (directions)	Positive, Negative, Neutral	60 familiar, 60 unfamiliar	1000 ms	*a'*	↓(*a'*)	↓(*a'*)	+
Komulainen (2018)	Escitalopram, 10 mg, 1 week	17	Current MDD	Placebo	15	Current MDD	Between-Subject, Experimental	Self; Neutral Control (directions)	Positive, Negative, Neutral	60 familiar, 60 unfamiliar	NR	*a'*, Accuracy, Misclassifications	-	-	+
Miskowiak (2007a)	Reboxetine, 4 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	NR	Accuracy, RT, *d'*	↑(RT)	↓(RT)	+
Norbury (2008)	Reboxetine, 4 mg, 1 week	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	Self	Positive, Negative	NR	500 ms	*a'*, Accuracy, RT	-	-	+
Pringle (2011)	Aprepitant, 125 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Exper	Self	Positive, Negative	60 familiar, 60 unfa	NR	Accuracy	↑ (Accuracy)	↓ (Accuracy)	+

							imental			milia r					
Pringle (2012)	Memantine, 10 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	NR	Misclassifications	-	↓ (Misclassifications)	+
Walsh (2018b)	Bupropion, 150 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Self	Positive, Negative	NR	NR	Misclassifications	↑ (Misclassifications)	↓ (Misclassifications)	+
Warren (2019)	St John's Wort, 200 mg 3/day, 1 week	23	Healthy	Placebo	24	Healthy	Between-Subject, Experimental	Self	Positive, Negative	60 familiar, 60 unfamiliar	500 ms	Accuracy, Misclassifications	↑ (Accuracy)	-	

NR = Not Reported

Image Recognition

Three studies measured recognition of emotional images (Table S7). Findings did not support the cognitive neuropsychological theory. Two studies reported ceiling effects across groups [54,55], and one study reported no significant differences between groups [59].

Table S7

Studies assessing the influence of antidepressants on recognition of emotional images. ↑ indicates enhanced emotional processing (e.g. better memory), ↓ indicates diminished emotional processing (e.g. worse memory), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Other Therapy**	**Task Design**					**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**			**Outcomes**	**Stimuli**	**Number**	**Time after categorisatio**	**Emotions**	**Positive**	**Negative**	
Gibbs (2013)	Reboxetine, 4 mg, Acute	57	Healthy [a]	Placebo	62	Healthy [a]	Between-Subject, Experimental	None	Accuracy	IAP	144 (72 familiar, 72 unfamiliar)	> 30 mins	Positive, Negative, Neutral	-	-	
Gibbs (2014)	Reboxetine, 4 mg, Acute	38	Healthy [b]	Placebo	34	Healthy [b]	Between-Subject, Experimental	None	Accuracy	IAP	144 (72 familiar, 72 unfamiliar)	> 30 mins	Positive, Negative, Neutral	-	-	
Tendolkar (2011)	Duloxetine, 60 mg, 14 days	18	Healthy	Placebo	18	Healthy	Crossover, Experimental	None	*d'*, RT	Photos	480 (240 familiar, 240 unfamiliar)	Immediately	Positive, Negative, Neutral	-	-	

IAP = International Affective Picture System

[a] Varied according to Alpha 2B Adrenoceptor Genotype, in this review we have grouped together

[b] Varied according to COMT Genotype, in this review we have grouped together

Other Memory

Two studies measured autobiographical memory; participants were presented with positive and negative cue words and were asked to recall specific personal memories based on these cues. One study also measured long-term recall of an emotional story (Table S8). No effects were reported in support of the cognitive neuropsychological theory. One study measuring autobiographical memory reported a main effect of faster reaction times in participants administered Reboxetine versus placebo [60]. The other study reported no effects of either Bupropion or Sertraline on autobiographical memory or recall of an emotional story [61].

Table S8

Studies assessing the influence of antidepressants on other measures of memory (autobiographical memory and long-term recall of an emotional story). ↑ indicates enhanced emotional processing (e.g., greater accuracy, faster reaction times), ↓ indicates diminished emotional processing (e.g., lower accuracy, slower reaction times), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Other Therapy**	**Task Design**			
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**			**Task**	**Stimuli**	**Trials**	**Max Retrieval**
Carvalho (2006)	Bupropion, 150 mg, Acute	35	Healthy	Placebo	35	Healthy	Between-Subject, Experimental	None	Autobiographical Memory	Words (Williams & Scott)	10	60 seconds
									Long-Term memory of an Emotional Story	Emotional Story	80	None
	Sertraline, 50 mg, Acute	35	Healthy						Autobiographical Memory	Words (Williams & Scott)	10	60 seconds
									Long-Term memory of an Emotional Story	Emotional Story	80	None
Papadatou-Pastou (2012)	Reboxetine, 4 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	None	Autobiographical Memory	Words (Williams & Broadbent)	12	20 seconds

Attention

Attentional Probe Task

Eighteen studies used an attentional probe task (Table S9). In keeping with the predictions of the cognitive neuropsychological theory, antidepressants were associated with reduced attention towards negative emotions in four studies [15,21,40,62] and increased attention towards positive emotions in one study [31]. One study reported that Buprenorphine decreased the number of first gazes towards emotional versus neutral faces, although this effect was heightened for faces expressing fear [50]. However, in contrast with the cognitive neuropsychological theory, one study reported decreased attention to positive words [33], four studies reported main effects in increasing processing of both positive and negative stimuli [2,9,19,29], and seven studies reported no effects of antidepressants [1,18,28,34,39,41,63].

Table S9

Studies assessing the influence of antidepressants on attention as measured by the dot-probe task. ↑ indicates enhanced emotional processing (e.g. increased attention), ↓ indicates diminished emotional processing (e.g. decreased attention), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**						**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Stimuli**	**Emotions**	**Presentation Length**	**Conditions**	**Trials**	**Outcomes**	**Positive**	**Negative**	
Bershad (2018)	Buprenorphine, 0.2 mg, Acute	38	Healthy	Placebo	38	Healthy	Crossover, Experimental	Faces	Neutral-Emotional	2000 ms	-	NR	First gaze	↓ (first gaze)	↓ (first gaze) [a]	
Browning (2007)	Citalopram, 20 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Words	Socially threatening-Neutral, Socially Relevant Positive-Neutral, Neutral-Neutral	500 ms, 14 ms	Unmasked, Masked	180	Median RT correct trials	↑ (Median RT)	-	
Bukala (2019)	Scopolamine, 1.5 mg, Acute	16	Healthy	Placebo	17	Healthy	Between-Subject, Experimental	NR	NR	NR	Unmasked, Masked	NR	Vigilance	-	-	+

Capitao (2015)	Fluoxetine, 20 mg, Acute	17	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Faces	Fear-Neutral, Happy-Neutral, Neutral-Neutral	100 ms, 1000 ms	Subliminal, Supraliminal	NR	Vigilance	-	-	+
Cerit (2015)	ARA290, 2 mg, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Pictures	Positive-Neutral, Negative-Neutral	500 ms	-	192	Vigilance	↑ (Vigilance)	↑ (Vigilance)	+
Chandra (2010)	Aprepitant, 125 mg, Acute	20	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Words	Socially threatening-Neutral, Socially Relevant Positive-Neutral, Neutral-Neutral	500 ms, 14 ms	Unmasked, Masked	180	Vigilance	↑ (Vigilance, Unmasked)	↑ (Vigilance, Unmasked)	+
Cooper (2015)	Tianeptine, 12.5 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Exper	Faces	Fear, Happy, Neutral	100 ms, 16 ms	Unmasked, Masked	192	Vigilance	↓ (Vigilance, Unm	-	

							imental							aske d)[b]		
Harmer (2011)	Agomelatine, 25 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Words	Positive, Negative, Neutral	500 ms, 14 ms	Unmasked, Masked	180	Vigilance	-	-	+
	Agomelatine, 50 mg, 1 week	16	Healthy											-	-	
Harmer (2013)	Citalopram, dose unclear, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	Words	Positive, Negative, Neutral	500 ms, 14 ms	Unmasked, Masked	180	Vigilance	-	-	
	GSK424887, dose unclear, Acute	18	Healthy											↑ (Vigilance, Unmasked)	↑ (Vigilance, Unmasked)	
Merens (2008)	Mixed type and dose, minimum 4 weeks	19	Remitted MDD	No Drug	20	Healthy	Between-Subject, Observational	Words	Threat-Neutral, Depression-Positive	500 ms	-	NR	Attentional Bias	-	-	+[c]
Murphy (2009b)	Citalopram, 20 mg, 1 week	14	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Faces	Fear-Neutral, Happy-Neutral, Neutral-	100 ms, 16 ms	Unmasked, Masked	192	Accuracy, RT, Vigilance	-	↓ (Vigilance, Unmasked)	
	Reboxetine, 4 mg	14	Healthy											-	-	

	b.i.d, 1 week								Neutral							
Pringle (2011)	Aprepitant, 125 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	Words	Socially threatening-Neutral, Positive-Neutral, Neutral-Neutral	500 ms, 14 ms	Unmasked, Masked	NR	Vigilance	-	↓ (Vigilance, Masked)	+
Pringle (2012)	Memantine, 10 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	Words	Socially threatening-Neutral, Positive-Neutral, Neutral-Neutral	500 ms, 14 ms	Unmasked, Masked	NR	Vigilance	-	-	+
Reed (2018)	Ketamine, 0.5 mg/kg, Acute	59	Healthy	Placebo	59	Healthy	Between-Subject, Experimental	Faces	Anger-Neutral, Happy-Neutral, Neutral-Neutral	500 ms		NR	RT, Vigilance	-	-	

Stein (2012)	Reboxetine, 20 mg, 1 week	12	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Faces	Sad-Neutral, Happy-Neutral	1000 ms	Infant, Adult	48	Vigilance	-	-	+[d]
	Citalopram, 4 mg b.i.d, 1 week	13	Healthy											-	-	
Walsh (2018)	Bupropion, 150 mg 7-10 days then 150 mg 2/day, 6 weeks	46	Current MDD	Placebo	42	Healthy	Within-subject Between-Subject groups experimental	Faces	Fear-Neutral, Happy-Neutral	NR	Unmasked, Masked	NR	RT	↑ (RT)	↑ (RT)	
Walsh (2018)	Bupropion, 150 mg 7-10 days then 150 mg 2/day, 6 weeks total	2 weeks: 43, 6 weeks: 60	Current MDD	None	46	Current MDD	Within-subject, Experimental	Faces	Fear-Neutral, Happy-Neutral	NR	Unmasked, Masked	NR	RT	↑ (RT)[e]	↑ (RT)[e]	
Walsh (2018b)	Bupropion, 150 mg, Acute	20	Healthy	Placebo	20	Healthy	Between-Subject, Experimental	Faces	Fear-Neutral, Happy-Neutral	NR	Unmasked, Masked	NR	Vigilance	-	↓ (Vigilance, Unmasked)	+
Warren (2019)	St John's Wort, 300	23	Healthy	Placebo	24	Healthy	Between-Subject,	Faces	Fear-Neutral, Happ	100 ms, 16 ms	Unmasked,	192	Vigilance	-	↓ (Vigilance,	

	mg 3/day, 1 week						Experimental		y-Neutral, Neutral-Neutral		Masked				Unmasked)	

Note: The vigilance outcome measure reflects differences between RTs in congruent trials (probe appears in the same position as emotion stimuli) versus incongruent trials (probe appears in opposite position from the emotional stimuli). Attentional bias scores indicate the difference between RTs for positive/neutral words from the RT for depressive/threatening words.

NR = Not Reported.

[a] Effects were most pronounced for fear

[b] Trend effect

[c] Threat-Neutral condition only

[d] Adult faces only

[e] Effects were reported for a group by time interaction with the MDD group showing a greater reduction in RTs compared to a healthy control group that did not receive an antidepressant, interaction effects with emotion were not reported.

Other Measures of Attention

A number of other tasks measuring attention were employed individually in five studies (Table S10). One observational study measuring eye movement during presentation of emotional images reported effects consistent with the cognitive neuropsychological theory. Individuals with depression taking an antidepressant showed longer gaze durations and greater fixations for positive images, and reduced fixations for negative images, compared to unmedicated participants experiencing depression [64]. Another study measuring eye movement during presentation of emotional faces in highly neurotic individuals reported somewhat inconsistent effects. Participants administered citalopram versus placebo showed main effects of increased scanning time for both positive and negative facial emotions, but also showed less spatial exploration of fearful faces of medium intensity [3]. The remaining studies measuring attention employed a rapid series visualisation task, an emotional oddball task, and a continuous performance task. No effects of antidepressants were reported [1,65,66].

Table S10

Studies assessing the influence of antidepressants on other measures of attention. ↑ indicates enhanced emotional processing (e.g. increased attention), ↓ indicates diminished emotional processing (e.g. decreased attention), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**						**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Type**	**Stimuli**	**Emotions**	**Duration**	**Trials**	**Outcomes**	**Positive**	**Negative**	
Capitao (2015)	Fluoxetine, 20 mg, Acute	17	Healthy	Placebo	18	Healthy	Between-subject Experimental	Rapid Series Visualisation	Words	Positive, Negative, Neutral	100 ms	NR	Accuracy	-	-	
Di Simplicio (2014)	Citalopram, 20 mg, 1 week	16	High Neuroticism	Placebo	17	High Neuroticism	Between-subject, experimental	Eye Tracking Task	Faces	Fear, Happiness, Neutral	500 ms	168	Scanpath length, Scanning time, Gaze Maintenance	↑ (Scanning time)	↓ (Scanpath length) ↑ (Scanning time)	
Robertson (2007)	Bupropion, 150 mg 1 week, followed by 300-450 mg based on clinical response for 7 weeks	8	Current MDD	None	10	Current MDD	Within-subject, observational	Emotional Oddball Task	Pictures (IAP)	Negative, Neutral	NR	NR	RT, 'Performance'	-	-	

Simmons (2009)	Escitalopram, 5 mg 3 days then 10 mg 18 days	15	Healthy	Placebo	15	Healthy	Crossover, experimental	Continuous Performance Task	Pictures (IAP)	Positive, Negative	2000 ms	34	RT, Accuracy	-	-	
Wells (2014)	Mixed [a]	21	Current MDD	None	47	Healthy	Between-subject, observational	Eye Tracking Task	Pictures (IAP)	Dysphoric, threat, positive, neutral	3000 ms	12	Total gaze time, Number of fixations, Fixation duration	-	-	
				None	26	Current MDD								↑ (Total gaze duration, Number of fixations)	↓ (Number of fixations)	

[a] Primary medication: n = 4 Venlafaxine, n = 3 duloxetine, n = 4 citalopram, n = 3 bupropion, n = 3 escitalopram, n = 1 sertraline, n = 1 paroxetine, n = 2 missing

Response Inhibition

The effect of antidepressants on response inhibition was assessed in eight studies (Table S11). Tasks included a Go/No-Go task (k = 2); an Implicit Association test (k = 1); and a Stroop task (k = 5). Most studies indicated that antidepressants did not impact response inhibition [39,67–70]. Two studies reported main effects of antidepressants; Psilocybin increased reaction times [71] and Duloxetine increased misses [72] across both positive and negative stimuli. Only one study reported effects consistent with the cognitive neuropsychological theory; healthy volunteers showed a greater increase in reaction times and error rates for negative words relative to positive after administration of Psilocybin versus placebo [16].

Table S11

Studies assessing the influence of antidepressants on response inhibition. ↑ indicates enhanced emotional processing (e.g. better response inhibition), ↓ indicates diminished emotional processing (e.g. worse response inhibition) , - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Other Therapies**	**Design**	**Task**						**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**			**Type**	**Stimuli**	**Emotions**	**Trials**	**Presentation Duration**	**Outcomes**	**Positive**	**Negative**	
Constant (2005)	Sertraline, 50 mg, 7 weeks [a]	20	Current MDD	None (baseline)	20	Current MDD [b]	Yes	Before/After, Experimental	Stroop	Words	Depression, Anxiety, Neutral	244	10,000 ms (supraliminal); 14 ms (subliminal)	RT	↑ (RT)	↑ (RT)	
Duschek (2014)	Unclear: 'taking antidepressants'	8	Females with Fibromyalgia	None	19	Females with Fibromyalgia [b]	Yes	Between-Subject, Observational	Stroop	Words	Positive, Negative, Neutral	120	2000 ms	Interference Score [d]	-	-	+
Fu (2015)	Duloxetine, initially 60 mg increased to a max of 120 mg, 12 weeks	24	Current MDD	None (baseline)	32	Current MDD [b]	No	Before/After, Experimental	Stroop	Words	Negative, Neutral	80	700 ms	RT, Accuracy	-	-	
Greer (2014)	Duloxetine, 30-120 mg, 12 weeks	21	Current MDD	None	21	Current MDD	No	Within-subject, Experimental	Go/No-Go	Words	Positive, Negative, Neutral	NR	NR	RT correct trials, Misses, False	↑ (Misses)	↑ (Misses)	

														Alarms			
Kometer (2012)	Psilocybin, 215 ug.kg, Acute [c]	17	Healthy	Placebo	17	Healthy	No	Crossover, Experimental	Go/No-Go	Words	Positive, Negative, Neutral	480	280 ms	RT, Error Rate	↑ (RT, Error Rates)[c]	↓ (RT, Error Rates)[c]	
McCabe (2009)	Aprepitant, 125 mg, Acute	12	Healthy	Placebo	12	Healthy	No	Between-Subject, Experimental	Stroop	Words	Positive, Negative, Neutral	160	1500 ms	RT, Accuracy	-	-	+
Merens (2008) Study 1	Varied, Minimum 4 weeks	19	Remitted MDD	None	20	Healthy	None	Between-Subject Observational	Implicit Association Test	NR	Self; Insects; Positive; Negative	NR	NR	Median RT Correct Trials	-	-	+
Merens (2008) Study 2	Varied, minimum 4 weeks	20	Remitted MDD	None	21	Healthy	No	Between-Subject, Observational	Stroop	Words	Positive, Depression, Neutral	NR	NR	RT, Interference Score[d]	-	-	

NR = Not Reported.

[a] One participant received 75 mg

[b] This study also included a healthy comparator group who did not receive the intervention. As we are interested in the effects of the intervention we have restricted analyses to those reporting the most suitable comparison between experimental and comparator conditions.

[c] Effects were not modulated by pre-treatment with Ketanserin (50 mg) versus placebo.

[d] Difference between the time taken to name the colors of affectively valenced words and the time needed to name the colors of neutral words

Matching

In eleven studies participants were presented with an emotional stimulus (faces or images) and asked to select or indicate matching emotional stimuli (Table S12). Only one study reported effects consistent with the cognitive neuropsychological theory; following scopolamine administration negative biases in patients with MDD, consisting of faster reactions to sad faces were no longer observed. However, no effect was observed for healthy controls [73]. Another study reported similar results in greater reaction times following administration of psilocybin although this was also observed for neutral faces and a positive condition was not used making the specificity of these effects unclear [74]. The remaining studies found no behavioural effects of antidepressants [75–81] or did not report behavioural results [82,83].

Table S12

Studies assessing the influence of antidepressants on matching of emotional stimuli. ↑ indicates enhanced emotional processing (e.g. greater accuracy, faster reaction times) ↓ indicates diminished emotional processing (e.g. worse accuracy, slower reaction times), - indicates no effect.

Study	**Experimental Condition**			**Control Condition**			**Design**	**Task**				**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Stimuli**	**Emotions**	**Matching Stimuli Presentation**	**Outcomes**	**Positive**	**Negative**	
Arce (2008)	Escitalopram, 5 mg/day 3 days then 10 mg/day 18 days	13	Healthy	Placebo	13	Healthy	Crossover, experimental	Faces	Anger, Fear, Happiness	Simultaneous	Accuracy, RT	-	-	
Fales (2009)	Escitalopram, 10-20 mg/day[a]	23	Current MDD	None (baseline)	23	Current MDD[b]	Within-subject, experimental	Faces	Fear, Neutral	Simultaneous	Accuracy, RT	-	-	
Furey (2015)	Scopolamine, 4.0 µg/kg, 3 sessions of 15 minutes	16	Current MDD	Placebo	16	Current MDD	Crossover, experimental	Faces	Happiness, Sadness	Simultaneous	RT, Bias	-	↓ (RT)	
	Scopolamine, 4.0 µg/kg, 3 sessions of 15 minutes	15	Healthy	Placebo	15	Healthy					RT, Bias	-	-	
Gorka (2019)	Varied[c]	28	Current MDD	None	59	Current MDD	Within-subject, observational	Faces	Anger, Fear, Happiness, Sadness	Simultaneous	Accuracy, RT	-	-	
Kraehenmann (2015)	Psilocybin, 0.16 mg/kg, Acute	25	Healthy	Placebo	25	Healthy	Crossover, experimental	Pictures from IAP	Negative, Neutral	Simultaneous	RT	-	↓ (RT)[d]	

Ma (2015)	Citalopram, 30 mg, Acute	46	Healthy	Placebo	46	Healthy	Crossover, experimental	Faces	Fear, Happiness, Neutral	Delayed	Accuracy, RT	-	-	
Norbury (2009)	Citalopram, 20 mg, 7-10 days	16	Healthy	Placebo	12	Healthy	Between-subject, experimental	Faces	Happiness, Fear	Simultaneous	Accuracy, RT	-	-	
Van Marle (2011)	Duloxetine, 60 mg, 14 days	19	Healthy	Placebo	19	Healthy	Crossover, experimental	Faces	Anger, Fear	Simultaneous	Accuracy, RT	-	-	
Windischberger (2010)	Citalopram, 20 mg, 10 days	18	Healthy	Placebo	18	Healthy	Crossover, experimental	Faces	Happiness, Anger, Fear, Sadness, Surprise, Disgust, Neutral/ Calm	Simultaneous	RT	-	-	
	Escitalopram, 10 mg, 10 days	18	Healthy									-	-	

Note: Bias scores reflect the difference in RT under implicit and explicit emotion-processing conditions

[a] 10 mg initially and then optional increase based on clinical response, end dose was 20 mg (n =9) and 10 mg (n = 8)

[b] This study also included a healthy comparator group who did not receive the intervention. As we are interested in the effects of the intervention we have restricted analyses to those reporting the most suitable comparison between experimental and comparator conditions.

[c] "The SSRI dosing schedule was designed to match real-world psychiatric practice. Therefore, as per standard guidelines, the assigned study physician determined, based on the participant's prior medication treatment history, which SSRI to initiate and a personalized, flexible dosing schedule for that individual (sertraline 50-200mg; citalopram 10-40mg; escitalopram 5-20mg; paroxetine 10-60mg; fluoxetine 10-80mg) with the goal of reaching a target dose at week 8 to be maintained for the final 4 weeks of the study. Of the 28 participants who received SSRIs, 14 received sertraline, 6 received escitalopram, 5 received fluoxetine, 2 received paroxetine, and 1 received citalopram. All participants reached target therapeutic dose by week 8"

[d] Represents a main effect as positive condition not measured, increased reaction times for both neutral and negative

[e] 50 mg for 3 days, than 100 mg for 3 week, then increase or decrease dependent on clinical indication. All p.s received a stable dose for at least 4 weeks before the final test session. At post-treatment session mean setraline dose was 105 mg (range 50 - 200mg)

Emotion-Potentiated Startle Response

Ten studies assessed the influence of antidepressants on emotion-potentiated startle responses (Table S13). Supportive of the cognitive neuropsychological theory, three studies reported that antidepressants were associated with a reduced response to negative images relative to control conditions [1,6,40]. Additionally, one study reported that participants administered Agomelatine showed a decreased startle response in the negative condition and enhanced startled response during the positive condition compared to placebo. Although this effect was only observed in participants administered the lower dose of the drug [28]. However, in contrast with these findings, one study reported an increase in the startle response for negative images [41], three studies observed main effects [7,26,31], and two studies reported no effects [9,21].

Table S13

Studies assessing the influence of antidepressants on emotion-potentiated startle responses. ↑ indicates enhanced emotional processing (e.g. greater startle responses) ↓ indicates diminished emotional processing (e.g. reduced startle responses), - indicates no effect.

Study	Experimental Condition			Comparator Condition			Design		Stimuli				Probe			Drug Effect		MA
	Drug	N	Sample	Drug	N	Sample		Outcome	Type	Emotions	No.	Presentation Duration	Type	Duration	Delivery Time	Positive	Negative	
Arnone (2009)	Mirtazapine, 15 mg, Acute	15	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	↓ (EBRM)	↓ (EBRM)	
Browning (2007)	Citalopram, 20 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	↑[a] (EBRM)	↑[a] (EBRM)	
Capitao (2015)	Fluoxetine, 20 mg, Acute	17	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	-	↓ (EBRM)	
Harmer (2004)	Citalopram, 20 mg, 1 week	14	Healthy	Placebo	14	Healthy	Between-Subject,	EBRM	IAP	Positive, Negative,	63	13 seconds	Acoustic, 95 dB burs	50 ms	1.5, 4.5 or 7.5 sec	-	↓[b] (EBRM)	

	Reboxetine, 4 mg b.i.d., 1 week	14	Healthy				Experimental			Neutral			ts white noise		onds	-	-	
Harmer (2008)	Duloxetine, 60 mg, Acute	12	Healthy	Placebo	12	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	↓[c] (EBRM)	↓[c] (EBRM)	
Harmer (2011)	Agomelatine, 25 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	↑ (EBRM)	↓ (EBRM)	
	Agomelatine, 50 mg, 1 week	16	Healthy													-	-	
Harmer (2013)	Citalopram, dose unclear, Acute	18	Healthy	Placebo	18	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	-	-	
	GSK424887, dose unclear, Acute	18	Healthy													-	-	
Pringle (2011)	Aprepitant, 125 mg, 1 week	16	Healthy	Placebo	16	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	-	↓[d] (EBRM)	

Pringle (2012)	Memantine, 10 mg, Acute	16	Healthy	Placebo	15	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	-	↑ (EBRM)	
Warren (2019)	St John's Wort, 300 mg 3 x a day, 1 week	23	Healthy	Placebo	24	Healthy	Between-Subject, Experimental	EBRM	IAP	Positive, Negative, Neutral	63	13 seconds	Acoustic, 95 dB bursts white noise	50 ms	1.5, 4.5 or 7.5 seconds	-	-	

EBRM = Eye-blink reflex magnitude

[a] Only found for raw scores, z-scores indicated no effect.

[b] Placebo and Reboxetine groups showed greater blink reflex magnitude for Negative vs. Positive & Neutral pictures, this was abolished in drug group.

[c] Trend Effect

[d] Effect of emotion observed in the placebo group, was abolished in the group receiving Aprepitant

Interpretative Biases

Two studies assessed threat-related interpretative biases through homophone tasks (Table S14). One study reported a reduction in the percentage of spellings related to threat interpretations of homophones following antidepressant treatment in patients with generalised anxiety disorder [84]. However, the other study reported no effects of either Citalopram or Reboxetine on response to threat-related words [85].

Table S14

Studies assessing the influence of antidepressants on interpretative biases. ↑ indicates enhanced emotional processing (e.g. greater biases) ↓ indicates diminished emotional processing (e.g. reduced biases), - indicates no effect.

Study	**Experimental Condition**			**Comparator Condition**			**Design**	**Task**					**Drug Effect**		**MA**
	Drug	**N**	**Sample**	**Drug**	**N**	**Sample**		**Type**	**Stimuli**	**Emotions**	**Trials**	**Outcomes**	Positive	Negative	
Mogg (2004)	Paorxetine/Citalopram, 20 mg, 4 weeks	16	Current GAD	None (baseline)	19	Current GAD[a]	Within-subject, Experimental	Homophone Task	Words	Threat-related, nonthreat-related, neutral filler words	28	Interpretative Bias Index (IBI)	-	↓ (IBI)	
Murphy (2009b)	Citalopram, 20 mg, 1 week	14	Healthy	Placebo	14	Healthy	Between-Subject, Experimental	Homograph Primed Lexican Decision Task	Words	Threat-related, neutral	80	RT, Accuracy	-	-	
	Reboxetine, 4 mg b.i.d, 1 week	14	Healthy										-	-	

IBI = Interpretative Bias Index, the percentage of spellings that corresponded to threat-related interpretations

[a] This study also included a healthy comparator group who did not receive the intervention. As we are interested in the effects of the intervention, we have restricted analyses to those reporting the most suitable comparison between experimental and comparator conditions

Olfactory Biases

One study asked participants to smell eight different odours and evaluate their pleasantness, familiarity, and intensity of the emotion evoked. Patients rated unpleasant odours as more pleasant following clinical improvement from antidepressant treatment [12].

Table S15

Results from multilevel meta-analyses examining differences in processing of positive and negative stimuli in participants administered antidepressants compared to a control condition (placebo or no drug) based on outcomes secondly most commonly reported with estimated standardised mean differences overall and by cognitive domain

	k		***n***		**SMD**	**SE**	**95% CI**	***p***
	Studies	**Effects**	**Experimental**	**Comparator**				
Positive [a]	20	50	454	468	0.01	0.05	-0.09, 0.12	0.791
Moderators [b]								
Word Categorisation (Accuracy)	12	12	183	178	-0.06	0.11	-0.28, 0.16	0.562
Dot Probe (Accuracy Vigilance)	3	3	63	47	0.33	0.20	-0.08, 0.73	0.110
FERT (RT)	16	18	309	357	-0.04	0.08	-0.21, 0.12	0.604
Word Recall (Misclassifications)	8	9	149	145	-0.02	0.12	-0.26,	0.846

	Word Recognition (RT)	8	8	141	124	0.18	0.13	0.22 - 0.07, 0.44	0.158
Negative [c]		20	50	451	463	-0.02	0.05	-0.12, 0.08	0.688
Moderators [d]									
	Word Categorisation (Accuracy)	12	12	183	178	0.09	0.11	-0.12, 0.30	0.400
	Dot Probe (Accuracy Vigilance)	3	3	64	47	0.00	0.20	-0.39, 0.39	0.997
	FERT (RT)	16	18	305	351	-0.07	0.08	-0.23, 0.10	0.422
	Word Recall (Misclassifications)	8	9	149	145	-0.10	0.12	-0.34, 0.13	0.377

Word Recognition (RT)	8	8	141	124	0.02	0.12	-0.23, 0.27	0.864

[a] Test for heterogeneity: Q(49) = 58.10, p = 0.175, σ Level 3 = 0.01, σ Level 2 = 0.00, I^2 Level 3 = 4.00%, I^2 Level 2 = 0.00%

[b] Test of moderators: F(5, 45) = 1.07, p = 0.388

[c] Test for heterogeneity: Q(49) = 41.29, p = 0.775, σ Level 3 = 0.00, σ Level 2 = 0.00, I^2 Level 3 = 0.00%, I^2 Level 2 = 0.00%

[d] Test of moderators: F(5, 45) = 0.441, p = 0.818

SMD = Standardised Mean Difference calculated using Hedge's g, SE = Standard Error, CI = Confidence Interval, RT = Reaction Times, Accuracy Vigilance = Vigilance scores calculated by subtracting accuracy in congruent trials (probe appears in same position as emotional stimuli) from incongruent trials (probe appears in opposite position from the emotional stimuli)

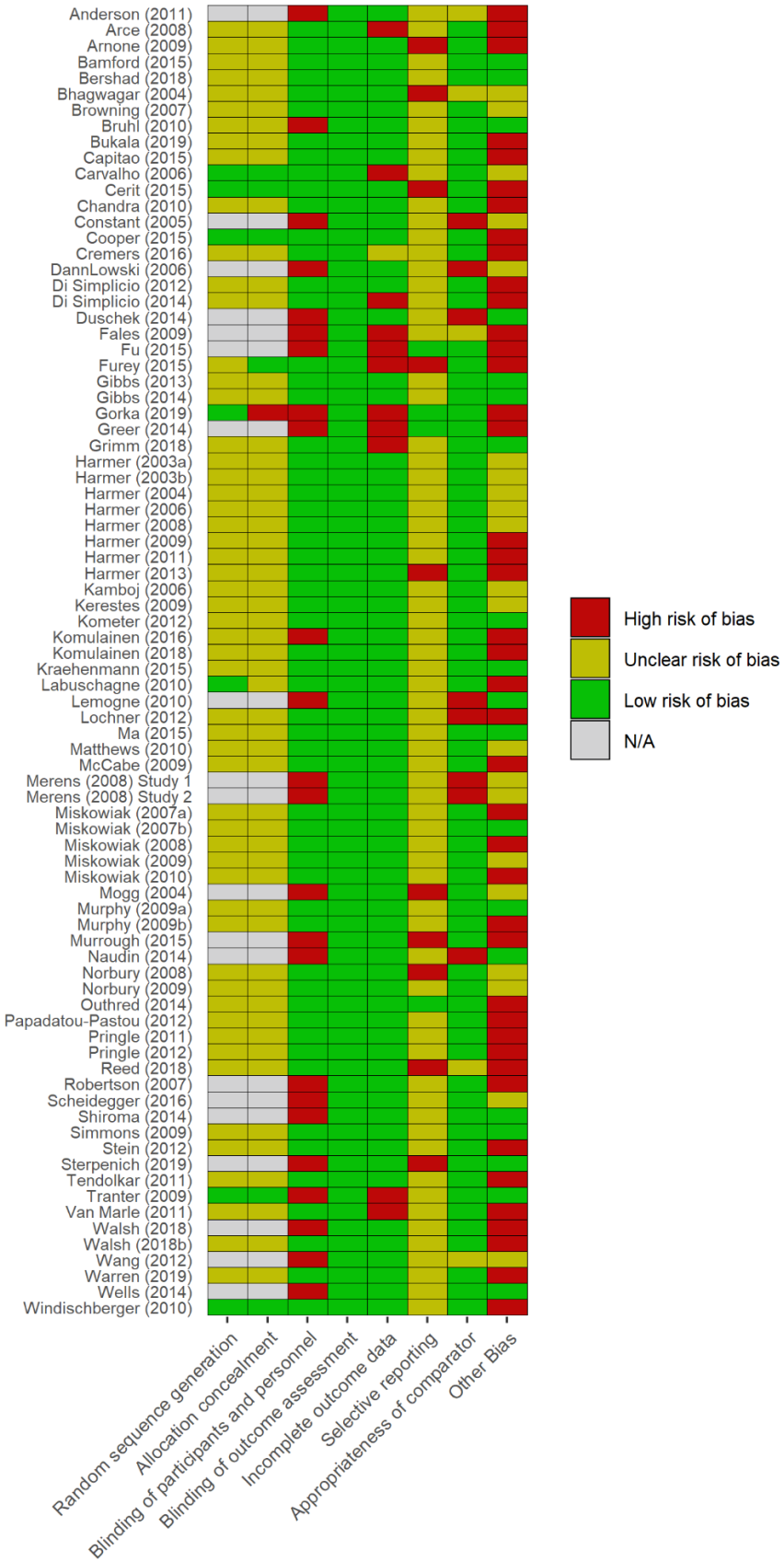


Figure S1

Risk of Bias judgements according to domain of bias and study

References

1. Capitão LP, Murphy SE, Browning M, Cowen PJ, Harmer CJ. 2015 Acute fluoxetine modulates emotional processing in young adult volunteers. *Psychol. Med.* **45**, 2295–2308. (doi:10.1017/S0033291715000240)

2. Chandra P, Hafizi S, Massey-Chase RM, Goodwin GM, Cowen PJ, Harmer CJ. 2010 NK1 receptor antagonism and emotional processing in healthy volunteers. *J. Psychopharmacol.* **24**, 481–487. (doi:10.1177/0269881109103101)

3. Di Simplicio M, Doallo S, Costoloni G, Rohenkohl G, Nobre AC, Harmer CJ. 2014 'Can you look me in the face?' Short-term SSRI administration reverts avoidant ocular face exploration in subjects at risk for psychopathology. *Neuropsychopharmacology* **39**, 3059–3066. (doi:10.1038/npp.2014.159)

4. Harmer CJ, Hill SA, Taylor MJ, Cowen PJ, Goodwin GM. 2003 Toward a neuropsychological theory of antidepressant drug action: Increase in positive emotional bias after potentiation of norepinephrine activity. *Am. J. Psychiatry* (doi:10.1176/appi.ajp.160.5.990)

5. Harmer CJ, Bhagwagar Z, Perrett DI, Völlm BA, Cowen PJ, Goodwin GM. 2003 Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology* **28**, 148–152. (doi:10.1038/sj.npp.1300004)

6. Harmer CJ, Shelley NC, Cowen PJ, Goodwin GM. 2004 Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *Am. J. Psychiatry* **161**, 1256–1263. (doi:10.1176/appi.ajp.161.7.1256)

7. Harmer CJ, Heinzen J, O'Sullivan U, Ayres RA, Cowen PJ. 2008 Dissociable effects of acute antidepressant drug administration on subjective and emotional processing measures in healthy volunteers. *Psychopharmacology (Berl).* **199**, 495–502. (doi:10.1007/s00213-007-1058-7)

8. Harmer CJ, O'Sullivan U, Favaron E, Massey-Chase R, Ayres R, Reinecke A, Goodwin GM, Cowen PJ. 2009 Effect of acute antidepressant administration on negative affective bias in depressed patients. *Am. J. Psychiatry* **166**, 1178–1184. (doi:10.1176/appi.ajp.2009.09020149)

9. Harmer CJ *et al.* 2013 Combined NK1 antagonism and serotonin reuptake inhibition: Effects on emotional processing in humans. *J. Psychopharmacol.* **27**, 435–443. (doi:10.1177/0269881112472558)

10. Miskowiak K, Inkster B, Selvaraj S, Wise R, Goodwin GM, Harmer CJ. 2008 Erythropoietin improves mood and modulates the cognitive and neural processing of emotion 3 days post administration. *Neuropsychopharmacology* **33**, 611–618. (doi:10.1038/sj.npp.1301439)

11. Murphy SE, Norbury R, O'Sullivan U, Cowen PJ, Harmer CJ. 2009 Effect of a single dose of citalopram on amygdala response to emotional faces. *Br. J. Psychiatry* **194**, 535–540. (doi:10.1192/bjp.bp.108.056093)

12. Naudin M, Carl T, Surguladze S, Guillen C, Gaillard P, Belzung C, El-Hage W, Atanasova B. 2014 Perceptive biases in major depressive episode. *PLoS One* **9**, e86832. (doi:10.1371/journal.pone.0086832)

13. Shiroma PR, Thuras P, Johns B, Lim KO. 2014 Emotion recognition processing as early predictor of response to 8-week citalopram treatment in late-life depression. *Int. J. Geriatr. Psychiatry* **29**, 1132–1139. (doi:10.1002/gps.4104)

14. Tranter R, Bell D, Gutting P, Harmer C, Healy D, Anderson IM. 2009 The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *J. Affect. Disord.* **118**, 87–93. (doi:10.1016/j.jad.2009.01.028)

15. Walsh AEL, Huneke NTM, Brown R, Browning M, Cowen P, Harmer CJ. 2018 A dissociation of the acute effects of bupropion on positive emotional processing and reward processing in healthy volunteers. *Front. Psychiatry* **9**. (doi:10.3389/fpsyt.2018.00482)

16. Kometer M, Schmidt A, Bachmann R, Studerus E, Seifritz E, Vollenweider FX. 2012 Psilocybin biases facial recognition, goal-directed behavior, and mood state toward positive relative to negative emotions through different serotonergic subreceptors. *Biol. Psychiatry* **72**, 898–906. (doi:10.1016/j.biopsych.2012.04.005)

17. Murrough JW *et al.* 2015 Regulation of neural responses to emotion perception by ketamine in individuals with treatment-resistant major depressive disorder. *Transl. Psychiatry* **5**, e509–e509. (doi:10.1038/tp.2015.10)

18. Stein A, Murphy S, Arteche A, Lehtonen A, Harvey A, Craske MG, Harmer C. 2012 Effects of reboxetine and citalopram on appraisal of infant facial expressions and attentional biases. *J. Psychopharmacol.* **26**, 670–676. (doi:10.1177/0269881111421970)

19. Cerit H, Veer IM, Dahan A, Niesters M, Harmer CJ, Miskowiak KW, Rombouts SARB, Van der Does W. 2015 Testing the antidepressant properties of the peptide ARA290 in a human neuropsychological model of drug action. *Eur. Neuropsychopharmacol.* **25**, 2289–2299. (doi:10.1016/j.euroneuro.2015.09.005)

20. Lochner C, Simmons C, Kidd M, Chamberlain SR, Fineberg NA, van Honk J, Ipser J, Stein DJ. 2012 Differential effects of escitalopram challenge on disgust processing in obsessive-compulsive disorder. *Behav. Brain Res.* **226**, 274–280. (doi:10.1016/j.bbr.2011.09.029)

21. Warren MB, Cowen PJ, Harmer CJ. 2019 Subchronic treatment with St John's wort produces a positive shift in emotional processing in healthy volunteers. *J. Psychopharmacol.* **33**, 194–201. (doi:10.1177/0269881118812101)

22. Bhagwagar Z, Cowen PJ, Goodwin GM, Harmer CJ. 2004 Normalization of Enhanced Fear Recognition by Acute SSRI Treatment in Subjects with a Previous History of Depression. *Am. J. Psychiatry* **161**, 166–168. (doi:10.1176/appi.ajp.161.1.166)

23. Harmer CJ, Mackay CE, Reid CB, Cowen PJ, Goodwin GM. 2006 Antidepressant Drug Treatment Modifies the Neural Processing of Nonconscious Threat Cues. *Biol. Psychiatry* **59**, 816–820. (doi:10.1016/j.biopsych.2005.10.015)

24. Miskowiak K, O'Sullivan U, Harmer CJ. 2007 Erythropoietin Reduces Neural and Cognitive Processing of Fear in Human Models of Antidepressant Drug Action. *Biol. Psychiatry* **62**, 1244–1250. (doi:10.1016/j.biopsych.2007.01.011)

25. Miskowiak KW, Favaron E, Hafizi S, Inkster B, Goodwin GM, Cowen PJ, Harmer CJ. 2010 Erythropoietin modulates neural and cognitive processing of emotional information in biomarker models of antidepressant drug action in depressed patients. *Psychopharmacology (Berl).* **210**, 419–428. (doi:10.1007/s00213-010-1842-7)

26. Arnone D, Horder J, Cowen PJ, Harmer CJ. 2009 Early effects of mirtazapine on emotional processing. *Psychopharmacology (Berl).* **203**, 685. (doi:10.1007/s00213-008-1410-6)

27. Bamford S, Penton-Voak I, Pinkney V, Baldwin DS, Munafò MR, Garner M. 2015 Early effects of duloxetine on emotion recognition in healthy volunteers. *J. Psychopharmacol.* **29**, 634–641. (doi:10.1177/0269881115570085)

28. Harmer CJ, De Bodinat C, Dawson GR, Dourish CT, Waldenmaier L, Adams S, Cowen PJ, Goodwin GM. 2011 Agomelatine facilitates positive versus negative affective processing in healthy volunteer models. *J. Psychopharmacol.* **25**, 1159–1167. (doi:10.1177/0269881110376689)

29. Walsh AEL, Browning M, Drevets WC, Furey M, Harmer CJ. 2018 Dissociable temporal effects of bupropion on behavioural measures of emotional and reward processing in depression. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170030. (doi:10.1098/rstb.2017.0030)

30. Kamboj SK, Curran H V. 2006 Scopolamine induces impairments in the recognition of human facial expressions of anger and disgust. *Psychopharmacology (Berl).* **185**, 529–535. (doi:10.1007/s00213-006-0332-4)

31. Browning M, Reid C, Cowen PJ, Goodwin GM, Harmer CJ. 2007 A single dose of citalopram increases fear recognition in healthy subjects. *J. Psychopharmacol.* **21**, 684–690. (doi:10.1177/0269881106074062)

32. Anderson IM *et al.* 2011 State-dependent alteration in face emotion recognition in depression. *Br. J. Psychiatry* **198**, 302–308. (doi:10.1192/bjp.bp.110.078139)

33. Cooper CM, Whiting DA, Cowen PJ, Harmer CJ. 2015 Tianeptine in an experimental medicine model of antidepressant action. *J. Psychopharmacol.* **29**, 582–590. (doi:10.1177/0269881115573810)

34. Bukala BR, Browning M, Cowen PJ, Harmer CJ, Murphy SE. 2019 Overnight transdermal scopolamine patch administration has no clear effect on cognition and emotional processing in healthy volunteers. *J. Psychopharmacol.* **33**, 255–257. (doi:10.1177/0269881118817161)

35. Cremers H, Lee R, Keedy S, Phan KL, Coccaro E. 2016 Effects of escitalopram administration on face processing in intermittent explosive disorder: An fMRI Study. *Neuropsychopharmacology* **41**, 590–597. (doi:10.1038/npp.2015.187)

36. Grimm O, Kraehenmann R, Preller KH, Seifritz E, Vollenweider FX. 2018 Psilocybin modulates functional connectivity of the amygdala during emotional face discrimination. *Eur. Neuropsychopharmacol.* **28**, 691–700. (doi:10.1016/j.euroneuro.2018.03.016)

37. Kerestes R, Labuschagne I, Croft RJ, O'Neill B V., Bhagwagar Z, Phan KL, Nathan PJ. 2009 Evidence for modulation of facial emotional processing bias during emotional expression decoding by serotonergic and noradrenergic antidepressants: An event-related potential (ERP) study. *Psychopharmacology (Berl).* **202**, 621–634. (doi:10.1007/s00213-008-1340-3)

38. Labuschagne I, Croft RJ, Phan KL, Nathan PJ. 2010 Augmenting serotonin neurotransmission with citalopram modulates emotional expression decoding but not structural encoding of moderate intensity sad facial emotional stimuli: An event-related potential (ERP) investigation. *J. Psychopharmacol.* **24**, 1153–1164. (doi:10.1177/0269881108097878)

39. Merens W, Booij L, Van Der Does AJW. 2008 Residual cognitive impairments in remitted depressed patients. Study 1. *Depress. Anxiety* **25**, E27–E36. (doi:10.1002/da.20391)

40. Pringle A, McTavish SF, Williams C, Smith R, Cowen PJ, Harmer CJ. 2011 Short-term NK1 receptor antagonism and emotional processing in healthy volunteers. *Psychopharmacology (Berl).* **215**, 239–246. (doi:10.1007/s00213-010-2133-z)

41. Pringle A, Parsons E, Cowen LG, McTavish SF, Cowen PJ, Harmer CJ. 2012 Using an experimental medicine model to understand the antidepressant potential of the N-Methyl-D-aspartic acid (NMDA) receptor antagonist memantine. *J. Psychopharmacol.* **26**, 1417–1423. (doi:10.1177/0269881112446535)

42. Dannlowski U, Kersting A, Donges US, Lalee-Mentzel J, Arolt V, Suslow T. 2006 Masked facial affect priming is associated with therapy response in clinical depression. *Eur. Arch. Psychiatry Clin. Neurosci.* **256**, 215–221. (doi:10.1007/s00406-005-0628-0)

43. Komulainen E, Heikkilä R, Nummenmaa L, Raij TT, Harmer CJ, Isometsä E, Ekelund J. 2018 Short-term escitalopram treatment normalizes aberrant self-referential processing in major depressive disorder. *J. Affect. Disord.* **236**, 222–229. (doi:10.1016/j.jad.2018.04.096)

44. Matthews SC, Simmons AN, Strigo IA, Arce E, Stein MB, Paulus MP. 2010 Escitalopram attenuates posterior cingulate activity during self-evaluation in healthy volunteers. *Psychiatry Res. - Neuroimaging* **182**, 81–87. (doi:10.1016/j.pscychresns.2010.02.003)

45. Di Simplicio M, Norbury R, Harmer CJ. 2012 Short-term antidepressant administration reduces negative self-referential processing in the medial prefrontal cortex in subjects at risk for depression. *Mol. Psychiatry* **17**, 503–510. (doi:10.1038/mp.2011.16)

46. Komulainen E *et al.* 2016 A single dose of mirtazapine attenuates neural responses to self-referential processing. *J. Psychopharmacol.* **30**, 23–32. (doi:10.1177/0269881115616384)

47. Lemogne C, Mayberg H, Bergouignan L, Volle E, Delaveau P, Lehéricy S, Allilaire JF, Fossati P. 2010 Self-referential processing and the prefrontal cortex over the course of depression: A pilot study. *J. Affect. Disord.* **124**, 196–201. (doi:10.1016/j.jad.2009.11.003)

48. Miskowiak K, Papadatou-Pastou M, Cowen PJ, Goodwin GM, Norbury R, Harmer CJ. 2007 Single dose antidepressant administration modulates the neural processing of self-referent personality trait words. *Neuroimage* **37**, 904–911. (doi:10.1016/j.neuroimage.2007.05.036)

49. Norbury R, Mackay CE, Cowen PJ, Goodwin GM, Harmer CJ. 2008 The effects of reboxetine on emotional processing in healthy volunteers: An fMRI study. *Mol. Psychiatry* **13**, 1011–1020. (doi:10.1038/sj.mp.4002091)

50. Bershad AK, Ruiz NA, De Wit H. 2018 Effects of buprenorphine on responses to emotional stimuli in individuals with a range of mood symptomatology. *Int. J. Neuropsychopharmacol.* **21**, 120–127. (doi:10.1093/ijnp/pyx077)

51. Scheidegger M, Henning A, Walter M, Lehmann M, Kraehenmann R, Boeker H, Seifritz E, Grimm S. 2016 Ketamine administration reduces amygdalo-hippocampal reactivity to emotional stimulation. *Hum. Brain Mapp.* (doi:10.1002/hbm.23148)

52. Sterpenich V *et al.* 2019 Increased Reactivity of the Mesolimbic Reward System after Ketamine Injection in Patients with Treatment-resistant Major Depressive Disorder. *Anesthesiology* **130**, 923–935. (doi:10.1097/ALN.0000000000002667)

53. Brühl AB, Kaffenberger T, Herwig U. 2010 Serotonergic and noradrenergic modulation of emotion processing by single dose antidepressants. *Neuropsychopharmacology* **35**, 521–533. (doi:10.1038/npp.2009.159)

54. Gibbs AA, Bautista CE, Mowlem FD, Naudts KH, Duka T. 2013 Alpha 2B adrenoceptor genotype moderates effect of Reboxetine on negative emotional memory bias in healthy volunteers. *J. Neurosci.* **33**, 17023–17028. (doi:10.1523/JNEUROSCI.2124-13.2013)

55. Gibbs AA, Bautista CE, Mowlem FD, Naudts KH, Duka DT. 2014 Catechol-O-methyltransferase val158met genotype determines effect of reboxetine on emotional memory in healthy male volunteers. *J. Psychiatry Neurosci.* **39**, E24. (doi:10.1503/jpn.130131)

56. Outhred T, Das P, Felmingham KL, Bryant RA, Nathan PJ, Malhi GS, Kemp AH. 2014 Impact of acute administration of escitalopram on the processing of emotional and neutral images: A randomized crossover fMRI study of healthy women. *J. Psychiatry Neurosci.* **39**, 267. (doi:10.1503/jpn.130118)

57. Wang Y, Xu C, Cao X, Gao Q, Li J, Liu Z, Sun N, Ren Y, Zhang K. 2012 Effects of an antidepressant on neural correlates of emotional processing in patients with major depression. *Neurosci. Lett.* **527**, 55–59. (doi:10.1016/j.neulet.2012.08.034)

58. Miskowiak KW, Favaron E, Hafizi S, Inkster B, Goodwin GM, Cowen PJ, Harmer CJ. 2009 Effects of erythropoietin on emotional processing biases in patients with major depression: An exploratory fMRI study. *Psychopharmacology (Berl).* **207**, 133–142. (doi:10.1007/s00213-009-1641-1)

59. Tendolkar I, Van Wingen G, Urner M, Jan Verkes R, Fernández G. 2011 Short-term duloxetine administration affects neural correlates of mood-congruent memory. *Neuropsychopharmacology* **36**, 2266–2275. (doi:10.1038/npp.2011.114)

60. Papadatou-Pastou M, Miskowiak KW, Williams JMG, Harmer CJ, Reinecke A. 2012 Acute antidepressant drug administration and autobiographical memory recall: A functional magnetic resonance imaging study. *Exp. Clin. Psychopharmacol.* **20**, 364. (doi:10.1037/a0027969)

61. Carvalho AF, Köhler CA, Cruz EP, Stürmer PL, Reichman BP, Barea BM, Izquierdo I, Chaves MLF. 2006 Acute treatment with the antidepressants bupropion and sertraline do

not influence memory retrieval in man. *Eur. Arch. Psychiatry Clin. Neurosci.* **256**, 320–325. (doi:10.1007/s00406-006-0640-z)

62. Murphy SE, Norbury R, O'Sullivan U, Cowen PJ, Harmer CJ. 2009 Effect of a single dose of citalopram on amygdala response to emotional faces. *Br. J. Psychiatry* **194**, 535–540. (doi:10.1192/bjp.bp.108.056093)

63. Reed JL, Nugent AC, Furey ML, Szczepanik JE, Evans JW, Zarate CA. 2018 Ketamine normalizes brain activity during emotionally valenced attentional processing in depression. *NeuroImage Clin.* **20**, 92–101. (doi:10.1016/j.nicl.2018.07.006)

64. Wells TT, Clerkin EM, Ellis AJ, Beevers CG. 2014 Effect of antidepressant medication use on emotional information processing in major depression. *Am. J. Psychiatry* **171**, 195–200. (doi:10.1176/appi.ajp.2013.12091243)

65. Robertson B, Wang L, Diaz MT, Aiello M, Gersing K, Beyer J, Mukundan S, McCarthy G, Doraiswamy PM. 2007 Effect of bupropion extended release on negative emotion processing in major depressive disorder: A pilot functional magnetic resonance imaging study. *J. Clin. Psychiatry* **68**, 261–267. (doi:10.4088/JCP.v68n0212)

66. Simmons AN, Arce E, Lovero KL, Stein MB, Paulus MP. 2009 Subchronic SSRI administration reduces insula response during affective anticipation in healthy volunteers. *Int. J. Neuropsychopharmacol.* **12**, 1009–1020. (doi:10.1017/S1461145709990149)

67. Duschek S, Werner NS, Limbert N, Winkelmann A, Montoya P. 2014 Attentional bias toward negative information in patients with fibromyalgia syndrome. *Pain Med. (United States)* **15**, 603–612. (doi:10.1111/pme.12360)

68. Fu CHY *et al.* 2015 Multimodal functional and structural neuroimaging investigation of major depressive disorder following treatment with duloxetine. *BMC Psychiatry* **15**, 1–11. (doi:10.1186/s12888-015-0457-2)

69. McCabe C, Cowen PJ, Harmer CJ. 2009 NK1 receptor antagonism and the neural processing of emotional information in healthy volunteers. *Int. J. Neuropsychopharmacol.* **12**, 1261–1274. (doi:10.1017/S1461145709990150)

70. Merens W, Booij L, Van Der Does AJW. 2008 Residual cognitive impairments in remitted depressed patients. Study 2. *Depress. Anxiety* **25**, E27–E36. (doi:10.1002/da.20391)

71. Constant EL, Adam S, Gillain B, Seron X, Bruyer R, Seghers A. 2005 Effects of sertraline on depressive symptoms and attentional and executive functions in major depression. *Depress. Anxiety* **21**, 78–89. (doi:10.1002/da.20060)

72. Greer TL, Sunderajan P, Grannemann BD, Kurian BT, Trivedi MH. 2014 Does duloxetine improve cognitive function independently of its antidepressant effect in patients with major depressive disorder and subjective reports of cognitive dysfunction? *Depress. Res. Treat.* (doi:10.1155/2014/627863)

73. Furey ML, Drevets WC, Szczepanik J, Khanna A, Nugent A, Zarate CA. 2015 Pretreatment differences in BOLD response to emotional faces correlate with Antidepress ant response to scopolamine. *Int. J. Neuropsychopharmacol.* **18**. (doi:10.1093/ijnp/pyv028)

74. Kraehenmann R, Preller KH, Scheidegger M, Pokorny T, Bosch OG, Seifritz E, Vollenweider FX. 2015 Psilocybin-induced decrease in amygdala reactivity correlates with enhanced positive mood in healthy volunteers. *Biol. Psychiatry* **78**, 572–581. (doi:10.1016/j.biopsych.2014.04.010)

75. Arce E, Simmons AN, Lovero KL, Stein MB, Paulus MP. 2008 Escitalopram effects on insula and amygdala BOLD activation during emotional processing. *Psychopharmacology (Berl).* **196**, 661–672. (doi:10.1007/s00213-007-1004-8)

76. Fales CL, Barch DM, Rundle MM, Mintun MA, Mathews J, Snyder AZ, Sheline YI. 2009 Antidepressant treatment normalizes hypoactivity in dorsolateral prefrontal cortex during emotional interference processing in major depression. *J. Affect. Disord.* **112**, 206–211. (doi:10.1016/j.jad.2008.04.027)

77. Gorka SM *et al.* 2019 Emotion-based brain mechanisms and predictors for SSRI and CBT treatment of anxiety and depression: a randomized trial. *Neuropsychopharmacology* **44**, 1639–1648. (doi:10.1038/s41386-019-0407-7)

78. Ma Y, Li B, Wang C, Zhang W, Rao Y, Han S. 2015 Allelic variation in 5-HTTLPR and the effects of citalopram on the emotional neural network. *Br. J. Psychiatry* **206**, 385–392. (doi:10.1192/bjp.bp.114.150128)

79. Norbury R, Taylor MJ, Selvaraj S, Murphy SE, Harmer CJ, Cowen PJ. 2009 Short-term antidepressant treatment modulates amygdala response to happy faces. *Psychopharmacology (Berl).* **206**, 197–204. (doi:10.1007/s00213-009-1597-1)

80. van Marle HJF, Tendolkar I, Urner M, Verkes RJ, Fernández G, van Wingen G. 2011 Subchronic duloxetine administration alters the extended amygdala circuitry in healthy individuals. *Neuroimage* **55**, 825–831. (doi:10.1016/j.neuroimage.2010.12.051)

81. Windischberger C *et al.* 2010 Area-specific modulation of neural activation comparing escitalopram and citalopram revealed by pharmaco-fMRI: A randomized cross-over study. *Neuroimage* **49**, 1161–1170. (doi:10.1016/j.neuroimage.2009.10.013)

82. Furey ML, Drevets WC, Hoffman EM, Frankel E, Speer AM, Zarate CA. 2013 Potential of pretreatment neural activity in the visual cortex during emotional processing to predict treatment response to scopolamine in major depressive disorder. *JAMA Psychiatry* **70**, 280–290. (doi:10.1001/2013.jamapsychiatry.60)

83. Victor TA, Furey ML, Fromm SJ, Öhman A, Drevets WC. 2010 Relationship between amygdala responses to masked faces and mood state and treatment in major depressive disorder. *Arch. Gen. Psychiatry* **67**, 1128–1138. (doi:10.1001/archgenpsychiatry.2010.144)

84. Mogg K, Baldwin DS, Brodrick P, Bradley BP. 2004 Effect of short-term SSRI treatment on cognitive bias in generalised anxiety disorder. *Psychopharmacology (Berl).* **176**, 466–470. (doi:10.1007/s00213-004-1902-y)

85. Murphy SE, Yiend J, Lester KJ, Cowen PJ, Harmer CJ. 2009 Short-term serotonergic but not noradrenergic antidepressant administration reduces attentional vigilance to threat in healthy volunteers. *Int. J. Neuropsychopharmacol.* **12**, 169–179. (doi:10.1017/S1461145708009164)

Chapter 4 Supplementary Materials

Supplementary materials for Hobbs, C., Murphy, S. E., Wright, L., Carson, J., Van Assche, I., O'Brien, J., ... & Button, K. S. (2020). Effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers. *BJPsych open*, *6*(6).

Supplementary Methods

Detailed Description of Cognitive Tasks

Study materials are available on Open Science Framework (https://osf.io/nhjvs/). Scripted text used by researchers to instruct participants for each cognitive task are available in the case report form, which is publicly available on Open Science Framework.

Prisoner's Dilemma

Cooperative behaviours were measured using an iterated Prisoner's Dilemma task (1,2). Participants were instructed that they had to choose to work with (cooperate) or against (defect) the computer to win points. On each trial, participants were shown a 2 x 2 matrix demonstrating the possible outcomes of players' choices. Participants made their choice, before being shown the other's choice and the resulting outcome. Points were structured so that the highest amount of points was won by defecting whilst the other chose to cooperate, followed in descending order by both players choosing to cooperate, both players choosing to defect, and choosing to cooperate whilst the other defected.

Participants completed two blocks of 26 trials, which differed according to social context. In one block the other player initially cooperated, whereas in a second block the other player initially defected. Two 'other' personas were used for each block. Order of social context and assignment of 'other' personae were counterbalanced. After the first trial, the other followed a 'tit for tat' strategy, mirroring the choice of the participant in the previous trial.

The proportion of cooperative choices and reaction times for cooperative choices were recorded.

Social Evaluation Learning

Learning of social evaluations was measured using a reinforcement learning task (3,4). Participants were told that they had to learn how much the computer liked themselves, a friend, and a stranger. Separate blocks were completed for each referential condition. On each trial, participants were presented with positive-negative personality word pairs and were

asked to select the word that represented the computer's attitude. No time limit was imposed. Participants were given feedback on their selection (presented for 2000ms). For each referential condition, participants learnt varying levels of positive 'like' rules (60-80% of positive words correct) and negative 'dislike' rules (20-40% of positive words correct). Referential condition and rule order were randomised. 24 trials were completed per block.

Learning was assessed through the number of errors made before reaching the criterion of eight consecutive rule-congruent responses. This was averaged across each level of the positive 'like' and negative 'dislike' rules. Bias scores were calculated by subtracting errors to criterion made when learning the dislike rule from the like rule. A positive value indicates a negative bias, as fewer errors are made learning the dislike rule compared to the like rule. Conversely, a negative value indicates a positive bias, as fewer errors are made learning the like rule compared to the dislike rule.

Emotional Categorisation and Recall

Positive and negative words were selected from a dataset of personality trait descriptors rated for likability (5). Two lists of 20 positive and 20 negative words were created matched according to word length. At the beginning of the task participants were asked to enter the first name of a familiar other to allow personalised task instructions. Participants then encoded personality traits to the self or other by categorising whether presented positive and negative words described themselves/the other ("In this task we will ask you to indicate whether each word describes [self/other]"). Separate blocks were completed for each referential condition, with order and list assignment randomised. Participants were instructed to press a key to indicate if the word described the person ('yes') or did not describe the person ('no'). 'j' and 'k' keys were used for input, with key assignment for each response randomised. Immediately following categorisation of personality traits participants were asked to recall as many of the presented characteristics as they could in two minutes, using the keyboard to enter their responses.

The total number of positive and negative words categorised as describing the self and other were recorded. The total number of positive and negative words correctly recalled were recorded according to referential condition.

Self-Esteem Go/No-Go Association

Inhibitory control when responding to affective words in relation to the self and others was measured using a go/no-go association task (6). This task was used to measure affective processing occurring in interaction with referential processing. Participants were instructed that they had to categorise words by pressing the spacebar. Participants categorised words relating to two referential conditions (self or others), and two emotional conditions (positive or negative characteristics). In each trial a word belonging to one of these categories was briefly displayed (600 ms) at the centre of the screen. Participants pressed the spacebar if the presented word belonged to a specified paired referential-emotion category (e.g. Self-Positive, Self-Negative, Other-Positive, Other-Negative). Four randomised blocks were completed relating to each referential-emotion combination, with 16 practice trials and 48 test trials per block. Response timeouts of 600 ms were applied.

Discriminative accuracy (d') was calculated through applying z-score transformations and subtracting hit z-scores from false alarm z-scores for each referential-emotion combination. Z-scores were adjusted by adding or subtracting 0.005 if the values were 0 or 1 to remove extreme values.

Associative Learning

Simple associative learning of abstract shape pairings with self, reward and emotion was measured using three tasks (7,8). This was used to measure how self-reference, emotional valence, and reward independently influence simple associative learning. Previous work in healthy controls has found that associative learning is prioritised for the self, positive stimuli, and high levels of reward (7,8). Nine practice trials and two blocks of 60 testing trials were completed per task.

In each task, participants were told that they had to match shapes with words or pictures. Shapes and stimuli varied according to each task; in the self-task, shapes were matched with the words 'self' 'friend' and 'stranger'; in the emotion task, shapes were matched with happy, neutral and sad cartoon faces; and in the reward task, shapes were matched with high (£9), medium (£3) and low (£1) monetary rewards. Shape-stimuli pairings were randomly assigned.

Participants were presented with a combination of the stimuli-shape pairings and pressed the 'n' or the 'm' key to indicate whether the presented pairings matched with the previously learnt associations. In each trial a fixation point was displayed for 2000 ms, followed by a

stimuli-shape pairing. Stimuli-shape pairings were presented for 100 ms for the self and reward associative tasks, and 150 ms for the emotion task due to the greater visual complexity of stimuli. Participants were asked to provide a response within 1100 ms and were then given feedback on their response for 500 ms. At the end of each block participants were informed of their accuracy (% correct).

For the reward task only, participants received a monetary reward based on the proportion of correct trials per category of reward stimuli.

Detailed Description of Statistical Models for Cognitive Tasks

Unless otherwise stated, analyses were pre-registered and confirmatory. For all models, subject was entered as a random effect to account for the repeated measures elements of the cognitive tasks. The citalopram group was used as the reference category in all analyses.

Prisoners' Dilemma

Proportion of cooperative behaviours was the outcome, drug group, social context and the interaction between drug group and social context were predictors.

Social Evaluation Learning

Bias score was the outcome, and drug group, referential condition and the interaction between drug group and referential condition were predictors. Exploratory analyses were conducted examining the effect of drug group (the predictor) on bias score (the outcome) separately for each referential condition.

To understand whether effects on bias scores were driven by learning within a particular rule (e.g. better learning of 'dislike' or worse learning of 'like'), we re-ran the models using errors to criterion as the outcome, and referential condition, rule, drug group and the interaction between these variables, as predictors. Exploratory analyses were conducted separately for each referential condition examining the effect of drug group, rule, and the interaction between drug group and rule on errors to criterion.

Referential Categorisation and Recall

The number of words categorised as descriptive ("yes") was the outcome, and drug group, referential condition, valence, and the interaction between these were predictors. This model was repeated with the total number of words correctly recalled as the outcome. As the citalopram group had a higher proportion of participants that did not speak English as a first language, we conducted a sensitivity analysis repeating this model with a binary variable

representing whether English was spoken as a first language (yes/no) included as an additional predictor variable.

Exploratory analyses were conducted separately for each referential condition, with total number of words correctly recalled as the outcome and valence, drug group and the interaction between drug group and valence as predictors.

Self-Esteem Go/No-Go Task

D' was the outcome, and drug group, referential condition, valence, and the interaction between these were predictors. To assess whether effects on d' may be driven by changes in hits or false alarms, these models were repeated separately with hits and false alarms as the outcome.

Associative Learning

Separate models were conducted for each task (self, emotion, reward) with accuracy (% correct) as the outcome. Stimuli, drug group, and the interaction between stimuli and drug group were predictors. These models were repeated with reaction times as the outcome.

Supplementary Tables

S1

Follow-up Contrasts exploring the effect of Drug Group at each Timepoint on VAS ratings of Happiness and Sadness

	Estimate	**SE**	**DF**	**t**	**p**
Happiness					
Baseline	-8.22	3.94	57.8	-2.09	0.041
Post-Drug	-0.95	3.94	57.8	-0.24	0.810
Post-Testing	-2.26	3.94	57.8	-0.57	0.568
Sadness					
Baseline	8.86	3.75	66	2.36	0.021
Post-Drug	3.58	3.75	66	0.95	0.343
Post-Testing	4.57	3.75	66	1.22	0.227

Note: Citalopram used as the reference category for drug group.

S2

Results from a mixed-effects regression analysis examining the effect of drug group, referential condition and rule on measures of learning in the Social Evaluation Learning Task

	β	**95% CI**	**p**
Bias Scores			
Intercept	-4.53	-7.20, -1.85	0.001
Drug group	1.95	-1.78, 5.69	0.308
Referential Condition			0.738
Self	Reference		
Friend	-1.18	-3.91, 1.56	0.403
Stranger	0.35	-2.39, 3.09	0.803
Drug group * Referential Condition			0.387
Self	Reference		
Friend	2.10	-1.72, 5.93	0.284
Stranger	-0.33	-4.15, 3.50	0.868
Errors to Criterion			
Intercept	4.13	2.37, 5.89	< 0.001
Drug group	1.71	-0.75, 4.17	0.175
Referential Condition			0.753
Self	Reference		
Friend	-0.58	-2.89, 1.74	0.627
Stranger	1.10	-1.21, 3.42	0.353
Rule	4.53	2.21, 6.84	< 0.001
Drug group * Referential Condition			0.585
Self	Reference		
Friend	-0.07	-3.30, 3.17	0.967
Stranger	-1.46	-4.70, 1.78	0.379
Drug group * Rule	-1.95	-5.19, 1.28	0.238

Referential Condition * Rule			0.845
Self	Reference		
Friend	1.18	-2.10, 4.45	0.483
Stranger	-0.35	-3.62, 2.93	0.834
Drug group*Referential Condition*Rule			0.821
Self	Reference		
Friend	-2.10	-6.68, 2.48	0.369
Stranger	0.33	-4.25, 4.91	0.889

Note: Citalopram used as the reference category for drug group.

S3

Results from a mixed-effects regression analysis examining the effect of drug group, referential condition and valence on total number of words categorised as descriptive and total number of words correctly recalled

	β	95% CI	p
Total Categorisations as Descriptive ("yes")			
Intercept	16.10	14.81, 17.40	<0.001
Drug Group	-0.20	-2.01, 1.61	0.833
Referential Condition	0.35	-1.45, 2.15	0.704
Valence	-11.70	-13.50, -9.90	<0.001
Drug Group * Referential Condition	0.60	-1.92, 3.12	0.640
Drug Group * Valence	0.89	-1.63, 3.41	0.490
Referential Condition * Valence	-1.25	-3.80, 1.30	0.339
Drug Group &* Referential Condition * Valence	0.01	-3.55, 3.58	0.995
Total Words Correctly Recalled			
Intercept	6.10	5.06, 7.14	<0.001
Drug Group	-0.96	-2.41, 0.50	0.200
Referential Condition	-0.05	-1.12, 1.02	0.927
Valence	-1.30	-2.37, -0.23	0.019
Drug Group * Referential Condition	-0.57	-2.07, 0.93	0.459
Drug Group * Valence	0.49	-1.01, 1.99	0.523
Referential Condition * Valence	-1.30	-2.82, 0.22	0.096
Drug Group &* Referential Condition * Valence	1.92	-0.20, 4.04	0.079

Note: Citalopram used as the reference category for drug group, self used as reference category for referential condition, positive ('likeable') used as reference category for valence.

S4

Results from a mixed-effects regression analysis examining the effect of drug group, referential condition and valence on discriminative accuracy (D'), hits (%), and false alarms (%) in a Go/No-Go Association Self-Esteem Task

	β	95% CI	p
D'			
Intercept	1.31	1.03, 1.60	< 0.001
Drug Group	0.20	-0.20, 0.59	0.339
Referential Condition	-0.47	-0.84, -0.11	0.012
Valence	-0.36	-0.73, 0.00	0.052
Drug Group * Referential Condition	-0.24	-0.74, 0.27	0.362
Drug Group * Valence	-0.41	-0.92, 0.09	0.112
Referential Condition * Valence	0.65	0.14, 1.16	0.015
Drug Group &* Referential Condition * Valence	0.45	-0.27, 1.16	0.223
Hits (%)			
Intercept	72.22	65.66, 78.79	< 0.001
Drug Group	8.77	-0.37, 17.90	0.064
Referential Condition	-10.00	-17.09, -2.91	0.007
Valence	-8.33	-15.42, -1.24	0.024
Drug Group * Referential Condition	-4.58	-14.45, 5.29	0.365
Drug Group * Valence	-7.29	-17.16, 2.58	0.151
Referential Condition * Valence	16.94	6.92, 26.97	0.001
Drug Group * Referential Condition * Valence	5.71	-8.25, 19.67	0.425
False Alarms (%)			

Intercept	28.61	21.69, 35.54	< 0.001
Drug Group	-0.49	-10.12, 9.15	0.922
Referential Condition	7.22	-0.31, 14.75	0.063
Valence	1.39	-6.14, 8.92	0.719
Drug Group * Referential Condition	1.11	-9.37, 11.59	0.836
Drug Group * Valence	9.03	-1.45, 19.51	0.095
Referential Condition * Valence	-6.94	-17.59, 3.70	0.204
Drug Group &* Referential Condition * Valence	-5.56	-20.37, 9.26	0.464

Note: Citalopram used as the reference category for drug group, self used as reference category for referential condition, positive used as reference category for valence.

S5

Results from a mixed-effects regression analysis examining the effect of drug group and stimuli on accuracy (%) and reaction times (ms) in associative learnings tasks of self, emotion and reward

	Accuracy (%)			**Reaction Times (ms)**		
	β	**95% CI**	**p**	**β**	**95% CI**	**p**
Self						
Intercept	89.96	85.03, 94.89	< 0.001	696.36	664.61, 728.10	< 0.001
Drug Group	1.30	-5.59, 8.18	0.713	-11.51	-55.87, 32.85	0.613
Stimuli			< 0.001			< 0.001
Self	Reference					
Friend	-5.59	-10.36, -0.82	0.024	28.12	10.71, 45.53	0.002
Stranger	-10.59	-15.36, -5.82	< 0.001	39.60	22.19, 57.01	< 0.001
Drug Group * Stimuli			0.503			0.812
Self	Reference					
Friend	-3.76	-10.43, 2.90	0.272	7.79	-16.54, 32.12	0.532
Stranger	-0.98	-7.65, 5.69	0.774	3.16	-21.17, 27.49	0.800
Emotion						
Intercept	80.81	73.81, 87.81	< 0.001	731.27	687.21, 775.34	< 0.001
Drug Group	0.08	-9.71, 9.86	0.988	-8.59	-70.16, 52.98	0.786
Stimuli			< 0.001			< 0.001
Happy	Reference					
Neutral	-12.61	-18.59, -6.64	< 0.001	48.11	25.63, 70.58	< 0.001
Sad	-12.45	-18.43, -6.47	< 0.001	45.79	23.32, 68.27	< 0.001

Drug Group * Stimuli			0.716			0.463
Happy	Reference					
Neutral	-2.71	-11.07, 5.64	0.526	-19.45	-50.85, 11.95	0.228
Sad	-3.14	-11.50, 5.21	0.463	-10.45	-41.85, 20.95	0.516
Reward						
Intercept	81.41	75.32, 87.51	<0.001	707.51	670.14, 744.88	< 0.001
Drug Group	-4.01	-12.53, 4.51	0.359	-23.29	-75.50, 28.93	0.386
Stimuli			0.080			0.008
High (£9)	Reference					
Medium (£3)	-5.15	-11.35, 1.06	0.108	25.44	1.30, 49.59	0.042
Low (£1)	-4.01	-10.22, 2.20	0.209	10.20	-13.95, 34.34	0.410
Drug Group * Stimuli			0.372			0.750
High (£9)	Reference					
Medium (£3)	-0.56	-9.23, 8.11	0.900	12.52	-21.21, 46.25	0.469
Low (£1)	4.98	-3.69, 13.66	0.264	8.31	-25.42, 42.04	0.630

Note: Citalopram used as the reference category for drug group.

Group assignment guesses and certainty ratings made by participants and researchers according to drug group

	Citalopram (N = 19)[a]	Placebo (N = 21)	p
Participant			
Group Guess, N (%)			$X^2(1) = 12.07$, $p < .001$
Citalopram	14 (74)	3 (14)	
Placebo	5 (26)	18 (86)	
Certainty, M (SD)			
Citalopram	53.74 (31.49)	24.38 (20.14)	$t(30.09)=3.47$, $p = 0.002$
Placebo	30.74 (29.63)	55.14 (22.99)	$t(33.89)=-2.89$, $p = 0.007$
Researcher			
Group Guess, N (%)			$X^2(1) = 0.06$, $p = 0.806$
Citalopram	8 (42)	7 (33)	
Placebo	11 (58)	14 (66)	
Certainty, M (SD)			
Citalopram	29.32 (27.62)	28.86 (19.35)	$t(31.88) = 0.06$, $p = 0.952$
Placebo	22.79 (23.65)	33.24 (22.73)	$t(37.25) = -1.42$, $p = 0.164$

[a] Data for one participant unavailable due to a technical error, total N = 20

S7

Results from a mixed-effects regression analysis examining the effect of drug group and timepoint on self-reported side effects

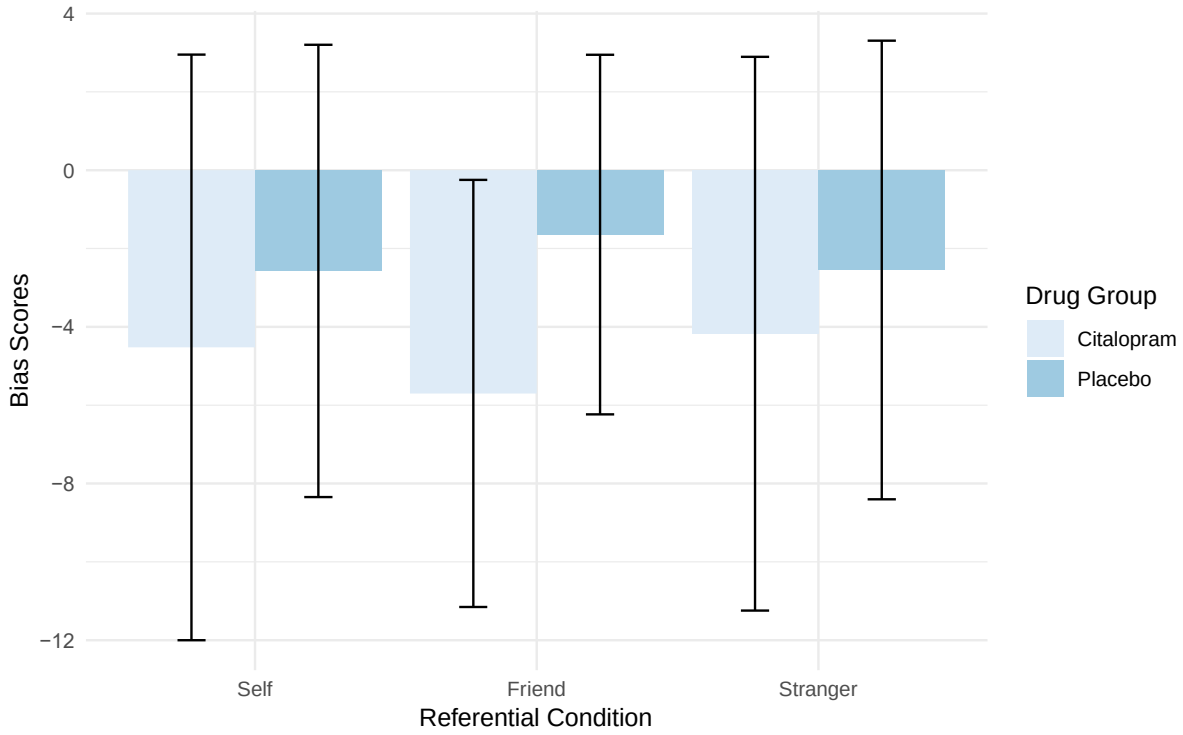
	β	95% CI	p
Nausea			
Intercept	1.00	0.80, 1.20	< 0.001
Drug Group	0.00	-0.27, 0.27	1.000
Timepoint			< 0.001
Baseline	Reference		
Post-Drug	0.68	0.45, 0.92	< 0.001
Post-Testing	0.47	0.24, 0.70	< 0.001
Drug Group * Timepoint			< 0.001
Baseline	Reference		
Post-Drug	-0.64	-0.95, 0.32	< 0.001
Post-Testing	-0.47	-0.79, -0.16	0.005
Dizziness			
Intercept	1.05	0.86, 1.24	< 0.001
Drug Group	-0.01	-0.27, 0.26	0.970
Timepoint			< 0.001
Baseline	Reference		
Post-Drug	0.53	0.31, 0.74	< 0.001
Post-Testing	0.32	0.10, 0.53	0.005
Drug Group * Timepoint			0.012
Baseline	Reference		
Post-Drug	-0.43	-0.73, -0.14	0.005
Post-Testing	-0.32	-0.61, - 0.02	0.039
Dry Mouth			
Intercept	1.32	1.10, 1.53	< 0.001
Drug Group	-0.17	-0.47, 0.13	0.262
Timepoint			0.410
Baseline	Reference		

Post-Drug	-0.11	-0.34, 0.13	0.377
Post-Testing	-0.21	-0.44, 0.02	0.080
Drug Group * Timepoint			0.522
Baseline	Reference		
Post-Drug	0.15	-0.17, 0.47	0.353
Post-Testing	0.16	-0.16, 0.48	0.323
Headache			
Intercept	1.00	0.84, 1.16	< 0.001
Drug Group	0.05	-0.18, 0.27	0.679
Timepoint			0.078
Baseline	Reference		
Post-Drug	0.26	0.06, 0.47	0.013
Post-Testing	0.21	0.01, 0.41	0.046
Drug Group * Timepoint			0.474
Baseline	Reference		
Post-Drug	-0.17	-0.45, 0.11	0.246
Post-Testing	-0.12	-0.40, 0.17	0.424
Alertness			
Intercept	2.26	1.85, 2.68	< 0.001
Drug Group	0.12	-0.45, 0.69	0.687
Timepoint			0.467
Baseline	Reference		
Post-Drug	-0.21	-0.58, 0.16	0.266
Post-Testing	0.00	-0.37, 0.37	1.000
Drug Group * Timepoint			0.928
Baseline	Reference		
Post-Drug	-0.03	-0.54, 0.48	0.916
Post-Testing	-0.10	-0.60, 0.41	0.715
Agitation			
Intercept	1.00	0.84, 1.16	< 0.001
Drug Group	0.00	-0.23, 0.23	1.000
Timepoint			0.011

Baseline	Reference		
Post-Drug	0.21	0.01, 0.41	0.040
Post-Testing	0.11	-0.09, 0.30	0.300
Drug Group * Timepoint			0.100
Baseline	Reference		
Post-Drug	-0.12	-0.39, 0.16	0.411
Post-Testing	0.18	-0.09, 0.45	0.199

Note: Citalopram used as the reference category for drug group.

Supplementary Figures



F1

Mean bias scores according to referential condition and drug group in the Social Evaluation Learning task. Greater bias scores indicate relatively better learning of the negative versus positive rule. Error bars represent standard deviations.

References

1. Wood RM, Rilling JK, Sanfey AG, Bhagwagar Z, Rogers RD. Effects of tryptophan depletion on the performance of an iterated Prisoner's Dilemma game in healthy adults. Neuropsychopharmacology. 2006;31(5):1075–84.

2. Rilling JK, Gutman DA, Zeh TR, Pagnoni G, Berns GS, Kilts CD. A neural basis for social cooperation. Neuron. 2002;35(2):395–405.

3. Button KS, Browning M, Munafò MR, Lewis G. Social inference and social anxiety: Evidence of a fear-congruent self-referential learning bias. J Behav Ther Exp Psychiatry. 2012;43(4):1082–7.

4. Button KS, Kounali D, Stapinski L, Rapee RM, Lewis G, Munafò MR. Fear of negative evaluation biases social evaluation inference: Evidence from a probabilistic learning task. PLoS One. 2015;10(4):e0119456.

5. Anderson NH. Likeableness Ratings of 555 Personality-Trait Words. J Pers Soc Psychol. 1968;9(3):272.

6. Gregg AP, Sedikides C. Narcissistic fragility: Rethinking its links to explicit and implicit self-esteem. Self Identity. 2010;9(2):142–61.

7. Sui J, Ohrling E, Humphreys GW. Negative mood disrupts self- and reward-biases in perceptual matching. Q J Exp Psychol. 2016;69(7):1438–48.

8. Stolte M, Humphreys G, Yankouskaya A, Sui J. Dissociating biases towards the self and positive emotion. Q J Exp Psychol. 2017;70(6):1011–22.

Chapter 5 Supplementary Materials

Hobbs, C., Beck, M., Denham, F., Pettit, L., Faraway, J., Munafò, M.R., Sui, J., Kessler, D., & Button, K. S. (2022). The relationship between change in social evaluation learning and mood in early antidepressant treatment: a prospective cohort study in primary care. *Manuscript in Preparation.*

Additional measures collected prior to the COVID-19 pandemic

Additional questionnaire measures and cognitive tasks were completed by participants prior to the COVID-19 pandemic. However, to allow for remote data collection and to reduce potential fatigue effects, we removed the measures outlined below.

Associative Learning

We used three tasks to examine different aspects of associative learning (self, reward and valence) occurring independently (Hobbs et al., 2021; Stolte et al., 2016; Sui and Humphreys, 2015). Participants completed the tasks sequentially, however the order was counterbalanced across time-points and participants. At the beginning of each task, participants were instructed to remember specified stimuli-abstract shape pairings. Stimuli consisted of the labels 'self', 'friend' and 'stranger' for the self task, the labels '£9', '£3' and '£1' for the reward task, and a happy, neutral and sad cartoon face for the valence task. In each trial a fixation point was initially presented for 200 ms, followed by a stimuli-shape pairing for 100-150 ms. Participants used the keyboard to indicate whether the presented pairings matched with the pairings specified at the beginning of the task. Two blocks of 60 trials were completed per task. Accuracy and reaction times were recorded.

Facial Emotion Recognition

Participants completed a facial emotion recognition task at each time-point, measuring accuracy in recognising six emotions (happiness, sadness, fear, anger, disgust, surprise) (Griffiths et al., 2015). In each trial a fixation cross appeared in the centre of the screen for 1500-2500 ms. This was then replaced by a face stimulus for 250ms, followed by a mask of visual noise for 250ms. Text labels for the six possible emotions then appeared on the screen in a circle formation. Participants were asked to select the emotion that they thought was displayed. This was a forced choice response although no time limit was imposed. Participants completed one block of 96 trials. Accuracy and reaction times were recorded.

Word Categorisation and Recall

Participants categorised whether 40 presented positive and negative traits described themselves or a familiar other (Hobbs et al., 2020). Traits were presented for 500 ms, and participants used the keyboard to indicate their response (yes/no). Participants were then given two minutes to recall these traits, using the keyboard to enter their responses. Separate blocks were completed for each referential condition in a randomised order. Recalled traits were recorded.

Self-Report Measures of Mood

At each timepoint participants completed the Brief Fear of Negative Evaluation Scale (Leary, 1983) as a measure of social anxiety, the Rosenberg Self-Esteem Scale (Rosenberg, 1965) as a measure of self-esteem, and the 24 item version of the Dysfunctional Attitudes Scale (Power et al., 1994) as a measure of maladaptive self-schema. To measure change in state affect during testing participants completed the positive and negative affect scale (PANAS) (Watson et al., 1988) before and after completion of the cognitive tasks at each timepoint.

Supplementary Table S1

Treatment Characteristics by Timepoint

	Baseline	2-weeks	6-weeks	8-weeks	6-months
N	29	23	21	22	11
Current Antidepressant, N (%)	28 (96.6)	23 (100.0)	19 (95.0)	20 (95.2)	8 (72.7)
Antidepressant Medication, N (%)					
Sertraline	15 (53.6)	11 (47.8)	9 (47.4)	11 (55.0)	7 (87.5)
Citalopram	9 (32.1)	9 (39.1)	8 (42.1)	8 (40.0)	1 (12.5)
Fluoxetine	2 (7.1)	2 (8.7)	2 (10.5)	0 (0)	0 (0)
Mirtazapine	2 (7.1)	0 (0.0)	0 (0.0)	0 (0)	0 (0)
Mirtazapine and sertraline	0 (0)	1 (4.4)	0 (0)	1 (5.0)	0 (0)
Change in Antidepressant Treatment [a], N (%)					
Discontinuation	-	1 (4.4)	1 (5.0)	0 (0)	3 (27.3)
Medication [b]	-	2 (9.1)	0 (0)	0 (0)	0 (0)
Dose [c]	-	1 (4.8)	2 (11.1)	7 (36.8)	3 (37.5)
Adherence [d]					
Taken tablets everyday	27 (96.4)	18 (78.3)	12 (63.2)	14 (70.0)	6 (75.0)
Taken nearly all of tablets	1 (3.6)	5 (21.7)	7 (36.8)	6 (30.0)	2 (25.0)
Side Effects, N (%)	14 (50.0)	11 (47.8)	6 (31.6)	4 (20.0)	2 (25.0)
Current Psychological Therapy, N (%)	2 (6.9)	2 (8.7)	2 (9.5)	4 (19.1)	0 (0)
Other Medication, N (%)	7 (24.1)	7 (30.4)	5 (23.8)	5 (23.8)	2 (18.2)
Data Collection, N (%)					
Face to face	19 (65.5)	15 (65.2)	13 (61.9)	15 (71.4)	4 (36.4)
Remote	10 (34.5)	8 (34.8)	8 (38.1)	6 (28.6)	7 (63.6)

[a] Change from previous timepoint

[b] One participant had treatment augmented with an additional antidepressant (mirtazapine to mirtazapine & sertraline), one participant switched from sertraline to fluoxetine

[c] Aside from one participant at 8-week follow-up who had a reduction in dose, all changes in doses of medication were an increase.

[d] Possible options to the statement 'As regards your antidepressants, I have...' were (1) 'Taken my tablets everyday' (2) 'Taken nearly all of my tablets', (3) 'Taken more than half of my tablets', (4) 'Taken less than half of my tablets', (5) 'Taken hardly any of my tablets', (6) 'Not taken any of my tablets'. No participants selected options 3-6.

Note: Percentages reflect proportions of participants within timepoints. For antidepressant medication, change in antidepressant treatment, adherence, and reported side effects only participants that reported currently taking an antidepressant responded, percentages therefore reflect the proportions of patients currently receiving an antidepressant within timepoints.

Supplementary Table S2

Mixed-effect linear regression models examining the association between change in biased learning of social evaluations (predictors) and change in GAD-7 scores (outcome) adjusting for change in PHQ-9 and BDI-II scores

	b	***b* 95% CI**	**β**	**β 95% CI**	**p**
Adjusting for change in PHQ-9 scores					
Intercept	0.49	-1.94, 2.91	-0.04	-0.31, 0.23	0.697
Bias Scores Change					
Self	0.13	0.04, 0.21	0.25	0.08, 0.41	0.004
Friend	0.17	0.08, 0.27	0.28	0.13, 0.44	0.001
Stranger	0.02	-0.07, 0.11	0.04	-0.12, 0.21	0.600
Baseline GAD-7	-0.15	-0.32, 0.02	-0.14	-0.30, 0.02	0.095
Session					0.630
Baseline to 2-weeks	Reference				
2 to 6-weeks	0.65	-1.15, 2.44	0.14	-0.26, 0.54	0.485
6 to 8-weeks	-0.07	-1.86, 1.72	-0.02	-0.41, 0.38	0.936
PHQ-9 Change	0.58	0.40, 0.75	0.57	0.39, 0.76	< .001
Adjusting for change in BDI-II scores					
Intercept	0.37	-2.13, 2.87	-0.17	-0.44, 0.09	0.773
Bias Scores Change					
Self	0.12	0.03, 0.21	0.23	0.06, 0.41	0.011
Friend	0.19	0.09, 0.30	0.30	0.14, 0.46	0.001
Stranger	0.08	-0.03, 0.19	0.13	-0.05, 0.31	0.163
Baseline GAD-7	-0.20	-0.38, -0.03	-0.18	-0.34, -0.02	0.029
Session					0.003
Baseline to 2-weeks	Reference				
2 to 6-weeks	2.62	0.88, 4.35	0.56	0.18, 0.94	0.005
6 to 8-weeks	-0.02	-1.85, 1.81	0.00	-0.40, 0.40	0.984
BDI-II Change	0.27	0.18, 0.35	0.52	0.35, 0.69	< .001

b = unstandardised regression coefficients, β = standardised regression coefficient.

References

Griffiths S, Penton-Voak IS, Jarrold C, et al. (2015) No own-age advantage in children's recognition of emotion on prototypical faces of different ages. *PLoS ONE* 10(5): e0125256. DOI: 10.1371/journal.pone.0125256.

Hobbs C, Murphy SE, Wright L, et al. (2020) Effect of acute citalopram on self-referential emotional processing and social cognition in healthy volunteers. *BJPsych Open* 6(6). DOI: 10.1192/bjo.2020.107.

Hobbs C, Sui J, Kessler D, et al. (2021) Self-processing in relation to emotion and reward processing in depression. *Psychological Medicine*: 1–13. DOI: 10.1017/s0033291721003597.

Leary MR (1983) A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin* 9(3). Sage PublicationsSage CA: Thousand Oaks, CA: 371–375. DOI: 10.1177/0146167283093007.

Power MJ, Katz R, McGuffin P, et al. (1994) The Dysfunctional Attitude Scale (DAS). A Comparison of Forms A and B and Proposals for a New Subscaled Version. *Journal of Research in Personality* 28(3). Academic Press: 263–276. DOI: 10.1006/jrpe.1994.1019.

Rosenberg M (1965) Rosenberg self esteem scle. *Personality and Individual Differences*. DOI: 10.1007/s12671-015-0407-6.

Stolte M, Humphreys G, Yankouskaya A, et al. (2016) Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology* 70(6): 1011–1022. DOI: 10.1080/17470218.2015.1101477.

Sui J and Humphreys GW (2015) The interaction between self-bias and reward: Evidence for common and distinct processes. *Quarterly Journal of Experimental Psychology* 68(10): 1952–1964. DOI: 10.1080/17470218.2015.1023207.

Watson D, Clark LA and Tellegen A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54(6): 1063. DOI: 10.1037//0022-3514.54.6.1063.

Chapter 6 Supplementary Materials

Supplementary materials for Hobbs, C., Vozarova, P., Sabharwal, A., Shah, P., Button, K.S. (2022). Is depression associated with reduced optimistic belief updating? *Royal Society Open Science, 9.*

Supplementary Table 1

List of Negative Life Events used as Stimuli

Event	Probability of happening	Reference
Abnormal heart rhythm	25 [a]	Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D JM, Beiser A, Wolf PA, Benjamin EJ. Lifetime risk for de Framingham Heart Study. Circulation. 2004 Aug 31;110(9
Age-Related Blindness	12	Evans JR, Fletcher AE, Wormald RP, Ng ES, Stirling S, S M, Jones D, Tulloch A. Prevalence of visual impairment i Britain: results from the MRC trial of assessment and man community. British Journal of Ophthalmology. 2002 Jul 1
Alzheimer's disease	15 [a]	Seshadri S, Wolf PA. Lifetime risk of stroke and dementia the Framingham Study. The Lancet Neurology. 2007 Dec
Arteries hardening (narrowing of blood vessels)	67 [a]	Leening M, Ferket B, Steyerberg E, Kavousi M, Deckers J lifetime risk and first manifestation of cardiovascular disea cohort study. BMJ. 2014;349:5992-5992
Artificial joint	19 [a]	Culliford DJ, Maskell J, Kiran A, Judge A, Javaid MK, Co of total hip and knee arthroplasty: results from the UK gen Osteoarthritis and cartilage. 2012 Jun 1;20(6):519-24.
Asthma	34	To T, Wang C, Guan J, McLimont S, Gershon AS. What i diagnosed asthma in Ontario, Canada?. American journal medicine. 2010 Feb 15;181(4):337-43.
Back Pain	70 [b]	World Health Organization. Priority diseases and reasons
Being cheated on by your husband/wife	20 [a]	Blow AJ, Hartnett K. Infidelity in committed relationships Marital and Family Therapy. 2005; 31(2):217-33.
Being convicted of a crime	11	Skardhamar T. Lifetime conviction risk—a synthetic coho Studies in Criminology and Crime Prevention. 2014 Jan 2
Being fired	45	Association of Accounting Technicians. Work – In Numbe Accounting Technicians. [updated 2015 November 4; cite https://www.aat.org.uk/news/article/work-numbers
Bone Fracture	45	Scholes S, Panesar S, Shelton NJ, Francis RM, Mirza S, M lifetime fracture prevalence in England: a population study Age and Ageing Journal. 2013; 43(2):234-40.

Cancer (of digestive system/lung/prostate/breast/skin)	50	Cancer Research UK Statistical Information Team. Lifetin Research UK; 2018.
Car Stolen	19	Koppel H. Lifetime likelihood of victimization. C: US De Statistics; 1987 Mar.
Card Fraud	15	Harrell E, Langton L. Victims of identity theft, 2012. US I Justice Programs, Bureau of Justice Statistics; 2013 Dec.
Death before age 80	43	Wright R. National life tables: 2012 to 2014. United King 2015.
Dementia	18 [a]	Seshadri S, Wolf PA. Lifetime risk of stroke and dementia the Framingham Study. The Lancet Neurology. 2007 Dec
Diabetes (type 2)	38	Magliano DJ, Shaw JE, Shortreed SM, Nusselder WJ, Lie Lifetime risk and projected population prevalence of diabe 1;51(12):2179-86
Divorce	28 [c]	Heines N. Divorces in England and Wales: 2017. United K Statistics; 2018.
Drug Abuse	35	Flatley J. Drug Misuse: Findings from the 2017/18 Crime Statistical Bulletin 14/18. London: Home Office; 2018 p. 2
Eye cataract (clouding of the lens of the eye)	30	Reidy A, Minassian DC, Vafidis G, Joseph J, Farrow S, W of serious eye disease and visual impairment in a north Lo cross sectional study. Bmj. 1998 May 30;316(7145):1643-
Gallbladder stones	60	Behari A, Kapoor VK. Asymptomatic Gallstones (AsGS)- Surgery. 2012 Feb 1;74(1):4-12.
Gluten Intolerance	13	Aziz I, Lewis NR, Hadjivassiliou M, Winfield SN, Rugg N assessing the population prevalence of self-reported gluten characteristics to secondary care. European Journal of Gas 26(1):33-9.
Having a stroke	18 [a]	Seshadri S, Beiser A, Kelly-Hayes M, Kase CS, Au R, Ka of stroke: estimates from the Framingham Study. Stroke. 2
Heart Failure	20	Bui AL, Horwich TB, Fonarow GC. Epidemiology and ris Reviews Cardiology. 2011 Jan;8(1):30.
Hernia	15 [a]	Primatesta P, Goldacre MJ. Inguinal hernia repair: inciden readmission and mortality. International Journal of Epiden
Infertility	11 [a]	Datta J, Palmer MJ, Tanton C, Gibson LJ, Jones KG, Mac Field N, Mercer CH, Johnson AM. Prevalence of infertilit women and men. Human Reproduction. 2016 Aug 19;31(9

Irritable bowel syndrome (disorder of the gut)	11	Canavan C, West J, Card T. The epidemiology of irritable epidemiology. 2014;6:71.
Kidney Stones	13 [a]	National Institute for Health and Care Excellence. Informa stones: Assessment and management. London: National In Excellence; 2017.
Knee Osteoarthritis	45	Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor WD, Luta G, Jordan JM. Lifetime risk of symptomatic kn Research: Official Journal of the American College of Rhe 13.
Migraine	31 [a]	Stewart WF, Wood C, Reed ML, Roy J, Lipton RB. Cumu women and men. Cephalalgia. 2008 Nov;28(11):1170-8.
Obesity	29 [a]	NHS Digital. Statistics on Obesity, Physical Activity and Kingdom: National Health Services; 2019.
Osteoporosis (reduced bone density)	26 [a]	International Osteoporosis Foundation. The Silent burden and healthcare system. Broken Bones, Broken Lives: A Re Crisis in the United Kingdom. Switzerland: International O
Severe teeth problems when old	40	Fuller E, Steele J, Watt R, Nuttall N. Oral health and funct Health Survey 2009. London: Health and Social Care Info
Skin burn	66	Alanazi A, Alanazi A, Alanazi M, Alenezi N, Qaisy F, Asiri A, Injuries Prevalence, Causes, Complications and Improvem Emergency Medicine and Critical Care. 2019; 3(6): 383-39
Ulcers	16 [b]	Shiotani A, Graham DY. Pathogenesis and therapy of gast Medical Clinics. 2002; 86(6):1447-66.
Victim of mugging	30	Koppel H. Lifetime likelihood of victimization. C: US De Statistics; 1987 Mar.
Victim of violence at home	22	Boyle A, Todd C. Incidence and prevalence of domestic v department. Emergency Medicine Journal. 2003 Sep 1;20(
Victim of violence by a stranger	37	(1) Koppel H. Lifetime likelihood of victimization. C: US Justice Statistics; 1987 Mar. (2) Morgan R E, Oudekerk B Bureau of Justice Statistics, U.S. Department of Justice. 2(
Victim of violence by an acquaintance	46	(1) Koppel H. Lifetime likelihood of victimization. C: US Justice Statistics; 1987 Mar. (2) Morgan R E, Oudekerk B Bureau of Justice Statistics, U.S. Department of Justice. 2(
Witness a traumatizing accident	70	Benjet C, Bromet E, Karam E, Kessler R, McLaughlin K, traumatic event exposure worldwide: results from the Wor Psychological Medicine. 2015; 46(2):327-343.

Training Events		
Death before age 70	22	Patel V. Deaths registered in England and Wales: 2017. U Statistics; 2018.
Herpes [d]	46	Schiffer JT, Corey L. New concepts in understanding disease reports. 2009 Nov 1;11(6):457-64.

[a] Average across genders calculated

[b] Upper estimate used

[c] Probability adjusted to reflect the proportion of people that are married or in a civil partnership (66%)

[d] Average across genders and ethnicities calculated

Supplementary Table 2

List of Negative Life Events not used as Stimuli

Event	Probability of happening	Reference
Autoimmune Disease	5.29	Eaton WW, Rose NR, Kalaydjian A, Pedersen MG, Morte autoimmune diseases in Denmark. Journal of autoimmunit
Appendicitis	8.6% males, 6.7% females	Addiss DG, Shaffer N, Fowler BS, Tauxe RV. The epiden appendectomy in the United States. American journal of e 25.
Bicycle Theft	Lifetime probability not available	
Blood clot in vein	8.1	Bell EJ, Lutsey PL, Basu S, Cushman M, Heckbert SR, Ll risk of venous thromboembolism in two cohort studies. Th 2016 Mar 1;129(3):339-e19.
Chronic high blood pressure	90	Vasan RS, Beiser A, Seshadri S, Larson MG, Kannel WB, lifetime risk for developing hypertension in middle-aged w Heart Study. Jama. 2002 Feb 27;287(8):1003-10.
Chronic ringing sound in ear (tinnitus)	Lifetime probability not available	
Death before age 60	7.45 [a]	Wright R. National life tables: 2012 to 2014. United Kingd 2015.
Death by infection	7.7	McPherson D, Griffiths C, Williams M, Baker A, Klodaw Sepsis-associated mortality in England: an analysis of mul 2010. BMJ open. 2013 Aug 1;3(8):e002586.
Disease of the spinal cord	Lifetime probability not available	
Domestic Burglary	72	Koppel H. Lifetime likelihood of victimization. C: US Dep Statistics; 1987 Mar.
Dying before age 90	78.4 [a]	Patel V. Deaths registered in England and Wales: 2017. U Statistics; 2018.
Epilepsy	5 [b]	Neligan A, Sander JW. The incidence and prevalence of ep Neurology. 2009.
Fraud when buying something on the internet	Lifetime probability not available	
Genital Warts	4.2 [a]	Sonnenberg P, Tanton C, Mesher D, King E, Beddows S, J Johnson AM. Epidemiology of genital warts in the British vaccination programmes. Sexually transmitted infections.

Glaucoma	2.4	Wolfs RC, Klaver CC, Ramrattan RS, van Duijn CM, Hof primary open-angle glaucoma: population-based familial a ophthalmology. 1998 Dec 1;116(12):1640-5.
Having fleas/lice	Lifetime probability not available	
Hepatitis A or B	(1) Hepatitis A: Lifetime probability in the UK not available but only 452 laboratory reports in 2018. Lifetime probability is therefore expected to be low. (2) Hepatitis B: 0.4%	(1) Public Health England. Laboratory reports of hepatitis 2018. London: Public Health England; 2019. (2) Ramsay M Control of hepatitis B in the United Kingdom. Vaccine. 19
House vandalized	Lifetime probability not available	
Household accident	Lifetime probability not available	
Limb Amputation	2.66	Bhuvaneswar CG, Epstein LA, Stern TA. Reactions to am Primary care companion to the Journal of clinical psychiat
Liver disease	Lifetime probability not available	
More than £30000 of debts	Lifetime probability not available	
Parkinson's disease	4.6 [a]	Licher S, Darweesh SK, Wolters FJ, Fani L, Heshmatollah J, Leening MJ, Ikram MK, Ikram MA. Lifetime risk of co elderly population. J Neurol Neurosurg Psychiatry. 2019 F
Restless legs syndrome	Lifetime probability not available	
Serious hearing problems	Lifetime probability not available	
Severe injury due to accident (traffic or house)	Lifetime probability not available	
Sport-related accident	Lifetime probability not available	
Theft from person	99	Koppel H. Lifetime likelihood of victimization. C: US Dep Statistics; 1987 Mar.
Theft from vehicle	Lifetime probability not available	
Victim of violence with need to go to A&E	Lifetime probability not available	

[a] Average across genders calculated

[b] Upper estimate used

Supplementary Table 3

List of Positive Life Events used as Stimuli

Event	Probability of happening	References
Fully paying off your mortgage before turning 55	59	Cherowbrier J. Share of adults with a mortgage in the United Kingdom (UK) 2017, by age. Understanding the financial lives of UK adults. London: Financial Conduct Authority; 2018 p. 116
Celebrate your 20th wedding anniversary	40 [a]	The National Archives. What percentage of marriages end in divorce? Part of Divorces in England and Wales. United Kingdom: Office for National Statistics; 2013.
Celebrate your 70th birthday	14	Clarke D. Mid-year population estimate of the United Kingdom in 2018, by age group. United Kingdom: Office for National Statistics; 2019.
Celebrate your diamond (60 years) wedding anniversary	11 [a]	The National Archives. What percentage of marriages end in divorce? Part of Divorces in England and Wales. United Kingdom: Office for National Statistics; 2013.
Create a start-up (a new business) that is successful for 5 or more years	10 [b]	Shaw B. The South West continues to have the highest five-year survival rate. Business demography, UK: 2017. United Kingdom: Office for National Statistics; 2018. Ipsos. Entrepreneurialism. The Emergence of Social Entrepreneurialism to Compete with Business Entrepreneurialism. Ipsos; 2018.
Earn more than £13 an hour	50	European Statistical Office. Low wage earners in the EU. Earning Statistics. Luxembourg: European Statistical Office; 2016.
Earn more than £30,000 or more a year (before tax)	17	Office for National Statistics. Distribution of total income before and after tax by gender. United Kingdom: Office for National Statistics; 2019.
Fall in love with someone you met on a dating site	28	The Statistics Portal Research Department. Experiences and attitudes regarding dating websites or apps according to users in the United Kingdom (UK) in June 2017. Online Dating Survey. United Kingdom: The Statistics Portal; 2017.

Get engaged on Valentine's Day	16 [a]	The Statistics Portal Research Department. When is the most popular time to propose? United Kingdom: The Statistics Portal; 2016.
Getting married in the summer	28 [a]	The Statistics Portal Research Department. Most popular months for weddings in the United Kingdom (UK) in 2018. Marriage and weddings in the United Kingdom. United Kingdom: The Statistics Portal; 2018 p. 28
Getting married or having a civil partner	66	Knipe E. Population estimates by marital status and living arrangements, England and Wales: 2002 to 2014. United Kingdom: Office for National Statistics; 2015.
Go abroad on holiday for more than 2 weeks	13	Office for National Statistics. Number of holiday visits abroad by United Kingdom (UK) residents in 2018, by length of stay. Travel Trends 2018. United Kingdom: Office for National Statistics; 2019.
Go on an all-inclusive holiday	15	Association of British Travel Agents. Leading types of holiday in the United Kingdom (UK) in 2018. Holiday Habits Report 2018. United Kingdom: Association of British Travel Agents; 2018 p. 4
Go on holiday abroad for Christmas	23	The Statistics Portal Research Department. Have you ever gone on holiday abroad over Christmas? United Kingdom: The Statistics Portal; 2016.
Go on holiday to USA/Canada	12	Luty J. World regions travellers intend to visit on their main holiday (4 nights or more) from the United Kingdom (UK) in 2018. United Kingdom: The Statistics Portal; 2018.
Go skiing	10	Vanat L. Number of people who ski in Europe as of 2018, by country (in 1,000). 2019 International Report on Snow & Mountain Tourism. United Kingdom: The Statistics Portal; 2018.
Going on a spontaneous holiday	17	The Statistics Portal Research Department. How far in advance will or have you book(ed) your main holiday? Holiday Travel in the United Kingdom. United Kingdom: The Statistics Portal; 2018 p. 23
Have a higher income or the same income in retirement compared to when you were working	15	The Statistics Portal Research Department. Thinking about your income in retirement, how does it compare with your income immediately before you retired? The Future of

		Retirement - Life after work. United Kingdom: The Statistics Portal; 2013.
Have a job you enjoy very much	34	Clarke D. How would you judge your personal job situation? Standard Eurobarometer 89, Annex. Brussels: European Commission; 2018.
Have over £50,000 in savings	12	The Statistics Portal Research Department. Share of savings among adults in the United Kingdom (UK) as of 2014. United Kingdom: The Statistics Portal; 2014.
Having a honeymoon outside of Europe	41 [a]	The Statistics Portal Research Department. Most popular destinations for honeymoons in the United Kingdom (UK) as of 2018. Marriage and weddings in the United Kingdom. United Kingdom: The Statistics Portal; 2018 p. 38
Hold an investment	28	Investment Statistics. How many Brits are investing in stocks and shares? London: Investment Statistics; 2018.
Live abroad	23	Clarke D. Share of people with international experience in Great Britain in 2018. Global Inclusiveness Survey. London: Ipsos Public Affairs; 2018 p. 51
Living in a house with 8 or more rooms	13	Office for National Statistics. Census - Number of rooms by number of bedrooms - Merged local authorities. United Kingdom: Office for National Statistics; 2018.
Own 2 or more cars	22	The Statistics Portal Research Department. Number of cars owned by households in the United Kingdom (UK) from 2015 to 2017. United Kingdom: The Statistics Portal; 2018.
Own a second home	10	Resolution Foundation. 21st Century Britain has seen a 30 per cent increase in second home ownership. London: Resolution Foundation; 2017.
Owning a house	53	Resolution Foundation. Housing tenure by age. Home ownership in the UK. London: Resolution Foundation; 2019.
Owning a house before turning 35	11 [c]	Ministry of Housing, Communities and Local Government. Home ownership. United Kingdom: Government Digital Service; 2018.
Owning a pet	45	Sabanoglu T. Leading pets, ranked by household ownership in the United Kingdom (UK) in 2017/18. United Kingdom: The Statistics Portal; 2018.
Playing a musical instrument	12	Luty J. Share of adults who played a musical instrument in the last year in England from 2005/06 to 2016/17. United

		Kingdom: Department for Digital, Culture, Media and Sport; 2017.
Repaying your student loan in full	30	Bolton P. Student loan statistics. Commons Briefing Papers. London: House of Commons; 2019.
Retire early (before 65)	39	Knight D. Pension Trends – Chapter 4: The Labour Market and Retirement. United Kingdom: Office for National Statistics; 2013.
Speaking two or more foreign languages fluently	15	European Statistical Office. 65% know at least one foreign language in the EU. Luxembourg: European Statistical Office; 2018.
Spend £20,000 or more on your wedding	20 [a]	Sabanoglu T. Most common budget ranges for weddings in the United Kingdom (UK) in 2018. United Kingdom: The UK Wedding Report; 2018 p. 12
Staying in a 4 or 5 star hotel on holiday	27	The Statistics Portal Research Department. What type of accommodation are you most likely to stay in on your main holiday? Holiday Travel in the United Kingdom. United Kingdom: The Statistics Portal; 2018 p. 24
Sticking to your wedding budget	17 [a]	Sabanoglu T. Share of individuals who agree with the following statements about their wedding budget planning in the United Kingdom (UK) in 2018. United Kingdom: The UK Wedding Report; 2018 p. 14
Taking part in a running race	10	Sports Marketing Surveys INC. UK´s running population reaches remarkable 10.5m says sports marketing surveys. United Kingdom: Sports Marketing Surveys INC; 2014.
Visit Asia	11	Luty J. World regions travellers intend to visit on their main holiday (4 nights or more) from the United Kingdom (UK) in 2018. United Kingdom: The Statistics Portal.; 2018.
Visit South America	14	Luty J. World regions travellers intend to visit on their main holiday (4 nights or more) from the United Kingdom (UK) in 2018. United Kingdom: The Statistics Portal; 2018.
Visit Spain	28	Forte F. Key inbound travel markets for Spain from 2000 to 2018, by number of arrivals. Survey of Tourist Movements at Borders Frontur. Spain: Subdirección General de Conocimiento y Estudios Turísticos & Ine; 2018.
Training events		

Living in a house with 4 or more bedrooms	19	Office for National Statistics. Census - Number of rooms by number of bedrooms - Merged local authorities. United Kingdom: Office for National Statistics; 2018.
Going on a holiday abroad in the winter	23	The Statistics Portal Research Department. Number of monthly overseas holiday visits from the United Kingdom (UK) from January 2017 to December 2018. Holiday Travel in the United Kingdom. United Kingdom: The Statistics Portal; 2018 p. 17

[a] Probability adjusted to reflect the proportion of people that are married or in a civil partnership (66%)

[b] Probability adjusted to reflect the proportion of people that have started a business (22%)

[c] Probability adjusted to reflect the proportion of people who own a house (53%)

Supplementary Table 4

Belief Updating Task Procedure Block Completion Order

Block Number	**Trial Type**	**Valence**[a]	**Event List**[a]	**Framing**[a]
1	**First Estimate & Actual Probability**	Positive or Negative	One or Two (20 events)	Happening or Not Happening
2	**Second Estimate**			
3	**First Estimate & Actual Probability**		One or Two (20 events)	Happening or Not Happening
4	**Second Estimate**			
5	**Memory Estimates**		One or Two (20 events; as in Block 1/2)	Happening or Not Happening (as in Block 1/2)
6	**Memory Estimates**		One or Two (20 events; as in Block 3/4)	Happening or Not Happening (as in Block 3/4)
7	**Ratings**		One and Two (40 events)	N/A
8	**First Estimate & Actual Probability**	Positive or Negative	One or Two (20 events)	Happening or Not Happening
9	**Second Estimate**			
10	**First Estimate & Actual Probability**		One or Two (20 events)	Happening or Not Happening
11	**Second Estimate**			
12	**Memory Estimates**		One or Two (20 events; as in Block 8/9)	Happening or Not Happening (as in Block 8/9)
13	**Memory Estimates**		One or Two (20 events; as in Block 10/11)	Happening or Not Happening (as in Block 10/11)
14	**Ratings**		One and Two (40 events)	N/A

[a] Counterbalanced

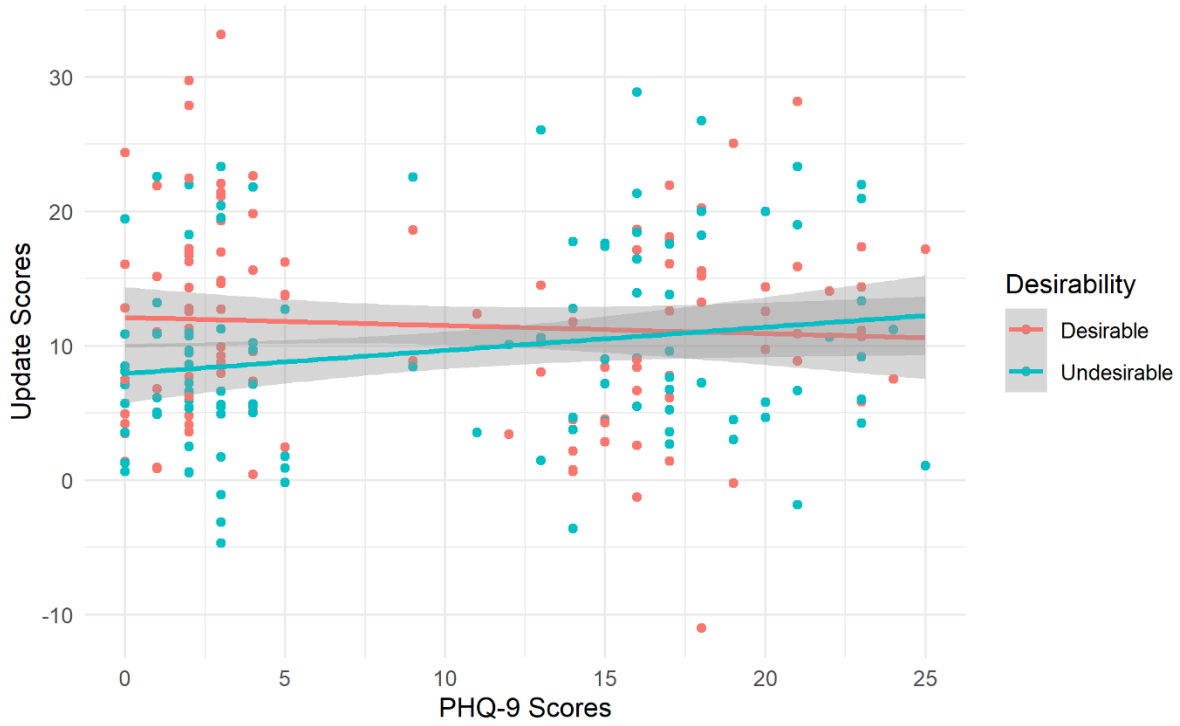
Supplementary Table 5 is an excel file not suitable for print, it is available for download at: https://www.dropbox.com/s/9os2aaoic790hr3/Supplementary%20Table%205%20Optimism%20Bias%20Depression%20Hobbs%20et%20al.xlsx?dl=0

Supplementary Table 6

Results from mixed-effects linear regression models examining the relationship of two continuous measures of depression and desirability with update scores

	β	95% CI	p
BDI-II			
Hypothesis 1 – Negative Life Events			
Intercept	0.09	-0.29, 0.46	< .001
BDI-II	-0.07	-0.42, 0.28	0.702
Desirability	-0.25	-0.49, -0.01	0.029
Group	0.08	-0.58, 0.73	0.818
BDI-II x Desirability	0.15	-0.09, 0.40	0.224
Hypothesis 2 – Positive and Negative Life Events			
Intercept	0.10	-0.23, 0.42	< .001
BDI-II	-0.18	-0.49, 0.13	0.262
Valence	-0.43	-0.66, -0.19	0.019
Desirability	-0.25	-0.48, -0.02	0.022
Group	0.32	-0.23, 0.86	0.257
BDI-II x Valence	0.02	-0.22, 0.25	0.877
BDI-II x Desirability	0.15	-0.08, 0.39	0.204
Valence x Desirability	0.35	0.02, 0.68	0.218
BDI-II x Valence x Desirability	0.01	-0.32, 0.34	0.948
PHQ-9			
Hypothesis 1 – Negative Life Events			
Intercept	0.20	-0.22, 0.61	< .001
PHQ-9	0.00	-0.39, 0.40	0.989
Desirability	-0.25	-0.49, -0.01	0.005
Group	-0.14	-0.90, 0.61	0.707
PHQ-9xDesirability	0.25	0.01, 0.49	0.046
Hypothesis 2 – Positive and Negative Life Events			
Intercept	0.14	-0.22, 0.50	< .001
PHQ-9	-0.17	-0.51, 0.18	0.337
Valence	-0.43	-0.66, -0.20	0.006
Desirability	-0.25	-0.48, -0.02	0.004
Group	0.23	-0.40, 0.85	0.479
PHQ-9 x Valence	0.07	-0.16, 0.31	0.531

PHQ-9 x Desirability	0.25	0.01, 0.48	0.038
Valence x Desirability	0.35	0.02, 0.68	0.064
PHQ-9 x Valence x Desirability	-0.12	-0.45, 0.21	0.480



Supplementary Figure 1

Relationship between total PHQ-9 scores and update scores according to desirability

Supplementary Table 7

Results from a Mixed-Effects Linear Regression Model examining differences in absolute memory errors by group, valence, and desirability

	β	95% CI	p
Intercept	-0.15	-0.41, 0.11	< .001
Group	0.51	0.14, 0.88	0.007
Valence	-0.15	-0.44, 0.15	0.329
Desirability	0.01	-0.29, 0.30	0.959
Group x Valence	-0.43	-0.85, -0.01	0.047
Group x Desirability	-0.28	-0.70, 0.14	0.187
Valence x Desirability	0.43	0.02 0.85	0.043
Group x Valence x Desirability	0.28	-0.32, 0.87	0.365

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Supplementary Table 8

Results from mixed-effects linear regression models adjusting hypothesis one and two for absolute memory errors

	β	95% CI	p
Hypothesis 1 – Negative Life Events			
Intercept	0.23	-0.03, 0.49	< .001
Group	-0.20	-0.57, 0.18	0.302
Desirability	-0.60	-0.93, -0.26	0.001
Memory Errors	-0.12	-0.26, 0.02	0.086
Group x Desirability	0.67	0.20, 1.15	0.007
Hypothesis 2 – Positive and Negative Life Events			
Intercept	0.36	0.11, 0.62	< .001
Group	-0.20	-0.57, 0.16	0.279
Valence	-0.63	-0.96, -0.31	< .001
Desirability	-0.60	-0.92, -0.27	< .001
Memory Errors	-0.12	-0.21, -0.02	0.020
Group x Valence	0.34	-0.13, 0.80	0.156
Group x Desirability	0.68	0.21, 1.14	0.004
Valence x Desirability	0.65	0.19, 1.10	0.006
Group x Valence x Desirability	-0.47	-1.13, 0.18	0.154

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Supplementary Table 9

Results from mixed-effects linear regression models adjusting hypothesis one and two for trait levels of optimism as measured by the LOT-R

	β	95% CI	p
Hypothesis 1 – Negative Life Events			
Intercept	0.16	-0.14, 0.47	< .001
Group	-0.08	-0.58, 0.42	0.753
Desirability	-0.60	-0.93, -0.27	< .001
LOT-R	0.12	-0.10, 0.34	0.286
Group x Desirability	0.71	0.24, 1.18	0.004
Hypothesis 2 – Positive and Negative Life Events			
Intercept	0.28	-0.01, 0.58	< .001
Group	-0.07	-0.53, 0.40	0.781
Valence	-0.62	-0.94, -0.29	< .001
Desirability	-0.60	-0.92, -0.27	< .001
LOT-R	0.13	-0.05, 0.31	0.167
Group x Valence	0.39	-0.08, 0.85	0.104
Group x Desirability	0.71	0.24, 1.17	0.003
Valence x Desirability	0.60	0.14, 1.06	0.011
Group x Valence x Desirability	-0.51	-1.16, 0.15	0.130

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Supplementary Table 10

Results from mixed-effects linear regression models adjusting hypotheses one and two (with update scores as the outcome) and hypotheses three and four (with initial estimates as the outcome) for framing of life events as happening versus not happening

	β	95% CI	p
Hypothesis 1 – Negative Life Events	**Update Scores**		
Intercept	0.28	0.07, 0.50	< .001
Group	-0.18	-0.46, 0.10	0.199
Desirability	-0.46	-0.70, -0.22	< .001
Framing	-0.17	-0.34, 0.00	0.054
Group x Desirability	0.48	0.14, 0.82	0.006
Hypothesis 2 – Positive and Negative Life Events	**Update Scores**		
Intercept	0.27	0.07, 0.47	< .001
Group	-0.18	-0.46, 0.10	0.198
Valence	-0.45	-0.70, -0.21	< .001
Desirability	-0.46	-0.70, -0.22	< .001
Framing	0.03	-0.09, 0.15	0.622
Group x Valence	0.29	-0.06, 0.64	0.101
Group x Desirability	0.48	0.14, 0.83	0.007
Valence x Desirability	0.44	0.10, 0.78	0.013
Group x Valence x Desirability	-0.33	-0.82, 0.16	0.189
Hypothesis 3 – Negative Life Events	**Initial Estimates**		
Intercept	-0.77	-0.98, -0.56	< .001
Group	0.91	0.63, 1.18	< .001
Framing	0.65	0.50, 0.80	< .001
Hypothesis 4 – Positive Life Events	**Initial Estimates**		
Intercept	0.05	-0.18, 0.28	< .001
Group	-0.63	-0.94, -0.32	< .001
Framing	0.51	0.36, 0.65	< .001

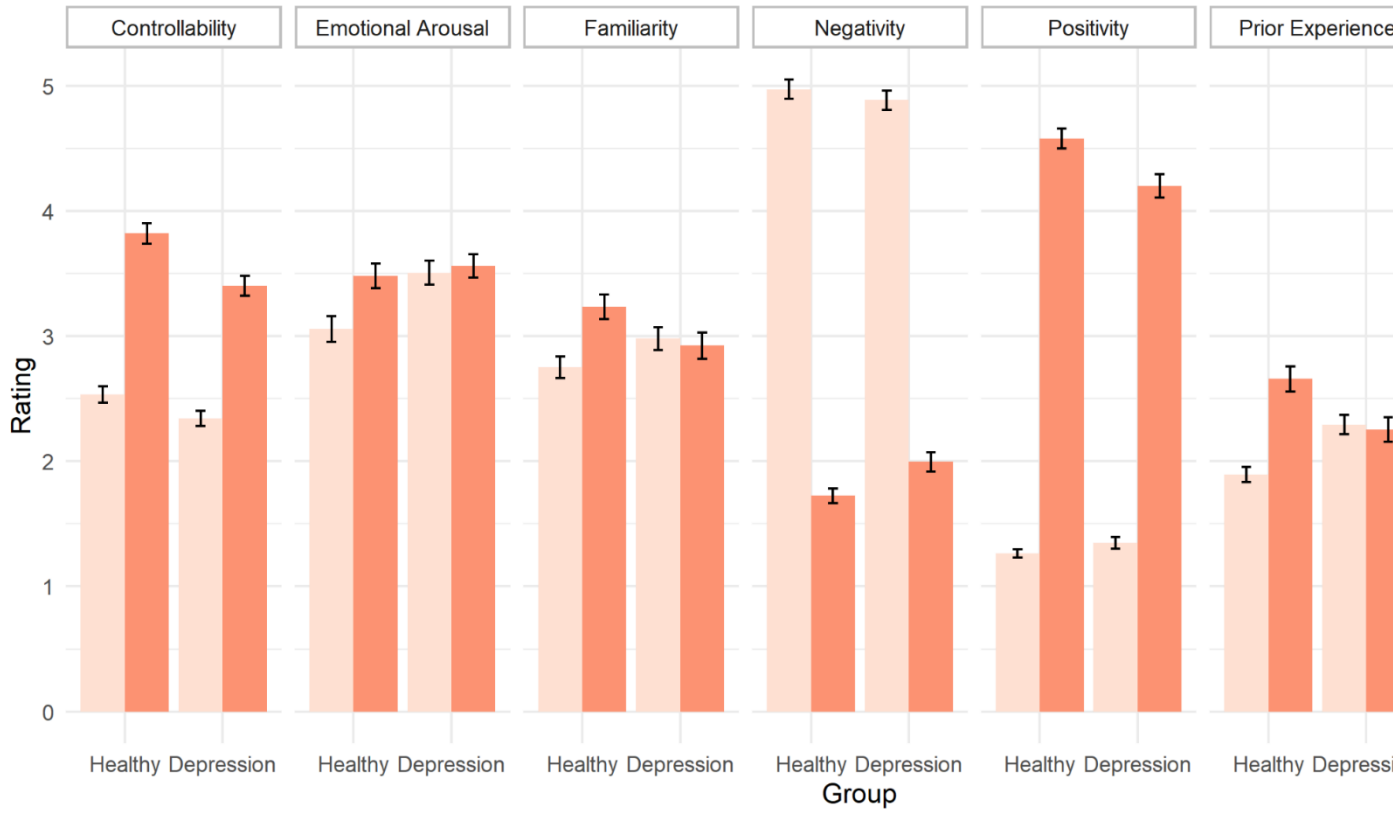
Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable, Framing – Happening.

Supplementary Table 11

Results from mixed-effects linear regression models adjusting hypotheses one and two (with update scores as the outcome) and hypotheses three and four (with initial estimates as the outcome) for valence order

	β	95% CI	p
Hypothesis 1 – Negative Life Events	**Update Scores**		
Intercept	0.30	0.00, 0.60	< .001
Group	-0.27	-0.64, 0.10	0.155
Desirability	-0.60	-0.93, -0.27	0.001
Valence Order	-0.08	-0.36, 0.21	0.594
Group x Desirability	0.71	0.24, 1.18	0.004
Hypothesis 2 – Positive and Negative Life Events	**Update Scores**		
Intercept	0.47	0.19, 0.76	< .001
Group	-0.27	-0.64, 0.09	0.147
Valence	-0.62	-0.94, -0.29	< .001
Desirability	-0.60	-0.92, -0.27	< .001
Valence Order	-0.17	-0.40, 0.07	0.166
Group x Valence	0.39	-0.08, 0.85	0.104
Group x Desirability	0.71	0.24, 1.17	0.003
Valence x Desirability	0.60	0.14, 1.06	0.011
Group x Valence x Desirability	-0.51	-1.16, 0.15	0.130
Hypothesis 3 – Negative Life Events	**Initial Estimates**		
Intercept	-0.28	-0.56, 0.00	< .001
Group	1.04	0.72, 1.35	< .001
Valence Order	-0.42	-0.74, -0.11	0.009
Hypothesis 4 – Positive Life Events	**Initial Estimates**		
Intercept	0.17	-0.14, 0.48	< .001
Group	-0.69	-1.04, -0.34	< .001
Valence Order	0.32	-0.03, 0.67	0.073

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable, Valence Order – Negative Events Completed First.



Supplementary Figure 2

Mean ratings for life events by group and valence. Error bars represent standard errors.

Supplementary Table 12

Results from mixed-effects linear regression models adjusting hypotheses one and two (with update scores as the outcome) and hypotheses three and four (with initial estimates as the outcome) for participant ratings of life events

	β	95% CI	p
Hypothesis 1 – Negative Life Events	**Update Scores**		
Intercept	0.24	-0.03, 0.50	0.043
Group	-0.24	-0.63, 0.15	0.227
Desirability	-0.58	-0.92, -0.24	0.001
Ratings			
Controllability	-0.07	-0.22, 0.07	0.332
Emotional Arousal	-0.03	-0.21, 0.15	0.751
Familiarity	0.11	-0.07, 0.29	0.234
Negativity	0.05	-0.12, 0.21	0.574
Positivity	0.14	-0.02, 0.31	0.096
Prior Experience	-0.05	-0.23, 0.13	0.554
Vividness	-0.09	-0.29, 0.12	0.399
Group x Desirability	0.70	0.22, 1.18	0.005
Hypothesis 2 – Positive and Negative Life Events	**Update Scores**		
Intercept	0.49	0.14, 0.84	0.005
Group	-0.23	-0.61, 0.15	0.229
Valence	-0.85	-1.43, -0.28	0.004
Desirability	-0.59	-0.92, -0.26	0.001
Ratings			
Controllability	-0.04	-0.18, 0.10	0.567
Emotional Arousal	-0.03	-0.16, 0.10	0.686
Familiarity	0.05	-0.09, 0.20	0.480
Negativity	0.16	-0.10, 0.41	0.237
Positivity	0.33	0.03, 0.63	0.033
Prior Experience	-0.03	-0.16, 0.11	0.721
Vividness	-0.06	-0.22, 0.09	0.435
Group x Valence	0.39	-0.09, 0.87	0.113
Group x Desirability	0.70	0.23, 1.17	0.004

Valence x Desirability	0.57	0.07, 1.08	0.027
Group x Valence x Desirability	-0.52	-1.18, 0.14	0.126
Hypothesis 3 – Negative Life Events	**Initial Estimates**		
Intercept	-0.37	-0.60, -0.15	< .001
Group	0.76	0.42, 1.11	< .001
Ratings			
Controllability	-0.25	-0.42, -0.08	0.004
Emotional Arousal	-0.07	-0.28, 0.13	0.487
Familiarity	-0.02	-0.22, 0.18	0.848
Negativity	0.04	-0.16, 0.24	0.687
Positivity	0.02	-0.19, 0.22	0.886
Prior Experience	0.24	0.04, 0.44	0.019
Vividness	0.16	-0.08, 0.40	0.196
Hypothesis 4 – Positive Life Events	**Initial Estimates**		
Intercept	0.19	-0.04, 0.41	0.222
Group	-0.38	-0.73, -0.03	0.032
Ratings			
Controllability	0.26	0.07, 0.45	0.008
Emotional Arousal	0.41	0.20, 0.62	< .001
Familiarity	0.18	-0.06, 0.41	0.137
Negativity	-0.04	-0.23, 0.14	0.654
Positivity	0.05	-0.18, 0.28	0.665
Prior Experience	0.05	-0.15, 0.25	0.632
Vividness	-0.15	-0.39, 0.09	0.227

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Supplementary Table 13

Results from mixed-effects linear regression models repeating hypotheses one and two (with update scores as the outcome) and hypotheses three and four (with initial estimates as the outcome) with event valence categorised according to participants' positivity and negativity ratings

	β	95% CI	p
Hypothesis 1 – Negative Life Events	**Update Scores**		
Intercept	0.11	-0.15, 0.37	< .001
Group	-0.17	-0.54, 0.20	0.370
Desirability	-0.35	-0.66, -0.05	0.025
Group x Desirability	0.61	0.17, 1.04	0.007
Hypothesis 2 – Positive and Negative Life Events	**Update Scores**		
Intercept	0.22	-0.04, 0.48	< .001
Group	-0.16	-0.53, 0.21	0.400
Valence	-0.37	-0.69, -0.05	0.026
Desirability	-0.33	-0.65, -0.01	0.044
Group x Valence	0.34	-0.12, 0.80	0.148
Group x Desirability	0.57	0.11, 1.03	0.015
Valence x Desirability	0.22	-0.23, 0.68	0.334
Group x Valence x Desirability	-0.56	-1.20, 0.09	0.094
Hypothesis 3 – Negative Life Events	**Initial Estimates**		
Intercept	-0.43	-0.67, -0.19	< .001
Group	0.88	0.54, 1.22	< .001
Hypothesis 4 – Positive Life Events	**Initial Estimates**		
Intercept	0.32	0.06, 0.57	< .001
Group	-0.64	-1.00, -0.28	0.001

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Supplementary Table 14

Means and standard deviations of participants' ratings for individual life events by group

Negative Life Events														
	Healthy (n = 56)													
	Controllability		Emotional Arousal		Familiarity		Negativity		Positivity		Prior Experience		Vividness	
Life Event	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Abnormal heart rhythm	1.76	0.88	2.74	1.37	2.30	1.37	4.74	1.29	1.33	0.62	2.11	1.52	2.67	1.51
Age-Related Blindness	1.48	0.68	2.98	1.51	2.43	1.50	4.95	1.63	1.13	0.34	1.31	0.61	2.66	1.64
Alzheimer's disease	1.63	0.73	3.78	1.71	2.94	1.71	5.61	0.95	1.06	0.23	1.94	1.41	3.34	1.51
Arteries hardening (narrowing of blood vessels)	2.96	1.40	2.40	1.38	2.15	1.20	4.98	1.42	1.24	0.72	1.40	1.06	2.50	1.47
Artificial joint	2.28	1.22	2.29	1.25	2.14	1.23	3.54	1.64	2.55	1.58	1.36	0.81	2.54	1.45
Asthma	2.35	1.48	2.31	1.26	3.31	1.91	4.43	1.38	1.26	0.54	2.51	1.86	3.21	1.72
Back Pain	2.98	1.14	3.15	1.41	4.45	1.30	4.75	1.37	1.31	0.68	4.18	1.58	4.11	1.33
Being cheated on by your husband/wife	2.24	1.18	4.38	1.82	2.86	1.87	5.63	0.84	1.13	0.34	1.93	1.66	3.74	1.29
Being convicted of a crime	4.34	1.73	3.28	1.91	2.17	1.52	5.51	1.17	1.12	0.45	1.09	0.36	2.68	1.40
Being fired	3.42	1.65	3.79	1.73	2.41	1.71	5.34	1.18	1.19	0.58	1.60	1.26	3.50	1.48
Bone Fracture	2.76	1.42	3.09	1.58	3.39	1.79	4.26	1.56	1.50	1.03	2.53	1.88	4.00	1.39
Cancer (of digestive system/lung/prostate/breast/skin)	1.65	0.75	4.14	1.81	3.59	1.72	5.70	1.14	1.12	0.45	2.03	1.67	3.67	1.63
Car Stolen	2.76	1.48	3.61	1.66	2.63	1.83	5.21	1.32	1.15	0.44	1.57	1.33	3.23	1.51
Card Fraud	3.02	1.20	3.26	1.65	3.09	1.50	5.19	1.04	1.07	0.33	2.87	1.91	3.21	1.60
Death before age 80	2.24	1.10	3.52	1.70	2.95	1.66	4.95	1.36	1.35	0.64	2.02	1.58	3.29	1.50
Dementia	1.93	1.05	3.72	1.64	3.40	1.66	5.41	1.11	1.23	0.84	2.20	1.60	3.38	1.43

Diabetes (type 2)	3.33	1.72	2.33	1.28	3.05	1.62	5.17	1.08	1.13	0.41	1.48	1.04	2.69	1.47
Divorce	3.37	1.34	4.10	1.68	3.48	1.92	5.40	0.90	1.67	1.19	1.87	1.63	3.86	1.50
Drug Abuse	3.96	1.85	3.08	1.64	2.68	1.61	5.61	1.06	1.13	0.45	1.32	0.77	2.79	1.30
Eye cataract (clouding of the lens of the eye)	1.73	0.72	2.59	1.50	2.49	1.52	4.35	1.78	1.06	0.23	1.41	1.01	2.73	1.61
Gallbladder stones	2.00	1.11	2.10	1.19	1.79	1.16	4.45	1.47	1.18	0.45	1.39	0.92	2.07	1.32
Gluten Intolerance	2.40	1.50	1.89	1.17	2.14	1.37	3.98	1.49	1.36	0.57	1.43	1.15	2.24	1.32
Having a stroke	2.19	1.25	3.83	1.64	2.75	1.64	5.35	1.37	1.06	0.23	1.57	1.21	3.21	1.63
Heart Failure	2.29	1.21	3.53	1.62	2.68	1.46	5.53	0.96	1.27	0.59	1.58	1.13	2.95	1.40
Hernia	2.02	1.13	2.02	1.08	2.58	1.53	4.34	1.76	1.14	0.35	1.54	1.33	2.41	1.55
Infertility	1.64	0.87	3.54	1.68	2.46	1.80	4.90	1.45	1.46	1.12	1.51	1.32	3.11	1.58
Irritable bowel syndrome (disorder of the gut)	2.49	1.38	2.23	1.27	2.29	1.56	4.89	1.24	1.24	0.49	1.89	1.58	2.53	1.36
Kidney Stones	1.91	0.95	2.12	1.42	2.20	1.29	4.72	1.34	1.20	0.46	1.28	0.63	2.35	1.46
Knee Osteoarthritis	2.09	1.16	2.40	1.35	2.00	1.35	5.03	1.05	1.16	0.45	1.37	0.88	2.44	1.46
Migraine	2.14	1.18	2.78	1.49	3.75	1.84	5.15	1.10	1.20	0.46	3.33	1.97	4.03	1.60
Obesity	4.38	1.47	3.07	1.55	3.46	1.77	5.00	1.30	1.27	0.55	2.00	1.40	3.76	1.59
Osteoporosis (reduced bone density)	2.31	1.11	2.32	1.47	2.35	1.46	4.79	1.17	1.31	0.77	1.51	1.17	2.62	1.62
Severe teeth problems when old	3.36	1.48	2.79	1.61	2.18	1.21	4.97	1.42	1.22	0.48	1.88	1.40	3.15	1.58
Skin burn	4.00	1.40	2.93	1.49	4.48	1.55	4.62	1.57	1.39	0.97	4.18	1.60	3.87	1.51
Ulcers	2.93	1.44	2.16	1.31	3.09	1.51	4.09	1.64	1.25	0.49	2.31	1.73	2.89	1.57
Victim of mugging	1.83	1.06	3.84	1.94	2.49	1.74	5.36	1.29	1.03	0.17	1.49	1.16	3.08	1.58
Victim of violence at home	2.08	1.11	3.58	1.89	2.61	1.84	5.74	0.75	1.05	0.32	1.56	1.34	3.24	1.82

Victim of violence by a stranger	1.81	1.33	3.79	1.87	2.26	1.48	5.44	1.37	1.14	0.35	1.64	1.35	3.53	1.74
Victim of violence by an acquaintance	1.95	1.14	3.76	1.75	2.08	1.35	5.18	1.36	1.30	0.88	1.42	0.98	2.94	1.50
Witness a traumatizing accident	1.49	1.01	4.03	1.82	2.92	1.72	5.33	1.20	1.12	0.40	2.38	1.77	3.68	1.65

	Depression (n = 54)													
	Controllability		Emotional Arousal		Familiarity		Negativity		Positivity		Prior Experience		Vividness	
Life Event	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Abnormal heart rhythm	1.97	1.09	3.27	1.61	2.93	1.69	4.42	1.57	1.33	0.67	2.53	1.71	3.41	1.58
Age-Related Blindness	1.59	0.80	3.55	1.74	2.64	1.51	4.91	1.50	1.30	0.97	1.70	1.40	3.23	1.65
Alzheimer’s disease	1.50	0.82	4.49	1.53	3.64	1.57	5.56	0.80	1.13	0.40	2.16	1.63	4.16	1.51
Arteries hardening (narrowing of blood vessels)	2.68	1.04	2.85	1.50	2.53	1.61	5.19	1.01	1.23	0.77	1.63	1.13	2.84	1.44
Artificial joint	2.00	1.00	2.91	1.71	2.48	1.66	3.76	1.67	2.18	1.36	1.65	1.32	2.93	1.54
Asthma	1.90	1.26	2.89	1.45	3.59	1.70	4.40	1.45	1.54	0.90	2.65	1.90	3.52	1.60
Back Pain	2.62	1.19	3.54	1.58	5.08	1.14	5.00	1.08	1.55	1.01	4.97	1.20	4.76	1.28
Being cheated on by your husband/wife	2.08	1.13	4.61	1.53	2.91	1.70	5.32	1.14	1.13	0.52	2.23	1.62	3.86	1.78
Being convicted of a crime	4.24	1.46	3.44	1.73	1.97	1.32	4.97	1.59	1.32	0.91	1.31	1.03	3.13	1.69
Being fired	3.39	1.45	3.85	1.66	2.84	1.66	4.84	1.57	1.51	0.91	1.97	1.44	3.71	1.64
Bone Fracture	2.58	1.30	3.43	1.63	3.61	1.74	4.51	1.52	1.63	1.13	2.51	1.93	3.95	1.54
Cancer (of digestive system/lung/prostate/breast/skin)	1.76	0.94	4.87	1.49	4.31	1.47	5.56	1.05	1.17	0.59	2.41	1.74	4.56	1.40
Car Stolen	2.50	1.31	3.56	1.79	2.35	1.60	5.18	1.29	1.38	1.11	1.83	1.63	3.56	1.42
Card Fraud	2.63	1.17	3.59	1.32	3.68	1.72	4.97	1.54	1.31	0.80	3.20	2.04	4.45	1.31
Death before age 80	2.26	1.16	3.61	1.86	2.77	1.75	4.50	1.64	1.72	1.24	1.81	1.51	3.85	1.79
Dementia	1.25	0.44	4.73	1.56	3.89	1.84	5.61	0.86	1.12	0.51	2.46	1.74	4.21	1.56

Diabetes (type 2)	3.43	1.26	3.24	1.63	3.72	1.71	4.75	1.38	1.24	0.55	2.29	1.71	3.75	1.66
Divorce	2.88	1.28	4.36	1.65	3.05	2.02	4.59	1.66	1.43	0.86	2.15	1.94	3.68	1.61
Drug Abuse	3.38	1.36	3.47	1.59	2.57	1.53	5.08	1.25	1.28	0.63	1.71	1.08	3.31	1.62
Eye cataract (clouding of the lens of the eye)	1.71	0.96	2.95	1.56	2.29	1.51	4.66	1.42	1.24	0.54	1.58	1.08	3.46	1.46
Gallbladder stones	2.11	0.97	2.71	1.36	2.42	1.54	4.54	1.35	1.30	0.71	1.74	1.29	2.58	1.40
Gluten Intolerance	2.26	1.41	2.34	1.40	2.68	1.61	3.44	1.58	1.64	0.90	2.45	1.66	2.97	1.38
Having a stroke	2.03	0.84	3.82	1.63	3.14	1.71	4.92	1.61	1.18	0.60	1.77	1.29	3.74	1.44
Heart Failure	2.19	0.98	3.89	1.56	2.66	1.76	5.77	0.48	1.10	0.38	1.70	1.42	3.75	1.64
Hernia	1.88	1.11	2.46	1.45	2.34	1.46	4.33	1.51	1.12	0.40	1.70	1.30	3.00	1.63
Infertility	1.43	0.78	3.76	1.81	2.58	1.57	4.54	1.82	1.39	0.95	1.89	1.58	3.68	1.72
Irritable bowel syndrome (disorder of the gut)	2.39	1.20	3.22	2.00	3.80	1.89	4.89	1.18	1.43	0.74	3.15	2.06	3.70	1.78
Kidney Stones	2.43	1.08	2.69	1.67	2.41	1.65	4.68	1.61	1.27	0.56	1.97	1.56	2.68	1.33
Knee Osteoarthritis	2.00	1.00	2.70	1.49	2.56	1.79	4.30	1.79	1.40	0.79	2.10	1.62	2.79	1.47
Migraine	2.11	1.06	3.70	1.49	4.78	1.56	5.16	0.99	1.30	0.81	4.41	1.79	4.95	1.29
Obesity	4.11	1.37	3.75	1.60	4.44	1.53	4.86	1.40	1.44	1.03	3.44	1.99	4.44	1.30
Osteoporosis (reduced bone density)	2.09	1.17	2.92	1.30	2.59	1.71	4.50	1.34	1.28	0.61	1.72	1.14	3.05	1.58
Severe teeth problems when old	3.44	1.29	3.44	1.46	2.97	1.40	5.16	0.86	1.19	0.49	2.17	1.40	3.67	1.44
Skin burn	3.66	1.53	3.31	1.62	4.03	1.58	4.73	1.30	1.67	1.12	4.21	1.63	4.38	1.35
Ulcers	2.31	1.20	3.03	1.36	3.15	1.66	4.74	1.31	1.49	0.91	2.73	1.76	3.48	1.65
Victim of mugging	1.89	0.93	4.18	1.57	2.82	1.56	5.52	0.97	1.20	0.82	1.89	1.47	4.13	1.36
Victim of violence at home	2.10	1.45	4.60	1.58	3.18	1.88	5.44	1.29	1.07	0.34	2.28	1.71	3.98	1.78

Victim of violence by a stranger	1.89	1.19	4.33	1.88	3.24	1.92	5.45	1.06	1.11	0.39	2.78	1.99	4.20	1.71
Victim of violence by an acquaintance	2.14	1.34	4.25	1.93	2.63	1.80	5.28	1.43	1.21	0.66	2.41	1.89	3.90	1.68
Witness a traumatizing accident	1.22	0.58	4.77	1.37	3.29	1.90	5.58	1.03	1.11	0.52	2.76	1.94	4.42	1.41

Positive Life Events

	Healthy (n = 56)													
	Controllability		Emotional Arousal		Familiarity		Negativity		Positivity		Prior Experience		Vividness	
Life Event	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Celebrate your 20th wedding anniversary	3.64	1.60	4.29	1.42	3.59	1.95	1.32	0.88	5.34	1.05	2.36	2.07	3.85	1.65
Celebrate your 70th birthday	2.85	1.42	4.16	1.48	2.84	1.72	1.61	1.28	5.30	0.94	1.46	0.99	3.36	1.57
Celebrate your diamond (60 years) wedding anniversary	2.40	1.31	3.95	1.82	2.23	1.43	1.18	0.66	5.23	1.18	1.46	0.98	2.86	1.54
Create a start-up (a new business) that is successful for 5 or more years	3.17	1.44	3.18	1.57	2.75	1.60	1.94	1.35	4.38	1.55	2.09	1.69	3.08	1.63
Earn more than £13 an hour	3.80	1.42	3.68	1.51	4.38	1.51	1.38	0.82	5.31	1.03	4.04	2.01	4.38	1.29
Earn more than £30,000 or more a year (before tax)	3.71	1.55	3.55	1.62	3.88	1.75	1.45	1.08	5.23	0.87	3.21	2.20	4.22	1.44
Fall in love with someone you met on a dating site	2.54	1.20	3.16	1.71	2.27	1.50	2.00	1.36	3.43	1.68	1.87	1.61	2.89	1.49
Fully paying off your mortgage before turning 55	3.46	1.27	3.80	1.42	3.16	1.94	1.13	0.41	5.25	1.13	2.09	1.78	3.65	1.56
Get engaged on Valentine's Day	3.63	1.88	3.09	1.79	2.45	1.62	2.12	1.55	3.46	1.63	1.46	1.34	3.10	1.51
Getting married in the summer	4.53	1.26	3.98	1.46	3.18	1.53	1.25	0.79	5.00	1.18	2.29	1.76	4.04	1.33
Getting married or having a civil partner	4.28	1.45	4.70	1.36	4.57	1.69	1.51	0.84	5.26	1.04	3.53	2.32	4.90	1.34
Go abroad on holiday for more than 2 weeks	4.62	1.16	4.30	1.53	4.14	1.82	1.36	0.92	5.29	0.87	4.18	1.99	4.55	1.41
Go on an all-inclusive holiday	4.33	1.35	3.17	1.51	3.74	1.76	2.00	1.29	4.24	1.54	3.40	1.96	3.95	1.51
Go on holiday abroad for Christmas	4.68	1.30	3.84	1.81	3.49	1.96	2.07	1.46	4.45	1.39	3.12	2.20	4.13	1.36
Go on holiday to USA/Canada	4.36	1.25	3.47	1.64	3.86	1.73	1.57	0.96	4.68	1.35	3.59	2.22	4.21	1.65

Go skiing	3.91	1.59	2.72	1.57	2.70	1.85	2.20	1.55	3.57	1.72	2.2	1.8	3.13	1.69
Going on a spontaneous holiday	4.19	1.57	4.17	1.56	4.05	1.71	1.49	1.12	5.05	1.08	3.3	1.7	4.14	1.40
Have a higher income or the same income in retirement compared to when you were working	3.00	1.55	3.48	1.52	2.17	1.54	1.72	1.28	5.16	1.26	1.63	1.39	3.38	1.51
Have a job you enjoy very much	3.82	1.45	4.15	1.41	3.98	1.64	1.65	1.25	5.37	0.83	3.93	1.97	4.42	1.40
Have over £50,000 in savings	3.71	1.52	4.28	1.52	3.62	1.86	1.32	0.94	5.67	0.52	2.56	1.79	4.05	1.38
Having a honeymoon outside of Europe	4.27	1.33	4.19	1.74	3.18	1.93	1.42	0.95	4.80	1.39	2.38	2.03	4.05	1.66
Hold an investment	3.73	1.68	2.79	1.52	3.93	1.86	1.57	0.99	4.34	1.61	3.50	1.94	3.41	1.76
Live abroad	4.24	1.48	3.84	1.76	3.37	1.95	1.77	1.06	4.95	1.14	3.22	2.23	4.30	1.31
Living in a house with 8 or more rooms	2.93	1.25	3.17	1.59	2.66	1.77	2.11	1.51	4.08	1.34	2.02	1.67	3.48	1.57
Own 2 or more cars	3.69	1.47	2.16	1.34	3.18	1.91	1.90	1.23	3.40	1.42	2.76	2.09	3.18	1.60
Own a second home	3.47	1.32	3.32	1.72	2.71	1.58	1.68	1.14	4.33	1.47	1.71	1.45	3.43	1.58
Owning a house	4.53	1.27	4.37	1.56	4.76	1.70	1.35	0.69	5.54	1.07	3.79	2.34	4.87	1.35
Owning a house before turning 35	3.63	1.41	4.28	1.55	4.12	1.66	1.39	0.95	5.41	1.00	3.39	2.25	4.49	1.42
Owning a pet	4.98	1.19	4.55	1.45	5.07	1.53	1.42	0.55	5.32	1.14	5.10	1.48	5.00	1.24
Playing a musical instrument	4.60	1.17	3.72	1.65	4.05	1.76	1.36	0.78	4.45	1.52	3.70	2.01	4.00	1.67
Repaying your student loan in full	3.22	1.51	2.77	1.69	2.74	1.88	1.82	1.39	4.64	1.66	2.14	1.91	2.85	1.81
Retire early (before 65)	3.42	1.50	3.66	1.64	3.03	1.94	1.27	0.52	5.00	1.38	1.63	1.41	3.33	1.78
Speaking two or more foreign languages fluently	4.33	1.31	3.08	1.78	3.58	2.00	1.40	0.95	5.13	1.18	2.75	2.08	3.95	1.79
Spend £20,000 or more on your wedding	4.33	1.44	2.80	1.54	2.10	1.43	3.60	1.81	2.43	1.33	1.33	0.94	2.40	1.29
Staying in a 4 or 5 star hotel on holiday	4.50	1.32	3.67	1.45	4.39	1.33	1.42	1.03	5.18	1.05	4.02	1.79	4.61	1.13
Sticking to your wedding budget	4.24	1.24	2.89	1.55	3.46	1.79	1.35	0.79	4.69	1.05	3.31	2.11	3.64	1.53

Taking part in a running race	4.64	1.31	2.70	1.58	3.34	2.07	2.33	1.68	3.55	1.73	3.08	2.06	3.58	1.57
Visit Asia	3.98	1.46	3.56	1.76	2.86	1.82	1.56	0.79	4.62	1.34	2.48	2.02	4.18	1.60
Visit South America	3.65	1.72	3.25	1.79	2.18	1.48	1.56	1.14	4.59	1.50	1.39	1.17	3.32	1.57
Visit Spain	4.65	1.03	3.73	1.37	4.55	1.58	1.59	0.85	5.11	1.06	4.88	1.88	4.41	1.44
	Depression (n = 54)													
	Controllability		Emotional Arousal		Familiarity		Negativity		Positivity		Prior Experience		Vividness	
Life Event	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Celebrate your 20th wedding anniversary	3.03	1.22	3.84	1.68	2.51	1.65	1.61	1.08	4.32	1.73	1.61	1.44	3.19	1.56
Celebrate your 70th birthday	2.43	1.38	3.62	1.51	2.76	1.78	2.41	1.63	4.00	1.73	1.61	1.38	2.90	1.58
Celebrate your diamond (60 years) wedding anniversary	1.78	0.95	3.54	1.99	2.03	1.30	1.72	1.32	4.18	1.67	1.12	0.32	2.68	1.72
Create a start-up (a new business) that is successful for 5 or more years	2.82	1.37	2.93	1.67	2.21	1.49	2.28	1.68	4.00	1.71	1.59	1.17	2.97	1.63
Earn more than £13 an hour	3.09	1.44	4.20	1.47	3.80	2.00	1.43	0.90	5.00	1.26	2.85	2.15	4.12	1.48
Earn more than £30,000 or more a year (before tax)	3.08	1.40	4.08	1.55	3.13	1.99	1.61	1.24	4.35	1.65	2.60	2.06	3.95	1.60
Fall in love with someone you met on a dating site	2.71	1.33	3.51	1.87	3.06	1.79	2.74	1.81	3.37	1.68	2.27	1.90	3.00	1.76
Fully paying off your mortgage before turning 55	3.03	1.16	3.90	1.69	2.08	1.26	1.70	1.42	4.78	1.65	1.32	0.96	2.95	1.47
Get engaged on Valentine's Day	2.79	1.74	3.14	1.69	1.79	1.34	2.39	1.78	2.87	1.53	1.18	0.51	2.89	1.71
Getting married in the summer	4.23	1.58	3.79	1.63	3.00	1.88	1.76	1.01	3.90	1.67	1.78	1.62	3.61	1.66
Getting married or having a civil partner	3.46	1.56	4.57	1.66	4.03	1.98	2.24	1.58	4.44	1.71	2.47	2.09	4.34	1.55
Go abroad on holiday for more than 2 weeks	3.72	1.37	3.68	1.72	3.50	1.98	1.67	1.21	4.75	1.40	3.16	2.16	4.03	1.69
Go on an all-inclusive holiday	4.05	1.59	3.58	1.65	3.33	1.90	2.00	1.36	4.24	1.61	3.38	2.12	3.85	1.81
Go on holiday abroad for Christmas	3.87	1.46	3.45	1.67	2.61	1.65	2.05	1.50	3.84	1.46	1.76	1.46	3.56	1.63
Go on holiday to USA/Canada	3.89	1.35	4.09	1.31	4.06	1.77	2.15	1.37	4.46	1.43	3.26	2.25	4.32	1.31

Go skiing	3.55	1.67	2.45	1.53	2.561	1.821	2.321	1.781	3.351	1.781	1.55	1.19	3.051	1.621
Going on a spontaneous holiday	3.03	1.48	4.00	1.31	3.132	1.522	2.141	1.231	4.301	1.281	2.70	1.73	4.051	1.361
Have a higher income or the same income in retirement compared to when you were working	2.16	1.13	3.23	1.73	1.576	0.999	1.955	1.391	4.321	1.781	1.21	0.88	2.731	1.591
Have a job you enjoy very much	2.97	1.30	4.33	1.54	3.431	1.771	1.641	1.281	5.231	1.171	3.18	1.83	4.141	1.401
Have over £50,000 in savings	3.00	1.36	4.17	1.70	2.311	1.441	1.701	1.531	4.971	1.461	1.65	1.38	3.541	1.421
Having a honeymoon outside of Europe	3.71	1.54	3.89	1.71	2.621	1.701	1.721	1.301	4.381	1.691	1.33	1.06	3.781	1.701
Hold an investment	3.30	1.47	3.15	1.73	2.541	1.691	2.151	1.421	3.771	1.711	2.37	1.85	3.111	1.751
Live abroad	3.28	1.43	3.46	1.78	2.721	1.951	2.161	1.401	3.841	1.571	2.07	1.84	3.581	1.601
Living in a house with 8 or more rooms	2.50	1.27	3.03	1.53	2.591	1.911	2.231	1.481	3.851	1.511	1.90	1.75	2.971	1.551
Own 2 or more cars	3.22	1.42	2.28	1.50	3.052	1.902	2.531	1.741	3.381	1.654	2.44	1.95	2.981	1.751
Own a second home	2.56	1.37	3.41	1.64	1.771	1.293	2.031	1.421	3.761	1.541	1.38	1.03	2.601	1.401
Owning a house	3.28	1.64	4.46	1.50	3.642	2.051	1.951	1.381	5.081	1.501	2.97	2.20	4.371	1.731
Owning a house before turning 35	3.00	1.41	3.94	1.64	3.061	1.851	2.081	1.501	4.501	1.561	2.29	1.90	3.611	1.611
Owning a pet	4.85	1.17	5.19	1.06	5.311	1.201	1.881	1.141	5.561	0.751	5.08	1.65	5.441	1.101
Playing a musical instrument	4.61	1.20	3.42	1.65	4.251	1.891	1.551	1.151	4.531	1.491	3.77	2.07	4.471	1.531
Repaying your student loan in full	2.62	1.41	3.15	1.87	2.511	1.821	2.111	1.561	3.951	1.881	1.55	1.33	3.001	1.601
Retire early (before 65)	2.67	1.24	3.86	1.68	2.121	1.561	1.731	1.101	4.811	1.491	1.22	0.85	3.311	1.661
Speaking two or more foreign languages fluently	3.86	1.40	2.89	1.59	2.771	1.511	1.501	1.221	4.231	1.817	2.18	1.48	3.051	1.601
Spend £20,000 or more on your wedding	4.28	1.49	2.89	1.66	2.231	1.421	3.371	1.761	2.391	1.531	1.27	1.03	2.411	1.481
Staying in a 4 or 5 star hotel on holiday	3.67	1.55	4.06	1.41	4.131	1.791	1.201	0.471	4.801	1.185	4.09	1.79	4.741	1.451
Sticking to your wedding budget	4.54	1.37	3.37	1.62	2.641	1.861	2.001	1.331	4.361	1.427	2.08	1.87	3.331	1.761

Taking part in a running race	4.05	1.67	2.58	1.37	3.20	1.88	2.67	1.78	3.26	1.79	2.50	1.91	3.43	1.65
Visit Asia	3.63	1.48	3.31	1.79	2.94	2.00	1.66	1.00	3.85	1.72	2.23	2.01	3.12	1.56
Visit South America	3.82	1.70	2.98	1.68	2.14	1.59	2.29	1.62	3.81	1.64	1.70	1.52	2.95	1.69
Visit Spain	4.06	1.47	3.61	1.73	4.34	1.96	1.70	1.00	4.46	1.60	3.51	2.22	4.11	1.59

Supplementary Table 15

Results from mixed-effects linear regression models adjusting hypotheses one to four for age

	β	95% CI	p
Hypothesis 1 – Negative Life Events	**Update Scores**		
Intercept	0.26	0.00, 0.52	< .001
Group	-0.29	-0.66, 0.09	0.133
Desirability	-0.60	-0.93, -0.26	0.001
Age	-0.06	-0.20, 0.08	0.413
Group x Desirability	0.71	0.23, 1.19	0.004
Hypothesis 2 – Positive and Negative Life Events	**Update Scores**		
Intercept	0.39	0.13, 0.65	< .001
Group	-0.29	-0.66, 0.08	0.127
Valence	-0.62	-0.94, -0.29	< .001
Desirability	-0.60	-0.92, -0.27	< .001
Age	-0.07	-0.19, 0.05	0.242
Group x Valence	0.39	-0.08, 0.85	0.104
Group x Desirability	0.71	0.24, 1.17	0.003
Valence x Desirability	0.60	0.14, 1.06	0.011
Group x Valence x Desirability	-0.51	-1.16, 0.15	0.130
Hypothesis 3 – Negative Life Events	**Initial Estimates**		
Intercept	-0.53	-0.76, -0.30	< .001
Group	1.08	0.75, 1.41	< .001
Age	0.07	-0.10, 0.23	0.411
Hypothesis 4 – Positive Life Events	**Initial Estimates**		
Intercept	0.41	0.18, 0.64	< .001
Group	-0.84	-1.17, -0.51	< .001
Age	-0.37	-0.54, -0.20	< .001

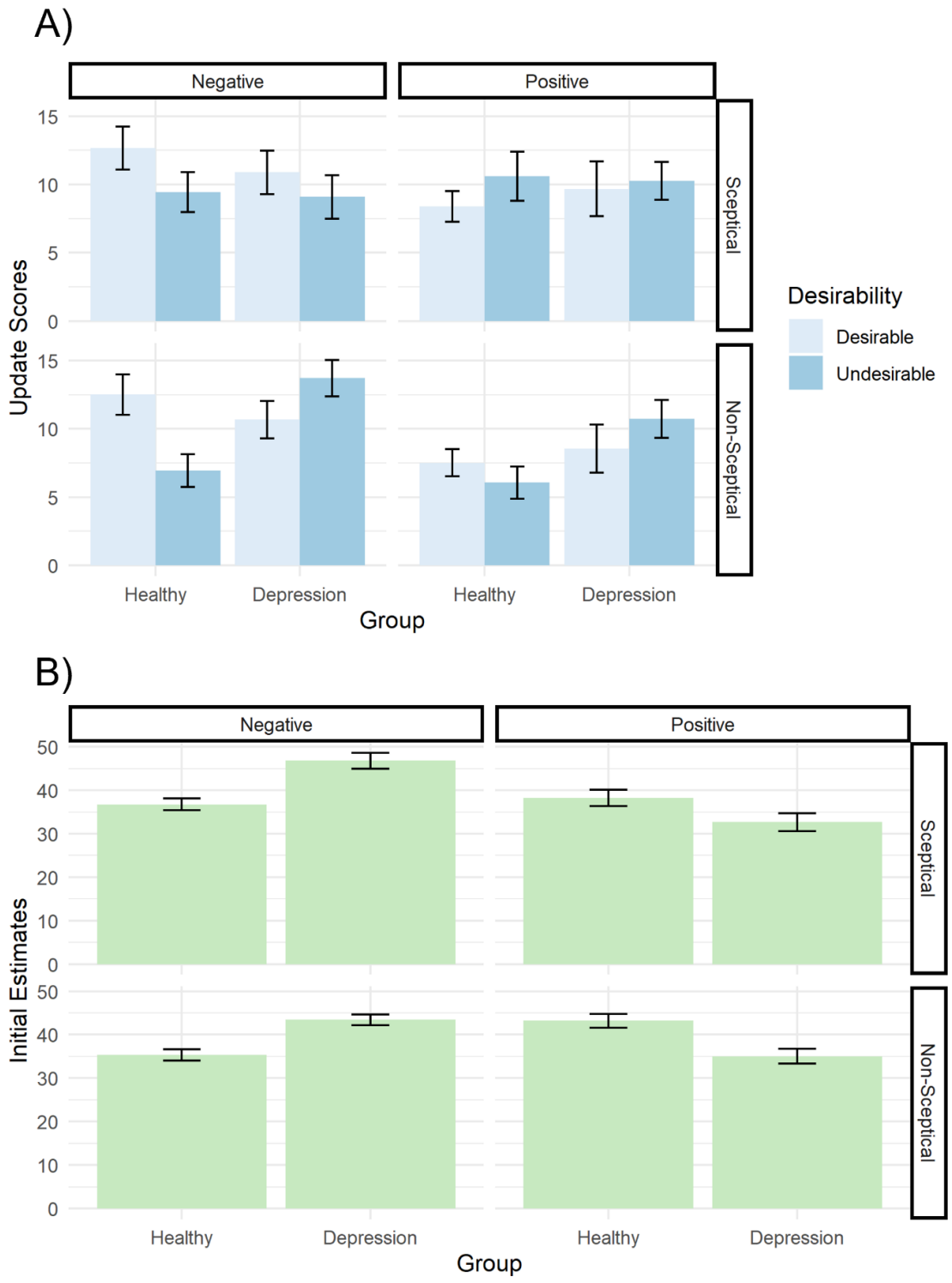
Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.

Supplementary Table 16

Results from mixed-effects linear regression models repeating hypotheses one and two (with update scores as the outcome) and hypotheses three and four (with initial estimates as the outcome) with participants sceptical as to the validity of presented probabilities of experiencing a life event excluded (n included in these models = 48)

	β	95% CI	p
Hypothesis 1 – Negative Life Events	**Update Scores**		
Intercept	0.20	-0.13, 0.54	< .001
Group	-0.25	-0.72, 0.23	0.313
Desirability	-0.72	-1.18, -0.26	0.003
Group x Desirability	1.13	0.47, 1.80	0.001
Hypothesis 2 – Positive and Negative Life Events	**Update Scores**		
Intercept	0.38	0.05, 0.71	< .001
Group	-0.25	-0.72, 0.23	0.314
Valence	-0.63	-1.07, -0.20	0.005
Desirability	-0.72	-1.16, -0.28	0.001
Group x Valence	0.38	-0.25, 1.00	0.235
Group x Desirability	1.13	0.51, 1.76	< .001
Valence x Desirability	0.53	-0.08, 1.15	0.089
Group x Valence x Desirability	-0.69	-1.57, 0.20	0.129
Hypothesis 3 – Negative Life Events	**Initial Estimates**		
Intercept	-0.49	-0.80, -0.18	< .001
Group	1.01	0.57, 1.45	< .001
Hypothesis 4 – Positive Life Events	**Initial Estimates**		
Intercept	0.39	0.06, 0.71	< .001
Group	-0.80	-1.27, -0.33	0.001

Reference groups: Group – Healthy, Valence – Negative, Desirability – Desirable.



Supplementary Figure 3

Mean update scores (A) and initial estimates (B) by group, valence, desirability and scepticism of the presented probability of experiencing a life event. Error bars represent standard error