Effect of valenced stimuli on hippocampus-dependent spatial memory in depression

Janice Williams

DClinPsy Thesis (Volume 1), 2017

University College London

THESIS DECLARATION FORM

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Name: Janice Williams

Date: 23rd June 2017

OVERVIEW

This thesis focused on neurobiological changes in depression. **Part 1** reviewed research investigating the relationship between trauma in childhood, depression in adulthood and dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis. The reviewed studies provided some evidence for associated changes to the HPA axis, however, there was inconsistency in findings as to whether this was linked to childhood trauma or depression in adulthood independently, or an interaction of the two. The studies provided tentative evidence for a specific sub-type of depression, associated with a history of childhood trauma, but study heterogeneity and lack of replication mean further research is needed.

Part 2 reported an investigation into the effects on hippocampus-dependent spatial memory of using valenced stimuli (neutral, negative) for both people with depression and healthy controls. No influence of valence on memory performance was found, counter to the hypotheses. However, some interesting effects of sex were found, warranting further investigation. Potential reasons for the null results were explored, along with suggestions for further research to allow these reasons to be examined. This investigation formed part of a joint study with Line Sagfors (trainee clinical psychologist, UCL; Sagfors, 2017).

Part 3 reflected on some of the methodological issues and learnings from the major research project, including adapting an existing spatial memory test and working with a mental health service and clinical population. It explored the advantages of conducting joint DClinPsy research, both practical and emotional, and ended with some brief reflections on the challenges of the literature review.

TABLE OF CONTENTS

TABLES AND FIGURES

Part 1: LITERATURE REVIEW

PART 2: EMPIRICAL PAPER

ACKNOWLEDGEMENTS

I would like to thank everybody who contributed to this research project. This includes all the people who participated in this research. In particular, I am grateful to the people struggling with considerable levels of depression who gave me their time to contribute to the understanding of depression, rather than for any immediate and direct benefit to themselves.

Two people deserve special thanks. Dr John King, my supervisor, who always made himself available to provide timely and thoughtful feedback and counsel, as well as reassurance and encouragement when needed. Also, my joint researcher, Line Sagfors. Without Line, the journey to the end of this thesis would have been a long, stressful and lonely one. Her organisational and motivational skills were particularly valued, as well as her ego-free attitude to just getting things done.

I would like to thank my husband and friends for their support and encouragement, and my children for their forbearance and understanding while their mother was otherwise engaged.

Lastly, I would like to thank the NHS for paying for my training and employing me while I trained. Without this, it would have been incredibly difficult for me to fulfil my goal of becoming a clinical psychologist.

PART 1: LITERATURE REVIEW

Relationship between depression in adulthood, childhood trauma and HPA axis functioning

1 ABSTRACT

Aims: Adults with depression show differences in hypothalamic-pituitary-adrenal (HPA) axis functioning, as do people with a history of childhood trauma. This review aimed to explore evidence for an interaction between childhood trauma and depression in adulthood and its effects on basal activity and reactivity of the HPA axis, as well as possible moderators of this relationship.

Method: A systematic review of PSYCHInfo, MedLine, and Embase was completed to identify peer-reviewed studies measuring depression, childhood trauma and HPA axis functioning up to August 2016. Fifteen studies were included in the review.

Results: There was heterogeneity within the studies in terms of design, methodology and analysis which made a comparison of studies problematic. There was evidence to suggest childhood trauma and depression were associated with changes to the HPA axis, both in terms of basal activity and reactivity. However, findings were inconsistent in terms of what these effects were and whether they were driven by a history of childhood trauma, depression or an interaction of the two. The heterogeneity also made it difficult to assess the influence of potential moderators, including sex and co-morbid PTSD.

Conclusions: Tentatively, the evidence pointed to a specific subtype of depression, associated with childhood trauma, one that may benefit from different treatment approaches. Given the potential importance of this, there is an urgent need for replication studies in this area, structured to also explore potential moderators, to construct a more reliable picture that may then be used to influence interventions.

2 INTRODUCTION

Stress occurs when a situation's perceived demands outweigh perceived coping ability. Stressful life events, including, but not limited to, physical and sexual assault, have been shown to be important predisposing factors in the development of depression, as well as triggers for specific depressive episodes (Kendler, Karkowski & Prescott, 1999). What is more, experiencing trauma in childhood significantly increases the risk of developing depression in adulthood (Heim et al., 2000; Fergusson, Swain-Campbell, & Horwood, 2002; Cohen, Brown& Smailes, 2001). Childhood trauma (CT) also influences the course of depression and treatment response (Nemeroff et al., 2003; Miniati et al., 2010). Studies found 35% of people with depression had a history of CT, compared to 15% of healthy controls (Fellitti et al., 1998; Young, Abelson, Curtis & Nesse, 1997). It is suggested this is due to the cumulative effects of the resultant stress response over time. While the first episode of depression may have a clear trigger, triggers for further episodes become increasingly unclear. This suggests people with depression become increasingly sensitised to stress, reducing the threshold to trigger an episode (Post, 1992).

Key neurobiological systems have been hypothesised to mediate this environmental vulnerability to depression following CT: the corticotropin-releasing factor (CRF) neuronal systems and the hypothalamic-pituitary-adrenal (HPA) axis. These are activated when a stressful stimulus is encountered and regulate the response. Initially, the release of CRF by the hypothalamus is stimulated. This precipitates the release of adrenocorticotropic hormone (ACTH) by the pituitary gland, which in turn stimulates the release of cortisol by the adrenal cortex (see Whitnall, 1993 for a review). Studies have shown CT can result in changes to CRF systems and HPA axis during development, altering their responses. These changes persist into adulthood and are not seen in adults without

CT (Heim et al., 2000; Shea, Walsh, MacMillan, & Steiner, 2005; Mello, Mello, Carpenter, & Price, 2003).

Adults with depression also show differences in HPA axis functioning (see Holsboer, 2000 for a review). There is evidence to suggest depression is associated with elevated cortisol levels, reduced feedback control, increased levels of corticotropin-releasing hormone (CRH) and alterations in response to CRH and ACTH (Gold, Drevets & Charney, 2002). However, findings are inconsistent. This may reflect clinical characteristic heterogeneity between studies, for example severity of illness (Appelhof et al., 2016), but may also reflect the dependence of differences in HPA axis dysregulation on whether people also have a CT history (Harkness, Stewart & Wynne-Edwards, 2011). Studies investigating the impact of CT on HPA axis functioning in adulthood have also been inconsistent. Both hyper and hypo HPA axis activation of basal activity and stress reactivity have been found (for a review see Baes, Tofoli, Martins, & Juruena, 2012). Taken together, these inconsistencies suggest it is necessary to consider the interaction of CT and depression in adulthood when investigating HPA axis dysregulation.

2.1 Measuring the HPA axis

There is a need to distinguish effects on basal activity and HPA axis reactivity. Studies have used both to assess the impact of CT and depression.

2.1.1 Basal activity

Basal activity has a circadian rhythm. Cortisol levels peak in the morning, and are at their lowest during night-time sleep (for a review see Walker, Terry & Lightman, 2010). Levels are typically measured in saliva or plasma at different time points, either across a day or a shorter space of time, and a mean calculated (e.g. Joyce et al., 2007). More recently analysis of cortisol accumulated in hair has been used to retrospectively assess longer term levels (Hinkleman et al., 2013). While some studies have found raised cortisol

levels with depression, this is not a consistent finding (see Steckler, Holsboer & Reul, 1999 for a review).

Basal HPA axis activity can also be reliably assessed using the cortisol awakening response (CAR; Mangold, Marino, & Javors, 2011). CAR is the acute response of cortisol on waking, peaking after 30-45 mins (Pruessner et al., 1997; Clow, Hucklebridge, Stalder, Evans & Thorn, 2010). Findings for depression are inconsistent. Some studies have found enhanced CAR, (Bhagwagar, Hafizi, & Cowen, 2003, Vreeburg et al., 2004), others blunted (Stetler & Miller, 2005). For CT in healthy controls, studies have shown enhanced CAR (Lu et al., 2013).

2.1.2 Reactivity

Impaired feedback inhibition is one way HPA axis reactivity is thought to be affected in depression (Pariante & Miller, 2001). Cortisol produced when the HPA axis is active mediates the system. There are two types of receptors, glucocorticoid receptors (GR) in the hypothalamus and pituitary gland, which are less sensitive to cortisol, and mineralcorticoid receptors (MR) in the hippocampus, which are more sensitive to cortisol. These 'turn off' the HPA axis when levels of circulating cortisol reach a certain level. The two types of receptors are believed to serve different functions. MR are believed to be involved in regulating daily circadian cortisol rhythms, and GR in the stress response (de Kloet, Vreugdenhil, Oitzl, & Joels, 1998; Young, Lopez, Murphy-Weinberg, Watson, & Akil, 1997). The balance between the two is therefore important for overall HPA axis functioning (de Kloet et al., 1998).

The dexamethasone suppression test (DST) was the first test of feedback inhibition (Carroll, Curtis, & Mendels, 1976). Failure to suppress endogenous cortisol release after administration of dexamethasone (a synthetic glucocorticoid) indicates reduced feedback sensitivity. Dexamethasone has a greater affinity for GR, and is believed to indicate

reduced GR sensitivity. People with depression show non-suppression with the DST, that is impaired feedback inhibition (e.g. Carroll et al., 1976). However, the DST has been criticised for not being sufficiently sensitive or specific, and only measuring GR activity (Pariante et al., 2002).

More recently, Pariante et al., (2002) developed a suppression test using prednisolone (PST). The PST is considered to be more sensitive as prednisolone mimics cortisol more accurately than dexamethasone, and is believed to detect reduced GR and MR sensitivity (Pariante et al., 2002). Studies using the PST with people with depression (Jurena et al., 2009; Jurena et al., 2010) have suggested no feedback impairment when compared with healthy controls. The authors suggest differences in findings versus the DST are due to MR compensating for altered GR function (Jurena et al., 2010).

HPA axis reactivity can also be measured by its response to stimulation with CRF (stimulating ACTH and subsequently cortisol) and ACTH (stimulating cortisol release directly).

Challenge tests are regarded as the most sensitive way of measuring HPA axis activity (Heuser, Yassoridis & Holsboer, 1994; Ising et al., 2005; Watson, Gallagher, Smith, Ferrier & Young, 2006). They combine stimulation with exogenous CRF and suppression using the DST (von Bardeleben & Holsboer, 1991). Dexamethasone is administered one night and CRF the next morning. After the injection of CRF, secretion of cortisol continues to be suppressed in healthy controls. Depressed patients show an 'escape' from suppression (Heim et al., 2008a) with elevated cortisol and ACTH concentrations. It has been suggested reduced GR sensitivity reduces suppression after dexamethasone is administered, increasing the effects of stimulation with exogenous CRF (von Bardeleben, & Holsboer, 1991; Heuser et al., 1994). This test has the same drawbacks as the DST in terms of only assessing GR function.

Stressor tests use a variety of paradigms to assess the HPA axis response to current stress. One of the best validated is the Trier Social Stressor Test (Kirschbaum, Pirke & Hellhammer, 1993). This uses a naturalistic situation (giving a speech or completing arithmetic tests in front of others) to induce similar neurobiological system changes as encountering real-life stressors. Healthy adults with a CT history have shown lower cortisol reactivity to the test (Carpenter et al., 2007).

2.2 PTSD: a potential moderator?

Depression and anxiety disorders, including PTSD, often occur together for people with a CT history (Saunders, Villeponteaux, Lipovsky, Kilpatrick & Veronen, 1992). CT can lead directly to PTSD, but also increases the risk of PTSD developing if trauma is experienced in adulthood (Bremner, Southwick, Johnson, Yehuda & Charney, 1993). PTSD has been found to be associated with changes to the HPA axis, including lower cortisol levels (Bierer et al., 2006; Carpenter et al., 2007; Heim, Newport, Bonsall, Miller & Nemeroff, 2001; Meewisse, Reitsma, de Vries, Gersons & Olff, 2007) and CAR (Yehuda, 2006; Rohleder, Joksimovic, Wold & Kirschbaum, 2004; Wessa, Rohleder & Kirschbaum, 2006). In contrast to depression, with PTSD supersuppression has been found, suggesting enhanced feedback sensitivity (Stein, Yehuda, Koverola & Hanna, 1997; Yehuda, Halligan & Bierer, 2002; Yehuda et al., 1993; Yehuda, Boisoneau, Lowy & Giller, 1995). An interaction of PTSD with CT has also been found in terms of increased cortisol levels in urine and feedback sensitivity (Lemieux & Coe, 1995; Stein et al., 1997).

2.3 Sex effects

Depression is twice as prevalent in women than men (Young, 1998; Weiss, Longhurst & Mazure, 1999). There may also be a stronger link with CT for women. One study found 69% of female participants with a CT history had depression in adulthood, compared to 27% of males (Carmen, Rieker, & Mills, 1984). Animal and human studies

suggest enhanced HPA axis stress reactivity in women (Ogilvie & Rivier, 1996; Handa, Burgess, Kerr & O'Keefe, 1994; Kant et al., 1983; Spinedi et al., 1994; Greenspan, Rowe, Maitland, McAloon-Dyke, & Elahi, 1993). A hypothesised link to the sex steroids (Burgess & Handa, 1992) requires further research (Heim, Newport, Mletzko, Miller & Nemeroff, 2008a). The research may suggest different neurobiological changes in response to CT for men and women, in part accounting for women's greater vulnerability to depression (Weiss et al., 1999).

2.4 Previous reviews

There have been several reviews in this area. Two recent ones looked more generally at CT's impact on the HPA axis and the relationship with later psychopathology, but neither included a systematic review of literature relating to depression (Struber, Struber & Roth, 2014; Nemeroff, 2016). Heim et al's (2008a) review looked at the relationship between CT, depression and HPA axis dysregulation, but restricted itself to the work of their own group. Baes et al. (2012) conducted a systematic review of GR and MR dysregulation studies only. They also included studies looking solely at people with depression and solely at healthy people with CT histories, as well as those investigating an interaction of the two. To the author's knowledge, there has been no systematic review looking at how the interaction of depression and CT affects measures of HPA axis functioning.

2.4.1 Aims of this review

The aim was to provide a systematic review of the literature investigating the relationship between CT, depression in adulthood and HPA axis functioning. The primary aim was to look at evidence for an interaction between CT and depression in adulthood, and its effect on basal cortisol levels and reactivity of the HPA axis. The secondary aim was

to consider evidence for potential moderating factors on any relationship, including sex, depression severity, types of trauma and co-morbid PTSD.

3 METHOD

3.1 Literature search and selection

A literature search was conducted using three electronic databases (Medline, PSYCHinfo and EMBASE) on 31st August 2016. An initial consideration of articles in this area suggested the terms for the systematic search (Table 1).

Table 1.

Literature search terms

HPA axis	History of childhood trauma	Depression
Cortisol	Child* abuse	Depress*
Hydrocortisol	Childhood trauma	MDD
Glucocorticoid*	Early life experiences	
Hypothalamic pituitary adrenal	Early life stress	

*truncated to allow for multiple endings of words

First duplicates were excluded. Abstracts were initially screened, then a full text search conducted for the remaining articles. Original research, published in English, involving human adults and a clinical group with current depression as the primary organising diagnosis, were included. Studies were only included if they investigated an interaction of the three key variables: cortisol or HPA axis activity; CT; and current depression status. Review papers were excluded but reference lists of ten relevant reviews identified were hand checked.

3.2 Assessment of methodological quality

Methodological quality was assessed using the Crowe Critical Appraisal Tool (CCAT; Version 1.4) (Crowe & Sheppard, 2011). The CCAT is a validated tool with good inter-rater reliability (.74; Crowe, Sheppard & Campbell, 2012). It provided a framework to assess the studies papers under eight headings, with a total score of 40 (five marks for each section). There were no norms for a 'good' score. For this review, a core percentage score was calculated from marks for Design, Sampling, Data Collection and Results sections.

3.3 Analysis

The qualitative analysis primarily focused on findings of the studies in terms of the relationship between depression, CT and measures of HPA axis activity/reactivity. Reference to comparator groups was made where possible, even where this was not the primary focus of the study. Any additional analysis included in the studies, for example levels of PTSD symptomology, was included only where this related to the review aims.

4 RESULTS

4.1 Study selection

The search identified 357 citations once duplicates were deleted (Figure 1). The review of titles and abstracts resulted in 30 full-text articles being retrieved. No additional articles were identified through the hand search of key review article references lists. Fifteen articles (published from 2000 to 2016) met the inclusion criteria and were included in the qualitative synthesis.

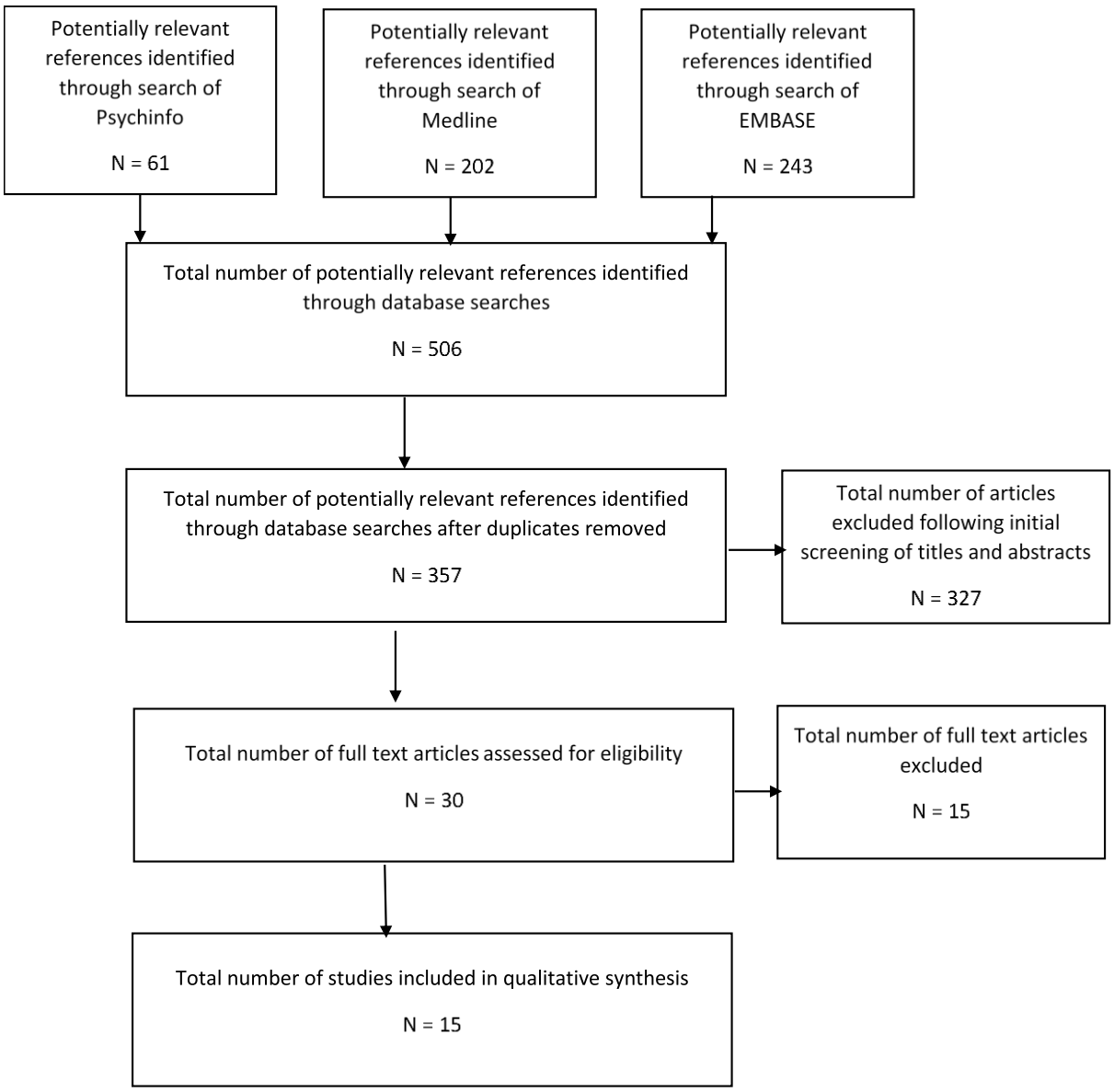


Figure 1: Flow chart for qualitative synthesis study selection

4.2 Study quality

The range of total CCAT scores was 50% to 75% (mean = 63%). For core CCAT scores, the range was 50% to 70% (mean = 62%; median = 60%). Table 2 gives the breakdown of each studies' core score, along with an assessment of strengths and weaknesses. The core scores were included in the qualitative synthesis summary (Table 4).

Table 2

Breakdown of core CCAT score with strengths and weaknesses by study

Study	Core score and breakdown		Strengths	Weaknesses
Hinkleman et al., 2013	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 3 3 **12 (60%)**	✓ 4 groups ✓ Male and female sample and sex included in analysis ✓ PTSD measured and used in analysis ✓ Short and long-term measures of cortisol	× No power analysis or effect sizes × No demographic breakdown × No analysis by depression severity × No analysis by CT subtype × Small CT only group
Joyce et al., 2007	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	2 3 4 3 **12 (60%)**	✓ Male and female sample and sex included in analysis ✓ Subtypes of CT measured and used in analysis ✓ Parental bonding measured and used in analysis ✓ Good sample size ✓ Depression severity included in analysis	× 2 groups – no CT only or HC × No power analysis or effect sizes × Unclear treatment of PTSD
Lopes et al., 2010	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 2 3 4 **12 (60%)**	✓ DHEAS measured as well as cortisol ✓ PTSD measured and used in analysis ✓ Depression severity included in analysis	× 3 groups – no CT only × Female only sample No power analysis or effect sizes × Relatively small sample size × No analysis by CT subtype
Peng et al., 2004	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 3 3 **12 (60%)**	✓ 4 groups ✓ Male and female sample ✓ Measured depression severity and used in analysis ✓ Excluded co-morbid PTSD ✓ Depression severity included in analysis	× No power analysis or effect sizes × No demographic breakdown × No sex analysis × No analysis by CT subtype
Lu, Goa, Huang, Li & Xu, 2016	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	4 3 4 4 **14 (70%)**	✓ 4 groups ✓ Males and female sample and sex included in analysis ✓ Excluded co-morbid PTSD ✓ Analysis by CT subtype	× No power analysis or effect sizes × Depression severity not included in analysis
Newport, Heim, Bonsall, Miller & Nemeroff, 2004	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 3 3 **12 (60%)**	✓ 4 groups ✓ PTSD measured and used in analysis ✓ Stress in adulthood measured	× No power analysis or effect size × Female only sample × Relatively small sample × Stress in adulthood not included in analysis × Depression severity not included in analysis
Baes, Martins, Tofoli & Jurena, 2014	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 2 4 3 **12 (60%)**	✓ Male and female ✓ Measured MR and GR activity separately ✓ Depression severity included in analysis	× No power analysis or effect sizes × No sex analysis × 3 groups – no CT only × Depression only group small × PTSD not included in analysis × Did not exclude bipolar
Jurena et al., 2009	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 3 3 **12 (60%)**	✓ Male and female ✓ Analysis by treatment response ✓ Used PST – more accurate measure ✓ Depression severity included in analysis	× No power analysis or effect sizes × Only 3 groups × No sex analysis × Small depression only group × Inpatient clinical population × PTSD present but not separately analysed

Study	Core score and breakdown		Strengths	Weaknesses
Heim et al., 2000	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 4 3 **13 (65%)**	✓ Power analysis included ✓ Validated stressor test used ✓ Stress in adulthood measured	× Female only × No effect sizes × Relatively small sample × Co-morbid PTSD present but not separately analysed × Depression severity and stress in adulthood not included in analysis × No demographic breakdown × Did not look at subtypes of abuse
Heim et al., 2002	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 4 4 **14 (70%)**	✓ Power analysis included ✓ Validated stressor test used ✓ PTSD, depression severity, stress in adulthood included in analysis	× Female only × No effect sizes × Relatively small sample × Did not look at subtypes of abuse
Brand et al., 2010	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	2 2 4 3 **11 (55%)**	✓ PTSD included as moderator ✓ Good overall sample size ✓ Measured stress in adulthood	× Female (mothers) only × Only two groups – no CT only or healthy control groups × No power analysis or effect sizes × Uneven group sizes × Non-standardised stressor test used × Did not look at abuse subtypes
Suzuki, Poon, Papdopoulos, Kumari & Cleaere, 2014	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 3 4 4 **14 (70%)**	✓ Male and female sample ✓ 4 groups ✓ Effect sizes given ✓ Specific trauma based stressor test ✓ Used CT subscales in analysis ✓ Inpatient and outpatient populations analysed separately ✓ Measured PTSD, depression severity and used as covariates	× No power analysis × Male sample relatively small × Not a validated stressor test × No sex analysis
Heim, Newport, Bonsall, Miller & Nemeroff, 2001	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	3 2 4 3 **12 (60%)**	✓ 4 groups ✓ Measured PTSD and current SLEs	× No power analysis or effect sizes × Female only sample × Relatively small sample × Did not analyse PTSD and SLEs separately × Did not look at subtypes of abuse × Depression severity not included in analysis
Heim, Mletzko, Purselle & Nemeroof, 2008b	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	4 3 4 3 **14 (70%)**	✓ 4 groups ✓ Power analysis included ✓ Measured PTSD and used as covariate in analysis ✓ CT subscales analysed ✓ Stress in adulthood measured	× Males only × Relatively small sample (depression only = 6) × No effect sizes × Depression severity and stress in adulthood not analysed
Watson et al., 2007	Design (/5) Sampling (/5) Data collection (/5) Results (/5) **Total (/20)**	2 2 3 3 **10 (50%)**	✓ Male and female sample ✓ Types and levels of CT analysed ✓ Limited sex analysis ✓ Measured family history, limited analysis	× 3 groups (depression, bipolar, HC) × Clinical group analysed together × Restricted separate depression group analysis × No power analysis or effect sizes × No assessment of PTSD × Control group not specific to study

4.3 Characteristics of the studies

4.3.1 Study samples

There were differences in study samples. In terms of sex: eight used men and women; six women only; and one men only. In terms of numbers of participant groups: nine had four groups (current depression and CT history; CT without current depression; current depression without a CT history; and healthy controls with no history of depression or CT); three had three groups (no CT only group); and two had two groups (depression groups only; see Table 4).

4.3.2 Trauma measures

Eight studies used the Childhood Trauma Questionnaire (CTQ; Pennebaker & Susman, 1988). Other trauma measures used were: Parental Bonding Instrument (PBI; Parker, Tupling, & Brown, 1979); Early Trauma Inventory (ETI; Bremner, Vermetten & Mazure, 2000), Childhood Experience of Care and Abuse Questionnaire (CECA-Q; Bifulco, Bernazzani, Moran & Jacobs, 2005). Four studies also used measures of current stress and life events: Life Events Survey (LES; Sarason, Johnson & Siegel, 1978); Perceived Stress Scale (PSS; Cohen, Kamarck & Mermelstein, 1983) and Daily Hassles Scale (DHS; Kanner, Coyne, Schaefer & Lazarus, 1981) (See Table 4).

4.3.3 Clinical measures

The most common measure of depression symptomology was the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960), used by 12 studies. Two used the Beck Depression Inventory (BDI-II; Beck, Steer & Brown, 1996) and one the Quick Inventory on Depression Symptomology (QIDS-SR; Rush et al., 2003). Three studies also measured PTSD symptomology using the PTSD Checklist Civilian Version (PCL-C; Blanchard, Jones-Alexander, Buckley & Forneris, 1996) and the Clinician Administered PTSD Scale (CAPS; Blake et al., 1995) (See Table 3).

4.3.4 *Measures of HPA axis*

Several different measures of the HPA axis were used, with the number of studies using each summarised in Figure 2. To facilitate the qualitative synthesis, studies were grouped according to the primary method of HPA axis assessment.

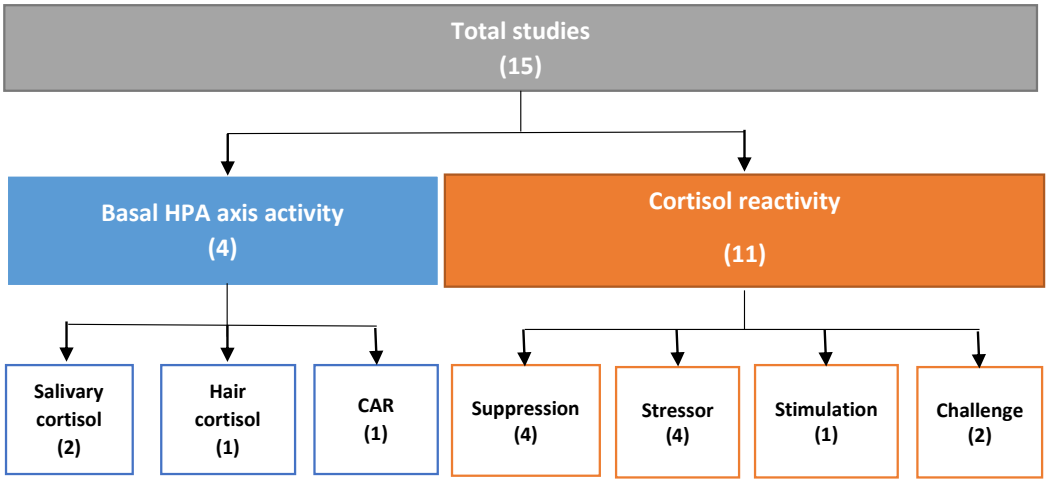


Figure 2: Grouping of studies for qualitative synthesis according to the primary method of assessing the HPA axis used

4.3.5 *PTSD and other co-morbidity*

Table 3 details whether samples included participants with co-morbid PTSD and other psychiatric disorders, if symptomology was measured and whether they were included in the analysis.

4.4 Synthesis

Table 4 provides a summary of the studies. The table and synthesis are grouped by the method used to measure the activity/reactivity of the HPA axis as detailed in Figure 2.

Table 3

Details of presence, measurement and inclusion in analysis of co-morbid PTSD and other psychiatric disorders.

Authors	Co-morbid PTSD	Other co-morbid disorders
Hinkleman et al., 2013	Present Included in analysis	Some excluded
Joyce et al., 2007	Not stated	Some excluded
Lopes et al., 2010	Present. Symptoms measured. Included in analysis	Axis I excluded
Peng et al., 2004	Excluded	Axis I and II excluded
Lu et al., 2016	Excluded	All other Axis I and II excluded
Newport et al., 2004	Present. Symptoms measured. Included in analysis	Some excluded (psychosis, bipolar disorder)
Baes et al., 2014	Present Not included in analysis	Psychotic symptoms excluded Other Axis I allowed but not separated out
Jurena et al., 2009	Present Only analysed as part of Axis I disorders, not separately	Some excluded Other Axis I and II allowed
Heim et al., 2000	Present No separate analysis	Psychosis, bipolar, eating disorders excluded
Heim et al., 2002	Present No separate analysis	Psychosis, bipolar, eating disorders excluded
Brand et al., 2010	Present Included in analysis	Psychosis, bipolar disorder excluded
Suzuki et al., 2014	Present Separated in analysis	Not excluded Anxiety measured
Heim et al., 2001	Present No separate analysis	Psychosis, bipolar, eating disorders excluded
Heim et al., 2008b	Present. Symptoms measured. Used as covariate	Psychosis, bipolar, eating disorders excluded
Watson et al., 2007	Not stated	Present – bipolar disorder group

Table 4

Summary of studies for qualitative synthesis

Authors	Aim	Sample	Age (M± SD; years)	Clinical measure	Trauma measure	Cortisol measure/paradigm	Statistical analysis	Core quality	Summary of main results
				1. BASAL	**ACTIVITY**				
Hinkleman et al., 2013	To examine retrospective cortisol levels according to the presence or absence of childhood trauma in patients with depression and healthy control subjects	N = 84 ♀ = 53 ♂ = 31 D+ = 24 D- = 19 CT = 7 HC = 34	41.7 ±10.5	HDRS	CTQ	Diurnal salivary cortisol (4 time points) Hair cortisol	Mixed ANOVA	60%	Hair and salivary cortisol: D+ = CT < HC = D-
Joyce et al., 2007	To examine whether childhood variables including physical and sexual abuse, influences plasma cortisol levels in major depression	N = 192 ♀ = 108 ♂ = 84 D+ = 34 D- = 155	31.8± 11.3	HDRS MADRS MSE	PBI	Mean afternoon saliva (5 time points)	One-way ANOVA Multiple linear regression	60%	Cortisol: D+ < D-
Lopes et al., 2010	To investigate the effects of a history or early life stress and PTSD symptoms on HPA axis function over the day for women with recurrent MDD matched with healthy controls	♀ = 53 D+ = 16 D- = 22 HC = 15	D+ 41.06 ±6.46 D- 39.73 ±10.12 HC 37.07 ±5,64	BDI PCL-C	CTQ	Diurnal salivary cortisol levels (3 measures) Salivary DHEAS	One-way ANOVA Post-hoc group analyses	60%	Cortisol and DHEAS: D+ = D- < HC
Peng et al., 2004	To explore the impact of childhood neglect on HPA axis functioning, dysfunctional attitudes and the severity of depression	N= 109 ♀ = 51 ♂ = 58 D+ = 28 D- = 30 CT = 22 HC = 29	D+ 28.87 ±6.28 D- 28.37 ±8.27 CT 28.37 ±5.28 HC 27.87 ±4.28	HDRS SSRS DAS	CTQ	CAR Δ salivary cortisol	Two-way ANOVA Post-hoc group analyses Correlational analysis	60%	Δ salivary cortisol: D+ = CT = HC > D-

Authors	Aim	Sample	Age (M± SD; years)	Clinical measure	Trauma measure	Cortisol measure/paradigm	Statistical analysis	Core quality	Summary of main results
			2.REACTIVITY						
Lu et al., 2016	To examine HPA axis activity with respect to histories of childhood trauma in both depression and normal controls	N=80 ♀ = 46 ♂ = 34 D+ = 18 D- = 17 CT = 22 HC = 23	D+ 23.70 ±4.13 D- 23.90 ±3.75 CT 21.50 ±3.91 HC 21.80 ±3.01	HDRS ZSRDS	CTQ	CAR: Δ salivary cortisol Suppression: DST (low dose) Suppression of salivary and plasma cortisol	Mixed repeated measures ANCOVA Correlational analysis	70%	CAR: D+ = CT D+ = D- CT > D- CT > HC DST: D+ < CT D+ < D- D+ < HC
Newport et al, 2004	To administer standard and low dose dexamethasone to abuse survivors with and without depression or PTSD to help elucidate the neurobiological mechanisms whereby childhood trauma results in a persistent vulnerability to a psychiatric disorder	♀ = 64 D+ = 16 D- = 10 CT = 19 HC = 19	D+ 32.40 ±7.80 D- 33.90 ±8.30 CT 33.30 ±6.40 HC 28.50 ±6.90	HDRS CAPS	ETI LES PSS	Suppression: DST (standard and low dose) Suppression of plasma ACTH and cortisol	Two-way ANCOVA with post hoc Tukey-Kramer pairwise comparisons	60%	Standard dose: D+ = D- = CT = HC Low dose: D+ > D- = CT = HC
Baes et al., 2014	To better understand the role of early life stress on MR function in depressive patients	N = 30 ♀ = 15 ♂ = 5 (data not available for control group) D+ = 13 D- = 7 HC = 10	D+ 39.50 ±2.70 D- 37.40 ±4.30 HC 29.40 ±8.00	HDRS	CTQ	CAR after suppression with: Placebo (P) Dexamethasone (D) Fludocortisone (F) Suppression of salivary cortisol	Mixed repeated measure one-way ANOVA with Tukey corrected post hoc tests Correlational analysis	60%	P, D, F: D+ = D- = HC D+ D>P, F>P, D>F D- D>P, P=F, D=F HC D>P, D>F, P=F

Authors	Aim	Sample	Age (M± SD; years)	Clinical measure	Trauma measure	Cortisol measure/paradigm	Statistical analysis	Core quality	Summary of main results
Jurena et al., 2009	To understand more about the role of the HPA axis in severe depression, specifically in the aetiology of treatment resistance	N = 91 ♀ = 51 ♂ = 58 D+ = 31 D- = 14 HC = 46	D+ 50.90 D- ±1.50 HC 46.50 ±2.10	RLCQ HDRS MADRS IDS BDI-II BAI	CECA	Suppression: PST Suppression of salivary cortisol	Mixed repeated measures ANOVA Correlational analysis	60%	Suppression: D+ = D- = HC
Heim et al., 2000	To determine whether early life stress results in a persistent sensitisation of the HPA axis to mild stress in adulthood, thereby contributing to vulnerability to psychopathological conditions	♀ = 49 D+ = 13 D- = 10 CT = 14 HC = 12	D+ 32 95% CI 27-36 D- 34.60 95% CI 29-41 CT 30 95% CI 27-33 HC 29 95% CI 24-34	HDRS	ETI LES DHS	Stressor: TSST Δ plasma ACTH and cortisol	Two-way repeated measure ANCOVA Post-hoc contrasts	65%	ACTH: D+ = CT > D- = HC Cortisol: D+ > CT = D- = HC Vs HC D+ ↑ ACTH ↑ Cortistol CT ↑ ACTH = Cortistol D- = ACTH = Cortistol
Heim et al., 2002	To evaluate the relative role of early adverse experience vs stress experience adulthood in the prediction of neuroendocrine stress reactivity in women	♀ = 49 D+ = 13 D- = 10 CT = 14 HC = 12	D+ 32 95% CI 27-36 D- 34.60 95% CI 29-41 CT 30 95% CI 27-33 HC 29 95% CI 24-34	HDRS	ETI LES DHS	Stressor: TSST Maximum plasma ACTH and cortisol	Multiple linear regression analyses	70%	**Predictors of maximum ACTH (35% of var):** Age (-) Childhood abuse history (+) No. childhood trauma experiences (-) No. adult trauma experiences (+) Depression symptoms (+) Childhood trauma x adult trauma (+) **Predictors of maximum cortisol (18% of var):** History of childhood abuse (+) No. childhood trauma experiences (-) Daily life hassles (-) Depression symptoms (+)

Authors	Aim	Sample	Age (M± SD; years)	Clinical measure	Trauma measure	Cortisol measure/paradigm	Statistical analysis	Core quality	Summary of main results
Brand et al., 2010	To investigate the association between maternal history of child abuse and maternal cortisol levels in a clinical sample of postpartum women and to explore whether depressive symptoms and stressful life events moderate this relationship	♀ = 126 D+ = 38 D- = 88	34±4.00	BDI-II	CTQ	Stressor: Specific infant/mother stressor paradigm Δ salivary cortisol	ANCOVA	55%	Baseline cortisol: D+ = D- Post stressor cortisol: D+ < D- Δ cortisol: D+ ↓ D- =
Suzuki et al., 2014	To examine stress reactivity in individuals with and without a history of childhood trauma and with or without depression by measuring cortisol responses to the passive viewing of stressful images, including images relevant to childhood trauma	N = 80 ♀ = 52 ♂ = 28 D+ = 20 D- = 18 CT = 17 HC = 24	D+ 52.10 ±12.00 D- 51.70 ±10.80 CT 44.30 ±12.50 HC 45.30 ±13.70	QIDS-SR	CTQ	Stressor: Specific psychosocial stressor test Δ salivary cortisol	One-way repeated measures ANOVA plus post hoc analysis Three-way mixed ANOVA Two-way ANCOVA	70%	Stress reactivity: D- > D+ = CT = HC
Heim et al., 2001	To evaluate pituitary-adrenal responses to standard HPA axis challenge tests in adult female survivors of childhood abuse with and without major depressive disorder	♀ = 66 D+ = 15 D- = 11 CT = 20 HC = 20	D+ 32.60 ±6.50 D- 34.50 ±6.90 CT 31.10 ±6.40 HC 27.80 ±6.80	HDRS	ETI LES DHS	Stimulation: CRF Δ plasma ACTH and cortisol	Three-way repeated measure ANOVA with post hoc contrasts	60%	CRF stimulation vs HC CT ↑ ACTH resp. D+ ↓ ACTH resp. D- ↓ ACTH resp. CT ↓ stim. cortisol levels D+ ↓ stim. cortisol levels D- = stim. cortisol levels

Authors	Aim	Sample	Age (M± SD; years))	Clinical measure	Trauma measure	Cortisol measure/paradigm	Statistical analysis	Core quality	Summary of main results
Heim et al., 2008b	To determine the impact of childhood trauma on dexamethasone/CRF test results in patients with major depressive disorder	= 49 D+ = 15 D- = 6 CT = 14 HC = 14	D+ 32.30 ±8.70 D- 30.00 ±9.50 CT 31.40 ±8.00 HC 29.10 ±9.00	HDRS CAPS	ETI LES	Challenge: DST/CRF stimulation Δ plasma ACTH and cortisol	One-way ANCOVA Post hoc analysis Two-way repeated measure ANCOVA Correlational analysis	70%	Cortisol and ACTH response: D+ > HC D+ > D-
Watson et al., 2007	To examine the relationship between early adverse life experience, family history and HPA axis function for patients with MDD and bipolar disorder	N = 58 = 11 = 19 (data not available for control group) Patients: MDD = 10 BD = 30 Divided into high (EN+) and low (EN-) emotional neglect HC = 28	MDD 51.90 ±9.50 BD 47.50 ±7.20	HDRS	CTQ	Challenge: DST/CRF stimulation Δ plasma ACTH and cortisol	Mann-Whitney U and Krushkal-Wallis H test Fishers Exact test	50%	Cortisol response: Total patients: EN- > EN+ = HC MDD: EN- = EN+ BD: EN- > EN+

Sample:
D+ Depression and history of childhood trauma
D- Depression only
CT History of childhood trauma only
HC Healthy controls

Results:
= no statistical difference
< statistically significantly lower
> statistically significantly higher
↑ statistically significant increase
↓ statistically significant decrease

Clinical measures:
HDRS = Hamilton Depression Rating Scale
MADRS = Montgomery Asberg Depression Rating Scale
MSE = Mental State Examination for Melancholia
BDI = Beck Depression Inventory
ZSRD = Zung Self-Rating Depression Scale
QIDS-SR = Inventory on Depression Symptomology
PCL-C = PTSD Checklist Civilian Version
CAPS = Clinician Administered PTSD Scale
Cortisol measures:
CAR = Cortisol Awakening Response
DHEAS = Dehydroepiandrosterone
PST = Prednisolone Suppression Test
DST = Dexamethasone Suppression Test
TSST = Trier Social Stressor Test

Trauma measures:
CTQ = Childhood Trauma Questionnaire
PBI = Parental Bonding Instrument
ETI = Early Trauma Inventory
LES = Life Events Survey
PSS = Perceived Stress Scale
CECA-Q = Childhood Experience of Care and Abuse Questionnaire
SSRS = Social Support Rating Scale
DAS = Dysfunctional Attitudes Scale
RLCQ = Recent Life Changes Questionnaire
DHS= Daily Hassles Scale

4.4.1 Basal activity

a) Cortisol level

Three studies looked solely at cortisol levels (Winkelman et al., 2013; Joyce et al., 2007; Lopes et al., 2010). Another two included data for baseline cortisol levels as part of reactivity studies (Heim et al., 2001; Brand et al., 2010). Overall, studies found a reduction in cortisol levels when groups were compared to healthy controls. However, they differed in whether this reduction was associated with CT or depression.

Both Hinkleman et al. (2013) and Joyce et al. (2007) found CT was associated with a reduction in cortisol relative to healthy controls, as did Heim et al. (2001). In Hinkleman et al. (2013), the same reduction was observed for both diurnal salivary cortisol and hair cortisol (allowing a retrospective analysis of accumulated cortisol levels over several months). There was no interaction between depression and CT and no effect of depression, an effect that held when two participants with co-morbid PTSD were removed.

Joyce et al. (2007) looked at mean afternoon cortisol levels for a sample with depression; there was no healthy control group or CT only group. Participants with depression and a CT history had reduced cortisol levels when compared to the depression only group. There was no effect of depression severity, which could be seen as supporting the above, but there was an association with a melancholic subtype (observable retardation and agitation assessed using the Mental State Examination for Melancholia; Parker & Hadzi-Pavlovic, 1996). An analysis by CT subtype and parental bonding showed this effect held for both childhood sexual and physical abuse, and maternal affectionless control (low maternal care plus high maternal protection). A multiple regression found maternal affectionless control, childhood physical abuse, childhood sexual abuse, melancholia and sex explained 22% of the variance in afternoon cortisol levels. Maternal affectionless control was the best predictor, followed by physical abuse, sexual abuse and

melancholia. Sex was not a significant predictor (Joyce et al., 2007). Heim et al. (2001) found lower baseline cortisol levels for groups with a CT history, whether or not they had current depression, supporting the above.

In contrast, Lopes et al. (2010) found depression was associated with reductions in both cortisol and dehydroepiandrosterone (DHEAS; an adrenal steroid that antagonises various stress-related effects of glucocorticoids on peripheral tissues and the hippocampus and is therefore thought to have protective biological actions). There were no additional effects of CT. Both depression and PTSD symptomology were negatively correlated with cortisol levels and DHEAS. Brand et al. (2010) also found, for mothers with a history of depression, no difference in baseline cortisol between those with and without a CT history, although there was no control group in this study. Maternal depression symptoms, stressful life events and a history of PTSD did not moderate the relationship between a CT history and baseline cortisol. However, it should be noted that the mothers' levels of current depression symptomology were not high, and a number (not specified) would not have met diagnostic criteria for current depression (Brand et al., 2010).

The effects of co-morbid PTSD were unclear. While Hinkleman et al. (2013) and Brand et al. (2010) found no effect of removing those with PTSD, Lopes et al. (2010) did find a correlation between PTSD symptoms and cortisol levels.

Methodological differences, in addition to cortisol measure used, made a comparison of these studies difficult, although overall each study scored 60% on the core CCAT quality assessment. Hinkleman et al. (2013) and Joyce et al. (2007) used male and female samples, while Lopes et al. (2010), Heim et al. (2001) and Brand et al. (2010) used female participants only. It is worth noting, however, neither Hinkleman et al. (2013) nor Joyce et al. (2007) found a sex effect on cortisol levels. While Hinkleman et al. (2013), Lopes et al. (2010) and Brand et al. (2010) assessed additional effects of PTSD, PTSD was

not included in the analysis or participant description in Joyce et al. (2007). Heim et al. (2001) provided information about numbers in each group meeting PTSD diagnostic criteria (14/15 in the group with depression and CT history versus 4/11 for the CT only group), but performed no subanalysis.

Cortisol level summary

- Overall reduction in cortisol in trauma and clinical groups versus healthy controls
- Studies differed in whether this reduction was associated with CT or depression
- Unclear effect of co-morbid PTSD
- Studies differed in sex of samples and inclusion of co-morbid PTSD

b) Cortisol Awakening Response (CAR)

One study solely assessed basal activity using the CAR (Peng et al., 2014) and two in conjunction with other tests (Baes et al., 2014; Lu et al., 2016). Peng et al. (2014) and Lu et al. (2016) found CAR hyperreactivity was associated with CT but not depression when groups with a CT history (with or without depression) were compared to healthy controls and groups with depression only. Hyperreactivity increased with CTQ scores (Peng et al., 2014; Lu et al., 2016), but not depression severity (Peng et al., 2014). The depression only group showed hyporeactivity when compared to healthy controls, consistent with previous studies into effects of depression on the CAR (Gold et al., 2002). The sample sizes of both studies precluded analysis by trauma type. Both specifically excluded participants with co-morbid PTSD, as well as all other Axis I and II co-morbid disorders.

In contrast, Baes et al., (2014) found no difference in the CAR between depression groups, with or without a CT history, and healthy controls. However, the study did not include a group with CT only and did not appear to exclude or measure PTSD symptomology, which could have been hypothesised to contribute to the difference in results.

All three studies used mixed sex samples but only Lu et al. (2016) conducted an analysis by sex and found no effect. This also contributed to a higher core CCAT score of 70% for Lu et al. (2016), versus 60% for both Peng et al. (2014) and Baes et al. (2014).

CAR summary

- Inconsistent results
- Two studies found CAR hyperactivity associated with childhood trauma and not depression. Both excluded co-morbid PTSD
- One study found no difference in CAR between depression groups and healthy controls but did not have a childhood trauma only group and handling of co-morbid PTSD unclear

4.4.2 Reactivity

Eleven studies looked at HPA axis reactivity and the synthesis was divided by the measurement method used (Figure 2).

a) Suppression

Four studies assessed HPA axis feedback regulation using suppression paradigms. Different paradigms were used to look at differences in suppression of endogenous cortisol release following administration of agonists of GR and/or MR. There were inconsistencies in differences in suppression found between groups. Co-morbid PTSD and depression severity appear to be important factors in accounting for these inconsistencies.

GR agonist

Findings of the two studies using the DST, which uses a GR agonist, appeared to suggest its effects depended on the presence of co-morbid PTSD. Groups with depression and a CT history but no PTSD indicated reduced suppression, that is a reduction in glucocorticoid feedback inhibition when compared to healthy controls, groups with depression only, and groups with a CT history only (Lu et al., 2016). There were no sex

effects in the mixed sample. All female groups with depression and a CT history plus PTSD showed an increase in suppression, when compared to the same three groups (Newport et al., 2004). Indeed, the pattern of results stayed the same when PTSD was used as the organising diagnosis, rather than depression (Newport et al., 2004), consistent with research looking at PTSD and suppression (e.g. Stein et al., 1997). Where there was a CT history but no current depression, there was no evidence of altered negative feedback versus healthy controls (Newport et al., 2004). The Lu et al. (2016) study received a higher CCAT core score (70% v 60%) as it used a mixed sample and included sex in the analysis, and for its clear exclusion of co-morbid PTSD.

MR & GR agonists

The two studies looking at suppression of both MR and GR used different paradigms and found contradictory patterns of results. Baes et al. (2014) investigated MR and GR suppression separately by assessing the effect on the CAR after administration of dexamethasone (the GR agonist used in the DST), fludrocortisone (a MR agonist) and a placebo. The group with depression and a CT history showed suppression with both GR and MR agonists. Higher CTQ scores and depression severity were associated with increased MR malfunction. Healthy controls and the group with depression only showed suppression with the GR agonist only. This suggested CT was associated with an impairment of MR function. It is worth noting there was no group with a CT history only. Levels of co-morbid PTSD in depression samples were also unknown.

Jurena et al. (2009) used the PST. Prednisolone is thought to be a MR and GR agonist and the PST therefore tests the effect of suppression of both. The study found no difference in suppression between those with depression, with or without a CT history (including when separated into emotional, physical and sexual abuse), and healthy controls. This was in contrast with findings from Baes et. al. (2014), and DST studies

(Newport et al., 2004; Lu et al., 2016). Jurena et al. (2009) hypothesised intact negative feedback was due to MR compensating for GR impairment. They also noted tests, such as used by Baes et al. (2014), could not purely suppress GR or MR. Jurena et al. (2009) used an inpatient population, assessed as being moderately treatment resistant, who may have had more severe depression than other studies using outpatient samples. The study investigated a specific intervention for people with treatment resistant depression. The sample was divided into responders and non-responders to the intervention based on changes in depression symptomology. Non-responders had impaired suppression compared to responders and the authors suggested there may be a subgroup within depression with impaired suppression and this may affect treatment response. However, differences in CT history between responders and non-responders were not investigated. One in nine of Jurena et al.'s (2009) sample with depression had co-morbid PTSD, but no separate analysis was undertaken.

Given the suggested impact of PTSD on findings using a GR agonist alone (Lu et al., 2016; Newport et al., 2004), the lack of information about, and analysis of, co-morbid PTSD within the samples of both Jurena et al. (2009) and Baes et al. (2014) meant this may have contributed to the apparently contradictory results. Both studies received core CCAT scores of 60%. While both studies used mixed samples, neither included sex in their analysis. Neither included a CT only group, again hindering a comparison of results.

Suppression summary

- Inconsistent methodologies and results
- For GR, those with depression, CT and PTSD showed enhanced reactivity, in line with PTSD findings. With no co-morbid PTSD, reduced reactivity was found, in line with depression studies
- For MR, findings differed by whether MR suppression was measured separately (impairment found) or with GR (no impairment found)

b) Stressors

There were inconsistent findings in studies using paradigms designed to assess HPA axis reactivity to a stressor. However, different paradigms were used and Suzuki et al., (2014) suggested these may be tapping into different levels of the HPA axis. The Heim studies (2000; 2002) used a validated stressor test, the TSST (Kirschbaum et al., 1993). Brand et al. (2010) used a specially developed mother/infant paradigm, and Suzuki et al. (2014) also used a specially developed paradigm, which included viewing images linked to CT as the stressor.

The two Heim studies (Heim et al., 2000; Heim et al., 2002) used the same data and found a CT history was associated with an increase in ACTH (precursor of cortisol), that is an enhanced response to a stressor, when compared to healthy controls. For those with a CT history and depression, increases in both cortisol and ACTH were found post-stressor. Those with a CT history only showed an enhanced ACTH response to the stressor, but a cortisol response similar to healthy controls. For those with depression only, a normal stress response was observed. Both Brand et al. (2010) and Suzuki et al. (2014) found the opposite; a CT history was associated with reduced cortisol post-stressor. For Suzuki et al. (2014), this held whether or not participants had current depression and the depression only group showed the highest reactivity. However, Brand et al. (2010) found depressive symptoms, a history of PTSD or current stressful life events moderated the relationship between CT history and stress induced cortisol change, all with a small to medium effect size. When plotted against CT history (presence or absence), those with high depressive symptoms, a history of PTSD or current stressful life events had higher cortisol levels at the different time points than those without a history of trauma.

Heim et al. (2002) investigated predictors of maximum cortisol and ACTH using data from Heim et al. (2000). For cortisol and ACTH, a CT history was the strongest positive predictor, while number of CT experiences was a negative predictor. An interaction term of CT with trauma in adulthood was the best ACTH predictor but was not a predictor of maximum cortisol levels. For cortisol, but not ACTH, daily life hassles (negatively) and depression symptoms (positively) were predictors. PTSD and current stressful life events were not predictors of either maximum ACTH or cortisol levels.

In all three studies, some participants with depression had co-morbid PTSD. Brand et al. (2010) specifically looked at the effect of PTSD on stress reactivity, while Heim et al. (2002) used it as a predictor of maximum levels. The two studies found contrasting PTSD influences. It is also worth noting that, while the mothers included in the Brand et al. (2010) mother and infant dyad study had a history of depression, their current depression symptomology was low, and the authors acknowledged a number (not specified) would not have met criteria for current depression diagnosis. While Brand et al. (2010) and the two Heim et al. (2000; 2002) studies used female participants, Suzuki et al. (2014) used a mixed sample, but did not include sex in the analysis. The specialist nature of the sample and stressor test used contributed to Brand et al. (2010) receiving the lowest quality rating of these studies at 55%. The quality of the other studies was broadly similar (Table 2).

Stressors summary

- Different paradigms used with inconsistent findings
- Two studies found an enhanced stress response with childhood trauma, with or without depression
- Two studies found a blunted stress response with childhood trauma, but differed in the additional influence of depression
- Inconsistent influence of PTSD, either enhancing stress response or not acting as a predictor

c) Stimulation

Heim et al. (2001) looked at the effects of CRF and ACTH stimulation. In their female sample, there were differences between those with a CT history with and without depression and between stimulation with CRF and ACTH. When stimulated with exogenous CRF, those with a CT history only showed an enhanced ACTH response compared with healthy controls. Those with depression, either only or in addition to a CT history, showed a blunted response. In terms of cortisol levels following both CRF and ACTH stimulation, both groups with CT history showed lower baseline and stimulated cortisol levels, compared with healthy controls.

The group with depression and CT history had significantly higher levels of co-morbid PTSD (14 out of 15), than CT only (four out of 11). The depression plus CT group also had significantly higher mean negative life events and daily hassles scores than CT only, suggesting higher current and chronic stress. However, no separate analysis was conducted, meaning their potential role as confounds is unknown, which contributed to the core quality score of 60%.

Stimulation summary

- One study only
- Different responses to stimulation found between those with a CT history with, and without, current depression
- Different responses to stimulation with CRF and ACTH.

d) Challenge

Two studies used challenge tests, combining stimulation and suppression tests (Heim et al., 2008b and Watson et al., 2007). While overall the studies found an enhanced cortisol response associated with CT, findings differed depending on the addition of depression, PTSD and type and level of CT. Using a male sample, Heim et al., (2008b) found

CT history was associated with enhanced ACTH and cortisol responses to both stimulation and suppression. There were no differences between those with depression only and healthy controls, but also no difference between healthy controls and those with CT history only. Only those with a CT history and current depression showed changes in response to stimulation and suppression. When this study used PTSD symptoms as a covariate, the HPA axis's response was reduced. There was also a positive correlation between cortisol and ACTH responses and overall CT scores, sexual and physical abuse subscales, abuse severity, CT duration and age of onset.

The Watson et al., (2007) study compared responses to challenge tests of those with depression with a group with bipolar disorder, excluded by most other studies, plus healthy controls. The same enhanced response as in Heim et al. (2008b) was found for the clinical sample (i.e. depression and bipolar groups together) with a CT history when compared to healthy controls. However, this was not consistent across the different types and levels (high and low) of each CT type. When compared to healthy controls, the clinical group showed an enhanced response on emotional abuse, sexual abuse and physical neglect subscales for both high and low trauma groups. However, the response was only higher than the control group for those in the low trauma clinical group on physical abuse and emotional neglect subscales. There was no difference in cortisol responses for those with high levels of physical abuse and emotional neglect when compared to healthy controls, counter to the original hypotheses. When high and low levels of the subscales were compared, the only significant difference in cortisol response was for emotional neglect. The clinical group was divided by diagnosis for this subscale. There was no difference in cortisol response between those with depression with high and low levels of emotional neglect. For those with bipolar, the cortisol response was lower for those with high emotional neglect than low emotional neglect.

There were significant methodological limitations to the Watson et al. (2007) study, which reduced its usefulness for this analysis, and resulted in it receiving the lowest core quality score of 50%. Co-morbid PTSD did not appear to be formally assessed. The group with depression only had ten participants. The control group data was taken from published CTQ data and not a group recruited specifically. Analysis of the depression group (as opposed to total clinical group) was limited, as was sex analysis of the mixed sample.

Challenge summary

- Overall enhanced cortisol response with childhood trauma
- Inconsistent findings in relation to additional influence of depression, PTSD, type and level of trauma
- One study included participants with bipolar disorder in clinical group

5 DISCUSSION

Fifteen studies were reviewed to investigate links between depression, a CT history and the HPA axis. This section will first summarise the main findings before considering evidence relating to potential moderating factors, limitations of the studies, clinical implications of findings and making some suggestions for future research.

5.1 Summary of main findings

It was difficult to discern from the studies to date whether CT, depression or an interaction of the two were associated with changes in HPA axis, the review's primary aim. Comparability was hindered by methodological variation in the studies, a lack of replication, and effect sizes only being reported in one study (Brand et al., 2007).

5.1.1 Reduction in baseline cortisol activity

Studies reliably showed a reduction in baseline cortisol when compared to healthy controls. However, there were inconsistent findings as to whether this was driven by CT

(Hinkleman et al., 2013, Joyce et al., 2007, Heim et al., 2001) or depression (Lopes et al., 2010, Brand et al., 2010). None of the studies found an interaction between CT history and depression though, suggesting the effects of the two factors on cortisol levels are independent.

5.1.2 Enhanced reactivity

There was evidence of enhanced HPA axis reactivity in terms of CAR (Peng et al., 2014, Lu et al., 2016) and response to: suppression (Newport et al., 2004); a stressor (Heim et al., 2000; Heim et al., 2002); stimulation (Heim et al., 2001); and challenge tests (Heim et al., 2008b; Watson et al., 2007). But again, studies differed in terms of whether depression, CT or an interaction of the two was driving this enhancement. For the CAR, hyperreactivity was associated with CT, rather than depression (Peng et al., 2014, Lu et al., 2016). With suppression and challenge tests, increased suppression was associated with the interaction of CT history and depression (Newport et al., 2004; Heim et al., 2008b). For both stressor (Heim et al., 2000; Heim et al., 2002) and stimulation tests (Heim et al., 2001), effects differed depending on whether reactivity of ACTH (precursor of cortisol) or cortisol was investigated, which suggested the effects on these two products of HPA axis activity may be different.

5.1.3 Blunted reactivity or no difference

However, there was also evidence of a blunted HPA axis response to: suppression (Lu et al., 2016); a stressor (Suzuki et al., 2014 et al; Brand et al., 2010); and stimulation (Heim et al., 2001). Evidence from Lu et al. (2016), Brand et al. (2010) and Heim et al. (2001) suggested it was the interaction between depression and CT history that accounted for this hyporeactivity of cortisol levels. For ACTH reactivity to CRF stimulation, results were inconsistent, with Suzuki et al. (2014) finding the effect related to CT history alone and Heim et al. (2001) to depression, with or without CT. Furthermore, some studies

showed no difference with healthy controls (CAR: Baes et al., 2014; suppression: Jurena et al., 2009).

5.1.4 Differences between depression with and without CT history

It was also difficult to draw any firm conclusions about whether the HPA axis functioning of those with depression and CT history differed from those with depression only. There was some evidence for a difference in: cortisol levels (Hinkleman et al., 2013, Joyce et al., 2007); CAR (Peng et al., 2014, Lu et al., 2016); levels of suppression (Lu et al., 2016; Baes et al., 2014); responses to a stressor (Heim et al., 2000; Suzuki et al., 2014); and to challenge tests (Heim et al., 2008b). But it was not clear whether addition of a CT history to a diagnosis of depression was associated with enhanced (Peng et al., 2014; Lu et al., 2016; Heim et al 2000; Heim et al., 2002) or blunted (Hinkleman et al., 2013; Joyce et al., 2007; Lu et al., 2016; Suzuki et al., 2014) HPA axis functioning. Furthermore, other studies found no difference between these groups for cortisol levels (Lopes et al., 2010; Brand et al., 2010), response to suppression (Baes et al., 2014) or stimulation (Heim et al., 2001).

5.2 Potential moderating factors

The review's secondary aim was to consider potential moderating factors on the relationship between CT, depression and HPA axis functioning. This was hindered by heterogeneity in how far study design and sample sizes facilitated an investigation of these factors.

5.2.1 PTSD

PTSD has been hypothesised to be a moderating factor (de Kloet et al., 2007). There was some evidence co-morbid PTSD moderated the effects of depression and CT on the functioning of the HPA axis. Lopes et al. (2010) found a negative correlation between PTSD symptomology and cortisol levels, in line with evidence of lower cortisol levels with

PTSD (Yehuda, 1998). For the CAR, if PTSD was excluded, a hyper-reactivity with a CT history was found (Peng et al., 2014; Lu et al., 2016). If it was not, there was no effect (Baes et al., 2014). Taken together these results could suggest a moderating effect, given evidence for reduced CAR with PTSD (Lemieux & Cole, 1995; Stein et al., 1997). Suppression by a GR agonist was increased versus healthy controls when co-morbid PTSD was included (Newport et al., 2004), but reduced when it was specifically excluded (Lu et al., 2016). PTSD symptoms reduced the enhanced response to both a stressor (Brand et al., 2010) and a challenge paradigm (Heim et al., 2008b). These results accord with findings of supersuppression in PTSD, suggesting increased glucocorticoid feedback sensitivity (e.g. Stein et al., 1997).

However, not all studies found an effect of PTSD. Hinkleman et al. (2013) and Brand et al. (2010) found no effect of PTSD on cortisol levels. PTSD symptoms were not a predictor of maximum cortisol or ACTH responses to a stressor test (Heim et al., 2002). Additionally, many studies were either unclear about whether or how co-morbid PTSD was assessed, or did not specifically include PTSD in their analysis, meaning the influence of PTSD on these findings was unknown. It was therefore, once again, difficult to be anything but highly tentative about any conclusions about the influence of PTSD.

5.2.2 Depression type and severity

Three studies suggested depression type or severity may also influence the impact on HPA axis functioning, and differences in depression severity between studies may help to account for inconsistent results. As depression severity increased: cortisol levels decreased (Lopes et al., 2010); MR malfunction increased (Baes et al., 2014); and cortisol levels following a stressor did not fall as much (Brand et al., 2010). However, three other studies found no effect of depression severity on cortisol levels (Joyce et al., 2007; Brand et al., 2010) nor the CAR (Peng et al., 2014). The remaining nine did not look at effects of

depression severity. Joyce et al. (2007) did find a small amount of variance in a multiple regression predicting cortisol levels was explained by current melancholic symptoms, in line with evidence that melancholic depression was associated with a hyperactive, rather than hypoactive, HPA axis (Gold & Chrousos, 2002). Samples also differed in terms of depression history. Interestingly, with a sample of inpatients with severe current symptoms, Jurena et al. (2009) found no differences between healthy controls when the PST was used as a test of suppression. Once again, heterogeneity in samples and analyses meant it was difficult to draw firm conclusions.

5.2.3 Type of trauma

The effects of different types of trauma on functioning of the HPA axis was also not clear from the research to date as only five studies investigated these. For many of the others, sample sizes precluded this level of analysis. While Watson et al. (2007) looked at types of trauma, this was for the whole clinical group and therefore included people with bipolar disorder as well as (a smaller number) with unipolar depression. Heim et al. (2008b) found a positive correlation between cortisol responses to a CRF challenge test and sexual and physical abuse scales. Jurena et al. (2009) divided the sample by scores on emotional, physical and sexual abuse subscales and still found no difference between PST responses of healthy controls and those with CT histories. The impact of CT on the functioning of the HPA axis could also be hypothesised to be influenced by severity, duration, number of traumatic experiences as well as the age at which abuse started. Only two studies looked at these: Heim et al. (2002) found the number of CT experiences was a negative predictor of maximum cortisol and ACTH levels after a stressor; for Heim et al. (2008b), severity of abuse, duration and age of onset were positively correlated with responses to the challenge tests used.

5.2.4 Stress in adulthood

The rationale for these studies was based on the hypothesis that trauma in childhood leads to changes in the HPA axis and its functioning persisting into adulthood. Stress and trauma in adulthood also lead to changes in the HPA axis (See Lupien, McEwen, Gunnar, & Heim, 2009 for a review). Therefore, trauma in adulthood, as well as stressful life events, has been hypothesised as moderating the relationship, either in addition to, or separate from CT (Young and Breslau, 2004). It may have acted as an additional confound in the studies.

Three specifically included measures of adulthood stress. Once again, results were mixed. Some found an effect (Brand et al., 2010 found a reduction in cortical response to stressors; in Heim et al., 2002, daily life hassles were significant negative predictors of maximum cortisol levels following a stressor), and others no effect (Brand et al., 2010 on cortisol levels; in Heim et al, 2002, stressful life events were not a significant predictor of maximum cortisol levels following a stressor). While Heim et al. (2001) measured negative life events and daily hassles, there was no specific analysis using these variables. Once again, uncertainty around the influence of this potential confound remained.

5.2.5 Sex

It has been hypothesised that HPA axis changes resulting from CT may be different in males and females (Weiss et al., 1999). Eight studies in this review used mixed samples, six female only, and one male only. Four of the mixed sample studies included sex in the analysis. None of these found any effect of sex on: cortisol levels (Hinkleman et al., 2013; Joyce et al., 2007); CAR (Peng et al., 2004; Lu et al., 2016) and suppression using the DST (Lu et al., 2016). This may have suggested sex was not an important confound. However, the limited number of studies considering sex differences meant this suggestion was tentative.

5.3 Limitations

Lack of replication was the main limitation of the studies. Heterogeneity in sample characteristics and methodology greatly hindered comparison, making it difficult to say with any confidence what the effects were, and only one study reported effect sizes (Brand et al., 2010). Some of this heterogeneity in methods may be due to the development of more effective and accurate ways of measuring HPA axis responsiveness. For example, introduction of the PST meant both MR and GR suppression could be investigated, an advantage over the DST, but greatly hindered a comparison with previous studies using the DST. Differences in measures also precluded any form of meta-analysis, which would have been useful given the sample sizes of the individual studies.

Only three studies (Heim et al., 2000; Heim et al., 2002; Heim et al., 2008b) included a power analysis. With the other 12, it was therefore difficult to know whether null effects were in fact due to a lack of power. Sample sizes were small, precluding the subanalyses necessary to allow a comprehensive investigation of potential confounds and moderators, meaning the influence of these remained unclear. Nine studies had four groups: depression plus CT history, depression only, CT history, healthy controls. This was the most effective structure for looking at group differences but not all studies had this structure, which again made it difficult to compare results and effects.

Studies were also inconsistent in the potential confounding or moderating factors included in their analyses. There are, however, numerous other factors influencing functioning of the HPA axis that could be included in future studies. Joyce et al.'s (2007) multiple regression found maternal affectionless control, childhood physical abuse, childhood sexual abuse, melancholia and sex only explained 22% of variance in afternoon

cortisol levels. Factors such as family environment at the time of the trauma, attachment and other pathology, might be hypothesised to contribute to the other 78%.

All studies were cross sectional, so it was only possible to look at associations and not causality. There is a great need for longitudinal studies in this area, tracking those who have suffered abuse in childhood into adulthood.

One final limitation was whether a statistically significant difference between groups is clinically significant. For example, while there was a statistically significant difference in the CAR between groups, it is not clear whether this would this noticeably change an individual's presentation.

5.4 Clinical implications

The lack of confidence about the effects of depression and CT on the HPA axis of this review was frustrating from a clinical point of view. As outlined above, there was a tentative suggestion the stress response of the neuroendocrine system of someone with depression and CT history may be different to someone with depression without a CT history, even if the precise nature of those effects is still unclear. Heterogeneity in treatment response of people with depression, both to medication and talking therapies, is well documented (for a review see Nemeroff, 2007). A CT history may be one factor accounting for this heterogeneity and may also suggest a need to include CT history when assessing depression. It could potentially suggest a subtype of depression, emerging as a side-effect of chronic trauma reactions following CT and the effect of these on the HPA axis. This subtype of depression may require a different emphasis within interventions, for example incorporating elements of self-compassion, which has been found to be effective for adults with PTSD reactions to childhood trauma (Lee & James, 2013). The differences found by Jurena et al. (2009) in suppression between responders and non-responders to

treatment may also support this idea, although it was not clear how many non-responders had a CT history.

5.5 Suggestions for further research

There is a clear need for replication using consistent study structures and methodologies in this area, especially given the potential clinical implications. These study protocols should facilitate a clear contrast between those with depression and co-morbid PTSD, and without co-morbid PTSD. Depression severity and current life stressors should also be routinely measured so their influence can be more thoroughly and consistently assessed. It would be optimal for sample sizes of all future studies to be sufficiently large to allow an analysis by type of trauma and sex so these too can be more effectively assessed. It would also be desirable for effect sizes to be more routinely reported to facilitate a comparison of results.

There is some suggestion the research in this area has moved on, with many now looking for epigenetic explanations for changes (for a review see Silberman, Acosta & Zubilete, 2016), and also for resilience in those who experience CT but do not develop depression. It is recommended these studies are structured so they can also contribute to answering of some of the more fundamental questions around the nature of these changes.

Findings also suggest future studies in other areas using depression groups should screen for CT and perform subanalyses on groups with and without CT to explore differences in effects between groups. The differences found between groups with depression with and without a CT history may be found in other areas, particularly where to date research using a depression group has shown equivocal results.

5.6 Conclusion

This review suggested a CT history and depression were associated with changes in the HPA axis, both in terms of basal activity and reactivity. However, it was difficult to say with any certainty what these effects are and whether any effects are driven by a CT history, current depression or an interaction between the two. It was also unclear what role sex, co-morbid PTSD, depression severity, current stressors and a whole host of other factors played in moderating this relationship. This was due to heterogeneity in samples and methodology and a more general lack of replication.

These are important questions to answer. Some of the results could be taken to suggest a specific subtype of depression, associated with CT, perhaps with its genesis in untreated reactions to the trauma. This form of depression may be characterised by different dysregulation of the HPA axis and therefore reactions to current stressors than depression without a history of CT, and may require different treatment approaches.

6 REFERENCES

Appelhof, B. C., Huyser, J., Verweij, M., Brouwer, J. P., van Dyck, R., Fliers, E., ... & Schene, A. H. (2006). Glucocorticoids and relapse of major depression (dexamethasone/corticotropin-releasing hormone test in relation to relapse of major depression). *Biological Psychiatry*, *59*(8), 696-701.

Baes, C. W., Martins, C., Tofoli, S., & Juruena, M. F. (2014). Early life stress in depressive patients: HPA axis response to GR and MR agonist. *Frontiers in Psychiatry*, *5*(2).

Baes, C.W., Tofoli, S. M., Martins, C. M. S., & Juruena, M. F. (2012). Assessment of the hypothalamic–pituitary–adrenal axis activity: glucocorticoid receptor and mineralocorticoid receptor function in depression with early life stress–a systematic review. *Acta Neuropsychiatrica*, *24*(1), 4-15

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

Bhagwagar, Z., Hafizi, S., & Cowen, P. J. (2003). Increase in concentration of waking salivary cortisol in recovered patients with depression. *American Journal of Psychiatry*, *160*(10), 1890-1891.

Bierer, L. M., Tischler, L., Labinsky, E., Cahill, S., Foa, E., & Yehuda, R. (2006). Clinical correlates of 24-H cortisol and norepinephrine excretion among subjects seeking treatment following the World Trade Center Attacks on 9/11. *Annals of the New York Academy of Sciences*, *1071*(1), 514-520.

Bifulco, A., Bernazzani, O., Moran, P. M., & Jacobs, C. (2005). The childhood experience of care and abuse questionnaire (CECA. Q): validation in a community series. *British Journal of Clinical Psychology*, *44*(4), 563-581.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, *8*(1), 75-90.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*(8), 669-673.

Brand, S. R., Brennan, P. A., Newport, D. J., Smith, A. K., Weiss, T., & Stowe, Z. N. (2010). The impact of maternal childhood abuse on maternal and infant HPA axis function in the postpartum period. *Psychoneuroendocrinology*, *35*(5), 686-693.

Bremner, J. D., Southwick, S. M., Johnson, D. R., Yehuda, R., & Charney, D. S. (1993). Childhood physical abuse and combat-related posttraumatic stress disorder in Vietnam veterans. *The American Journal of Psychiatry*, *150*(2), 235.

Bremner, J. D., Vermetten, E., & Mazure, C. M. (2000). Development and preliminary psychometric properties of an instrument for the measurement of childhood trauma: The Early Trauma Inventory. *Depression and Anxiety*, *12*(1), 1-12.

Burgess, L. H., & Handa, R. J. (1992). Chronic estrogen-induced alterations in adrenocorticotropin and corticosterone secretion, and glucocorticoid receptor-mediated functions in female rats. *Endocrinology*, *131*(3), 1261-1269.

Carmen, E. H., Rieker, P. P., & Mills, T. (1984). Victims of violence and psychiatric illness. In *The Gender Gap in Psychotherapy* (pp. 199-211). Boston, MA: Springer US.

Carpenter, L. L., Carvalho, J. P., Tyrka, A. R., Wier, L. M., Mello, A. F., Mello, M. F., ... & Price, L. H. (2007). Decreased adrenocorticotropic hormone and cortisol responses to stress in healthy adults reporting significant childhood maltreatment. *Biological Psychiatry*, *62*(10), 1080-1087.

Carroll, B. J., Curtis, G. C., & Mendels, J. (1976). Neuroendocrine regulation in depression: II. Discrimination of depressed from nondepressed patients. *Archives of General Psychiatry*, *33*(9), 1051-1058.

Clow, A., Hucklebridge, F., Stalder, T., Evans, P., & Thorn, L. (2010). The cortisol awakening response: more than a measure of HPA axis function. *Neuroscience & Biobehavioral Reviews*, *35*(1), 97-103.

Cohen, P., Brown, J., & Smailes, E. (2001). Child abuse and neglect and the development of mental disorders in the general population. *Development and Psychopathology*, *13*(04), 981-999.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 385-396.

Crowe, M, & Sheppard, L. (2011). A general critical appraisal tool: an evaluation of construct validity. *International Journal of Nursing Studies, 48(*12), 1505-1516.

Crowe, M, Sheppard, L, & Campbell, A. (2012). Reliability analysis for a proposed critical appraisal tool demonstrated value for diverse research designs. *Journal of Clinical Epidemiology, 65*(4), 375-383.

De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., & Joels, M. (1998). Brain corticosteroid receptor balance in health and disease 1. *Endocrine Reviews*, *19*(3), 269-301.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., ... & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine*, *14*(4), 245-258.

Fergusson, D. M., Swain-Campbell, N. R., & Horwood, L. J. (2002). Does sexual violence contribute to elevated rates of anxiety and depression in females? *Psychological Medicine*, *32*(06), 991-996.

Gold, P. W., & Chrousos, G. P. (2002). Organization of the stress system and its dysregulation in melancholic and atypical depression: high vs low CRH/NE states. *Molecular Psychiatry*, *7*(3), 254.

Gold, P. W., Drevets, W. C., & Charney, D. S. (2002). New insights into the role of cortisol and the glucocorticoid receptor in severe depression. *Biological Psychiatry*, *52*(5), 381-385.

Greenspan, S. L., Rowe, J. W., Maitland, L. A., McAloon-Dyke, M., & Elahi, D. (1993). The pituitary-adrenal glucocorticoid response is altered by gender and disease. *Journal of Gerontology*, *48*(3), M72-M77.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*, *23*(1), 56-62.

Handa, R. J., Burgess, L. H., Kerr, J. E., & O'Keefe, J. A. (1994). Gonadal steroid hormone receptors and sex differences in the hypothalamic-pituitary-adrenal axis. *Hormones and Behavior*, *28*(4), 464-476.

Harkness, K. L., Stewart, J. G., & Wynne-Edwards, K. E. (2011). Cortisol reactivity to social stress in adolescents: role of depression severity and child maltreatment. *Psychoneuroendocrinology*, *36*(2), 173-181.

Heim, C., Mletzko, T., Purselle, D., Musselman, D. L., & Nemeroff, C. B. (2008b). The dexamethasone/corticotropin-releasing factor test in men with major depression: role of childhood trauma. *Biological Psychiatry*, *63*(4), 398-405.

Heim, C., Newport, D. J., Bonsall, R., Miller, A. H., & Nemeroff, C. B. (2003). Altered pituitary-adrenal axis responses to provocative challenge tests in adult survivors of childhood abuse. *American Journal of Psychiatry, 158,* 575-581.

Heim, C., Newport, D. J., Heit, S., Graham, Y. P., Wilcox, M., Bonsall, R., ... & Nemeroff, C. B. (2000). Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *Jama*, *284*(5), 592-597.

Heim, C., Newport, D. J., Mletzko, T., Miller, A. H., & Nemeroff, C. B. (2008a). The link between childhood trauma and depression: insights from HPA axis studies in humans. *Psychoneuroendocrinology*, *33*(6), 693-710.

Heim, C., Newport, D. J., Wagner, D., Wilcox, M. M., Miller, A. H., & Nemeroff, C. B. (2002). The role of early adverse experience and adulthood stress in the prediction of neuroendocrine stress reactivity in women: a multiple regression analysis. *Depression and Anxiety*, *15*(3), 117-125.

Heuser, I., Yassouridis, A., & Holsboer, F. (1994). The combined dexamethasone/CRH test: a refined laboratory test for psychiatric disorders. *Journal of Psychiatric Research*, *28*(4), 341-356.

Hinkelmann, K., Muhtz, C., Dettenborn, L., Agorastos, A., Wingenfeld, K., Spitzer, C., ... & Otte, C. (2013). Association between childhood trauma and low hair cortisol in depressed patients and healthy control subjects. *Biological Psychiatry*, *74*(9), e15-e17.

Holsboer, F. (2000). The corticosteroid receptor hypothesis of depression. *Neuropsychopharmacology*, *23*(5), 477-501.

Ising, M., Künzel, H. E., Binder, E. B., Nickel, T., Modell, S., & Holsboer, F. (2005). The combined dexamethasone/CRH test as a potential surrogate marker in depression.

Progress in Neuro-Psychopharmacology and Biological Psychiatry, *29*(6), 1085-1093.

Joyce, P. R., Williamson, S. A., McKenzie, J. M., Frampton, C. M., Luty, S. E., Porter, R. J., & Mulder, R. T. (2007). Effects of childhood experiences on cortisol levels in depressed adults. *Australian and New Zealand Journal of Psychiatry*, *41*(1), 62-65.

Juruena, M. F., Cleare, A. J., Papadopoulos, A. S., Poon, L., Lightman, S., & Pariante, C. M. (2010). The prednisolone suppression test in depression: dose–response and changes with antidepressant treatment. *Psychoneuroendocrinology*, *35*(10), 1486-1491.

Juruena, M. F., Pariante, C. M., Papadopoulos, A. S., Poon, L., Lightman, S., & Cleare, A. J. (2009). Prednisolone suppression test in depression: prospective study of the role of HPA axis dysfunction in treatment resistance. *The British Journal of Psychiatry*, *194*(4), 342-349.

Kanner, A. D., Coyne, J. C., Schaefer, C., & Lazarus, R. S. (1981). Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine*, *4*(1), 1-39.

Kant, G. J., Lenox, R. H., Bunnell, B. N., Mougey, E. H., Pennington, L. L., & Meyerhoff, J. L. (1983). Comparison of stress response in male and female rats: pituitary cyclic AMP and plasma prolactin, growth hormone and corticosterone. *Psychoneuroendocrinology*, *8*(4), 421-428.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. *American Journal of Psychiatry*, *156*(6), 837-841.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76-81.

Lee, D. A., & James, S. (2013). *The compassionate-mind guide to recovering from trauma and PTSD: Using compassion-focused therapy to overcome flashbacks, shame, guilt, and fear*. Oakland, CA: New Harbinger Publications.

Lemieux, A. M., & Coe, C. L. (1995). Abuse-related posttraumatic stress disorder: evidence for chronic neuroendocrine activation in women. *Psychosomatic Medicine*, *57*(2), 105-115.

Lopes, R. P., Grassi-Oliveira, R., de Almeida, L. R., Stein, L. M., Luz, C., Teixeira, A. L., & Bauer, M. E. (2011). Neuroimmunoendocrine interactions in patients with recurrent major depression, increased early life stress and long-standing posttraumatic stress disorder symptoms. *Neuroimmunomodulation*, *19*(1), 33-42.

Lu, S., Gao, W., Huang, M., Li, L., & Xu, Y. (2016). In search of the HPA axis activity in unipolar depression patients with childhood trauma: Combined cortisol awakening response and dexamethasone suppression test. *Journal of Psychiatric Research*, *78*, 24-30.

Lu, S., Gao, W., Wei, Z., Wu, W., Liao, M., Ding, Y., ... & Li, L. (2013). Reduced cingulate gyrus volume associated with enhanced cortisol awakening response in young healthy adults reporting childhood trauma. *PloS One*, *8*(7), e69350.

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience*, *10*(6), 434-445.

Mangold, D., Marino, E., & Javors, M. (2011). The cortisol awakening response predicts subclinical depressive symptomatology in Mexican American adults. *Journal of Psychiatric Research*, *45*(7), 902-909.

Meewisse, M. L., Reitsma, J. B., De Vries, G. J., Gersons, B. P., & Olff, M. (2007). Cortisol and post-traumatic stress disorder in adults. *The British Journal of Psychiatry*, *191*(5), 387-392.

Mello, A. D. A. F. D., Mello, M. F. D., Carpenter, L. L., & Price, L. H. (2003). Update on stress and depression: the role of the hypothalamic-pituitary-adrenal (HPA) axis. *Revista Brasileira de Psiquiatria*, *25*(4), 231-238.

Miniati, M., Rucci, P., Benvenuti, A., Frank, E., Buttenfield, J., Giorgi, G., & Cassano, G. B. (2010). Clinical characteristics and treatment outcome of depression in patients with and without a history of emotional and physical abuse. *Journal of Psychiatric Research*, *44*(5), 302-309.

Nemeroff, C. B. (2007). The burden of severe depression: a review of diagnostic challenges and treatment alternatives. *Journal of Psychiatric Research*, *41*(3), 189-206.

Nemeroff, C. B. (2016). Paradise lost: the neurobiological and clinical consequences of child abuse and neglect. *Neuron*, *89*(5), 892-909.

Nemeroff, C. B., Heim, C. M., Thase, M. E., Klein, D. N., Rush, A. J., Schatzberg, A. F., ... & Rothbaum, B. O. (2003). Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proceedings of the National Academy of Sciences*, *100*(24), 14293-14296.

Newport, D. J., Heim, C., Bonsall, R., Miller, A. H., & Nemeroff, C. B. (2004). Pituitary-adrenal responses to standard and low-dose dexamethasone suppression tests in adult survivors of child abuse. *Biological Psychiatry*, *55*(1), 10-20.

Ogilvie, K. M., & Rivier, C. (1996). Gender Difference in Alcohol-Evoked Hypothalamic-Pituitary-Adrenal Activity in the Rat: Ontogeny and Role of Neonatal Steroids. *Alcoholism: Clinical and Experimental Research*, *20*(2), 255-261.

Pariante, C. M., & Miller, A. H. (2001). Glucocorticoid receptors in major depression: relevance to pathophysiology and treatment. *Biological Psychiatry*, *49*(5), 391-404.

Pariante, C. M., Papadopoulos, A. S., Poon, L., Checkley, S. A., English, J., Kerwin, R. W., & Lightman, S. (2002). A novel prednisolone suppression test for the hypothalamic-pituitary-adrenal axis. *Biological Psychiatry*, *51*(11), 922-930.

Parker, G., & Hadzi-Pavlovic, D. (1996). *Melancholia: a disorder of movement and mood: a phenomenological and neurobiological review*. Cambridge: Cambridge University Press.

Parker, G., Tupling, H., & Brown, L. B. (1979). A parental bonding instrument. *British Journal of Medical Psychology*, *52*(1), 1-10.

Peng, H., Long, Y., Li, J., Guo, Y., Wu, H., Yang, Y., ... & Ning, Y. (2014). Hypothalamic-pituitary-adrenal axis functioning and dysfunctional attitude in depressed patients with and without childhood neglect. *BMC Psychiatry*, *14*(1), 1.

Pennebaker, J. W., & Susman, J. R. (1988). Disclosure of traumas and psychosomatic processes. *Social Science & Medicine*, *26*(3), 327-332.

Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology. *American Journal of Psychiatry*, *149*, 999-1010.

Pruessner, J. C., Wolf, O. T., Hellhammer, D. H., Buske-Kirschbaum, A., Von Auer, K., Jobst, S., ... & Kirschbaum, C. (1997). Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sciences*, *61*(26), 2539-2549.

Rohleder, N., Joksimovic, L., Wolf, J. M., & Kirschbaum, C. (2004). Hypocortisolism and increased glucocorticoid sensitivity of pro-Inflammatory cytokine production in Bosnian war refugees with posttraumatic stress disorder. *Biological Psychiatry*, *55*(7), 745-751.

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

Sarason, I. G., Johnson, J. H., & Siegel, J. M. (1978). Assessing the impact of life changes: development of the Life Experiences Survey. *Journal of Consulting and Clinical Psychology*, *46*(5), 932.

Saunders, B. E., Villeponteaux, L. A., Lipovsky, J. A., Kilpatrick, D. G., & Veronen, L. J. (1992). Child sexual assault as a risk factor for mental disorders among women a community survey. *Journal of Interpersonal Violence*, *7*(2), 189-204.

Shea, A., Walsh, C., MacMillan, H., & Steiner, M. (2005). Child maltreatment and HPA axis dysregulation: relationship to major depressive disorder and post traumatic stress disorder in females. *Psychoneuroendocrinology*, *30*(2), 162-178.

Silberman, D. M., Acosta, G. B., & Zubilete, M. A. Z. (2016). Long-term effects of early life stress exposure: Role of epigenetic mechanisms. *Pharmacological Research*, *109*, 64-73.

Spinedi, E., Salas, M., Chisari, A., Perone, M., Carino, M., & Gaillard, R. C. (1994). Sex differences in the hypothalamo-pituitary-adrenal axis response to inflammatory and neuroendocrine stressors. *Neuroendocrinology*, *60*(6), 609-617.

Steckler, T., Holsboer, F., & Reul, J. M. (1999). Glucocorticoids and depression. *Best Practice & Research Clinical Endocrinology & Metabolism*, *13*(4), 597-614.

Stein, M. B., Yehuda, R., Koverola, C., & Hanna, C. (1997). Enhanced dexamethasone suppression of plasma cortisol in adult women traumatized by childhood sexual abuse. *Biological Psychiatry*, *42*(8), 680-686.

Stetler, C., & Miller, G. E. (2005). Blunted cortisol response to awakening in mild to moderate depression: regulatory influences of sleep patterns and social contacts. *Journal of Abnormal Psychology*, *114*(4), 697.

Strüber, N., Strüber, D., & Roth, G. (2014). Impact of early adversity on glucocorticoid regulation and later mental disorders. *Neuroscience & Biobehavioral Reviews*, *38*, 17-37.

Suzuki, A., Poon, L., Papadopoulos, A. S., Kumari, V., & Cleare, A. J. (2014). Long term effects of childhood trauma on cortisol stress reactivity in adulthood and relationship to the occurrence of depression. *Psychoneuroendocrinology*, *50*, 289-299.

von Bardeleben, U., & Holsboer, F. (1991). Effect of age on the cortisol response to human corticotropin-releasing hormone in depressed patients pretreated with dexamethasone. *Biological Psychiatry*, *29*(10), 1042-1050.

Vreeburg, S. A., Hoogendijk, W. J., van Pelt, J., DeRijk, R. H., Verhagen, J. C., van Dyck, R., ... & Penninx, B. W. (2009). Major depressive disorder and hypothalamic-pituitary-

adrenal axis activity: results from a large cohort study. *Archives of General Psychiatry*, *66*(6), 617-626.

Walker, J. J., Terry, J. R., & Lightman, S. L. (2010). Origin of ultradian pulsatility in the hypothalamic–pituitary–adrenal axis. *Proceedings of the Royal Society of London B: Biological Sciences*, *277*(1688), 1627-1633.

Watson, S., Gallagher, P., Smith, M. S., Ferrier, I. N., & Young, A. H. (2006). The dex/CRH test—is it better than the DST? *Psychoneuroendocrinology*, *31*(7), 889-894.

Watson, S., Owen, B. M., Gallagher, P., Hearn, A. J., Young, A. H., & Ferrier, I. N. (2007). Family history, early adversity and the hypothalamic-pituitary-adrenal (HPA) axis: Mediation of the vulnerability to mood disorders. *Neuropsychiatric Disease and Treatment*, *3*(5), 647.

Weiss, E. L., Longhurst, J. G., & Mazure, C. M. (1999). Childhood sexual abuse as a risk factor for depression in women: psychosocial and neurobiological correlates. *American Journal of Psychiatry*, *156*(6), 816-828.

Wessa, M., Rohleder, N., Kirschbaum, C., & Flor, H. (2006). Altered cortisol awakening response in posttraumatic stress disorder. *Psychoneuroendocrinology*, *31*(2), 209-215.

Whitnall, M. H. (1993). Regulation of the hypothalamic corticotropin-releasing hormone neurosecretory system. *Progress in Neurobiology*, *40*(5), 573-629.

Yehuda, R. (1998). Neuroendocrinology of trauma and post-traumatic stress disorder. In Yehuda, R. (Ed.). *Psychological trauma*. Arlington, VA: American Psychiatric Association Publishing.

Yehuda, R. (2006). Advances in understanding neuroendocrine alterations in PTSD and their therapeutic implications. *Annals of the New York Academy of Sciences*, *1071*(1), 137-166.

Yehuda, R., Boisoneau, D., Lowy, M. T., & Giller, E. L. (1995). Dose-response changes in plasma cortisol and lymphocyte glucocorticoid receptors following dexamethasone administration in combat veterans with and without posttraumatic stress disorder. *Archives of General Psychiatry*, *52*(7), 583-593.

Yehuda, R., Halligan, S. L., & Bierer, L. M. (2002). Cortisol levels in adult offspring of Holocaust survivors: relation to PTSD symptom severity in the parent and child. *Psychoneuroendocrinology*, *27*(1), 171-180.

Yehuda, R., Southwick, S. M., Krystal, J. H., Bremner, D., Charney, D. S., & Mason, J. W. (1993). Enhanced suppression of cortisol following dexamethasone administration in posttraumatic stress disorder. *American Journal of Psychiatry*, *150*, 83-83.

Young, E. A. (1998). Sex differences and the HPA axis: implications for psychiatric disease. *The journal of gender-specific medicine: JGSM: the official journal of the Partnership for Women's Health at Columbia*, *1*(1), 21-27.

Young, E. A., Abelson, J. L., Curtis, G. C., & Nesse, R. M. (1997). Childhood adversity and vulnerability to mood and anxiety disorders. *Depression and Anxiety, 5*, 66–72.

Young, E. A., & Breslau, N. (2004). Saliva cortisol in posttraumatic stress disorder: a community epidemiologic study. *Biological Psychiatry*, *56*(3), 205-209.

Young, E. A., Lopez, J. F., Murphy-Weinberg, V., Watson, S. J., & Akil, H. (1997). Normal pituitary response to metyrapone in the morning in depressed patients: implications for circadian regulation of corticotropin-releasing hormone secretion. *Biological Psychiatry*, *41*(12), 1149-1155.

PART 2: EMPIRICAL PAPER

Effect of valenced stimuli on hippocampus-dependent spatial memory in depression

1 ABSTRACT

Aims: People with depression are biased towards processing emotionally-negative stimuli. They also show impairments in hippocampal volume and hippocampus-dependent spatial memory. Here, a cross-sectional design was used to explore, for the first time, effects of stimuli affective valence (neutral/negative) on hippocampus-dependent spatial memory performance.

Method: A 2X2 mixed design was used. Two groups (participants with moderate to severe depression; healthy controls) completed the Town Square Test of spatial memory, with neutral and negatively valenced stimuli. Participants' allocentric memory score (correct viewpoint-independent recognition of stimuli's original location) was used for analysis, plus data collected on the depression group's clinical history and current symptomology.

Results: Mixed analysis of variance (ANOVA) models were used. Essentially, no effects of valence on allocentric performance were found, for either depression or healthy control group. There were some sex effects: women with depression performed better than healthy women in the allocentric shifted view condition, though this was independent of valence; and being a female with depression predicted enhanced controlled allocentric performance both overall and with negative stimuli, as did increasing age. Neither current symptomology or depression characteristics influenced depression group performance.

Conclusions: Potential reasons for the null results were explored, including: confounding valence and arousal; lack of valence processing; subjectivity of valence; and the depression group's high levels of co-morbid anxiety. Further research is required, addressing some of these proposed confounds, to elucidate these potentially important mechanisms underlying depression.

2 INTRODUCTION

2.1 Depression

Depression is the most common mental health issue worldwide and an estimated one in five of the population will be affected at some point in their lives (Kessler et al., 2005). According to the Diagnostic Statistical Manual of Mental Disorders (5th Edition; American Psychiatric Association, 2013), to be diagnosed with depression someone must have persistent low mood and/or loss of pleasure in activities lasting for at least two weeks, plus five additional symptoms from a list including sleep problems, fatigue and concentration issues.

Cognitive models of depression (Beck, 1967) emphasise the role of negative schema of the self, the world/others, and the future in the development and maintenance of depression. Explicitly looking to challenge and change these negative schema forms the focus of cognitive therapeutic interventions (Beck, 1967, 1976). While 50%–70% no longer meet criteria for depression once cognitive therapy is completed (Craighead, Sheets, Brosse, & Ilardi, 2007), at least half of these will relapse within two years (Vittengl, Clark, Dunn, & Jarrett, 2007). Given the significant levels of disability caused by depression (Judd et al., 2000), finding out more about processes underlying the disorder is desirable so more effective treatments can be developed to improve both recovery and relapse rates.

2.2 Reductions in hippocampal volume in depression

Reductions in hippocampal volume in depression are a robust research finding. A recent meta-analysis of fMRI studies found significant reductions in hippocampal volume in depression when compared with healthy controls (Schmaal et al., 2015). This is in line with previous meta-analyses (Arnone, McIntosh, Ebmeier, Munafò & Anderson, 2012; Lorenzetti, Allen, Fornito & Yücel, 2009; Koolschijn, van Haren, Lensvelt-Mulders, Hulshoff-Pol & Kahn, 2009; McKinnon, Yucel, Nazarov & MacQueen, 2009). Within this overall effect

on hippocampal volume, Schmaal et al. (2015) found no associations with age, depression severity, medication, or co-morbid anxiety. However, recurrent depression, rather than a first episode, was associated with lower volumes versus healthy controls, as was early onset before the age of 21.

Reductions in hippocampal volume have been attributed to atrophy occurring as a result of the hypothalamic-pituitary-adrenal axis being chronically hyperactive in depression (Campbell & MacQueen, 2004). This hypercortisolaemia cytoxicity is thought to reduce levels of growth-promoting brain-derived neurotrophic factor (BDNF; Cosi, Spoerri, Comelli, Guidolin & Skaper, 1993) which in turn suppresses hippocampal neurogenesis (Camerson & McKay, 1999). Evidence from animal studies suggests antidepressant medication may in part reduce depression symptoms by increasing BDNF levels, and thereby neurogenesis (MacQueen, Ramakrishnan, Ratnasingan, Chen & Young, 2003). However, there is only preliminary evidence of this in humans (for a review see Willner, Scheel-Krüger & Belzung, 2013).

2.3 Deficits in autobiographical memory specificity in depression

There is robust evidence from both lesion (e.g. Scoville & Milner, 1957) and neuroimaging studies (see Martinelli, Sperduti & Piolino, 2012, for a review) that one of the main functions of the hippocampus is in episodic memory, including autobiographical memory for personally experienced events occurring at a particular time and place.

The presence of overgeneral memory (OGM) in depression is well documented (Williams et al., 2007). OGM refers to difficulties in retrieving specific autobiographical memories, coupled with the enhanced recall of categorical memories that also occurs in depression (see Williams et al., 2007 for a review). OGM not only impacts on recall of past events, but also imagining the outcomes of future ones (Tulving, 2001), and problem solving (Arie, Apter, Orbach, Yefet, & Zalzman, 2008). It is thought OGM might interfere

with cognitive therapy, leaving negative schemas unchallenged, contributing to the maintenance of depression.

Given the robust finding of reduced hippocampal volume in depression, studies have looked for an association with OGM. One group of researchers has consistently found increased activity in a hippocampus-centric network when people with depression successfully recall specific autobiographical memories. This is relative to healthy controls and when recall is unsuccessful (Young et al., 2013, 2014a, 2014b, 2012). This network includes the pre-frontal cortex, amygdala and hippocampus. The authors suggest the increase in hippocampal activity is necessary to compensate for the reduction in volume in depression, allowing specific memories to be successfully retrieved.

2.4 Deficits in spatial memory in depression

Another main hippocampal function is in spatial memory. The hippocampus processes the cognitive map, or spatiotemporal context, of autobiographical memories, during encoding and recall (O'Keefe & Nadel, 1978). This allows spatial memory to be independent of viewpoint, or allocentric, as opposed to viewpoint dependent or egocentric. While allocentric memory relies on the right hippocampus and medial temporal lobe, egocentric memory is dependent on the posterior parietal lobe (Burgess, 2008; Doeller, Barry & Burgess, 2012).

Allocentric memory allows for flexible spatial reasoning. At its most basic, this allows objects to be located even when viewpoint changes. Given the shared neurobiological features, the allocentric hippocampal cognitive map has been hypothesised to be key in providing a spatio-temporal stamp for autobiographical memories. These allocentric representations allow them to be contextualised and specifically recalled (Burgess, Maguire & O'Keefe, 2002) in a coherent order (Tulving, 2002). This is supported

by findings of links between OGM and hippocampal activity in depression (e.g. Young et al., 2012).

Recent findings suggest spatial memory may also be impaired in people with depression (Cornwell et al., 2010; Gould et al., 2007; Wong, 2015). Gould et al. (2007) used a virtual reality spatial memory test and Cornwell et al. (2010) a virtual version of the Morris water maze task. Both studies found impairments in spatial memory performance when compared to healthy controls. Cornwell et al. (2010) also found a concurrent reduction in hippocampal activation while people with depression were completing the task. Both of these tasks, however, require spatial navigation and therefore measure these cognitive processes in addition to allocentric performance.

The Town Square Test (King, Burgess, Hartley, Vargha-Khadem & O'Keefe, 2002) measures both allocentric and egocentric memory for object locations. In a virtual environment, participants' viewpoint is either kept the same between encoding and retrieval (testing egocentric memory) or shifted, so recognition takes place from another viewpoint (testing allocentric memory). The Town Square Test is considered to be a more precise measure of allocentric spatial memory, both because it does not require navigation-based processes and because of the instantaneous shift between presentation and test viewpoint in the shifted view condition. This minimises the use of egocentric processes to update participants' location during the move from presentation to test viewpoint (Burgess et al., 2002). As the shifted view condition may also use egocentric processes, participants' score in the egocentric same view condition can be subtracted from their score in the shifted viewpoint condition to give an indication of allocentric memory performance controlling for egocentric performance.

The Town Square Test has been structured so same and shifted view conditions are difficulty matched for healthy controls. There should, therefore, be no difference in performance between these conditions. Studies using the Test have broadly supported

this, although one study found deficits in allocentric performance for control group participants (Wilkins et al., 2013).

Participants with hippocampal lesions only show deficits in shifted viewpoint performance (King et al., 2002). These findings support both the hippocampus' role in allocentric spatial memory and use of the Town Square Test to assess hippocampal functioning. Using the test, Wong (2015) found allocentric memory deficits in depression, both relative to egocentric memory and healthy controls (using data from Smith, Burgess, Brewin & King, 2015), suggesting hippocampal functioning impairment, in line with reduced hippocampal volume (Schmaal et al., 2015).

2.5 Mood congruence in depression

2.5.1 Recognition memory

The general population shows enhanced recognition memory for positive stimuli versus neutral and negative stimuli (Dolcos, LaBar, & Cabeza, 2004). For people with depression, the opposite pattern is seen and recognition performance is only enhanced with negatively-valenced stimuli (Roberson-Nay et al., 2006). This difference is thought to be due to preferential encoding of mood congruent stimuli (Bower, 1981).

2.5.2 OGM

There is consistent evidence of mood congruence effects in OGM (see Williams et al., 2007 for a review). While people with depression show impaired autobiographical memory specificity overall compared to healthy controls, they are better able to produce specific memories to negative cue words than positive ones. The pattern is reversed for healthy controls (Williams & Scott, 1988).

2.5.3 Emotional attention and cognitive control

Mood congruent biases towards negative stimuli in depression have also been found in terms of attention (Mathews, Ridgeway & Williamson, 1996), with an enhanced

perigenual anterior cingulate cortex (pACC) response implicated (Mitterschiffthaler et al., 2003). People with depression also exhibit poorer cognitive control (see Castaneda et al., 2008 for a review), and are less able to suppress responses to negatively valenced stimuli where these are irrelevant, affecting task performance (e.g. Joorman & Gottlib, 2008).

2.6 Accounting for mood congruence effects

Two hypotheses account for when and how mood congruent stimuli may affect task performance of both healthy controls and people with depression.

2.6.1 Modulation hypothesis

According to the modulation hypothesis, encoding and consolidation of affective information within the medial temporal lobe (MTL), including the hippocampus, is modulated by the amygdala (McGaugh, 2002). It is suggested this enhanced amygdala-MTL interaction accounts for enhanced memory for affective stimuli (Dolcos et al., 2004).

For people with depression, it has been hypothesised the mood congruence effect on memory results from a lowering of the threshold for engagement of the amygdala in memory formation with negative stimuli. This overactivation of the neural pathway of the modulation hypothesis means the amygdala is more readily engaged than in people without depression (Roberson-Nay et al., 2006). Thus, while reduced memory for neutral and positive stimuli in depression is thought to result from reduced hippocampal volume, enhanced memory for negative stimuli is linked to enhanced connectivity between the amygdala and hippocampus. Supporting this, Hamilton and Gotlib (2008) found greater interaction between the amygdala and hippocampus in a depression group than a non-depressed group only during successful encoding of negative stimuli in a recognition memory task. Additional support comes from neuroimaging evidence that both the hippocampus and amygdala are overactivated by negative stimuli in depression (e.g. Fu et al., 2004), as is the pACC, connecting the two (see meta-analysis by Sacher et al., 2012).

Given the involvement of the hippocampus in OGM, modulation could account for OGM mood congruence effects in depression.

2.6.2 Affective Interference Hypothesis

There is also evidence that negatively valenced stimuli act as a distractor for people with depression when valence is irrelevant, disrupting task performance (Siegle, Steinhauer, Thase, Stenger & Carter, 2002; Joorman & Gottlib, 2008, 2010). According to the Affective Interference Hypothesis (Gottlib & Joorman, 2010), people with depression process emotional material first, even when task-irrelevant. This is the result of a combination of preferential emotional attention, leading to difficulties disengaging attention from negative stimuli (Wang et al., 2008), and deficits in cognitive control, making it difficult to suppress the response to valence when this is task irrelevant (Clark, Chamberlain, & Sahakian, 2009). This cognitive control failure has been linked to deficits in the dorsal anterior cingulate cortex (Posner & Rothbart, 2007) and the prefrontal cortex (e.g. Bench, Friston, Brown, Frackowiak & Dowan, 1993). In summary, if stimuli valence is mood congruent but task-irrelevant, it may interfere with task performance, rather than enhance it (Gottlib & Joorman, 2010).

2.7 Mood congruence and spatial memory

There is tentative support for affective interference in a spatial memory task. When asked to recall the location of a valenced picture after a delay, scores reduced as affective arousal of the images increased (Mather et al., 2006). While this study did not use a depression group, depression symptomology was measured. As depression scores increased, spatial memory for negative pictures decreased. The authors suggested this interference was due to an impairment in feature binding. In line with the affective interference hypothesis, emotional components were processed preferentially to the detriment of feature binding in working memory required to remember location. The

authors referred to this as 'mental rubbernecking', whereby mood congruent valence means the item is remembered, but not its location.

However, the study did not involve a viewpoint shift. It therefore measured egocentric spatial memory, using the posterior parietal lobe and not hippocampally-dependent allocentric memory. It is therefore not known whether affective interference still occurs when the hippocampus is involved and modulation may occur.

2.8 Rationale and aims of study

2.8.1 *Rationale*

While the reduction in hippocampal volume is a robust finding in depression (Schmaal et al., 2015), implications for day to day functioning of people with depression are less clear cut. It can lead to impairments in hippocampally-dependent memory (Williams et al., 2007; Wong, 2015), but there is also evidence of compensatory increases in hippocampal activity in some instances (e.g. Young, 2012).

Adding further complexity is the influence on performance of stimuli's emotional content in depression and what this might tell us about how processing valence influences hippocampal functioning, including connectivity with the amygdala. This area has not received much attention to date.

This study therefore looked to pull together these strands of research into memory processes and mood congruence effects in depression by investigating for the first time the impact of valenced stimuli on hippocampally-dependent allocentric memory. As neurobiological evidence suggests the same medial temporal structures are involved in both of these processes, establishing a link in terms of observable memory performance would be informative of mechanisms underlying these features of depression. For the non-depressed control group, this was also the first investigation of the effect of valenced stimuli on allocentric spatial memory.

The study built on the previous investigation of allocentric memory in depression (Wong, 2015) by using valenced stimuli in the Town Square Test. In Wong's (2015) study, and all previous research using the Town Square Test, stimuli were low resolution images. Stimulus valence had not been explicitly controlled, although there had been an assumption stimuli were not strongly valenced, positively or negatively. Here, both neutral and negatively valenced stimuli were used to allow effects on performance to be investigated. Words were used rather than images to allow valence to be more reliably manipulated, by using a normed database (Warriner, Kuperman & Brysbaert, 2013).

2.8.2 Primary aim and hypotheses

The primary aim of the research was therefore to investigate effects on allocentric memory performance of stimulus type (neutral or negative) and any differences between people with depression and healthy controls. In Wong's (2015) study, when (uncontrolled) neutral stimuli were used, people with depression showed impairments in allocentric performance, both in relation to healthy controls and egocentric performance. This research extended this by comparing this with performance when negative stimuli were used.

This was the first time the interaction between allocentric memory and the valence of stimuli had been studied. Modulation and affective interference hypotheses would make different predictions for the impact on performance of using negatively valenced stimuli. These alternative hypotheses were explored. For neutral stimuli, there was a single hypothesis as both modulation and affective interference hypotheses would make the same predictions.

Neutral stimuli

Hypothesis: The depression group will show the same impairment in the shifted view condition when compared to the same view condition as found in Wong's (2015) study. The control group will not show an impairment. The controlled allocentric memory

performance (i.e. when same view scores are subtracted from shifted view scores to control for egocentric memory performance) of people with depression will also be significantly worse with neutral stimuli than healthy controls'. This will be due to hippocampal impairment (Schmaal et al., 2015).

Negative stimuli

Enhancement hypothesis: Shifted view performance will be significantly better for people with depression, both when compared to their own same view performance, and healthy controls' shifted view performance. Controlled allocentric memory performance with negative stimuli will also be significantly better for people with depression than healthy controls and performance with neutrally valenced stimuli. This will be due to enhanced connectivity between the amygdala and hippocampus improving performance with negatively valenced stimuli in the hippocampally-dependent allocentric condition for people with depression.

Interference hypothesis: In the Town Square Test, valence is irrelevant to the task. For people with depression, processing negatively valenced words will interfere with processing location and therefore test performance. Given the enhanced cognitive demands of the shifted view condition, interference will be greater in this condition than the same view condition, and for controlled allocentric performance. For the control group, interference will not occur and performance will be better than the depression group with negative stimuli on same and shifted view conditions and for controlled allocentric performance.

2.8.3 Secondary aim

The secondary aim was to explore predictors of allocentric performance with negative stimuli in depression. According to the meta-analysis (Schmaal et al., 2015), early onset of depression (before the age of 21) and number of episodes are predictive of

hippocampal volume reductions, and so tentatively might be hypothesised to influence allocentric performance with negatively valenced stimuli. Similarly, this meta-analysis also found no association between hippocampal volume, symptom severity and co-occurring anxiety, and so, again tentatively, these may be hypothesised to not influence performance.

3 METHOD

3.1 Joint project arrangement

This was a joint project with another trainee clinical psychologist (Sagfors, 2017; see Appendix A). Data collection was shared as both projects used the same participants and there was overlap in measures. However, the other trainee's research question focused on links between OGM and allocentric spatial memory as an indicator of hippocampal function. Therefore, different primary variables were used for analysis. Only data for this project have been included here.

3.2 Research design

This cross-sectional study used mixed designs to investigate the primary research aim. For all designs, the between groups independent variable was group (two levels: depression; control) and the within group independent variable was stimulus valence (two levels: neutral; negative). A 2x2x2 mixed design added viewpoint (two levels: same; shifted) as a within group variable and used correct location recognition as the dependent variable. A 2x2 mixed design used controlled allocentric memory score (shifted view minus same view percentage correct) as the dependent variable to investigate any effects of stimuli valence for the two groups.

Multiple regression was used to explore predictors of controlled allocentric memory performance with negative stimuli for the depression group, which formed the secondary research aim.

3.3 Sample size: power analysis

This was the first time valenced stimuli had been used within the Town Square Test. Wong (2015) found a large effect size (.83) for differences in allocentric memory between a depression group and data for healthy controls from another study (Smith et al., 2015). A previous study using the Test (Smith et al., 2015), found a medium effect size (partial eta-sq. = .1) for the relationship between allocentric memory performance and posttraumatic stress disorder. Similarly, using the Test to investigate the relationship between performance and experience of induced trauma memories for a healthy sample, Bisby, King, Brewin, Burgess & Curran (2010), found a medium to large effect size for the negative relationship (r=-.43).

A power analysis conducted for the joint project using GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), with a power level of .80 and alpha of .05, suggested a sample size of 68 (34 in each group) to detect a small to medium effect. A subsequent sensitivity analysis used GPower 3.1 (Faul et al., 2007) for a repeated measures ANOVA (within-between interaction) with two groups and two measures, assuming sphericity and correlation between measures of .50. With a sample size of 68, at the same power and alpha levels, an effect size of at least .17 (small to medium) would be detectable.

3.4 Setting

Recruitment and testing of the depression group was based within an NHS Improving Access to Psychological Therapies (IAPT) service. Control group participants were recruited through two routes (see Figure 1) and testing took place at UCL's Department of Psychology.

3.5 Participants

A total of 72 participants were recruited, 35 for the control group and 37 for the depression group. All participants were 18-65 years old. Inclusion and exclusion criteria for both groups were designed to reduce the potential impact of confounding variables on spatial memory performance, including key co-morbid and neurological disorders (see Figure 1). For the depression group, onset at least 2.5 years ago was an inclusion criterion as this had been found to be associated with hippocampal damage (Schmaal et al. (2015). Co-morbid anxiety was allowed for the depression group as Schmaal et al.'s (2015) meta-analysis found no association between hippocampal volume and co-occurring anxiety, and as the two disorders commonly co-occur (Moffitt et al., 2007).

Participants who were UCL undergraduates were incentivised through credits towards a compulsory course requirement. All other participants were paid £12.50 at the end of participation as compensation for travel expenses.

3.6 Ethical considerations

Wong's (2015) study recruited from the same IAPT service and was granted ethical approval by the National Research Ethics Service (NRES) Committee East Midlands – Nottingham 2 in August 2014 (REC reference: 14/EM/0222). An amendment to this ethics for this project was obtained (see Appendix B for letter granting approval), as recruitment was from the same sites and cohort. The control group was recruited and tested under the project supervisor's existing UCL CEHP programme ethics approval (SHaPS-2014-JK-009).

In terms of managing risk, all testing for the depression group took place on IAPT sites and any concerns around risk were raised with the duty worker. There was one incident, managed in line with this protocol. For the control group, researchers assessed risk on an individual basis, deciding whether confidentiality needed to be broken (e.g. where potential for harm was high). There were no incidences.

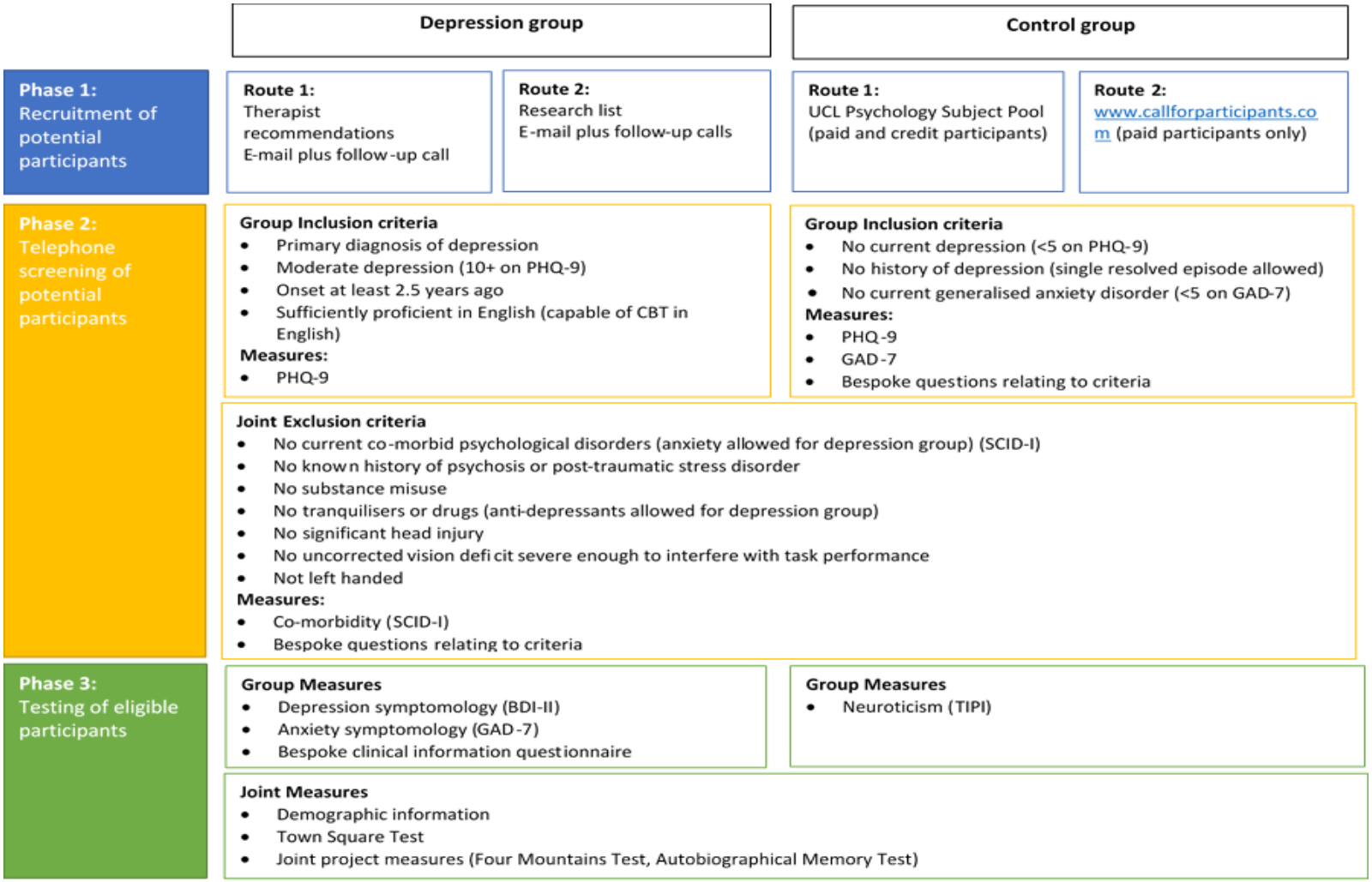


Figure 1: Summary of procedure phases and routes, eligibility criteria and measures used in each phase for depression and control groups

3.7 Procedure

See Figure 1 for a summary.

3.7.1 Phase one: Recruitment of potential participants

Depression group

The depression group was identified using two routes. Firstly, IAPT clinicians made referrals for screening based on inclusion and exclusion criteria, obtaining verbal consent for their clients to be contacted by study researchers by telephone. Two clinician referrals were received.

Secondly, researchers were provided with a list of the IAPT service's clients who had consented to being contacted for research purposes. Participants were e-mailed to assess willingness to participate if they met the following criteria: aged 18-65; scored ten or more on their latest Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001); were currently being seen by the service (i.e. not discharged); were low risk. Follow-up calls were made within a week. A total of 282 potential participants were contacted using this route.

Control group

Participants were recruited through two online facilities: UCL Psychology Subject Pool; and www.callforparticipants.com. Sites included information on the main inclusion and exclusion criteria. Potential participants signed up for telephone screening and testing slots.

3.7.2 Phase two: telephone screening

All participants were screened by telephone to ensure they met inclusion and exclusion criteria (Figure 1 summarises criteria and measures for each group). Standardised measures, including the PHQ-9 (Kroenke et al., 2001), Generalised Anxiety Disorder-7 (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006) and a subset of the Structure

Clinical Interview for DSM Disorders (SCID-I; First, Spitzer, Gibbon & Williams, 2002), were supplemented with specific questions. If participants did not meet criteria, they were told they were ineligible to participate and contact was terminated.

In the depression group, out of 60 people screened, 21 were ineligible. In the control group, of the 39 screened, two were ineligible.

3.7.3 Phase three: Testing of eligible participants

Eligible participants were given a testing appointment and e-mailed the relevant participant information sheet in advance. Written consent was obtained at the start of the testing session (see Appendix C). Relevant additional measures for each group were then administered (see Figure 1), followed by joint project measures. Participants then completed the Town Square Test before finally completing a questionnaire rating the valence of the words used in the Test. Participants were paid or given course credit as appropriate on completion. Thirty-six depression group and 35 control group participants completed testing.

3.8 Measures

Figure 1 summarises measures used during Phase two telephone screening.

3.8.1 Phase three: testing of eligible participants

Allocentric memory measure: Town Square Test (King et al., 2002)

For the Town Square Test, a virtual environment consisting of a courtyard with distinct buildings around the perimeter was observed on a laptop with a 15-inch screen. In the presentation phase, participants were shown stimuli in blocks of six, one at a time for three seconds each in one of 21 randomly located placeholders on the courtyard floor, with an interval of one second between presentation items. In the test phase, participants were shown each stimulus in its original location placeholder along with three identical foils in random placeholders. Each of the test stimuli was coded with a coloured square, and

participants used this to indicate the correct presentation location. There was no time limit. Presentation order within a block in the test phase was randomised to reduce predictability and use of mnemonic strategies.

Two viewpoints in opposite corners of the courtyard were used for presentation and test (Figure 2). There were two viewpoint conditions. In the same-view condition, viewpoint remained the same between presentation and test (Figure 2, a) and b)). In the shifted-view condition, presentation was at one viewpoint and test at the other, that is viewpoint shifted between presentation and test (Figure 2 c) and d)). The viewing rotation required to move from one view to the other was 140 degrees.

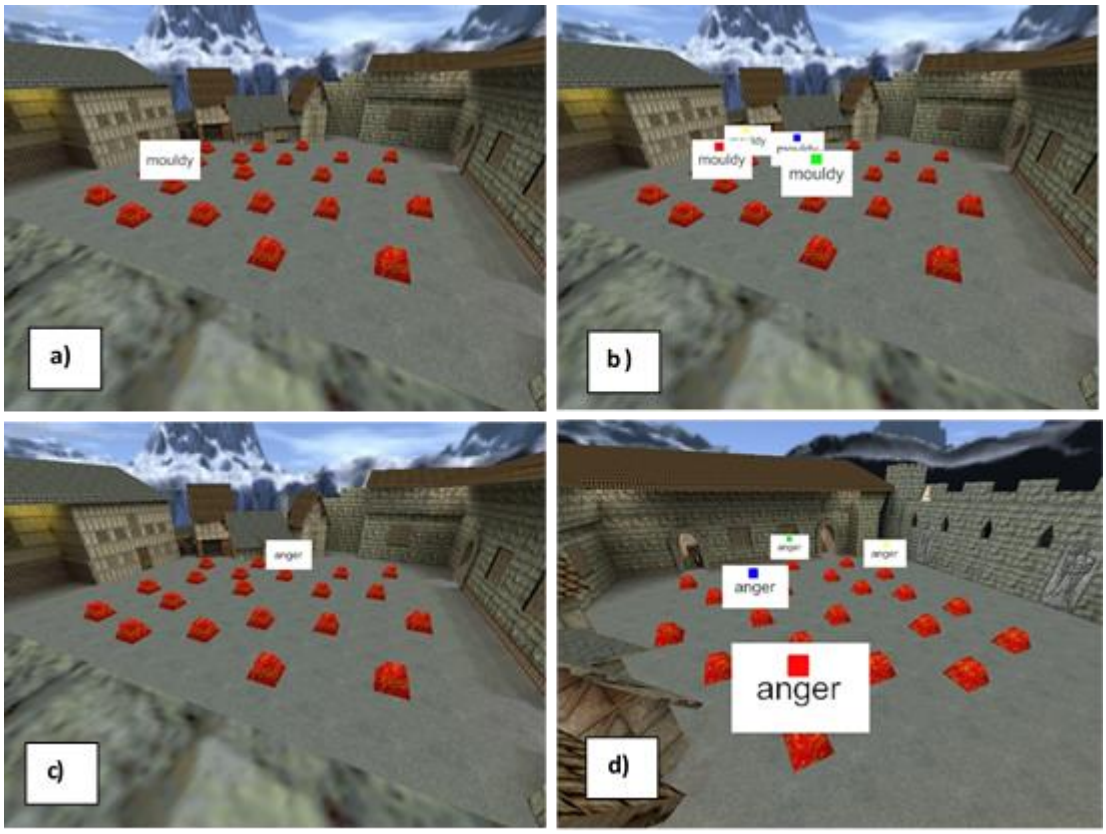


Figure 2: Town Square Test. Views overlooking the middle of the virtual environment courtyard. The same viewpoint condition is shown in a) (presentation) and b) (testing). The shifted viewpoint condition is shown in c) (presentation) and d) (testing), with the traffic cone in d) showing the point from which stimuli were viewed at presentation in c).

Same and shifted view conditions have been difficulty matched to reduce performance differences for control participants (King, Trinkler, Hartley, Vargha-Khadem, & Burgess. 2004). Foil locations were nearer the target in the same view condition to make it more difficult, and foil locations were spread across the courtyard for the shifted view condition to facilitate correct responses.

Previous versions of the Town Square test had used low resolution photographs of objects as stimuli (King et al., 2002). Images were chosen for their recognisability and valence was not specifically controlled. Stimuli valence may therefore have acted as a confound. To control stimulus valence in a reliable way, a word-based version was developed specifically for this project.

Neutral and negatively valenced words were taken from a database (Warriner et al., 2013) which provides valence norms for 13,915 English lemmas. Norms were based on responses of 1827 participants who rated their first and immediate affective reactions to words on a scale from one to nine, with one indicating feeling completely unhappy, nine completely happy and five if neutral about the word.

For this study, neutral words were selected from those scoring 5 +/- .08 (e.g. barrel, jigsaw). For negative words, lowest scoring words were selected (e.g. poison, traitor). To minimise distress or disturbance, no words were selected relating to: violence, weapons, sex, disease or words generally considered to be taboo. These words also showed significant differences in database ratings between males and females (Warriner et al., 2013).

A homogenous set was selected, controlled for word length (5-7 letters) and reading scores. Appendix D contains the words used with their valence ratings. For negative words, the range of valence ratings was 1.88 (pollute) to 2.52 (creepy), mean was 2.22 and standard deviation .19. For neutral words, the range was 4.92 (barrel) to 5.05

(various, e.g. canteen), mean was 5.00 and standard deviation .04. Words were presented in a black typeface on white background. Arial, a sans serif font, was chosen for maximum legibility at variable size as words were scaled with distance (Figure 2).

There was evidence the enhanced memory effect for negatively valenced stimuli with depression persisted after stimulus presentation (Siegle et al., 2002). To address this, neutral and negative words were presented in separate blocks with order of presentation randomised across participants. There were eight blocks of six for each valence condition, four blocks for each viewpoint condition. A total of 96 words was presented, 24 in each of the four combinations of viewpoint (same, shifted) and valence (neutral, negative).

Processing the words including their valence was not necessary to complete the task. Therefore, participants were told prior to starting they would be asked at the end for their affective reaction to the words and to bear this in mind when completing the task (see Appendix E).

For the analysis, a percentage correct score for same and shifted viewpoints for both the neutral and negatively valenced word conditions was calculated. A measure of allocentric performance, controlling for differences in egocentric performance was calculated as the difference between shifted (using allocentric and egocentric processes) and same view performance (using egocentric only).

Additional measures

For the depression group, the Beck Depression Inventory (BDI-II; Beck, Steer, Ball & Ranieri, 1996) was used to indicate depression severity, and the GAD-7 to assess presence and levels of co-morbid generalised anxiety disorder (GAD). Clinical information was collected using a bespoke questionnaire (see Appendix F).

For the control group, the Ten Item Personality Measure (TIPI; Gosling, Rentfrow & Swann, 2003) was administered to assess neuroticism as research had suggested neuroticism was a risk factor for depression (Kendler, Gardner, & Prescott, 2002; Kendler,

Kessler, Neale, Heath & Eaves, 1993). Individuals with high neuroticism had also been found to show similar memory biases towards negative stimuli as people with depression (Ruiz-Caballero & Bermudez, 1995). A neuroticism score was derived by summing the two neuroticism items.

For both groups, demographic information was collected using a bespoke questionnaire (see Appendix G), including age, sex, ethnicity and whether English was the participant's first language.

Joint project measures

Two measures were also administered specifically for the joint project: Four Mountains Test (Hartley et al., 2007); and Autobiographical Memory Test (Williams & Broadbent, 1986).

3.9 Data analysis procedures

To analyse the primary research question, firstly, a 2 (group: control; depression) x2 (valence: neutral; negative) x2 (viewpoint: same; shifted) mixed model mixed analysis of variance (ANOVA) was used with total percentage correct as the dependent variable. Because hippocampus-dependent memory performance reduces with age (Raz et al., 2010), age was controlled for by adding as a covariate. Sex was also added to this model as there is evidence of sex differences in spatial memory (Grön, Wunderlich, Spitzer, Tomczak & Riepe, 2000). To investigate any order effects, presentation order of valenced stimuli was also added.

A 2 (group: control; depression) x 2 (valence: neutral; negative) ANOVA was then used with controlled allocentric performance score (shifted minus same scores) as the dependent variable. Once again, age was a covariate, and sex and order were added to the model.

To explore predictors of performance for the depression group, hierarchical multiple regression was used. Block 1 contained potential demographic confounds, Block 2 symptomology variables and Block 3 depression characteristic variables.

4 RESULTS

4.1 Data preparation and exploration

Data analysis was performed using the Statistical Package for Social Sciences Version 24 (SPSS, IBM Corp. in Armonk, NY). Dependent variables were first investigated for outliers and adherence to assumptions of parametric tests.

Normality checks: the z-scores for kurtosis and skewness were calculated for each of the Town Square Test variables for both groups. None of the values were more than z=2.5 or p=.01. In addition, none of the Kolmogorov-Smirnov tests were significant at a threshold of p<.01. The distributions of the dependent variables for both groups were therefore assumed to meet assumptions of parametric testing.

Outliers: standardised z scores were calculated for the Town Square Test variables. A threshold of +/- 3 was applied. No outliers were identified in either control or depression group.

4.2 Participant demographics and clinical characteristics

Three depression group participants were excluded from the analysis; two because they did not complete the Town Square Test and one because they were left handed. The demographics of participants included in the analysis are detailed in Table 1. There were no differences between groups in terms of demographics (see Table 1 for test statistics). A Chi-squared test for ethnicity could not be conducted as it was not possible to combine

categories in a meaningful way to achieve cell sizes of more than five. Asian was the biggest ethnic group for the control group; White British for the depression group.

Table 1

Demographic characteristics of participants by group plus test statistics for group differences

	Control group N=35	Depression group N=34	Test statistic and p value
Age (years) M (SD, range)	31.66 (11.40, 18-62)	34.76 (12.72, 20-59)	t(66)=1.02, p=.31
Gender N (%)			
Male	13 (37)	17 (51.5)	χ^2 (1)=1.16, p=.28
Female	22 (63)	16 (48.5)	
Ethnicity N (%)			
White British	6 (17.1)	18 (54.5)	N/A
White Other	8 (22.9)	8 (24.2)	
Black/Black British	5 (14.3)	2 (6.1)	
Asian	12 (34.3)	2 (6.1)	
Asian British	0 (0)	3 (9.1)	
Mixed	3 (8.6)	2 (6.1)	
Other	1 (2.9)	0 (0)	
English as a first language N (%)			
Yes	22 (62.9)	26 (78.8)	χ^2 (1)=2.30, p=.13
No	13 (27.1)	7 (21.2)	
Employment status N (%)			
Employed	16 (45.7)	20 (60.6)	χ^2 (1)=1.78, p=.18
Unemployed/student	19 (54.3)	13 (39.4)	
Education N (%)			
Up to undergraduate	9 (25.7)	9 (27.3)	χ^2 (2)=5.07, p=.08
Undergraduate	11 (31.4)	13 (39.4)	
Postgraduate	15 (42.9)	11 (33.3)	
Income N(%)			
Up to £25,000	16 (45.7)	13 (39.4)	χ^2 (2)=5.07, p=.08
£26,000 - £50,000	17 (48.6)	10 (30.3)	
Over £50,000	2 (5.7)	8 (24.2)	

The depression group's clinical characteristics are summarised in Table 2. For BDI-II, categorises of minimal (0-9), mild (10-18), moderate (19-29) and severe (30-63) were used (Beck, Steer & Garbin, 1988). For GAD-7, categories of minimal (0-4), mild (5-9), moderate (10-14) and severe (>15) were used (Spitzer et al., 2006), as well as numbers meeting a diagnostic threshold for generalised anxiety disorder (score of ten or more; Spitzer et al., 2006). Participants were categorised as having either a first degree relative with a history of depression (parents, siblings), only a second degree relative (grandparent, aunt, uncle, nephew, niece or cousin), or none.

Table 2

Clinical characteristics of the depression group

	Depression group N=34
BDI	
Score M (SD, range)	29.18 (8.11, 11-48)
Category N (%)	
Mild	2 (6.1)
Moderate	14 (42.4)
Severe	17 (51.5)
GAD-7	
Score	11.15 (3.51, 4-21)
Category	
Minimal	1 (3.0)
Mild	11 (33.3)
Moderate	16 (48.5)
Severe	5 (15.2)
Cut-off	29 (87.9)
Depression history	
Years since first episode M (SD, range)	17.18 (12.15, 3-50)
Start of current episode (months) M (SD, range)	25.15 (38.14, 3-216)
Length of current episode (months) M (SD, range)	25.21 (37.51, 3-216)
Total number of episodes M (SD, range)	3.92 (1,44, 1-7)
Currently taking medication N (%)	18 (54.5)
Currently undergoing talking therapy N (%)	29 (87.9)
Number of sessions of talking therapy to date M (SD, range)	5.38 (4.08, 0-20)
Family history N (%)	
First degree relative	23 (69.7)
Second degree relative	3 (9.1)
None	6 (18.2)

4.3 Comparison of shifted and same viewpoint conditions

Figure 3 shows mean percentage total correct for each group by valence in same and shifted viewpoint conditions.

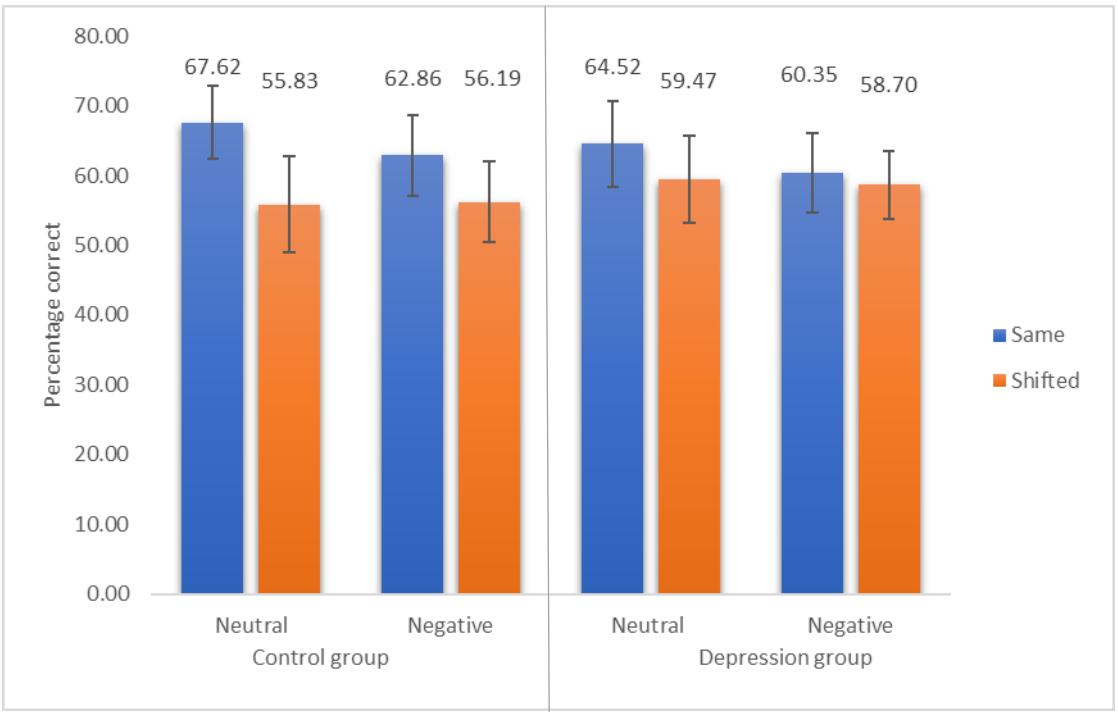


Figure 3. Mean percentage total correct for each group by valence in same and shifted viewpoint conditions. Error bars represent 95% CI for mean.

Controlling for age (*M*=33.16), a 2x2x2 mixed model ANOVA with group (control, depression) as between subjects factor and viewpoint (same, shifted) and valence (neutral, negative) as within subjects factors showed a main effect of viewpoint ($F(1.65)=6.15$, $p=.02$, $\eta_p^2= .09$). There were no significant main effects for valence ($F(1,65)=2.05$, $p=.16$, $\eta_p^2=.03$) or for group ($F(1,65)=.33$, $p=.57$, $\eta_p^2=.01$). The interactions between view and group ($F(1,65)=2.80$, $p=.09$, $\eta_p^2=.04$), valence and group ($F(1,65)=.13$, $p=.72$, $\eta_p^2=.01$), view and valence ($F(1,65)=1.78$, $p=.30$, $\eta_p^2=.02$), and valence, view and group ($F(1,65)=.68$, $p=.42$, $\eta_p^2=.01$) were all non-significant. If valence and group were ignored, participants did significantly better in the same view condition (*EMM* = 63.87) than the shifted view condition (*EMM* = 57.35). No post-hoc analysis was therefore conducted.

As the view by group interaction was approaching significance, the interaction plot was inspected (Figure 4). This suggested a trend towards people with depression performing better than the control group in the shifted view condition.

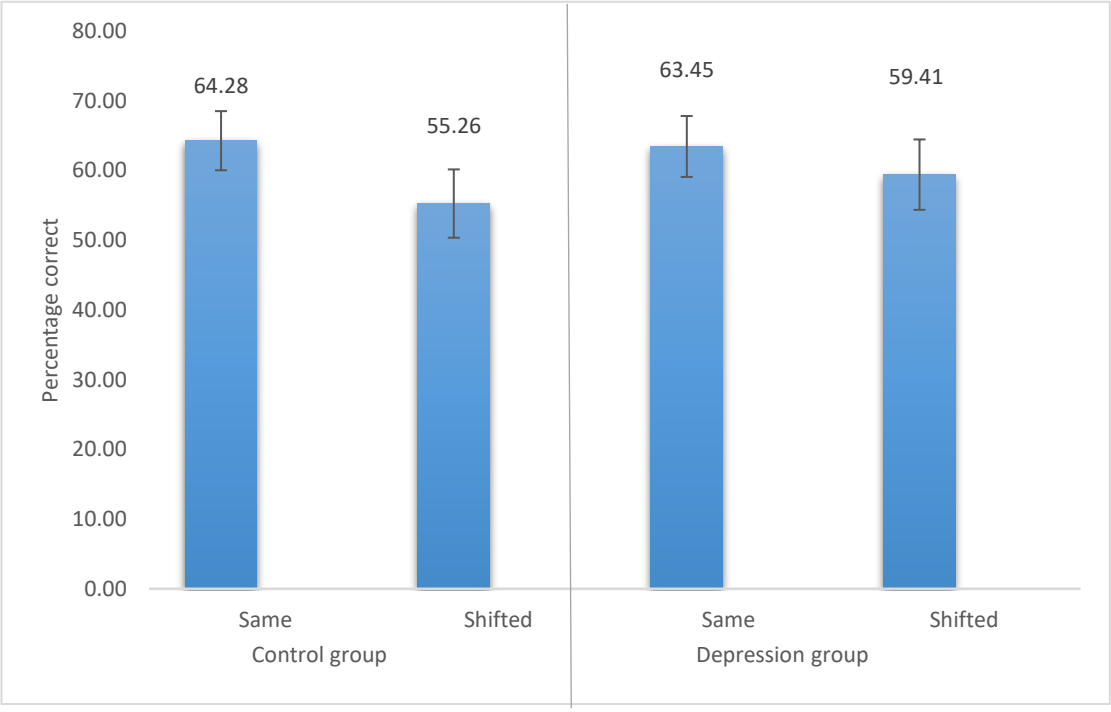


Figure 4. Mean percentage total correct for each group in same and shifted viewpoint conditions. Error bars represent 95% CI for mean.

4.3.1 Sex

As a way of investigating potential reasons for the trend above and given reported differences in terms of sex for spatial memory performance, sex was added to the model as a between subjects factor. There was a significant interaction between group, viewpoint and sex ($F(1,63)=5.12$, $p=.03$, $\eta_p^2=.08$). Inspection of the plots suggested this was driven by women with depression performing better in the shifted view condition than men with depression and male or female healthy controls (See Figure 5).

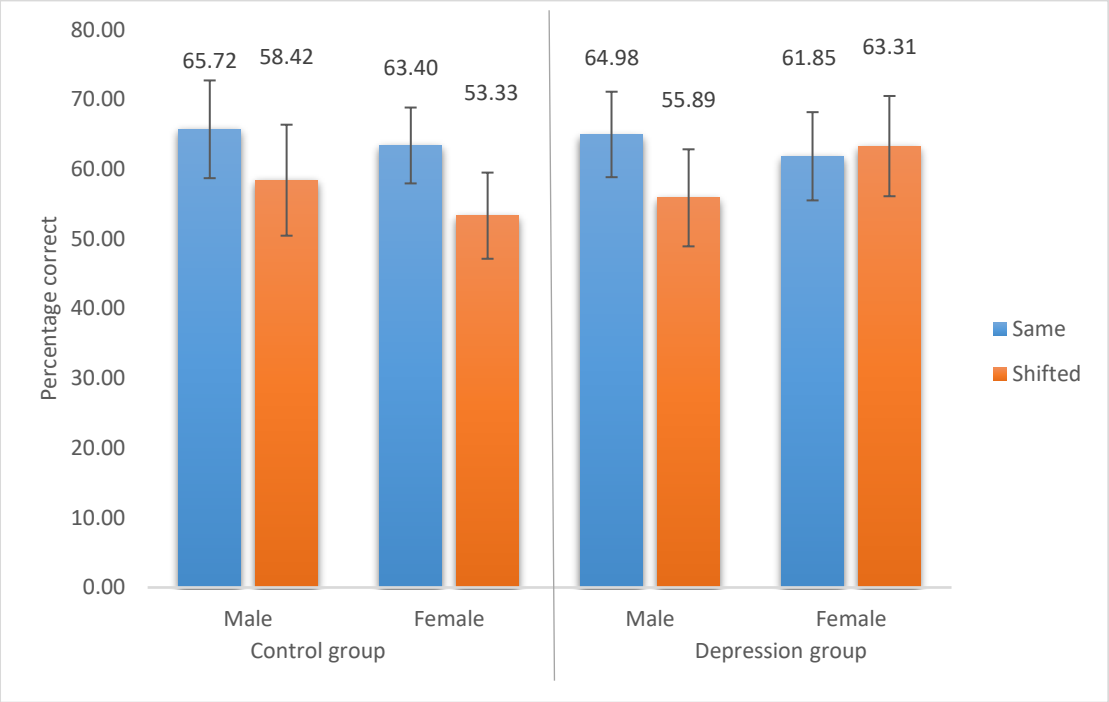


Figure 5. Mean percentage correct for each group by sex in same and shifted viewpoint conditions. Error bars represent 95% CI for mean.

A post-hoc analysis was therefore conducted for female participants only (control: N=22; depression: N=16). Overall shifted and same view performance was calculated as an average across valence factors. Initially, a 2 (between; group: control; depression) x 2 (within; viewpoint: same; shifted) mixed ANOVA with age as a covariate was conducted. There was no main effect of viewpoint ($F(1,35)=.65$, $p=.43$, $\eta_p^2=.02$) or group ($F(1,35)=1.41$, $p=.24$, $\eta_p^2=.04$) but there was an interaction between viewpoint and group, as might have been expected from the plots ($F(1,35)=9.65$, $p=.004$, $\eta_p^2=.22$). A between groups 2 (group: control; depression) x 2 (viewpoint: same; shifted) ANOVA with age as a covariate was then used to investigate this interaction. There was a main effect of group ($F(2,34)=5.02$, $p=.01$, $\eta_p^2=.23$). Analysis of simple effects revealed no viewpoint simple effect in the same view condition ($F(1.35)=.26$, $p=.62$, $\eta_p^2=.01$), that is no difference in performance between women with depression ($M=60.54$; $SD=16.90$) and in the control group (M=65.34, SD=10.30) when viewpoint was kept the same between presentation and test. However,

women with depression (*M*=62.24, *SD*=15.64) performed significantly better when viewpoint was shifted between presentation and test than healthy control females (*M*=54.92; *SD*=13.62; *F*(1.35)=5.92, *p*=.02, η_p^2=.15).

4.3.2 Order

To test for any effect of the order in which valenced stimuli were presented, order was added to the model (with age as a covariate but removing sex). There was no main effect of order (*F*(1,63)=.06, *p*=.82, η_p^2=.001). However, there was a significant interaction between valence and order (*F*(1,63)=4.21, *p*=.04, η_p^2=.06). In other words, if group and viewpoint were ignored, performance with neutral and negative stimuli were significantly different depending on which stimuli were presented first (Figure 6).

To test this interaction, overall performance with neutral and negative stimuli was calculated by averaging performance across viewpoint conditions. These figures were then adjusted for age. When neutral stimuli were presented first, a paired t-test showed no difference in overall performance with neutral stimuli (*M*=59.44, *SD*=15.51) and negative stimuli (*M*=59.80, *SD*=14.70; *t*(33)=.21, *p*=.83; 95%CI [-3.87-3.13]). When negative stimuli were presented first, there was a significant difference; participants performed better overall with neutral (*M*=64.27, *SD*=16.47) than negative stimuli (*M*=58.82, *SD*=12.25; *t*(33)=2.70, *p*=.01; 95%CI [1.34-9.57]).

Concerns about power meant no additional demographic and clinical variables were added to the model.

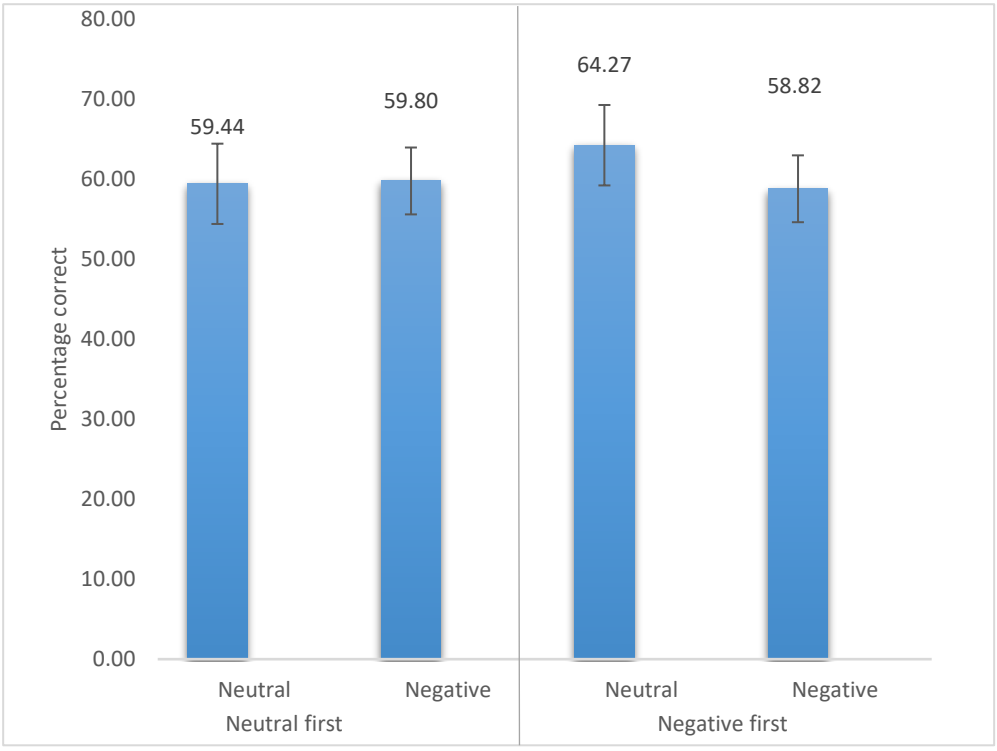


Figure 6. Mean percentage correct for valence condition by order of stimuli presentation. Error bars represent 95% CI for mean.

4.4 Controlled allocentric memory performance

The shifted-view minus same-view percentage correct score was used to give an indication of allocentric memory performance, controlling for egocentric memory performance. Figure 7 shows mean controlled allocentric performance by group and valence. This was used as the dependent variable in a mixed 2 (group: depression, control) x2 (valence: neutral, negative) mixed model ANOVA, once again controlling for age. The ANOVA did not show a main effect of group ($F(1,65)=2.80$, $p=.10$, $\eta_p^2=.04$) or valence ($F(1,65)=1.08$, $p=.3$, $\eta_p^2=.02$). There was also no interaction between group and valence ($F(1,65)=.66$, $p=.42$, $\eta_p^2=.01$). That is, there was no significant difference in controlled allocentric performance, both overall and across groups and valence conditions.

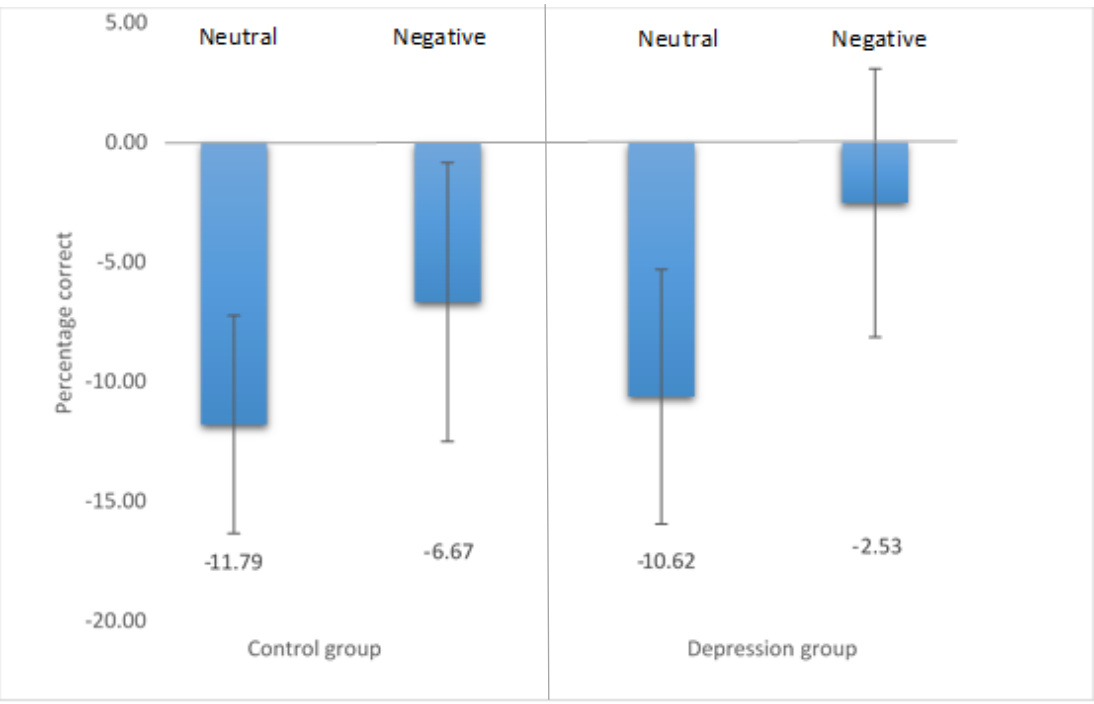


Figure 7. Mean controlled allocentric performance by group and valence condition. Error bars represent 95% CI for mean.

4.4.1 Sex & order

As for the 2x2x2 model, sex was added to the model. The main effect of sex was not significant ($F(1,63)=.76$, $p=.39$, $\eta_p^2=.01$), nor were any interactions (all $ps>.05$). Performance was similar, overall and across groups and valence conditions, for men and women. This was also true when presentation order was added to the model (without sex). The main effect was not significant ($F(1,63)=1.79$, $p=.19$, $\eta_p^2=.03$), nor were any interaction terms (all $ps>.05$). The order in which stimuli were presented did not make a difference to controlled allocentric performance, both overall and across groups and valence conditions.

4.5 Predictors of controlled allocentric performance for people with depression

This was an exploratory analysis to identify predictors of controlled allocentric performance for people with depression. Hierarchical multiple regressions were

conducted. Potential demographic confounds (age, sex, English as a first language, employment and education) formed the first block. Employment (employed; unemployed or student) and education (less than degree level; degree level or higher) were collapsed into two categories for ease of interpretation. The second block contained depression (BDI-I) and anxiety (GAD-7) symptomology scores. The third block contained depression characteristics (total number of depressive episodes, time since first episode, length of current episode, whether medication is being taken, whether undergoing talking therapy and whether a first or second degree relative had a history of depression). The multiple regression was conducted for overall controlled allocentric performance (calculated by averaging neutral and negative valence conditions), and for neutral and negative valence conditions separately.

Table 3

Results of hierarchical regression of controlled allocentric performance for people with depression

	Overall			With neutral stimuli			With negative stimuli		
	F/FΔ	*df*	$R^2/R^2\Delta$	*F/FΔ*	*df*	$R^2/R^2\Delta$	*F/FΔ*	*df*	$R^2/R^2\Delta$
Block 1 *Demographic variables*	4.37*	5,21	.51	2.14	5,21	.08	5.87*	5,21	.58
Block 2 *Symptomology variables*	.79	2,19	.06	.46	2,19	.05	2.05	2,19	.07
Block 3 *Depression characteristics*	.74	6,12	.12	1.20	6,12	.16	.64	6,12	.08
Reduced model *Age + sex*	6.76**	2,30	.31	-	-	-	10.91**	2,30	.42

Block 1: test statistics = F, R^2 Blocks 2 & 3: test statistics = $F\Delta$, $R^2\Delta$ (model change) *p<.01 **p<.001

Table 4

Reduced regression models of controlled allocentric performance for people with depression

	Overall			With negative stimuli		
	t	*p*	*β*	*t*	*p*	*β*
Age	2.35	.03	.36	3.12	.004	.43
Sex	2.75	.01	.42	3.37	.002	.47

Table 3 summarises the results of the hierarchical regression. For both overall controlled allocentric performance and performance with negative stimuli, only the model containing potential demographic confounds was a significant predictor, addition of the other blocks of variables did not significantly add to the model's predictive power. For controlled allocentric performance with neutral stimuli, it was not possible to predict controlled allocentric performance with neutral stimuli with any of the variables entered into the model. Other factors must have accounted for variability in performance in this condition.

For both overall controlled egocentric performance and with negative stimuli, the reduced model contained the two variables that significantly predicted performance: age and sex (Table 4). For both, sex had the strongest relationship with performance. There were no issues with collinearity in either model (tolerance >.1 and VIF <10 for both variables).

The reduced model regression equations were:

Overall controlled allocentric performance = -21.08 + .35 age + 10.34 sex

Controlled allocentric performance with negative stimuli = -28.33 + .54 age + 14.56 sex

For both, the negative constant implied controlled allocentric performance was impaired compared to egocentric. According to the equations, increasing age improved overall controlled allocentric performance by .35% per year, and by .54% per year of age with negative stimuli. Being female (coded as 1) increased overall performance by 10.34% and by 14.56% with negative stimuli, compared to being male. Effects on overall performance were the opposite to what would be expected from previous research (Raz, Ghisletta, Rodrigue, Kennedy & Lindenberger, 2010; Persson et al., 2013).

4.6 Associations with neuroticism

Pearson correlations were used to investigate associations between control group performance and control group neuroticism scores on the TIPI. There were no significant correlations between neuroticism scores and performance on the Town Square Test, by valence, viewpoint, valence and viewpoint, controlled allocentric performance by valence or total controlled allocentric performance (all *ps*>.05).

5 DISCUSSION

This study used a word-based version of the Town Square test to investigate the relative effects on hippocampus-dependent allocentric memory of using neutral and negatively valenced stimuli, for a depression group and control group. The aim was to bring together for the first time research showing memory biases towards negatively valenced stimuli and impairments in allocentric memory performance amongst people with depression to allow their interaction to be investigated. This section will first provide a summary of the results, relating these back to the study's aims and hypotheses, before looking to provide some insights into the pattern of results. It will then consider limitations, clinical implications and finally suggestions for future research.

5.1 Summary of research findings: primary aim

The primary aim of the research was to investigate any effects on Town Square Test performance of using neutral and negative stimuli and any differences between depression and control groups. As this was the first time this interaction had been studied, one tentative hypothesis was proposed for neutral stimuli. Following Wong's (2015) findings, it was predicted that, for people with depression, performance would be worse in the shifted view condition because of impairments in hippocampal functioning required for allocentric processing in the shifted view condition than the same view (Schmaal et al.,

2015). The control group would show no significant difference between shifted and same viewpoint performance. The impairment in shifted view for people with depression would also impact on allocentric processing when egocentric processing was controlled for, which would be worse than for healthy controls.

For negative stimuli, two alternative hypotheses were tested. The enhancement hypothesis, based on the modulation hypothesis (Dolcos et al., 2004), predicted people with depression would perform significantly better with negative stimuli in the shifted view condition, both when compared to the same view condition and to healthy controls. This would also be reflected in controlled allocentric performance, which would be better for people with depression than healthy controls.

The interference hypothesis, based on the affective interference hypothesis (Gottlib & Joorman, 2010), predicted people with depression would perform significantly worse in the shifted view condition with negative stimuli, both when compared to performance in the same view, and to healthy controls. Controlled allocentric performance would also be significantly worse for people with depression than healthy controls.

While it could be said there was evidence to support the hypothesised lack of difference between shifted and same view performance for control participants with neutral stimuli, there was no evidence to support other predictions. The only effect of valence found in any analysis was an order effect, whereby participants performed better overall with neutral than negative stimuli when negative stimuli were presented first. There were no other effects of valence for either group in terms of same view (egocentric), shifted view (uncontrolled allocentric) or controlled allocentric performance. A main effect of viewpoint was found, whereby participants performed worse on the shifted than the same view condition. However, there were no interactions between viewpoint, group and valence, that is viewpoint made a difference only at the overall performance across valence

conditions and groups level; there were no differences between groups and valence conditions.

There was also a trend towards people with depression performing better in the shifted viewpoint condition than same. This led to sex being added to the model and a sub-analysis looking only at women. Women with depression performed better than controls in the shifted viewpoint (uncontrolled allocentric) condition, but not in the same. However, once again, there were no valence effects.

For order, there were no differences between viewpoint conditions and no difference between groups. There was also no order effect for controlled allocentric performance. Stimuli order was randomised due to evidence of a carryover effect in affective bias research (Siegle et al., 2002). When people with depression are exposed to negative stimuli, the enhanced memory effect can persist beyond presentation, affecting performance on later trials. The order effect found whereby participants performed better overall with neutral than negative stimuli when negative stimuli were presented first, would suggest a similar carryover in this study. However, given the lack of interaction with either group or viewpoint plus the null results in terms of valence effects on performance elsewhere, it was difficult to provide an explanation or be more than tentative.

5.2 Summary of findings: secondary aim

The secondary aim was to explore predictors of controlled allocentric performance for people with depression. Only age and sex were found to predict any type of performance, predicting both overall performance and with negative stimuli. None of the clinical variables showed any influence on performance. The lack of association with number of depressive episodes did not support the recent meta-analysis (Schmaal et al., 2015), which found this predicted hippocampal damage. However, the lack of association with depression symptomology, co-morbid anxiety and use of anti-depressants did support

the findings (Schmaal et al., 2015). Wong (2015) also found no association between performance and depression severity or time since first episode, but the study did find an association with anxiety symptomology, unlike here. However, given the lack of valence effects in the main analysis, these supportive findings should be interpreted cautiously.

In terms of sex being a predictor, this supported the finding that females with depression performed better in the shifted view condition when compared to control group females. For both overall controlled allocentric performance and with negative stimuli, being female predicted better performance. It was not, however, a predictor of controlled allocentric performance with neutral stimuli.

The age effect was also interesting. In general population studies, hippocampal functioning has been found to decline with age (Raz et al., 2010), and indeed this is why age was included as a covariate in the main analyses. However, here, as age increased, so did overall controlled allocentric performance and with negative stimuli. Schmaal et al. (2015) did not find age was a moderator of reductions in hippocampal volume. Early onset before the age of 21 was associated with smaller hippocampal volumes, suggesting it was age factors relating to depression that impacted on hippocampal volume, rather than the person's age per se. This would allow for different age effects on hippocampally-dependent spatial functioning for people with depression than the general population.

5.3 Balancing alternative hypotheses

Two alternative sets of hypotheses for the interaction between valence and allocentric memory performance with negative stimuli were proposed with two different, and opposite, sets of predictions. The results of this study supported neither. There were three possible explanations for this. One was that both potential effects – enhancement and interference – occurred simultaneously, cancelling each other out and accounting for the null results. Another was that one or other was correct but confounds in the study

design meant effects were not seen. Finally, that neither was correct and the null effects were attributable to other factors.

5.4 Potential confound: use of words

To attempt to account for the null results, a number of potential confounds in the study were considered. The first, and perhaps the most important, was the use, for the first time, of words within the Town Square Test.

Clearly, reading words uses different cognitive processes to processing images. Reading English words uses the working memory's phonological loop and does not require visuo-spatial analysis to process images (Baddely, 1986). It has been suggested spatial location information is therefore more likely to be automatically encoded with memory representations of images than words (Pezdek, Romow & Sobolik, 1986) and there is evidence spatial memory for images is better than for English words (Pezdek et al., 1986). While both processes involve both sides of the hippocampus, there is a tendency for different areas to be involved to a greater or lesser extent. The right hippocampus has a central role in non-verbal visual memory (Milner, 1968) as well as spatial location. The left hippocampus is more involved in processing verbal information (O'Keefe and Nadel, 1978).

However, as words were used for all conditions in the Town Square Test, this could not account for the null results, and so confounds related to words used need to be considered.

5.4.1 Potential word confound: valence and arousal

There is evidence to suggest valence and arousal have different effects on memory, utilising different cognitive and neural processes. Kensinger and Corkin (2004) used a word recognition paradigm with healthy participants. Words were either neutral, arousing negative (e.g. rape, slaughter), or non-arousing negative (e.g. sorrow, mourning). Participants remembered more negative words than neutral, whether or not they were

arousing. However, using imaging, the study found different brain networks were used for arousing and non-arousing negative words.

For arousing words, and in line with the modulation hypothesis, an amygdala-hippocampus network was activated. With stimuli that were valenced, but not arousing, a different network of brain areas was involved, the prefrontal cortex and hippocampus. The amygdala-hippocampus network is relatively automatic and linked to an assessment of threat. A secondary encoding task competing for resources did not interfere with this effect. The prefrontal cortex-hippocampus network is involved in more controlled and elaborative encoding, such as autobiographical or semantic elaboration or extra rehearsal. It is therefore less automatic and the effect on memory was reduced with a concurrent task. The authors suggested confounding arousal and valence could help explain contradictory findings within the literature (e.g. Dolcos, Graham, LaBar & Cabeza, 2003; Hamann, Ely, Grafton & Kilts, 1999).

The importance of arousal is supported by Mather et al. (2006), when pictures were used in a spatial location memory test. Valence did not affect recognition memory for location, but arousal did, with location memory reducing as arousal increased.

Kensinger and Corkin's (2004) findings may have implications for the findings of this study, while noting theirs was a word recognition and not a spatial memory task and also involved healthy controls rather than depressed participants. In the current study, words were taken from a normed database of valence ratings (Warriner et al., 2013) and arousal was not controlled. The list included words that could be considered to be arousing and non-arousing. Arousal could therefore have acted as a confound.

Here, processing valence was task-irrelevant and so the Town Square Test itself could be seen as a competing task. With arousing, negatively valenced words, the concurrent demands of the spatial memory task would not interfere with the automatic

enhanced memory effect due to activation of the amygdala-hippocampus network. It could therefore be hypothesised, performance on the spatial memory test would be enhanced, as suggested by the modulation hypothesis. With non-arousing, negatively valenced words, the spatial memory task would interfere with the more elaborate encoding processes of the prefrontal cortex-hippocampus network, and no enhanced memory effect would occur, as suggested by the affective interference hypothesis. If there was a similar proportion of arousing and non-arousing stimuli, these two effects may have cancelled each other out, accounting for the null effects. This would certainly warrant further investigation in future studies, comparing arousal as well as valence conditions to see whether a negative, high arousal condition would produce results in line with the modulation hypotheses.

5.4.2 Potential word confound: valence not processed

Another potential reason for the lack of effects is that valence was not processed. Valence was irrelevant to the task. In an attempt to prompt people to process valence as well as location, participants were told they would be asked to rate word valence at the end of the task. This may not have been enough. How far this could attribute for the null findings is debatable. It has been suggested assessment of valence is pre-attentive, occurring automatically as part of our threat appraisal system and involving the amygdala (Öhman, 1992). When people were asked to physically encode stimuli (e.g. count vowels), valence only led to enhanced recognition memory with positive stimuli, there was no effect for negative stimuli. With more effortful, semantic processing, the bias disappeared (Ferré, 2003). The authors drew on mood-congruence at the time of encoding (Bower, 1981) to explain this finding. Because this bias is at the level of encoding and pre-attentive, it disappears with effortful, semantic processing.

This could account for the lack of valence effect within the control group, as only neutral and negative stimuli were used. Control participants may, therefore, not have processed valence, automatically or otherwise. While this research has not been replicated for a depression group, applying the same assumptions of mood congruence, it would be expected that the depression group processed valence with negative stimuli but not neutral. In this case, other potential confounds, such as arousal and valence may have come into play to account for the null results. Repeating the study, controlling for arousal, may allow these effects to be isolated.

5.4.3 Potential word confound: subjective valence

While participants were asked to rate word valence, it was not possible within the time limits of this study to input and analyse these data. What is more, the collection of participants' subjective valence ratings was only originally intended to be a manipulation to encourage participants to process stimuli valence, there would also be concerns about the quality of these data.

As words were supposed to be negative or neutral, responses would have been expected to be between 1 (completely unhappy) and 5/6 (neutral). However, a visual check suggested a broader range of responses, including positive ratings up to the maximum of nine. Word valence is of course subjective and will be influenced by experiences brought to mind by the word. It would be interesting to explore whether subjective valences of words rated as neutral within the database (Warriner et al., 2013) are influenced by depression.

As the depression group had been asked about their depression history at the start of testing, this may have been activated and have affected the rating of purportedly neutral words. Other studies with valenced stimuli used participant valence ratings as a variable in

the analysis (e.g. Kensinger & Corkin, 2004). It would be interesting to see in future studies if using individuals' own ratings of valence would produce the hypothesised effects.

5.4.4 Potential word confound: types of words used

The list of words used in the Town Square Test was a mixture of concrete nouns (e.g. barrel, maggot) and adjectives and adverbs (e.g. mouldy, cunning; see Appendix D). Clark and Teasdale (1985) found, for people with depression, bias towards negative words was strongest for negative trait adjectives than other kinds of words. Therefore, types of words used in the study could also have acted as a confound.

5.4.5 Potential word confound: English as a second language

While there was no difference between the groups in numbers with English as a first language, differences within groups were not tested. However, English as a first language was not a predictor of performance for people with depression. Researchers were occasionally asked during the valence rating task if it was acceptable for participants without English as a first language to translate words into their native language to assess their emotional reaction. It would be interesting to repeat the study with a group of native English speakers to assess any possible confounding effects.

5.5 Potential confound: co-morbid anxiety

Of the depression group, 80% met criteria for co-morbid GAD. There is a paucity of research looking at how the presentation of depression may differ with co-morbid anxiety. Wong (2015) found increased anxiety in the depression sample was associated with reduced controlled allocentric memory performance, and therefore an implied link to hippocampal impairment. However, Schmaal et al.'s (2015) findings suggested no impact of co-morbid anxiety on hippocampal volumes. Linked to the modulation hypothesis, an fMRI study with a valenced word memory paradigm (van Tol et al., 2012) found hyperactivation of the right hippocampus and amygdala during encoding of negative words

only for those with depression, and not with depression and anxiety. Given high levels of co-morbidity in this sample, this could have contributed to the lack of valence effects on performance.

High levels of co-morbid anxiety precluded any subanalysis. It would be worth samples for future studies having a better balance of people with and without co-morbid GAD as there may be differences in how valence affects performance that would be worth exploring.

5.6 Sex

This study produced three interesting sex differences warranting further investigation. Firstly, women with depression performed better in the shifted view condition than women in the control group. Secondly, being a woman with depression predicted better overall controlled allocentric performance and with negative stimuli. Thirdly, sex was not a predictor of performance with neutral stimuli. There are sex differences in depression. Women are twice as likely as men to have depression (Kessler, McGonagle, Swartz, Blazer & Nelson, 1993; Kuhnert, 2003) and show different symptomology, for example more insomnia, somatic symptoms, and increased appetite/weight gain (Schuch, Roest, Nolen, Penninx & de Jonge, 2014). Whether there are sex differences in memory or cognitive biases found for people with depression is under-researched. Some behavioural affective memory studies (e.g. Siegle et al., 2002) found no differences in terms of sex and there were no sex differences in hippocampal volume impairment in Schmaal et al.'s (2015) meta-analysis. However, many studies did not look at sex differences (e.g. Mather et al., 2006). For the general population, there is evidence of a male advantage in spatial memory (e.g. Persson et al., 2013), but some studies found women performed better in object location tests when there was also a verbal component

(Voyer, Postma, Brake, & Imperato-McGinley2007), as there was here, which may help to explain this finding.

The study's sample size precluded any subanalysis to see whether females with depression were different demographically to control group females, or clinically from depression group males. However, the lack of associations between variables other than age and sex with controlled allocentric performance may suggest, even if this analysis had been possible, differences in these variables would not account for the enhanced performance of females with depression. Future studies with larger sample sizes allowing sex differences to be more fully explored would therefore be desirable.

5.7 Additional limitations

5.7.1 Sample size

The sample size was in line with the power analysis and therefore the study should have been sufficiently powered to detect a small to medium effect. However, the sample size did preclude certain subanalyses that may have helped to develop explanations for the unexpected null effects, including investigation of sex effects.

5.7.2 Self-report of depression history

The study used self-report to gain participants' depression history. This information was not clinically corroborated. It was clear participants found questions such as number of previous depressive episodes difficult to answer accurately. This difficulty increased with the length of time taken to receive a formal diagnosis. For many, formal diagnosis confirmed their own assessment of subjective mood over a number of years. The estimate of episodes therefore included the pre-diagnostic period, and was not clinician endorsed. It is difficult to know how this might be addressed in future, but it certainly limited the accuracy of clinical data used in the analysis.

5.7.3 Use of mnemonic strategies

The use of words in the Town Square Test was also intended to reduce the use of mnemonic strategies to provide a purer indication of hippocampally-dependent performance. Use of strategies was not routinely asked about but some participants, who performed particularly well on the test, reported using strategies, for example using words to develop stories linked to their position relative to the four corners of the courtyard. While it is not known how far such strategies were used, they could have acted as a potential confound on performance. Future studies could consider specifically asking about strategy use so their influence on performance could be tested.

5.8 Clinical implications

The impact of valence on allocentric spatial memory is an important question to answer as it may help elucidate some of the mechanisms underlying depression, given links to hippocampal functioning and how this may account for memory effects, most notably OGM. Increasing knowledge may lead to more effective therapeutic approaches, improving rates of remission and relapse.

Unfortunately, the lack of valence effects, and confounds potentially accounting for the null results, means further research into the interaction between valenced stimuli and hippocampally-dependent spatial memory is necessary before any clinical implications can be explored.

5.9 Further studies

A number of suggestions have been made for further studies in this area. Perhaps most important would be to control for arousal and use participants' own subjective ratings of valence in the analysis. Finally, a larger sample size, one with more people without co-morbid anxiety, would allow this, and potential sex differences, to be explored.

5.10 Conclusion

While mood congruent effects in depression are well documented, the nature of these differs depending on a number of factors, including the task-relevance of valence, as does the theoretical approaches recruited to explain them. The aim of this study was to investigate the effects of valenced stimuli on hippocampally-dependent allocentric spatial memory to further elucidate these effects. However, no effects of valence were found on performance on the Town Square Test, for either a depression group or healthy controls. Hypothesised reasons for these null results included: confounding valence and arousal; a lack of processing of valence; the subjective nature of valence; and the presence of high levels of co-morbid anxiety in the depression group. For this reason, further research is required, so the influence of some of these potential confounds can be explored. This would be important to do given the potential clinical implications for research in this area, in terms of improving understanding of the cognitive and neurobiological processes underlying depression, which may in turn lead to improvements in treatment approaches.

6 REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Washington, DC: Author.

Arie, M., Apter, A., Orbach, I., Yefet, Y., & Zalzman, G. (2008). Autobiographical memory, interpersonal problem solving, and suicidal behavior in adolescent inpatients. *Comprehensive Psychiatry*, *49*(1), 22-29.

Arnone, D., McIntosh, A.M., Ebmeier, K.P., Munafò, M.R., Anderson, I.M. (2012). Magnetic resonance imaging studies in unipolar depression: systematic review and metaregression analyses. *European Neuropsychopharmacology, 22,* 1 –16.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Clarendon Press

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper and Row. Republished as Beck, A. T. (1970). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press

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

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588-597.

Bench, C. J., Friston, K. J., Brown, R. G., Frackowiak, R. S. J., & Dolan, R. J. (1993). Regional cerebral blood flow in depression measured by positron emission tomography: The relationship with clinical dimensions. *Psychological Medicine, 23,* 579-590.

Bisby, J. A., King, J. A., Brewin, C. R., Burgess, N., & Curran, H. V. (2010). Acute effects of alcohol on intrusive memory development and viewpoint dependence in spatial memory support a dual representation model. *Biological Psychiatry, 68,* 280-286.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*(2), 129.

Burgess, N. (2008). Spatial cognition and the brain. *New York Academy of Science, 1124*, 77-97

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Cameron, H.A., & McKay, R.D. (1999). Restoring production of hippocampal neurons in old age. *Nature Neuroscience, 2,* 894-7.

Campbell, S., & MacQueen, G. (2004). The role of the hippocampus in the pathophysiology of major depression. *Journal of Psychiatry Neuroscience, 29,* 417–426.

Castaneda, A. E., Suvisaari, J., Marttunen, M., Perälä, J., Saarni, S. I., Aalto-Setälä, T., ... & Tuulio-Henriksson, A. (2008). Cognitive functioning in a population-based sample of young adults with a history of non-psychotic unipolar depressive disorders without psychiatric comorbidity. *Journal of Affective Disorders*, *110*(1), 36-45.

Clark, L., Chamberlain, S. R., & Sahakian, B. J. (2009). Neurocognitive mechanisms in depression: implications for treatment. *Annual Review of Neuroscience*, *32*, 57-74.

Clark, D. M., & Teasdale, J. D. (1985). Constraints on the effects of mood on memory. *Journal of Personality and Social Psychology*, *48*(6), 1595.

Cornwell, B. R., Salvadore, G., Colon-Rosario, V., Latov, D. R., Holroyd, T., Carver, F. W., ... & Grillon, C. (2010). Abnormal hippocampal functioning and impaired spatial navigation in depressed individuals: evidence from whole-head magnetoencephalography. *American Journal of Psychiatry*, *167*(7), 836-844.

Cosi, C., Spoerri, P.E., Comelli, M.C., Guidolin, D., & Skaper, S.D. (1993). Glucocorticoids depress activity-dependent expression of BDNF mRNA in hippocampal neurones. *Neuroreport, 4*(5), 527–530

Craighead, W. E., Sheets, E. S., Brosse, A. L., & Ilardi, S. S. (2007). Psychosocial treatments for major depressive disorder. In P. E. Nathan & J. M. Gorman (Eds.), *A Guide to Treatments That Work* (3rd ed., pp. 289–308). New York: Oxford.

Doeller, C. F., Barry, C., & Burgess, N. (2012). From cell to systems: Grids and boundaries in spatial memory. *The Neuroscientist, 18,* 556-566.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*(5), 855-863.

Dolcos, F., Graham, R., LaBar, K., & Cabeza, R. (2003). Coactivation of the amygdala and hippocampus predicts better recall for emotional than for neutral pictures. *Brain and Cognition*, *51*(2), 221-223.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191.

Ferré, P. (2003). Effects of level of processing on memory for affectively valenced words. *Cognition and Emotion*, *17*(6), 859-880.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders: SCID-I.* Washington, DC: American Psychiatric Press

Fu, C. H., Williams, S. C., Cleare, A. J., Brammer, M. J., Walsh, N. D., Kim, J., ... & Bullmore, E. T. (2004). Attenuation of the neural response to sad faces in major depression by antidepressant treatment: a prospective, event-related functional magnetic resonance imaging study. *Archives of General Psychiatry*, *61*(9), 877-889.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285-312.

Gould, N. F., Holmes, M. K., Fantie, B. D., Luckenbaugh, D. A., Pine, D. S., Gould, T., . . . Zarate, C. A. (2007). Performance on a virtual reality spatial memory navigation task in depressed patients. *American Journal of Psychiatry, 164,* 516-519.

Grön, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., & Riepe, M. W. (2000). Brain activation during human navigation: gender-different neural networks as substrate of performance. *Nature Neuroscience*, *3*(4), 404-408.

Hamilton, J. P., & Gotlib, I. H. (2008). Neural substrates of increased memory sensitivity for negative stimuli in major depression. *Biological Psychiatry*, *63*(12), 1155-1162.

Hamann, S. B., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, *2*(3), 289-293.

Hartley T., Bird C. M., Chan D., Cipolotti L., Husain M., Vargha-Khadem, F., & Burgess N. (2007). The hippocampus is required for short-term topographical memory in humans. *Hippocampus, 17,* 34–48.

Joormann, J., & Gotlib, I. H. (2008). Updating the contents of working memory in depression: interference from irrelevant negative material. *Journal of Abnormal Psychology*, *117*(1), 182.

Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression: relation to cognitive inhibition. *Cognition and Emotion*, *24*(2), 281-298.

Judd, L. L., Akiskal, H. S., Zeller, P. J., Paulus, M., Leon, A. C., Maser, J. D., et al. (2000). Psychosocial disability during the long-term course of unipolar major depressive disorder. *Archives of General Psychiatry, 57*, 375–380.

Kendler, K.S., Gardner, C.O., & Prescott, C.A. (2002). Toward a comprehensive developmental model for major depression in women. *American Journal of Psychiatry, 159,* 1133–1145.

Kendler, K.S., Kessler, R.C., Neale, M.C., Heath, A.C., & Eaves L.J. (1993). The prediction of major depression in women: toward an integrated etiologic model. *American Journal of Psychiatry, 150,* 1139–1148.

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 3310-3315.

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

Kessler, R. C., McGonagle, K. A., Swartz, M., Blazer, D. G., & Nelson, C. B. (1993). Sex and depression in the National Comorbidity Survey I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders*, *29*(2), 85-96.

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). Human hippocampus and viewpoint dependence in spatial memory. *Hippocampus, 12*, 811-820.

King, J. A., Trinkler, I., Hartley, T., Vargha-Khadem, F., & Burgess, N. (2004). The hippocampal role in spatial memory and the familiarity-recollection distinction: A case study. *Neuropsychology*, *18*(3), 405.

Koolschijn, P.C.M.P., van Haren, N.E.M., Lensvelt-Mulders, G.J.L.M., Hulshoff-Pol, H.E., Kahn, R.S. (2009). Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Human Brain Mapping, 30,* 3719 –3735.

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*, 606- 613.

Kuehner, C. (2003). Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. *Acta Psychiatrica Scandinavica*, *108*(3), 163-174.

Lorenzetti, V., Allen, N.B., Fornito, A., Yücel, M. (2009). Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. *Journal of Affective Disorders, 117*, 1 –17.

MacQueen, G.M., Ramakrishnan, K., Ratnasingan, R., Chen, B., & Young, L.T. (2003). Desipramine treatment reduces the long-term behavioural and neurochemical sequelae of early maternal separation. *International Journal of Neuropsychopharmacology, 6*(4), 391-6.

Martinelli, P., Sperduti, M., & Piolino, P. (2013). Neural substrates of the self-memory system: New insights from a meta-analysis. *Human Brain Mapping*, *34*(7), 1515-1529.

Mather, M., Mitchell, K. J., Raye, C. L., Novak, D. L., Greene, E. J., & Johnson, M. K. (2006). Emotional arousal can impair feature binding in working memory. *Journal of Cognitive Neuroscience*, *18*(4), 614-625.

Mathews, A., Ridgeway, V., & Williamson, D. A. (1996). Evidence for attention to threatening stimuli in depression. *Behaviour Research and Therapy*, *34*(9), 695-705.

McGaugh, J. L. (2002). Memory consolidation and the amygdala: a systems perspective. *Trends in Neurosciences*, *25*(9), 456-461.

McKinnon, M.C., Yucel, K., Nazarov, A., MacQueen, G.M. (2009). A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. *Journal of Psychiatry Neuroscience, 34,* 41-54.

Milner, B. (1968). Visual recognition and recall after right temporal-lobe excision in man. *Neuropsychologia*, *6*(3), 191-209.

Mitterschiffthaler, M. T., Kumari, V., Malhi, G. S., Brown, R. G., Giampietro, V. P., Brammer, M. J., ... & Sharma, T. (2003). Neural response to pleasant stimuli in anhedonia: an fMRI study. *Neuroreport*, *14*(2), 177-182.

Moffitt, T. E., Harrington, H., Caspi, A., Kim-Cohen, J., Goldberg, D., Gregory, A. M., & Poulton, R. (2007). Depression and generalized anxiety disorder: Cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Archives of General Psychiatry, 64,* 651-660.

Öhman, A. (1992). Orienting and attention: Preferred preattentive processing of potentially phobic stimuli. *Attention and information processing in infants and adults. Perspectives from Human and Animal Research*, 263-295.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map.* Oxford, UK: Oxford University Press

Persson, J., Herlitz, A., Engman, J., Morell, A., Sjölie, D., Wikström, J., & Söderlund, H. (2013). Remembering our origin: Gender differences in spatial memory are reflected in gender differences in hippocampal lateralization. *Behavioural Brain Research*, *256*, 219-228.

Pezdek, K., Roman, Z., & Sobolik, K. G. (1986). Spatial memory for objects and words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12,* 530-537

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58,* 1-23

Raz N, Ghisletta P, Rodrgiue K, Kennedy K, Lindenberger U. 2010. Trajectories of brain aging in middle-age and older adults: regional and individual differences. *NeuroImage, 51,* 501–511.

Roberson-Nay, R., McClure, E. B., Monk, C. S., Nelson, E. E., Guyer, A. E., Fromm, S. J., ... & Pine, D. S. (2006). Increased amygdala activity during successful memory encoding in adolescent major depressive disorder: An FMRI study. *Biological Psychiatry*, *60*(9), 966-973.

Ruiz-Caballero, J. A., & Bermúdez, J. (1995). Neuroticism, mood, and retrieval of negative personal memories. *The Journal of General Psychology*, *122*(1), 29-35.

Sacher, J., Neumann, J., Funfstuck, T., Soliman, A., Villringer, A., & Schroeter, M. L. (2012). Mapping the depressed brain: A meta-analysis of structural and functional alterations in major depressive disorder. *Journal of Affective Disorders, 140*, 142-148.

Sagfors, L.F. (2017). *Autobiographical and hippocampus-dependent spatial memory in depression.* Unpublished doctoral dissertation. University College London, London.

Schmaal, L., Veltman, D. J., van Erp, T. G., Sämann, P. G., Frodl, T., Jahanshad, N., ... & Zurowski, B. (2015). Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Molecular Psychiatry*.

Schuch, J. J., Roest, A. M., Nolen, W. A., Penninx, B. W., & de Jonge, P. (2014). Gender differences in major depressive disorder: results from the Netherlands study of depression and anxiety. *Journal of Affective Disorders*, *156*, 156-163.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*, 11–21

Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A., & Carter, C. S. (2002). Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biological Psychiatry*, *51*(9), 693-707.

Smith, K. V., Burgess, N., Brewin, C. R., & King, J. A. (2015). Impaired allocentric spatial processing in posttraumatic stress disorder. *Neurobiology of Learning and Memory, 119*, 69-76.

Spitzer, R. L., Kroenke, K., & Williams, J. B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*, 1092-1097.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 356,* 1505-1515.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53,* 1-25.

van Tol, M. J., Demenescu, L. R., Van der Wee, N. J., Kortekaas, R., Marjan, N., Den Boer, J. A., ... & Veltman, D. J. (2012). Functional magnetic resonance imaging correlates of emotional word encoding and recognition in depression and anxiety disorders. *Biological Psychiatry*, *71*(7), 593-602.

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative metaanalysis of cognitive-behavioral therapy's effects. *Journal of Consulting and Clinical Psychology, 73*, 475–488.

Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review*, *14*(1), 23-38.

Wang, L., LaBar, K. S., Smoski, M., Rosenthal, M. Z., Dolcos, F., Lynch, T. R., ... & McCarthy, G. (2008). Prefrontal mechanisms for executive control over emotional distraction are altered in major depression. *Psychiatry Research: Neuroimaging*, *163*(2), 143-155.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.

Wilkins, L. K., Girard, T. A., King, J., King, M. J., Herdman, K. A., Christensen, B. K., & King, J. (2013). Spatial-memory deficit in schizophrenia spectrum disorders under viewpoint-independent demands in the virtual courtyard task. *Journal of Clinical and Experimental Neuropsychology*, *35*(10), 1082-1093.

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin, 133,* 122-148.

Williams J. M. G., & Broadbent K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology, 95,* 144–149.

Williams, J. M. G., & Scott, J. (1988). Autobiographical memory in depression. *Psychological Medicine*, *18*(3), 689-695.

Willner, P., Scheel-Krüger, J., & Belzung, C. (2013). The neurobiology of depression and antidepressant action. *Neuroscience & Biobehavioral Reviews*, *37*(10), 2331-2371.

Wong, Y.P. (2015). *An investigation of the links between depression and hippocampal-based allocentric spatial memory*. Unpublished doctoral dissertation. University College London, London.

Young, K. D., Bellgowan, P. S., Bodurka, J., & Drevets, W. C. (2013). Behavioral and neurophysiological correlates of autobiographical memory deficits in patients with depression and individuals at high risk for depression. *JAMA psychiatry*, *70*(7), 698-708.

Young, K. D., Bellgowan, P. S. F., Bodurka, J., & Drevets, W. C. (2014a). Neurophysiological correlates of autobiographical memory deficits in currently and formerly depressed subjects. *Psychological Medicine*, *44*(14), 2951-2963.

Young, K. D., Bellgowan, P. S. F., Bodurka, J., & Drevets, W. C. (2014b). Neurophysiological correlates of autobiographical memory deficits in currently and formerly depressed subjects. *Psychological Medicine*, *44*(14), 2951-2963.

Young, K. D., Erickson, K., Nugent, A. C., Fromm, S. J., Mallinger, A. G., Furey, M. L., & Drevets, W. C. (2012). Functional anatomy of autobiographical memory recall deficits in depression. *Psychological Medicine*, *42*(02), 345-357.

PART 3: CRITICAL APPRAISAL

1 Introduction

This critical appraisal will focus on the major research project but will include some brief reflections on the literature review at the end. It will cover the following areas: experiences of doing a joint project; developing a word-based version of the Town Square Test; recruitment issues; issues with screening and testing; and the researcher therapist divide. It will end with some brief reflections on both the major research project and the literature review.

2 Experiences of doing a joint project

I found doing my DClinPsy major research project jointly with another researcher incredibly effective. My joint researcher and I very quickly formed a good working alliance, one we managed to maintain throughout the project, even when we faced challenges with recruitment. I found there were many advantages to doing a joint project. From a practical point of view, sharing data collection meant we were able to have a bigger sample size and a specifically recruited control group, both improvements on the previous trainee's (Wong, 2015) project using the Town Square Test (King et al., 2002) with a depression group. It was always useful to be able to talk through decisions that needed to be made about measures, procedures and many other aspects, and developing the protocol for the study was very much a joint effort. The joint researcher's counsel was always valuable and useful.

When we were recruiting the depression group, the joint researcher and I had complementary placement days. This meant we could offer testing on more days increasing the chances of turning successful screening into completed testing opportunities. It also gave us more flexibility to offer to attend team meetings at the different locations of the Improving Access to Psychological Therapies (IAPT) service so we could present the project to clinicians. Some of the testing locations were quite difficult for

me to get to, but were closer to where the joint researcher lived, and so we agreed to split testing locations to ensure commuting times were not prohibitive for either of us. This is just one example of the pragmatic and ego-free working relationship we managed to forge.

The advantages of doing a joint project were not just practical. I believe my stress levels were significantly lower than they would have been if I had been doing the project alone. Not only was I able to share the tasks of the project, but I also had someone to voice and share frustrations with, as well as share the relief and achievement of meeting key milestones. I believe we motivated each other to keep going when recruitment for the depression group was proving long and arduous. Left to my own devices, I think I would probably have stopped earlier and settled for a sample size smaller than the one indicated in the power analysis as being the optimal one. But together we persevered and I am proud we managed to test all the participants in our original plan.

3 Developing a word-based version of the Town Square Test

When the decision was made for my project to investigate how differently valenced stimuli might affect performance on the Town Square Test, initially my supervisor and I planned to develop a new version with valenced images as images had been used on the Test to date. There are a number of image databases with normalised valenced ratings, including the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1999). Another University College London researcher who had used the Town Square Test before (Bisby, King, Brewin, Burgess & Curran, 2010) had piloted a version using valenced images taken from IAPS, but had not used it in a study. He told us he had found participants were not completely engaging with the valence of the images as this was secondary to task performance. In discussion with my supervisor, we identified several additional drawbacks with using valenced images. Within the Town Square Test, the placeholders stimuli are presented on are on the floor of the virtual environment. Some of these are at the back of

the Square. At this distance, my supervisor and I considered it would be difficult, especially at low image resolutions, to know with any level of certainty what the image was, including its valence. With previous versions, the nature of the image had not been important. However, with this version, it would be important for people to consistently and quickly recognise the image and its valence.

My supervisor then suggested using words rather than images. He found a normed database (Warriner, Kuperman, & Brysbaert, 2013) and I used this to develop a list of potential words. I did research into reading words as I did not want reading time for words to be a confound. Research by New, Ferrand, Pallier and Brysabaert (2006) found a U-shaped function when assessing reaction times in a naming task. Reaction times were the same for words between five and eight letters and so it was agreed to select words of this length. When the words were put onto the placeholders in the Square, words of eight letters sometimes overlapped when the four options were shown close together in the test condition, and so words of eight letters were removed from the set. As New et al. (2009) also found an inhibitory effect of 20m/s per syllable, disyllabic words were selected. My supervisor and I agreed to not formally control frequency, for example by using frequency ratings, but prioritised intuitively higher frequency words. In terms of the font used, there is evidence of no difference in reading speed between a fixed width font (e.g. Courier New) or a variable width font (e.g. Arial; Rayner, Slattery & Bélanger, 2010). Therefore, the font was chosen for clarity from the rear of the Town Square virtual environment.

When the selected words were put into the virtual environment, we considered them to readable even at the placeholders at the back of the Square. I highlighted, as the other researcher had found when piloting the valenced picture version, that the word itself was not vital to the task just its location, and there was therefore no need for people to process the valence of the word. We agreed the manipulation whereby people would be

told to pay attention to how the word made them feel before doing the Town Square Test so they could rate the words when given a list at the end.

It is difficult to know whether this manipulation was sufficient for people to process task-irrelevant word valence, or whether this in part accounts for the unexpected null results. We changed the Test in two ways concurrently: using words rather than images; and manipulating word valence. With hindsight, perhaps we should have investigated using words for the images used in the image based Town Square Test as the neutral word condition. Then we could have compared performance on this with the image based version. However, the valence of the original images was not controlled and we may have found the words for some of the images used were not have rated as sufficiently neutral on the normed word database (Warriner et al., 2013).

People's emotional reaction to words, and images, is subjective. We did ask people to rate the words after they had completed the Town Square Test. However, because recruitment of the depression group took longer than anticipated, it was not possible within the time constraints of the project to analyse these ratings separately. I did have a look at the rating sheets and they suggested people did not necessarily rate the words in line with normed valence ratings. If participants were rating the words in line with the norms, the maximum rating would be 6 (out of a possible 9, with 9 indicating a word making people feel completely happy). It was clear from looking at the rating sheets some participants rated some words as having clear positive valence, using the highest possible score of 9. It would be interesting to use people's subjective ratings of valence to classify the words and see if this made a difference to the analysis. Several other studies have used this approach (Kensinger & Corkin, 2004).

Words were also chosen to minimise any mnemonic strategies used by participants to remember the locations as these would mean performance was less reliant on

hippocampal processing. There was anecdotal evidence that some people were using strategies. One woman I tested, who did particularly well on the test, told me she used the words to develop stories linked to the four corners of the town square to help remember locations. The relatively large number of test items in predictable formats might encourage participants to develop such strategies to improve test performance. It is therefore difficult to know what changes could be made to further reduce the use of strategies that might interfere with the Town Square Test's ability to provide an indication of hippocampally-dependent allocentric memory.

Previous studies using the Town Square Test have used two list lengths of three and six items for the same and shifted view conditions (e.g. Wong, 2015; Smith, Burgess, Brewing & King, 2015). They have consistently found no difference in performance between the list lengths. This study had four conditions, with four combinations of valence and viewpoint (same-neutral, same-negative, shifted-neutral, shifted-negative), rather than the simpler two viewpoint conditions of previous studies. We therefore decided to have just one list length of six, to keep the total number of trials manageable. Even then, there were a total of 96 trials, 48 each for neutral and negative valence, with four blocks of six trials at same viewpoint and four of six for the shifted viewpoint.

4 Recruitment issues

When planning project recruitment, the joint researcher and I had anticipated recruitment of the control group might be more problematic than the group with depression. We had links with an IAPT service which had been used for UCL DClinPsy research before and felt confident we would meet our recruitment targets quickly. The depression group recruitment proved to be difficult and labour intensive, and took far longer than anticipated. We finally finished recruitment and testing of those with

depression at the end of March 2017. Our initial target had been the end of December 2016.

We had understood the last trainee had issues with referrals from clinicians which had held up recruitment of people with depression. We also understood, once the trainee had access to the service's list of people who had consented to being contacted for research purposes, she quickly recruited the final participants to her study. We therefore knew how important it was to gain early access to the research list.

There were a few delays to starting recruitment of the depression group. Firstly, the previous IAPT contact had left the service and so we had to identify someone else who could act as a sponsor for our research. We made contact with our new sponsor at the end of May 2016. The research had been presented by the previous contact and the service had agreed to support it in July 2015. However, when our new sponsor came on board it became apparent it would need to be re-presented at a research meeting. This finally happened in July 2016, where, once again, the service agreed to support the research.

Access to the research list was also delayed. We got access to the list at the end of September 2016. We were therefore initially reliant on referrals from clinicians. As agreed with our research sponsor, we presented our research at service team meetings (one in August 2016, but it was not possible to get a slot at another site until December 2016) and we made laminated cards summarising the research and criteria for participation for clinicians to keep on their desks as reminders. We had several meetings with the sponsor to discuss recruitment and to find out if there was anything more we could be doing to keep front of mind among clinicians at the different service sites. Overall, through this process, we received two referrals, one of which was not eligible. The previous trainee had received 17 referrals through this route, although seven were not eligible. I am not sure why we received so few clinician referrals. The sponsor implied there were a number of

concurrent research projects looking for clinician referrals and we might simply have been at the end of a long list. As these projects came to an end, we thought more clinician referrals would start to come through but this did not happen. The sponsor e-mailed certain clinicians who he thought might have appropriate clients, but even this direct contact did not increase clinician referrals.

This meant we were almost completely reliant on the list of people who had consented to being contacted for research purposes. Using this list turned out to be very labour intensive, involving several steps. We had a research assistant at the service who was responsible for providing us with lists taken from the service's central database. We provided the eligibility criteria for the research and discussed them on the phone. Lists were drawn at regular two to four week intervals across the recruitment period. When we received the lists, we needed to check them against inclusion criteria: aged under 65, with a PHQ-9 (Patient Health Questionnaire-9; Kroenke, Spitzer & Williams, 2001) score of more than 10. We then needed to ascertain where they were on the service's care pathway. Increasingly we found people on the lists had not yet been assessed or had not started therapy, or had been assessed but had been deemed to be inappropriate for the service and were in the process of being discharged. We had agreed, to effectively manage any emerging risk, we would not see people who had been discharged. We also did not contact people who scored more than 1 on question 9 of the PHQ-9 which asks about suicidality. The lists therefore held diminishing returns over time as we had to exclude more and more potential participants.

We then e-mailed potential participants with a summary of the study from my secure NHS e-mail account. Initially, we waited for people to contact us to say they were interested but increasingly realised very few people were spontaneously getting in touch, perhaps unsurprising given they had depression. We therefore started to make follow-up

calls to people we had e-mailed to ask them if they had seen the e-mail and to assess interest in participation. This was fruitful, but had a low call to contact ratio and was something neither I nor the joint researcher enjoyed doing.

Recruitment was sporadic. At times, we would become quite concerned about numbers but then we would have a flurry of contacts, successful screenings and testings, which we would find encouraging, only for our next set of e-mails to not generate as many contacts. In total, we e-mailed 282 people from the lists and screened 60, a response rate of 21%. Of these, 21 were not eligible, a rejection rate of 35%. Of the 39 who passed screening, we completed testing with 36, the other three cancelled appointments and it was not possible to rearrange. Of these, 33 were included in my analysis: one was left handed and two did not complete the Town Square Test. So, from initial e-mail contact to participant included in analysis, the response rate was 12%. Ultimately, it was a relief to meet our final recruitment target of end of March 2017.

5 Issues with screening and testing

At one point, the number of people with depression who did not meet inclusion criteria during telephone screening reached approximately 50%. This could be for a few reasons but most commonly because of co-morbid conditions, including psychosis. Given the length of time recruitment of this group was taking, this was frustrating as well as concerning. If the proportion being screened out had stayed the same, I don't believe it would have been possible to meet the numbers indicated by the power analysis within the necessary timeframe. The exclusion criteria in terms of co-morbidity had been set to minimise possible confounds. For example, a study (Smith et al., 2015) found people with post-traumatic stress disorder (PTSD) showed a similar deficit in performance in the Town Square Test versus healthy controls as found by Wong (2015) for people with depression. It was therefore important to ensure people with depression did not have co-morbid PTSD.

In many ways, it would have been desirable for the sample not to have co-morbid anxiety as this could also have acted as a confound. However, as 80% of the sample met clinical thresholds for GAD as measured by the GAD-7 (Generalised Anxiety Disorder-7; Spitzer, Kroenke & Williams, 2006), it would have been very difficult to recruit a depression sample without co-morbid GAD. Luckily, over time, the percentage of people being screened out fell and fewer were rejected at this stage. Overall 35% of those screened did not meet eligibility criteria and had to be excluded.

One potential participant complained about a screening call with the joint researcher. He did not meet some of the early criteria and so, according to our protocol, the joint researcher had ended the call after informing the participant they were not eligible to participate. The potential participant was unhappy with the way the call had been ended and e-mailed me to complain. The crux of his complaint appeared to be that the co-morbid disorder leading to him being excluded from our research was something that had resulted in him being excluded from many other things across his life. We informed our supervisor and the sponsor at the service and agreed I would call him to apologise. We hypothesised his concerns were not about how the joint researcher had conducted this specific call, but a broader issue of exclusion faced by the potential participant, and indeed many with mental health issues. This turned out to be the case when I spoke to the potential participant, and he was reassured and grateful for the call. After this, however, we did adjust our screening procedure for the depression group. We made sure we briefly explained why we have eligibility criteria in research (to reduce potential confounds) and that, if it became clear the potential participant did not meet criteria, we would end the call then and not take up more of their time than necessary. We had no further incidents.

At first, we did not include the need to be right handed in the initial e-mail we sent to potential depression group participants. This was an oversight. We had one participant who, it emerged during the screening call, was left handed. When we told her one of our criteria was that participants needed to be right handed, she queried this. We sought the advice of our supervisor and we agreed, given it was our oversight to not specify handedness in the initial e-mail, that we would include the person for testing. There is also some debate within the research community about the inclusion of left-handed participants in research using these kinds of tests (Willems, Van der Haegen, Fisher & Francks, 2014). Some suggest, given left handed people make up roughly 10% of the population, knowing how left-handed people perform in these tests is useful information (Willems et al., 2014). However, as this was the only left-handed person in the cohort, I excluded this person's data from the analysis. Left-handedness could have acted as a confound and there would be no opportunity for any subanalysis comparing the performance of left and right handers. We then ensured the initial e-mail included being right handed as an inclusion criterion.

In terms of issues with the testing of participants, one of the depression group participants told me, if she did the spatial memory tasks, she would be guessing as she had "no spatial awareness" and would therefore be performing at or around chance. Testing was therefore terminated. With another depression group participant tested by the joint researcher, the Town Square Test programme crashed and so it was not possible to complete the test. Given two of the tests involved using a laptop, we were fortunate to only have one where technical issues meant we could not complete data collection. Finally, one of the control group participants I was due to test, arrived at the testing slot clearly inebriated and so I had to tell him, as sensitively as possible, he would not be able to participate. I made sure I was in a public area in case he did not react well to being challenged about his fitness to participate, but in fact he simply apologised and left.

6 The researcher therapist divide

I was struck many times how difficult it can be to separate being a therapist from being a researcher. Some of these times occurred during the screening phase. We used the SCID-I (Structure Clinical Interview for DSM Disorders; First, Spitzer, Gibbon & Williams, 2002) to assess for co-morbid disorders. This provides a set of questions designed around specific mental health conditions plus follow-up questions. It was important to keep in mind that in this capacity I was a researcher and not a therapist. At times the therapist in me wanted to continue to ask questions, to formulate their difficulties, when it had become clear, as a researcher, they did not meet criteria and so were not eligible. The other time it was difficult to maintain separation was when people in the depression group were answering questions about their depression history. I was struck by how many of the people I tested had been struggling with depression for many years but had only recently sought help.

7 Reflections on the major research project

The project investigated the interaction of two research findings: that people with depression have a bias in memory towards negatively valenced material (e.g. Roberson-Nay et al., 2006); and that people with depression have impaired hippocampal function and might therefore be expected to show impaired performance on tests requiring the hippocampus, for example spatial memory tests, as Wong (2015) had found.

How these two phenomena might interact was much harder to predict and this is why alternative hypotheses were specified. Based on research, two different potential hypotheses with alternative predictions of the impact of valenced stimuli on allocentric memory performance were proposed. However, to find essentially no impact of valence for either depression or control group was unexpected and disappointing given the amount of time and effort put into the research. I had hoped my research would be able to

elucidate cognitive and memory processes in depression, knowledge that might eventually improve treatment effectiveness, but I ended the process with more questions than answers.

8 Reflections on literature review

When looking for a topic for my literature review, my initial idea to look in the simplest terms at the links between the hypothalamic-pituitary-adrenal (HPA) axis and depression as I knew a little about the long-term effects of cortisol on the hippocampus from undergraduate psychology lectures. However, I found a recent review had been done in this area (Stetler & Miller, 2011). I then considered reviewing the links between child trauma and depression as this is an area of interest for me. Again, a recent review had been published in this area (Carr, Martins, Stingel, Lemgruber & Juruena, 2013). I therefore decided to bring these two areas together and look at the relationship between the HPA axis, childhood trauma and depression in adulthood. Initial searches suggested, while there were a few research groups looking at this (e.g. Christine Heim, Charles Nemeroff and colleagues), there had not been a systematic review in this area.

The neurobiology of the HPA axis is not an area of expertise for me, nor are the many ways of testing its functioning. I spent a great deal of time researching and developing an understanding of the intricate and many levelled functioning of the HPA axis and ways of measuring it. Due to other commitments, both on the course and outside it, the time I could devote to the literature review was fragmented over several months. I was aware of having to spend time almost every time I came back to reviewing the papers identified by my search, re-educating myself about the tests and what they showed in terms of the functioning of the HPA axis. With hindsight, it would have been much more efficient and effective to have taken larger blocks of study leave so I could achieve more with the knowledge I had assimilated, rather than feeling increasingly frustrated with the

time I was having to spend re-acquainting myself with the literature each time I came back to it.

In spite of this, I am pleased with the finished review. It was ambitious, particularly given my lack of knowledge in this area. The literature is fragmented and there is a lack of replication or indeed overlap in terms of client groups and measures. But important questions are raised by this research, questions it is currently difficult to answer given the lack of consistency in methodology and findings. Some of the results potentially point to a sub-type of depression in adulthood linked to trauma in childhood and distinct from depression in the absence of a history of childhood trauma. This could be the result of untreated reactions to the trauma. Not only might this sub-type be characterised by different effects on the HPA axis, its existence might help explain why some people with depression do not respond as well as others to talking therapy, including cognitive behavioural therapy, and anti-depressant medication. If future research in this area could help to elucidate this sub-type, it may be possible to develop a more effective treatment for depression with a history of childhood trauma.

9 REFERENCES

Bisby, J. A., King, J. A., Brewin, C. R., Burgess, N., & Curran, H. V. (2010). Acute effects of alcohol on intrusive memory development and viewpoint dependence in spatial memory support a dual representation model. *Biological Psychiatry, 68,* 280-286.

Carr, C. P., Martins, C. M. S., Stingel, A. M., Lemgruber, V. B., & Juruena, M. F. (2013). The role of early life stress in adult psychiatric disorders: a systematic review according to childhood trauma subtypes. *The Journal of Nervous and Mental Disease*, *201*(12), 1007-1020.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders: SCID-I.* Washington, DC: American Psychiatric Press

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 3310-3315.

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

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). Human hippocampus and viewpoint dependence in spatial memory. *Hippocampus, 12*, 811-820.

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*, 606- 613.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The Center for Research in Psychophysiology, University of Florida*.

New, B., Ferrand, L., Pallier, C., & Brysabaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45-52.

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*, *17*(6), 834-839.

Sagfors, L.F. (2017). *Autobiographical and hippocampus-dependent spatial memory in depression.* Unpublished doctoral dissertation. University College London, London.

Smith, K. V., Burgess, N., Brewin, C. R., & King, J. A. (2015). Impaired allocentric spatial processing in posttraumatic stress disorder. *Neurobiology of Learning and Memory, 119*, 69-76.

Spitzer, R. L., Kroenke, K., & Williams, J. B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*, 1092-1097.

Stetler, C., & Miller, G. E. (2011). Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosomatic Medicine*, *73*(2), 114-126.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin, 133,* 122-148.

Willems, R. M., Van der Haegen, L., Fisher, S. E., & Francks, C. (2014). On the other hand: including left-handers in cognitive neuroscience and neurogenetics. *Nature Reviews Neuroscience*, *15*(3), 193-201.

Wong, Y.P. (2015). *An investigation of the links between depression and hippocampal-based allocentric spatial memory*. Unpublished doctoral dissertation. University College London, London.

APPENDICES

APPENDIX A: DESCRIPTION OF JOINT WORKING ARRANGEMENTS

Description of joint working arrangements

This study was conducted jointly with Line Sagfors (trainee clinical psychologist) as the participant groups, including eligibility criteria, were the same for both projects. Table 1 provides a summary of the stages of research, indicating which were conducted jointly and which were conducted individually. In terms of testing, some of the measures were shared and some were specific to the individual projects, and this is summarised in Table 2. Conducting the research as a joint project facilitated the recruitment of a sample sufficient to detect a small to medium effect size, according the power analyses conducted for the two projects, and also of a specific control group. It also meant we could combine our research expense allowances so it was possible for us to financially compensate all participants (apart from five UCL undergraduates in the control group who received course credit instead) to cover travel expenses and time.

Table 1.

Details of stages of research conducted individually and jointly

	Individual	Joint	Comments
Literature Review	X		
Empirical paper			
Research question	X		
Research design	X	X	Initially individual, with collaboration on the final design and procedures
Ethical approval		X	Tasks conducted jointly or shared
Liaison with IAPT services		X	Tasks conducted jointly or shared
Participant recruitment		X	Tasks shared
Participant testing		X	Tasks shared
Data scoring and input		X	Tasks shared
Data analysis	X		
Write-up	X		
Critical appraisal	X		

Table 2.

Details of joint and individual measures

	Joint	Individual: Janice Williams' project	Individual: Line Sagfors' project
BDI-II	✓		
GAD-7	✓		
Bespoke clinical information questionnaire	✓		
Demographic information	✓		
TIPI		✓	
Town Square Test (King et al., 2002)		✓	
Four Mountains Test (Hartley et al., 2007)			✓
AMT (Williams & Broadbent, 1986)			✓

BDI-II = Becks Depression Inventory (Beck, Steer, Ball & Ranieri, 1996)
GAD-7 = Generalised Anxiety Disorder =-7 (Spitzer, Kroenke, Williams, & Löwe, 2006)
TIPI = Ten Item Personality Measure (Gosling, Rentfrow & Swann, 2003)
Town Square Test (King, Burgess, Hartley, Vargha-Khadem, & O'Keefe, 2002)
Four Mountains Test (Hartley et al., 2007)
AMT = Autobiographical Memory Test (Williams & Broadbent, 1986)

REFERENCES

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588-597.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Hartley T., Bird C. M., Chan D., Cipolotti L., Husain M., Vargha-Khadem, F., & Burgess N. (2007). The hippocampus is required for short-term topographical memory in humans. *Hippocampus, 17,* 34–48.

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). Human hippocampus and viewpoint dependence in spatial memory. *Hippocampus, 12*, 811-820.

Spitzer, R. L., Kroenke, K., & Williams, J. B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*, 1092-1097.

Williams J. M. G., & Broadbent K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*, *95*, 144–149.

APPENDIX B: LETTER CONFIRMING ETHICAL APPROVAL

NHS
Health Research Authority

East Midlands - Nottingham 2 Research Ethics Committee

Royal Standard Place
Nottingham
NG1 6FS

22 March 2016

[redacted]
R&D (1st Floor, Maple House), Rosenheim Wing, Ground Floor
25 Grafton Way, London
WC1E 6DB

[redacted]

Study title:	**The Relationship Between Depressive Symptoms and Hippocampally-dependent Allocentric Memory**
REC reference:	**14/EM/0222**
Amendment number:	**third amendment**
Amendment date:	**13 February 2016**
IRAS project ID:	**148009**

The above amendment was reviewed at the meeting of the Sub-Committee held on 21 March 2016.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMP)	third amendment	13 February 2016
Other [Developing a word based version of the Town Square Task]	1.0	03 March 2016
Other [Autobiographical Memory Test]	1.0	03 March 2016
Participant consent form	5 - tracked	01 February 2016
Participant information sheet (PIS)	5 - tracked	01 February 2016
Research protocol or project proposal [study protocol (= Outline Rationale + Study Overview)]	5 - tracked	01 February 2016
Research protocol or project proposal	5 - tracked	01 February 2016
Summary CV for student [Line Sagfors]		
Summary CV for student [Janice Williams]		
Validated questionnaire [BDI - II]		

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

14/EM/0222:	**Please quote this number on all correspondence**

[redacted]

pp.

Professor Frances Game
Chair

E-mail: NRESCommittee.EastMidlands-Nottingham2@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

Copy to: [redacted]

APPENDIX C: PARTICIPANT INFORMATION SHEETS AND CONSENT FORMS

NHS TRUST LOGO

UCL

Understanding the Effects of Depression on Memory

This study has been approved by the University College London Ethics Committee (Project ID Number): SHaPS-2014-JK-009

Name, Address and Contact Details of Investigators:

Principal investigator:
[redacted]
Research Department of Clinical, Education and Health Psychology
University College London
1-19 Torrington Place, London, WC1E 7HB
United Kingdom
+44 (0)20 7679 5993
[redacted]

Researchers:
[redacted]
Doctorate in Clinical Psychology
Research Department of Clinical, Education and Health Psychology
University College London
[redacted]

We would like to invite you to participate in this research project.

We would like to invite you to take part in our research study. You should participate only if you want to do so. Before you decide whether to take part, we would like you to understand why the study is carried out, what you would be asked to do, and how the study will be conducted. Please take some time to read this sheet thoroughly, and to discuss it with other people if you wish. One of our team will go through the information sheet with you and answer any questions you have. Please feel free to ask any further questions about the study, or if you find anything on this sheet unclear.

Why is this study being done?

Little is currently known about the impact of depression on the way memories are processed in the brain, but we know that memory can be affected. To understand the effects of depression on memory, we are investigating memory in both people with depression and people without depression. **You would be tested as part of the group of people who are not currently suffering from depression.**
By investigating spatial memory (remembering where things are) - as well as autobiographical memory (remembering things we have experienced) - we hope to

gain a better understanding of depression and to develop more effective screening tools to measure the impact of depression.

Do I have to take part?

No. Your participation in the study is entirely voluntary. It is your choice whether or not you would like to participate. If you do give consent to take part in the study, you are still free to leave the study at any point, without giving a reason. If you leave, any information for the research that we have already collected from you will be destroyed.

What will happen if I decide to take part?

During a telephone call with one of our researchers, we will ask you some further questions to verify your eligibility. This will include questions about your mood and questions about whether you have had mental health problems in the past. During this time, any other questions you may have about your participation will be addressed. If you decide to participate and are suitable, the researcher will arrange a convenient time to meet with you at University College London.

You will then complete a series of tasks and questions, some of them using a computer. We will ask you till fill in a short questionnaire about personality and ask you give some basic information about your age, sex, occupation and education. There will be three tasks investigating memory. One of the tasks is an autobiographical memory test where the researcher will give you a number of cue words and ask you to recall events from the past that relate to these words. You will then perform a pen-and-paper task in which you will be asked to view some pictures of landscapes, then select the correct answer from a few options. The third task is a computerized memory task where you will view some scenes, then answer some questions to see how well you remember what you saw. Most people find these tasks quite interesting to do. At the end of the session, the study researcher will conduct a debriefing and address any other questions or concerns you may have. Your participation should take approximately 60 minutes.

What are the possible disadvantages and risks for taking part?

We do not anticipate any potential disadvantages or risks in your participation. However, we will support you if you become upset or distressed during the study. You will be given time at the end of the study to be fully debriefed with a member of the research team.

What are the possible benefits of taking part?

You may find it interesting to complete these tasks and the information gathered during this study will also help to inform our understanding of depression, which will hopefully be a step towards helping the development of novel screening tools in the future.

Will I be paid for taking part in this study?

We will give you £12.50 when you have completed the tasks and questionnaires to compensate you for your time and any travel expenses you may have incurred.

What will happen to the results of the research study?

The results will be written up in the form of reports to be submitted as a doctoral thesis project. As mentioned, you will not be identifiable from the results. On completion, you will be sent a summary of the study results if you request for it. You may contact one of the study researchers listed below via phone or email.

All data will be collected and stored in accordance with the Data Protection Act 1998. Only UCL researchers working with [redacted] will analyze these data.

If you wish to discuss participation either before or after taking part, you can contact us by email or phone using the contact details at the top of this document.

CONSENT FORM

Title of Project: **Understanding the Effects of Depression on Memory**

Name of Researchers: [redacted]

Please initial all boxes

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. I have received the contact details for the investigators, who I can contact if I wish to discuss my participation. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐

3. I agree to take part in the above study. ☐

Name of Participant	Date	Signature
Name of Person taking consent.	Date	Signature

NHS TRUST LOGO

UCL

Understanding the Effects of Depression on Memory

This study has been approved by the Research Ethics Committee for East Midlands – Nottingham 2 (Project ID Number): 14/EM/0222.

We would like to invite you to participate in this research project.

We would like to invite you to take part in our research study. You should participate only if you want to do so. Before you decide whether to take part, we would like you to understand why the study is carried out, what you would be asked to do, and how the study will be conducted. Please take some time to read this sheet thoroughly, and to discuss it with other people if you wish. One of our team will go through the information sheet with you and answer any questions you have. Please feel free to ask any further questions about the study, or if you find anything on this sheet unclear.

Why is this study being done?

Little is currently known about the impact of depression on the way memories are processed in the brain, but we know that memory can be affected. By investigating spatial memory (remembering where things are) - as well as autobiographical memory (remembering things we have experienced) - we hope to gain a better understanding of depression and to develop more effective screening tools to measure the impact of depression.

Why have you been invited to take part?

You have been invited to take part in this study because you have recently been referred to (or self-referred) for depression and assessed at one of the clinical services in collaboration with the research team. Alternatively, you may have completed an intake form indicating your consent to be contacted by a researcher affiliated with the [redacted] to contact you for research matters. There will be approximately 33 other participants recruited for this study.

Do I have to take part?

No. Your participation in the study is entirely voluntary. It is your choice whether or not you would like to participate. Deciding not to take part in the study will not affect the care you receive from services either now or in the future. If you do decide to participate, you will be given this information sheet to keep, and your therapist will ask you to sign a consent form stating that you wish to take part. You may either sign this consent form and return it immediately to your therapist or bring it back to

the psychology department within 24 hours. Alternatively, you may post it back to the psychology department. We will provide you with a stamp and envelope for postage. If you do give consent to take part in the study, you are still free to leave the study at any point, without giving a reason. This will not affect the care you are currently receiving, or will receive in the future. If you leave, any information for the research that we have already collected from you will be destroyed.

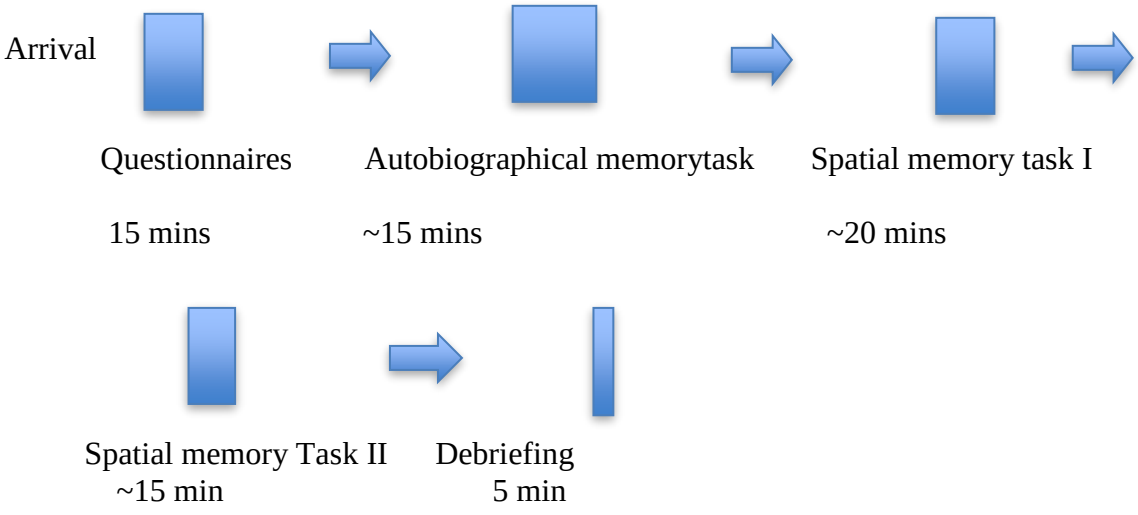
What will happen if I decide to take part?

Your psychological therapist at the clinical service will give you some details of the study during your first session. If you wish to take part in the study, your therapist will then ask you some questions to determine your eligibility to participate, as not everyone would be suitable. The participant information sheet will be handed to you. Your therapist will provide us with your contact details after you have provided consent to participate in the study. During a telephone call with one of our researchers, we will ask you some further questions to verify your eligibility. During this time, any other questions you may have about your participation will be addressed. If you decide to participate and are suitable, the researcher will arrange a convenient time to meet with you, preferably prior to your third therapy session with your therapist.

If you had previously indicated your consent on an intake form, the [redacted] will provide the researcher with your contact number. The researcher will contact you to find out if you are still interested to participate in the study. If so, the researcher will tell you about the current study and ask you some questions to verify your eligibility. A copy of the participant information sheet will be posted to you. The researcher will phone you in about a week to find out if you have received the participant information sheet. If you still express interest in the study, the researcher will arrange a convenient time to meet with you.

Study overview:

The session will take approximately 70 minutes as follows:



If you agree to participate in this study, you will be asked to come to a site most convenient for you, within the [redacted]. Upon your arrival, you will meet a member of the research team and you can ask any other questions you may have. Following this, you will be asked to complete questionnaires on depression and anxiety. The next task is an autobiographical memory test where the researcher will give you a number of cue words and ask you to recall events from the past that relate to these words. You will then perform a pen-and-paper task (Spatial memory task I), in which you will be asked to view some pictures of landscapes, then select the correct answer from a few options. Spatial memory task II is a computerized task. You will view some scenes, then answer some questions to see how well you remember what you saw. Most people find these tasks quite interesting to do. At the end of the session, the study researcher will conduct a debriefing and address any other questions or concerns you may have.

What are the possible disadvantages and risks for taking part?

We do not anticipate any potential disadvantages or risks in your participation. However, we will support you if you become upset or distressed during the study. You will be given time at the end of the study to be fully debriefed with a member of the research team. Your personal therapist will also be aware of your participation in the study and able to support you should you find discussing your experiences difficult.

What are the possible benefits of taking part?

You may find it interesting to complete these tasks and the information gathered during this study will also help to inform our understanding of depression, which will hopefully be a step towards helping the development of novel screening tools in the future.

Will I be paid for taking part in this study?

We will give you £12.50 when you have completed the tasks and questionnaires to compensate you for your time and any travel expenses you may have incurred.

Who will know you are taking part in the study?

We will inform your personal therapist and your GP of your participation. Information collected during all stages of the study will be kept strictly confidential. All information will only be viewed by members of the research team at University College London. However, if through the course of the study it was found that you are at immediate risk of harm to yourself or others, this information will be shared with your therapist or GP and if necessary, emergency services.

Will my taking part in the study be kept confidential?

Your consent form will be kept in a separate location from all your other data, ensuring that this remains anonymous. All data will be stored in secure location whereby a participant ID will be assigned to your data, with no identifiable personal information. Any data with your identifying details removed and stored on computers will be password protected. Any published data will also be entirely anonymous, meaning that individuals cannot be identified.

The data from this study will be stored in accordance with the UCL and NHS Data Protection and Records Management policies. We will follow ethical and legal practice and all information about you will be handled in confidence.

All data will be collected and stored in accordance with the Data Protection Act 1998.

What will happen to the results of the research study?

The results will be written up in the form of reports to be submitted as a doctoral thesis project. As mentioned, you will not be identifiable from the results. On completion, you will be sent a summary of the study results if you request for it. You may contact one of the study researchers listed below via phone or email.

What if there is a problem?

Every care will be taken in the course of this study. If you wish to complain, or have any concerns about any aspect of the way you have been approached or treated by members of staff you may have experienced due to your participation in the research, National Health Service or UCL complaints mechanisms are available to you. Please ask the researcher if you would like more information on this. In the unlikely event that you are harmed by taking part in this study, compensation may be available.

If you suspect that the harm is the result of the Sponsor's (University College London) negligence, then you may be able to claim compensation. After discussing with the researcher, please make the claim in writing [redacted], who is the Chief Investigator for the research and is based at UCL. The Chief Investigator will then pass the claim to the Sponsor's Insurers, via the Sponsor's office. You may have to bear the costs of the legal action initially, and you should consult a lawyer about this.

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the NRES Committee (East Midlands – Nottingham 2).

Contact Details

If you wish to contact the research team to discuss any of the information further or any concerns you have about the study, then please do so by getting in touch with the members of the research team listed below:

[redacted]
Research Department of Clinical, Education and Health Psychology
Room 440, 4th Floor
1-19 Torrington Place, London, WC1E 7HB
Telephone: 020 7679 5993 (x45993)
[redacted]

[redacted]
Research Department of Clinical, Education and Health Psychology
Room 442, 4th Floor
1-19 Torrington Place, London, WC1E 7HB
Telephone: 020 7679 1621 (x41621)
[redacted]

[redacted]
Research Department of Clinical, Education and Health Psychology
[redacted]

[redacted]
Research Department of Clinical, Education and Health Psychology
[redacted]

NHS TRUST LOGO

Centre Number:

Study Number:

Patient Identification Number for this trial:

CONSENT FORM

Title of Project: **Understanding the Effects of Depression on Sleep and Memory**

Name of Researcher: [redacted]

Please initial all boxes

4. I confirm that I have read and understand the information sheet dated 1st February 2016 (version 5) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

5. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐

6. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from University College London, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

7. I agree to my GP being informed of my participation in the study. ☐

8. I agree to take part in the above study. ☐

Name of Participant	Date	Signature
Name of Person taking consent.	Date	Signature

Please feel free to contact us if you have any further concerns.

[redacted]
Research Department of Clinical, Education and Health Psychology
Room 440, 4th Floor
1-19 Torrington Place, London, WC1E 7HB
Telephone: 020 7679 5993 (x45993)
[redacted]

[redacted]
Research Department of Clinical, Education and Health Psychology
Room 442, 4th Floor
1-19 Torrington Place, London, WC1E 7HB
Telephone: 020 7679 1621 (x41621)
[redacted]

[redacted]

[redacted]

APPENDIX D: LIST OF WORDS AND VALENCE RATINGS USED IN THE TOWN SQUARE TEST

Neutral and negatively valenced words used in the Town Square Test with mean valence ratings from Warriner et al. (2013)

Neutral		**Negative**	
Word	**Mean valence rating**	**Word**	**Mean valence rating**
Barrel	4.92	Afraid	2.25
Bedding	5.05	Anger	2.5
Boxers	5.05	Annoy	2.49
Brazen	5.00	Arrest	2.33
Buckle	5.05	Attack	2.00
Canteen	5.05	Awful	2.28
Caving	5.00	Breakup	2.19
Chemist	4.95	Coward	2.42
Clinic	4.95	Creepy	2.52
Collar	4.95	Crisis	2.05
Column	5.00	Deadly	1.90
Corny	5.00	Deathly	1.95
Council	5.00	Decay	2.4
Cunning	5.05	Dreaded	2.11
Depend	5.00	Failure	2.15
Fiddle	5.05	Greedy	2.10
Forty	5.00	Harmful	2.29
Gender	5.05	Hateful	1.90
Goggle	5.00	Hatred	2.38
Heady	5.00	Jealous	2.38
Iceberg	5.05	Kidnap	2.19
Induce	5.00	Lethal	2.48
Ingest	5.00	Maggot	2.00
Jigsaw	5.05	Mouldy	1.95
Jockey	4.95	Neglect	2.16
Kiosk	5.00	Obese	2.32
Loosen	5.00	Perish	2.22
Machine	5.00	Pervert	2.45
Metal	5.05	Pimple	2.11
Milky	5.00	Poison	2.16
Narrow	4.95	Pollute	1.88
Pageant	5.00	Prison	1.94
Project	5.00	Racist	2.05
Remark	5.05	Ransom	2.32
Retail	5.00	Rotten	2.43
Ringer	4.95	Sadness	2.40
Rumble	5.00	Seasick	1.89
Saddle	4.95	Sewage	2.33
Sewing	5.05	Suffer	2.05
Shorten	5.05	Suspect	2.39
Staple	5.00	Toxic	2.35
Storage	4.95	Traffic	2.18
Tighten	5.00	Traitor	2.39
Trendy	4.95	Upset	2.45
Untie	4.95	Vandal	2.32
Waiter	5.05	Victim	2.05
Widen	5.00	Widow	2.28
Wiring	5.05	Worry	2.10

APPENDIX E: PARTICIPANT WORD VALENCE RATING SHEET

Participant word valence rating sheet

PARTICIPANT NUMBER:

We would like to look at your responses to the words used in the study. You will use a scale to rate how you felt while reading each word. The scale ranges from 1 (unhappy) to 9 (happy). At one extreme of this scale, you are happy, pleased, satisfied, contented, hopeful. When a word made you feel completely happy you should indicate this by choosing rating 9. The other end of the scale is when you feel completely unhappy, annoyed, unsatisfied, melancholic, despaired, or bored. You can indicate feeling completely unhappy by selecting 1. Use the numbers in between to describe intermediate feelings of pleasure, eg if a word made you feel a little bit happy, score it a 6. If you felt completely neutral, neither happy nor sad, select the middle of the scale (rating 5).

Please work at a rapid pace and don't spend too much time thinking about each word. Don't worry if you can't remember how the word made you feel at the time, just rate how it makes you feel when you read it now.

1	2	3	4	5	6	7	8	9
Completely Unhappy				**Neutral Neither happy or sad**				**Completely happy**

Word	**Rating** **1 = completely unhappy** **5 = neutral** **9 = completely happy**
Afraid	
anger	
annoy	
arrest	
attack	
awful	
barrel	
bedding	

Word	**Rating** **1 = completely unhappy** **5 = neutral** **9 = completely happy**
boxers	
brazen	
breakup	
buckle	
canteen	
caving	
chemist	
Clinic	
Collar	
column	
Corny	
council	
coward	
creepy	
Crisis	
cunning	
deadly	
deathly	
decay	
depend	
dreaded	
failure	

Word	**Rating** **1 = completely unhappy** **5 = neutral** **9 = completely happy**
Fiddle	
Forty	
gender	
goggle	
greedy	
harmful	
hateful	
hatred	
heady	
iceberg	
induce	
ingest	
jealous	
jigsaw	
jockey	
kidnap	
Kiosk	
lethal	
loosen	
machine	
maggot	
metal	

Word	**Rating** **1 = completely unhappy** **5 = neutral** **9 = completely happy**
Milky	
mouldy	
narrow	
neglect	
obese	
pageant	
perish	
pervert	
pimple	
poison	
pollute	
prison	
project	
Racist	
ransom	
remark	
Retail	
ringer	
rotten	
rumble	
saddle	
sadness	

Word	**Rating** **1 = completely unhappy** **5 = neutral** **9 = completely happy**
seasick	
sewage	
sewing	
shorten	
staple	
storage	
suffer	
suspect	
tighten	
Toxic	
traffic	
traitor	
trendy	
Untie	
Upset	
vandal	
Victim	
waiter	
widen	
widow	
wiring	
worry	

APPENDIX F: PARTICIPANT CLINICAL INFORMATION QUESTIONS

Participant clinical information questions

CLINICAL INFORMATION

PARTICIPANT NUMBER:

1. When did this episode/period of depression start?

2. How many times have you had episodes/periods of depression before?

3. When did you have your first episode/period of depression?

4. When you have had depression in the past, which treatments have you used? Have you used medication? Have you used CBT or other talking therapies? Did you find them helpful? If so, how helpful did you find them (Improved my mood a little, improved my mood a lot)?

Treatment type	Used?	Helpful?	Improved a little? Improved a lot?
Medication ……			
CBT			
Other talking therapy …….			

5. Are you currently having talking therapy/CBT? If so, how many sessions have you had so far? Have you found it helpful? If so, how helpful (Improved my mood a little, improved my mood a lot)?

Treatment type	Number of sessions	Helpful?	Improved a little? Improved a lot?
CBT			
Other talking therapy …….			

6. Are you on medication for depression at the moment? If so, what is the name of the anti-depressant? How much do you take every day? How long have you been taking it for? Have you found it helpful? If so, how helpful have you found it (Improved my mood a little, improved my mood a lot)?

Medication name	Dosage	Length of time taken for	Helpful?	Improved a little? Improved a lot?

7. How long have you had depression this time?

8. Do you know if anyone else in your family has or has ever had depression?

APPENDIX G: PARTICIPANT DEMOGRAPHIC INFORMATION FORM

Participant Demographic Information Form

Participant number:

Age:

Sex:

How would you describe your ethnicity?

Is English your first language?

What is your current occupation?

If you are currently unemployed, what was your last occupation?

Please select the furthest level of education you have completed:

- ☐ None
- ☐ Secondary Education (GCSE/O-Levels)
- ☐ Post-Secondary Education (College, A-Levels, NVQ3 or below, or similar)
- ☐ Vocational Qualification (Diploma, Certificate, BTEC, NVQ 4 and above, or similar)
- ☐ Undergraduate Degree (BA, BSc etc.)
- ☐ Post-graduate Degree (MA, MSc etc.)
- ☐ Doctorate (PhD)

What is your approximate household income (before tax)?

- ☐ Less than £15.000
- ☐ £15.000-£25.000
- ☐ £26.000-£35.000
- ☐ £35.000-£50.000
- ☐ £50.000-£70.000
- ☐ £70.000 or more

APPENDIX H: LIST OF ABBREVIATIONS

List of abbreviations

ANOVA	Analysis of variance
BDI-II	Beck Depression Inventory
BDNF	Brain-derived neurotrophic factor
CECA-Q	Childhood Experience of Care and Abuse Questionnaire
CT	Childhood trauma
CTQ	Childhood Trauma Questionnaire
CAPS	Clinician Administered Post-Traumatic Stress Disorder Scale
CRF	Corticotropin-releasing factor
CRH	Corticotropin-releasing hormone
CAR	Cortisol awakening response
CCAT	Crowe Critical Appraisal Tool
DHS	Daily Hassles Scale
DST	Dexamethasone suppression test
ETI	Early Trauma Inventory
GAD-7	Generalised Anxiety Disorder-7
GR	Glucocorticoid receptors
HDRS	Hamilton Depression Rating Scale
HC	Healthy controls
HPA	Hypothalamic-pituitary-adrenal axis
IAPT	Improving Access to Psychological Therapies
IAPS	International Affective Picture System
LES	Life Events Survey
MTL	Medial temporal lobe

MR	Mineralcorticoid receptors
NRES	National Research Ethics Service
OGM	Overgeneral memory
PHQ-9	Patient Health Questionnaire-9
PBI	Parental Bonding Instrument
PSS	Perceived Stress Scale
pACC	Perigenual anterior cingulate cortex
PTSD	Post-traumatic stress disorder
PST	Prednisolone suppression test
PCL-C	PTSD Checklist Civilian Version
QIDS-SR	Quick Inventory on Depression Symptomology
SPSS	Statistical Package for Social Sciences
SLE	Stressful life event
SCID-I	Structured Clinical Interview for DSM Disorders
TIPI	Ten Item Personality Inventory
UCL	University College London