COVID-19 Impact Statement

Upon recommendation by the University of Manchester Doctoral Thesis Guidelines, a brief COVID-19 impact statement has been included below to inform examiners of changes made to the PhD project as a result of the campus closure (18/3/2020) and cessation of face-to-face testing at both the University and NHS sites (suspended on 13/03/2020).

Due to these restrictions, the recruitment target and research question for the behavioural study of affective cognition were revised, following discussions with my supervisory team. Together with the supervisory team, the joint decision was made to abandon recruitment of a separate schizophrenia group (only 6 participants in this group had been recruited at that time) and instead include only the healthy control and depressed groups, with high/low levels of early life stress. Furthermore, unbeknownst at the time initially, testing was not able to be resumed face-to-face in time to be feasible for my PhD project, hence the remaining groups were smaller than initially planned (N=52, including 27 healthy controls and 25 depressed participants). Results were nevertheless analysed as planned, and findings are discussed with a clear disclosure that these findings should be interpreted cautiously and are mainly exploratory in nature, aiming to elucidate new avenues for future research in this area. In regard to the 6 schizophrenia patients recruited, this information is included in the general discussion, drawing on previous evidence that emphasises the need for additional research in this area, and highlighting the possibilities for future studies.

Early Life Stress and Major Depressive Disorder: Neural and Cognitive Mechanisms

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2021

Franziska Goer

School of Biological Sciences
Division of Neuroscience and Experimental Psychology

Table of Contents

Early Life Stress and Major Depressive Disorder: Neural and Cognitive Mechanisms

Total word count: 54,593

Abbreviations

ACC: Anterior Cingulate Cortex
BDI-II: Beck Depression Inventory -II
CA: Cornu Ammonis
CGT: Cambridge Gambling Task
CSA: Childhood Sexual Abuse
CTQ: Childhood Trauma Questionnaire
DG: Dentate Gyrus
EA: Emotional Abuse
ELS: Early Life Stress
EN: Emotional Neglect
ERT: Emotional Recognition Task
FAGN: Face Affective Go No-Go
GMV: Grey Matter Volume
HC: Healthy Control
HRSD: Hamilton Rating Scale for Depression
M: Mean
MDD: Major Depressive Disorder
MDE: Major Depressive Episode
MINI: Mini International Neuropsychiatric Interview
ROI: Region of Interest
MIR: Monetary Incentive Reward
mOFC: Medial Orbitofrontal Cortex
PA: Physical Abuse
PCC: Posterior Cingulate Cortex
PD: Prisoners' Dilemma
PN: Physical Neglect
PRT: Progressive Raito Task
RA: Risk Adjustment
SA: Sexual Abuse
SD: Standard Deviation
TAQ: Traumatic Antecedents Questionnaire
UG: Ultimatum Game
vmPFC: Ventromedial Prefrontal Cortex

Abstract

There is increasing evidence that both major depressive disorder (MDD) and early life stress (ELS) have shared effects on cognitive and neurobiological measures, including affective (emotion-laden) cognition and brain structure. Despite the documented high prevalence of ELS in MDD and undisputed role of ELS as a key environmental risk factor for MDD, very few studies in depression have controlled for ELS. This has made it difficult to disentangle the relative effects of ELS and MDD (and possible interactions) on affective cognition and morphological measures.

My PhD addressed this gap in the literature by applying a variety of methods and types of studies, including a systematic review and cross-sectional neuropsychological and neuroimaging studies. The overarching hypothesis of the PhD was that ELS at least partially mediates the effects on several cognitive and neural markers often attributed to depression, specifically affective cognition and brain structure. Study 1 comprehensively and systematically assessed affective cognition in ELS and MDD using a novel validated test battery. Results emphasised the multifaceted nature of affective cognition and revealed that specific constructs of affective cognition were affected primarily by MDD diagnosis, while others were more sensitive to ELS. A systematic review of grey matter volume (GMV) changes in MDD and ELS was provided in Study 2. This highlighted the key role ELS appears to play in GMV reductions of several brain regions, in particular the hippocampus and its subfields, independent of MDD. However, many included studies suffered from low/absent levels of ELS in the healthy control groups, limiting the ability to draw firm conclusions about the independent effect of ELS on brain structure. Furthermore, despite emerging evidence for sensitive periods during brain development and differential neurobiological effects of specific types of ELS, included studies did not measure timing of ELS and the vast majority did not specify ELS subtypes. Study 3 directly addressed these issues by investigating GMV of the hippocampus and amygdala (and their subfields) in MDD and ELS (specifically childhood sexual abuse; CSA) in a highly controlled study. Results indicated significant volume reductions in several hippocampal head subfields in CSA, independent of MDD. On the other hand, Study 4 revealed that another morphological index, cortical thickness, appears more sensitive to MDD than CSA. Specifically, MDD, independently of CSA, was associated with significantly increased cortical thickness in the medial orbitofrontal cortex and insula.

Overall, studies 1-4 indicate that distinctive constructs of affective cognition and different morphological indices appear to be differentially and specifically affected by ELS or MDD. The findings of my doctoral research demonstrate the importance of measuring and controlling for ELS in studies of affective cognition and brain structure in depression to avoid misattribution of observed effects. Given that analyses all controlled for ELS and ensured matched healthy control groups with comparable ELS (or specifically CSA) levels were included, findings can be more firmly attributed to either MDD or ELS than has often been the case in previous studies. This distinction is important as it may inform mechanisms underlying ELS as a risk factor for depression and elucidate more targeted treatments/interventions.

Declaration and copyright statement

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

Acknowledgments

I would like to express my heartfelt gratitude to my family, who have supported me from day one and have been an ever-present source of motivation. Mama and Max, I would like to thank you for your wise counsel and encouragement, especially when times were difficult, and for always putting a smile on my face. Papa, I want to thank you for always being my biggest champion and never doubting that I would achieve my dream one day. Your sympathetic ear kept me going throughout my PhD, and I wish I could share this final moment with you. I miss you and I hope I have made you proud. Paddy, thank you from the bottom of my heart for your never-waning encouragement, patience, humour, and love, which have kept me going over the past 4 years. I could not have done this without you.

I would like to express my sincere gratitude to Rebecca Elliott, for being the best supervisor I could have asked for. Thank you for sharing your knowledge and guidance, and for your encouragement, support, and understanding throughout my PhD. Thank you also to Richard Drake and Joanna Neill for your guidance and support. I would also like to thank Diego Pizzagalli and Emily Belleau for the opportunity to return to McLean Hospital for my placement and for their knowledge, guidance, and supervision.

I would like to express my gratitude to the Medical Research Council Doctoral Training Partnership for funding my PhD and enabling the research presented in this thesis. Finally, I would like to thank all research participants who took part in these studies for their invaluable contribution.

The Author

I received a BSc in Psychology (awarded summa cum laude) from Yale University in 2013. The following two years (2013-2015) I worked as a research assistant for Professor Diego Pizzagalli at the Laboratory for Translational and Affective Neuroscience (LATN) at the Centre for Depression, Anxiety, and Stress Research at McLean Hospital, Boston. As a research assistant I was involved in every aspect of the research process, including study design, ethical approval, data collection, and data analysis and synthesis. I subsequently completed my MRes in Cognitive Neuroscience (awarded distinction) at UCL, supervised by Professor Jonathan Roiser and Dr. Oliver Robinson.

In 2016 I was granted a Medical Research Council (MRC) Doctoral Training Partnership (DTP) to complete my PhD at the Faculty of Biology, Medicine, and Health at the University of Manchester, supervised by Professor Rebecca Elliott (primary supervisory), Dr. Richard Drake, and Professor Joanna Neill. Given my previous research experience and interest in the cognitive neuroscience of mood disorders, I was thrilled to be given the opportunity to extend my knowledge and learn new research skills to address novel research questions.

I was awarded a competitive MRC DTP flexible training grant for a research placement at McLean Hospital (at the LATN, directed by Professor Diego Pizzagalli) in 2019, where I was able to learn new analysis skills and apply these to neuroimaging data I had collected previously as a research assistant.

Throughout my PhD I have delivered oral and poster presentations at national and international conferences, including the British Association for Psychopharmacology and the Organization for Human Brain Mapping, in addition to several presentations and talks presented at the University of Manchester as part of seminar series, department meetings, and PhD showcases. During my time at the University of Manchester I have also contributed to public engagement (such as the British Science Fair) and worked as a teaching assistant for undergraduate seminars in psychology and neuroscience.

Introduction

1. Background

1.1. Major Depressive Disorder

A 2011 report presented at the World Economic Forum estimated that in the US alone the economic cost of mental health disorders would account for a $16.1 trillion loss over the next 20 years, more than all other non-communicable diseases combined (including, amongst others, cancer, cardiovascular diseases and diabetes; Bloom et al., 2012). In addition to their staggering economic costs, mental illnesses are also among the most disabling disorders and severely impair quality of life. Surveys from 28 representative countries indicated that mental health disorders, including depression, are more severely disabling than any physical conditions examined, including widespread conditions such as headaches, diabetes, chronic pain and cancer (Kessler et al., 2011). In the UK, mental health disorders have an estimated lifetime prevalence of 25%, with major depressive disorder (MDD) being the most common diagnosis (Beddington et al., 2008; Green et al., 2005). Given the immense costs to society, as a whole as well as to individuals afflicted with mental health disorders such as depression, there is a clear need for improved understanding of the aetiology of the disorder, as well as identifying cognitive and neural markers associated with MDD in the hope of identifying novel treatment targets and avenues for intervention.

1.2 Early Life Stress

Prior to discussing the link between early life stress (ELS) and MDD, it is important to note that the nomenclature for ELS is inconsistent, with various studies interchangeably using different terms including childhood trauma, childhood adversity/abuse, child maltreatment, amongst others (Smith & Pollak, 2020). To further complicate comparisons of studies of ELS,

many different definitions are applied, with researchers considering different ages, severity of exposure, and types of maltreatment (which may be limited to abuse/neglect or include various additional factors such as parental separation, death of a family member, natural disaster, robbery, institutionalisation etc). Reflecting the vastly differing definitions, a large variety of measures have been used to assess ELS, and as such complicate interpretations of findings between studies and disciplines. For the purposes of this PhD, the term ELS will be consistently used to describe childhood abuse/neglect occurring until age 18, and consisting of 5 subtypes of abuse, frequently studied using the validated Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003), a 28-item retrospective self-report questionnaire. The 5 subtypes of ELS assessed include emotional abuse (EA), physical abuse (PA), sexual abuse (SA), emotional neglect (EN), and physical neglect (PN).

ELS is known to significantly increase the risk of various psychiatric disorders including depression (Lindert et al., 2014; McCrory et al., 2017). Estimates of the population attributable risk proportion for MDD accounted for by ELS range from 23% (Kessler et al., 2010) to 54% (Anda et al., 2002). Several prospective longitudinal studies have confirmed the heightened risk for MDD following significant ELS exposure, estimating an approximate 1.5 - 3 odds ratio of developing MDD in later life, which may differ slightly depending on type of abuse and exposure to multiple types of abuse (Danese et al., 2009; Norman et al., 2012; Scott et al., 2010; Widom et al., 2007). Furthermore, MDD in conjunction with ELS has been associated with higher rates of comorbidity with other psychiatric disorders, earlier age of MDD onset, (Widom et al., 2007) and increased likelihood of poor treatment response and outcomes (see Nanni et al., 2012 for a meta-analysis) compared to MDD individuals with no ELS exposure. Moreover, environmental risk factors, including ELS, may

play a larger role than genetics in the development of MDD, with twin studies suggesting relatively low levels of heritability (in particular for less severe MDD and later age of onset; Kendler et al., 2006; Kendler et al., 1992; Lyons et al., 1998). However, recent studies have highlighted that gene-environment interactions, such as with ELS, may play an important role in the development of depression, though findings are still inconclusive and hindered by various methodological issues (see Li et al., 2020 for a systematic review). Given the substantial increase in risk for developing MDD following ELS, it comes as no surprise that rates of ELS within MDD samples are significantly higher than in the general community - studies have found a greater than 2-fold increase of significant ELS exposure in MDD participants compared to HC with a prevalence between 50.5% to 62.5% of significant childhood trauma reported in MDD participants (Williams et al., 2016; Xie et al., 2018).

Theories diverge on whether ELS may represent a general and cumulative effect on later life psychopathology (Evans & Whipple, 2013), or may depend on specific types of stressors (McLaughlin & Sheridan, 2016). Increasing evidence suggests that ELS subtypes (EA, PA, SA, EN, and PN) may be associated with particular psychiatric disorders and cognitive/neurobiological changes. However, findings to date are somewhat inconsistent and may depend on measures used to assess ELS and MDD as well as age/gender of participants – a systematic review of ELS subtypes and adult psychopathology found strongest associations between PA, SA and neglect with mood disorders (Carr et al., 2013) while a longitudinal study in adolescents found only EA and EN/PN predicted later depressed mood (while correlations with PA and SA were no longer significant after controlling for EA, EN, and PN; Cohen et al., 2019). A systematic review of neuroimaging studies in ELS found that different subtypes (including SA, EA, and EN/PN) were associated

with specific structural and functional abnormalities, while other deficits (such as attenuated frontal cortex grey matter volume) were common to all subtypes of ELS (Cassiers et al., 2018). Importantly, these two theories may not be mutually exclusive, as emerging research suggests that both specific and cumulative effects of ELS have been observed on various aspects of neurobiology and cognition (Smith & Pollak, 2020). Research investigations should hence ideally incorporate both into their investigations, allowing for measures of both cumulative ELS and specific types of abuse using standardised measures (to allow for meaningful comparisons between studies).

1.3. Possible mechanisms underlying relationship between ELS and MDD

While ELS as a risk factor for psychiatric disorders, including MDD, is undisputed, the mechanisms underlying this link are still poorly understood (Smith & Pollak, 2020). Accumulating evidence suggests that there are likely multiple pathways by which ELS may contribute to the development of MDD, including various neurobiological, genetic, neuroendocrine, and cognitive changes associated with ELS (Pechtel & Pizzagalli, 2011; Smith & Pollak, 2020).

Several epigenetic modifications of stress response genes have been documented following ELS, and may constitute one of the pathways by which ELS confers risk for development of MDD (Caspi et al., 2003; Jiang et al., 2019; Penner-Goeke & Binder, 2019). While gene x environment interactions (in particular pertaining to the serotonin transporter gene and the FKBP5 gene) and epigenetic mechanisms (such as DNA methylation) have been associated with ELS and MDD, results to date are inconsistent and further research is needed to investigate these possible links (Heim & Binder, 2012; Jaworska-Andryszewska &

Rybakowski, 2019; Lin & Tsai, 2019). Another key mechanism that may contribute to heightened risk for MDD is a sensitisation of the neuroendocrine stress response, immune activation, and increased inflammation, following ELS (Heim et al., 2008). Several empirical studies and theoretical frameworks have further highlighted the key role of neurobiological and cognitive changes in the aetiology and maintenance of depression (Disner et al., 2011; Drevets et al., 2008; Roiser & Sahakian, 2013). Emerging evidence suggests that ELS may be associated with similar cognitive and neurobiological markers as depression, which may in fact represent a key link between ELS exposure and later life depression (Heim et al., 2004; Pechtel & Pizzagalli, 2011; Smith & Pollak, 2020). Many neurobiological and cognitive markers have been studied in MDD and ELS, including structural and functional brain measures and affective (emotion-laden) and non-affective (non-emotional) cognition.

Both structural and functional neuroimaging studies have implicated various brain structures in MDD and ELS including those involved in stress response, emotion processing and regulation, memory, reward and punishment learning, executive functioning, and others. In particular, altered functioning and structure has been reported in frontal lobe structures, the limbic system, and striatum, amongst others, in both ELS and MDD (Smith & Pollak, 2020; Pechtel & Pizzagalli, 2011). Due to practical limitations (both in terms of time and funding constraints) this PhD focusses on structural brain changes in MDD and ELS. One of the key findings consistently reported in MDD, a reduction in hippocampal grey matter volume (see Videbech & Ravnkilde, 2004 for a meta-analysis), has increasingly been reported in ELS (for systematic reviews and meta-analyses, see Calem et al., 2017; Paquola et al., 2016), and emerging evidence suggests that perhaps grey matter volume (GMV) changes following ELS may serve as a vulnerability to later life MDD (Rao et al., 2010). Other

areas have been similarly implicated, including the caudate, orbitofrontal cortex, and posterior cingulate cortex (see **Study 2** for a systematic review of the literature). Furthermore, other morphological indices, such as surface area, and cortical thickness, have been increasingly studied in MDD, and, to a lesser degree, ELS. While the literature is somewhat more limited and inconsistent to date (possible due to methodological and sampling differences), multiple brain regions have been implicated in both MDD and ELS (see **Study 4** for a more detailed discussion and empirical findings).

The plethora of brain regions identified highlight the complexity of underlying neural mechanisms, and emphasise the involvement of multiple cognitive and neurobiological mechanisms in the consequences of ELS and pathophysiology of depression. However, it should be noted that the vast majority of both structural and functional neuroimaging studies in humans to date have focused on either MDD or ELS in isolation, making it difficult to assess the relative contribution of each (or possible interactions) on neural markers.

Finally, disruptions in cognition, particularly affective cognition (relating to processing of emotional stimuli), are a hallmark of depression and are thought to play a key role in both the aetiology and maintenance of the disorder (Elliott et al., 2011; Roiser et al., 2012; Roiser & Sahakian, 2013; Sanislow et al., 2010). While depression has also been associated with deficits in non-affective cognition, including executive functioning, memory, and attention (Keefe and Harvey, 2012; Rock et al., 2014), deficits in affective cognition are thought to play a central role in the development and exacerbation of the disorder, and have been linked to both functional and clinical outcomes (Couture et al., 2006; Insel et al., 2010; Roiser et al., 2012; Sanislow et al., 2010). Multiple changes in affective cognition in MDD

have been observed, including mood-congruent processing biases, and deficits in motivation, reward and punishment processing/learning (Roiser & Sahakian, 2013). Though far fewer studies have investigated affective cognition in ELS, converging evidence points to several similarities in affective cognition changes observed in MDD, which may confer an increased risk for later life depression (Pechtel & Pizzagalli, 2011). However, as with neurobiological studies, affective cognition in MDD and ELS are rarely investigated in the same sample, and most studies of depression do not measure or control for ELS, and as such risk potentially misattributing several observed changes in cognition to MDD when ELS may represent a key mediating factor.

In summary, while ELS has clearly been implicated in the pathophysiology of depression, the exact mechanisms have not been fully established to date. Studies from a variety of disciplines, including psychology, epidemiology, genetics, biology, and cognitive neuroscience highlight the complex nature of mechanisms likely contributing to ELS exposure conferring risk for development of depression. There is also increasing evidence that these effects do not occur in isolation, and instead interact in complex ways to give rise to vulnerabilities for depression. For instance, evidence suggests that changes in the immune system and inflammatory response following ELS may be a result of neurobiological changes in neural circuits involved in stress response and emotional processes, such as the prefrontal cortex, hippocampus, amygdala, and striatum (Smith & Pollak., 2020). Due to practical limitations and feasibility, this PhD focussed on two of the key aforementioned proposed mechanisms underlying risk for depression following ELS exposure: cognitive (specifically affective cognition) and neurobiological markers (specifically structural brain changes).

1.4. Affective cognition

1.4.1. Measures of affective cognition

Affective (emotion-laden or 'hot') cognition may be broadly defined as cognitive processes occurring in emotional contexts that interact and shape subsequent behaviour (Elliott et al., 2011). One of the main issues impeding progress in this research area is the multitude of measures used to assess affective cognition due to a lack of consensus around which subcomponents constitute this multi-faceted construct and how to measure them comprehensively. Lack of standardisation and resulting diversity in methodology of assessment have in turn made it difficult to draw overarching conclusions from previous research (Elliott et al., 2011). While several standardised test batteries exist for non-affective ('cold') cognition such as the widely implemented and cited CANTAB tests (see www.cambridgecognition.com), measures for affective cognition are significantly sparser to date. One of the few affective cognition test batteries, the P1vital® Oxford Emotional Test Battery (ETB; https://www.p1vital.com) comprises 5 tasks (measuring emotion identification/categorisation and attention to and memory for affective information), and has been well-validated and successfully applied in clinical trials and studies investigating the effect of psychopharmacological interventions (such as antidepressants) on cognitive processing (e.g. Harmer et al., 2003; Harmer et al., 2011; Murphy et al., 2008). However, several proposed components of affective cognition, including motivation/reward and social cognition (Elliott et al., 2011; Roiser & Sahakian, 2013), are not assessed. Similarly, while the MATRICS Consensus Cognitive battery (which was developed primarily for use in schizophrenia), includes a social cognition task, the remaining 9 tasks are predominantly non-affective tasks (using non emotional stimuli) (Nuechterlein et al., 2008).

A novel, recently validated, neuropsychological test battery of affective cognition (EMOTICOM) was developed to address this gap by providing a comprehensive and standardised assessment of the multifaceted construct of affective cognition and enable comparisons between different patient populations to be drawn (Bland et al., 2016). EMOTICOM consists of 17 individual tasks assessing different constructs of affective cognition, which developers broadly classified as falling into four key domains: emotional processing, motivation/reward, social cognition, and impulsivity (Bland et al., 2016). As the most comprehensive test battery for affective cognition developed to date, we selected the EMOTICOM test battery to be used in this PhD project, with the aim of elucidating the effects of MDD versus ELS on numerous distinct constructs of affective cognition.

1.4.2. Affective cognition in MDD and ELS

Affective cognition in MDD and ELS is discussed in detail in **Study 1**. While changes in both affective cognition and 'cold' (emotion-independent or non-affective) cognition (Roiser et al., 2009) have been documented in MDD (for a review, see Roiser & Sahakian, 2013), and ELS (for a review, see Pechtel & Pizzagalli, 2011), abnormalities in affective cognition in particular may be longer lasting and thereby potentially heighten risk for psychopathology including MDD (Pechtel & Pizzagalli, 2011). Furthermore, changes in affective cognition in MDD have been associated with functional (Couture et al., 2006; Insel et al., 2010; Sanislow et al., 2010) and clinical (Bouhuys et al., 1999) outcomes. Numerous changes in affective cognition have been documented in MDD, in particular a mood-congruent processing bias (negative affective bias), and deficits in motivation, reward and punishment processing/learning (Roiser & Sahakian, 2013). Several of these findings have been similarly identified in ELS, though studies to date are less numerous (Elliott et al., 2011; Pechtel &

Pizzagalli, 2011). More research is needed, particularly in samples measuring both ELS and MDD, to comprehensively assess affective cognition and determine the relative effect of ELS versus MDD.

1.5. Brain structure

1.5.1. Morphological measures and methods

Similarly to ELS, comparisons in structural brain changes between studies can at times be impeded by differences in methodology (both during neuroimaging data acquisition and analysis) and choice of outcome variables. Commonly investigated measures of neural morphology include GMV, cortical thickness, and surface area. While cortical thickness and surface area/GMV are thought to be phenotypically independent, surface area and GMV may be more closely associated (Winkler et al., 2009). Furthermore, it is important to note that all three morphological indexes have been associated with age-related atrophy during normal ageing (Lemaitre, 2010), highlighting the need for studies to either control for age or ensure comparison groups are well-matched in age.

In terms of previous literature in MDD and ELS, GMV is indisputably the most frequently studied morphological variable to date, though more studies are beginning to emerge, in particular for cortical thickness (as discussed below). Various methods have been employed to analyse neuroimaging data for GMV/cortical thickness, including manual tracing, and various statistical/analyses packages such as FreeSurfer (https://surfer.nmr.mgh.harvard.edu) or Statistical Parametric Mapping (SPM; including the VBM toolbox; https://www.fil.ion.ucl.ac.uk/spm/). In addition to being less time consuming than manual tracing, semi-automatic processing streams such as voxel-based morphometry

(VBM) using SPM or volumetric segmentation of brain structures using the feely available FreeSurfer (Fishl, 2012), have the added benefit of ensuring higher degrees of standardisation in analyses methods across studies, thereby improving comparability. FreeSurfer's validated processes for segmentation of hippocampal subfields and amygdala nuclei offer a particular advantage over other methods, as it allows for a much more nuanced approach to volumetric analyses of these structures (Iglesias et al., 2015; Saygin et al., 2017). We hence selected FreeSurfer's analysis stream for both GMV and cortical thickness analyses in MDD and ELS (**Studies 3** and **4**, respectively).

1.5.2. Structural brain changes in MDD and ELS

Analogous to affective cognition, similarities in structural brain changes in MDD and following ELS exposure have begun to emerge. As previously mentioned, MDD has been consistently associated with reductions in hippocampal volume compared to HC (Bremner et al., 2000; Campbell et al., 2004; Videbech & Ravnkilde, 2004) and more recent analysis of hippocampal subfields suggest this may be particularly prominent for the cornu ammonis and dentate gyrus (Huang et al., 2013; Treadway et al., 2015; Lindqvist et al., 2014; Han et al., 2016). Interestingly, hippocampal GMV reductions have been similarly reported in ELS (Calem et al., 2017; Paquola et al., 2016). Few studies have assessed hippocampal GMV in samples measuring both ELS and MDD. Emerging evidence suggests that hippocampal volume reductions in MDD patients may be driven, at least partially, by ELS (Colle et al., 2017; Chaney et al., 2014; Opel et al., 2014; Vythilingam et al., 2002), though a few studies have failed to replicate this finding (Gerritsen et al., 2015). Furthermore, reduced GMV of cornu ammonis and DG subfields have been observed in ELS, independent of MDD diagnosis/symptoms (Aghamohammadi-Sereshki et al., 2021; Teicher et al., 2012), but a few

contradictory findings have also been reported (Mikolas et al., 2019). Though less frequently studied than the hippocampus, GMV alterations in other regions, previously associated with MDD, have also been reported in ELS, including, amongst others, the amygdala (Hamilton et al., 2008; Pechtel et al., 2014), caudate (Frodl et al., 2017), and orbitofrontal cortex (Lim et al., 2018; Saleh et al., 2017; Yang et al., 2017).

Other morphological indexes, such as cortical thickness, have also been investigated in MDD and ELS. Studies in ELS have pointed to widespread cortical thinning of various brain structures (e.g. Dannlowski et al., 2012; Gold et al., 2016; Bounoua et al., 2020). It should be noted, however, that studies differed significantly in regard to ELS definition and measures. This may be particularly problematic given that the effect of cortical thinning may depend on the type of abuse - for example, thinning of the parahippocampal gryrus, frequently implicated in ELS, may be driven specifically by childhood sexual abuse (CSA), independent of other forms of abuse or depressive symptoms (Heim et al., 2013). Interestingly, studies in MDD have also reported reduced cortical thickness in the parahippocampal gyrus in MDD compared to HC, however these studies did not measure or control for ELS (Papmeyer et al., 2015). Meta-analyses of cortical thickness in MDD have yielded heterogenous results, possibly due to differences in methodology and participants sampled (such as medicated versus unmedicated MDD; Li et al., 2020; Suh et al., 2019). Furthermore, none of these studies accounted for ELS. In fact, to our knowledge, only a single study of cortical thickness in MDD to date has controlled for ELS, though it did not have a HC comparison group with

ELS, making it difficult to fully analyse relative contributions of MDD versus ELS (Jaworska et al., 2014).[1]

Overall, studies of GMV and cortical thickness and ELS have shown considerable overlap in findings, highlighting the possibility that certain morphological changes may not be solely attributable to MDD. While few studies have specifically analysed this relationship in samples measuring both MDD and ELS, emerging evidence suggests that ELS may in fact be driving several GMV reductions observed in MDD. However, findings to date are sparse and at times inconsistent, which may be attributable to different definitions and measures of ELS, differences in MDD samples (e.g. medication status), and differences in neuroimaging acquisition and analysis techniques. Furthermore, even studies that included measures of both MDD and ELS frequently lack a HC group with high ELS levels, and as a result analyses of ELS can only be conducted within MDD. This makes it difficult to disentangle relative effects of MDD and ELS on brain structure. There is hence a clear need for further studies to investigate the relative contributions of ELS and MDD on morphological measures such as GMV and cortical thickness, using carefully designed studies addressing the aforementioned methodological concerns and including well-matched control groups (in particular inclusion of a high ELS HC group to match ELS in MDD). This is specifically addressed in **Study 3** (GMV) and **Study 4** (cortical thickness).

2. Summary of salient gaps in the literature

[1] For more detailed background on GMV and cortical thickness changes in ELS and MDD, please see **Study 2** (for a systematic review of GMV in MDD and ELS) and **Study 3** (GMV) and **Study 4** (cortical thickness).

Despite ELS being a well-documented risk factor for MDD and in turn being significantly overrepresented in MDD samples compared to healthy controls (HC), few studies of cognition and brain structure in depression have measured or controlled for ELS. This is particularly problematic given the overlap in findings (outlined above and discussed in more detail in **Studies 1-4**) in both ELS and MDD, suggesting that potentially some of the widely reported finings of abnormal affective cognition and brain structure in MDD may be at least partially explained by ELS. There is increasing awareness that ELS may mediate experimental variables studied in MDD, and some recent studies that have included measures of both MDD and ELS have pointed to ELS as a potential driving factor behind some of the changes previously attributed to MDD (such as reductions in hippocampal GMV; Chaney et al., 2014; Opel et al., 2014; Vythilingam et al., 2002). Fewer studies of affective cognition have included measures of both MDD and ELS, though some evidence suggests that certain components of affective cognition, such as decreased sensitivity to reward, may mediate the relationship between ELS and later life depression (Guyer et al., 2006; Miu et al., 2017). Disentangling the effect of ELS and MDD on affective cognition and brain structure has been further complicated by inconsistent methodologies and key differences in participant samples (including differing definitions and measures of ELS, lack of standardised or comprehensive testing of affective cognition, differences in MDD sample characteristics, such as antidepressant treatment and severity of depression, and a range of different assessment methods and inclusion criteria for MDD diagnosis/symptoms). Another issue apparent in the few studies that have included measures of ELS and MDD in the same investigation, is the frequent lack of a HC group with high ELS exposure, resulting in very skewed groups in which the HC has little/none ELS exposure, and the MDD group has significantly higher levels. This once against leads to an issue in which MDD and ELS are

conflated and independent effects are difficult to assess. Novel studies using a systematic and comprehensive approach are thus needed to assess the relative contributions of MDD and ELS on affective cognition and brain structure, in the hope of shedding further light on the mechanisms underlying the link between ELS and MDD. This is particularly important given the vast societal and individual cost of both ELS and MDD, and may highlight novel treatment targets and/or more effectively direct interventions where they may be most beneficial and effective.

3. Aims

The overarching aim of this PhD was to assess the relative contribution of ELS and MDD on aspects of affective cognition and neurobiology. The main research question was whether changes in affective cognition and brain structure previously reported in MDD are potentially, at least in part, explained by ELS.

To address this research question, we developed two central aims:

1) To comprehensively and systematically assess affective cognition in MDD and ELS using a standardised and validated test battery of affective cognition (EMOTICOM; Bland et al., 2016) and reliable measures of ELS (the Childhood Trauma Questionnaire, CTQ; Bernstein et al., 2003) and MDD (Mini-International Neuropsychiatric Interview, version 7.0 for DSM 5; MINI; Sheehan et al., 1998) and widely used validated symptom measures including the Beck Depression Inventory-II (BDI-II; Beck et al, 1996), and Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960).
2) To assess the relative impact of ELS and MDD on brain structure. Specifically:

a. To systematically review the literature of studies of GMV in samples measuring both depression and ELS to identify whether ELS/MDD may have independent or interactive effects on GMV, and to identify brain areas of interest.
b. To investigate whether MDD or ELS (or an interaction of the two) drive GMV changes identified through the systematic review (and improving on previous studies' methodological constraints, such as including high and low ELS groups for both MDD and HC, to allow for meaningful comparisons and statistical analyses).
c. To assess effects of ELS and MDD on cortical thickness in key brain areas implicated in previous studies of ELS/MDD (which have been predominantly studied in isolation thus far).

Importantly, in all original research studies, we aimed to ensure an even distribution of ELS in MDD and HC groups, allowing for careful control of ELS as a potential confounding factor and enable meaningful analyses about individual contributions of ELS and MDD (or identify potential interactions).

4. Hypotheses

The overarching hypothesis of the PhD, based on previously discussed literature indicating substantial overlap in affective cognition and neurobiological changes previously largely attributed to MDD but more recently also highlighted in ELS, was that ELS at least partially mediates these changes. Specifically, in regard to the aims set out above:

1)

a. In general, we hypothesised that different aspects of affective cognition would be differentially affected by MDD and ELS, with ELS driving some changes previously attributed to MDD.
b. More specifically, based on previous literature, we hypothesised that depression would be associated with negative affective biases while ELS would be more associated with a threat related (anger/fear) emotion processing bias.
c. We predicted that MDD would be characterised by reduced motivational/reward function, which may be, at least partially, accounted for by ELS.
d. We also hypothesised that MDD would demonstrate heightened negative moral emotions (e.g. shame, guilt, and self-blame), which may be, at least partially, mediated by ELS.
e. Specific hypotheses were not made for the remaining three tasks/constructs, including theory of mind (ToM) and social economic exchange games, due to inconsistencies in previous literature and limited data in MDD/ELS, respectively.

2)

a. Based on previous literature, we hypothesised that ELS partially or fully drives the reduction of GMV often observed in MDD, in particular in the hippocampus and its subfields.
b. We hypothesised that reductions in GMV in the hippocampus and its subfields (particularly the cornu ammonis and dentate gyrus) would be partially or fully mediated by childhood sexual abuse (CSA) . We did not have

any specific a priori predictions pertaining to the amygdala and its nuclei given the inconsistency of previous findings in MDD and ELS.

c. Based on previous evidence in studies of depression, we hypothesised that MDD would be associated with increased cortical thickness in the anterior and posterior cingulate cortices, medial orbitofrontal cortex and insula, and decreased cortical thickness in the parahippocampal gyrus.

d. We hypothesised that CSA would be associated with widespread cortical thinning, particularly the parahippocampal gyrus which has been specifically implicated in CSA.

e. Given the scarcity of studies investigating cortical thickness in samples measuring both depression and ELS, we had no specific a priori predictions on possible interactions between CSA and MDD on cortical thickness. However, given the overlap in findings, we predicted that cortical thinning of the parahippocampal gryrus may be driven primarily by CSA, independent of MDD.

5. Overview of thesis

This PhD thesis contains 4 studies designed to address the overarching aim of disentangling effects of MDD and ELS on affective cognition and brain structure. **Studies 1-4** represent the above mentioned aims (1, 2a, 2b, and 2c), respectively.

This thesis has been arranged in 'Journal Format' style, to reflect the individual studies performed and allow for submission to peer-reviewed publications. The four studies include one systematic review (**Study 2**) and three cross-sectional studies (**Studies 1, 3, 4**) using a

variety of research methods and data analyses. More details on individual studies are given below.

5.1. Studies and rationale

While affective cognition deficits have been documented in both MDD and ELS individually and have yielded several corresponding results, very few studies have investigated this in samples measuring both depression and ELS simultaneously. Moreover, to our knowledge, no study to date has systematically or comprehensively assessed affective cognition in MDD and ELS within a single balanced design. This is particularly problematic since, as discussed above, affective cognition is a multifaceted construct and as such cannot be reduced to single tasks. **Study 1** addressed this gap in the literature by employing the EMOTICOM test battery to assess 11 distinct constructs of affective cognition in MDD versus HC (while controlling for ELS exposure by ensuring comparable distribution of ELS in both groups). Additionally, correlations between ELS and experimental measures were examined, both across participants and within diagnostic groups, to identify affective cognition constructs that may be particularly sensitive to ELS exposure.

Studies 2-4 aimed to extend the research question from the first study, in which the relative contribution of ELS and MDD on behavioural measures (specifically affective cognition) was assessed, to neurobiological measures. These studies address the second main aim of the PhD pertaining to structural brain changes (GMV and cortical thickness) following ELS and MDD. In order to inform the research question and analyses of subsequent studies directly assessing morphological indexes, a systematic review of the existing literature of GMV changes in samples measuring both ELS and MDD was conducted (**Study 2**). Results

emphasised that ELS may drive GMV reductions frequently observed in MDD, particularly of the hippocampus.

However, studies included in the systematic review were largely limited by low levels of ELS in the HC group (or ELS wasn't even measured in the HC to begin with), making it difficult to draw firm conclusions on the effect of ELS on GMV independent of depression. Furthermore, studies differed greatly in regard to type of ELS included, and the vast majority of studies did not conduct any additional analyses on ELS subtypes. This may be particularly problematic since a systematic review of neuroimaging findings found that specific subtypes of ELS (including sexual abuse, emotional abuse, and neglect) differentially predicted structural brain changes (whilst others were common to all ELS subtypes; Cassiers et al., 2018). Additionally, none of the studies identified through the systematic review restricted ELS participants according to timing of abuse, or controlled for this in analyses. Emerging research suggests that neurobiological sequalae of ELS may be particularly deleterious (and affect specific brain regions) at particular times during development (Cross et al., 2017). For instance, a study of hippocampal GMV found that childhood sexual abuse (CSA) at ages 3-5 and 11-14 predicted greatest atrophy in the hippocampus, corpus callosum, and frontal cortex (Andersen et al., 2008), while another study identified ELS at ages 10-11 as a sensitive period for later-life amygdala GMV abnormalities (Pechtel et al., 2014). Finally, several studies included in the systematic review did not control for total intracranial volume (eTIV), meaning potential differences between participants in head size and total GMV were unaccounted for and could potentially influence results (Maksimovskiy, 2019a; Maksimovskiy, 2019b).

We hence wanted to design an experimental study of GMV in MDD and ELS that addressed these methodological concerns identified by the systematic review and would thereby enable improved analyses of the relative effects of MDD and ELS. This is done in **Study 3,** which uses FreeSurfer (version 7.0) software to analyse hippocampal and amygdala GMV (including of individual subfields and nuclei) in 75 unmedicated female participants. The hippocampus was chosen as it had been identified as a primary region in the aforementioned systematic review (**Study 2**), and the amygdala was included because this novel version of FreeSurfer software enabled amygdala nuclei segmentation for the first time. Given the mixed evidence base on amygdala GMV in MDD and ELS (Hamilton et al., 2008; Paquola et al., 2016), it was proposed that nuclei segmentation may provide needed specificity to help clarify these heterogenous findings. To our knowledge, only five previous studies had analysed hippocampal and/or amygdala subfields in samples including at least some measure of depression and ELS, and many of these have suffered from the same methodological limitations previously discussed. Therefore, **Study 3** aimed to address the research question in the most controlled method possible – participants were specifically recruited for 4 groups, including MDD with and without significant CSA, and HC with and without significant CSA. In order to address the issue of timing and type of ELS, we limited inclusion to individuals experiencing significant CSA (though other types of abuse were also permissible for ecological validity and practical limitations as CSA rarely occurs in a vacuum; Molnar et al., 2001) during ages 5-14 (previously identified as particularly vulnerable time periods for later life hippocampal volume; Andersen et al., 2008).

Study 3 served as an extension of the GMV study to analyse effects of MDD and ELS on another morphological index that has been much less frequently studied in MDD and ELS,

cortical thickness. FreeSurfer software package was used for processing neuroimaging data and generating cortical thickness estimates, which has shown good agreement with histologic measurements (Cardinale et al., 2014). FreeSurfer employs a surface-based computational paradigm which has shown comparable results to voxel-based methods for assessing cortical thickness (Clarkson et al., 2011). To date, only a single previous study has investigated cortical thickness in a sample assessing both MDD and ELS (Jaworska et al., 2014). However, this study had several methodological limitations (such as an absence of a HC control group in ELS analyses and lack of correction for multiple comparisons) hindering clear conclusions about relative contributions of ELS and MDD to cortical thickness to be drawn. **Study 4** used the same sample as **Study 3** and hence shared the advantages of a carefully implemented experimental design with four distinct participant groups (unmedicated MDD/HC with high/absent ELS), validated measures of both MDD and ELS, and specifying type (CSA) and timing (ages 5-14) of ELS. This enabled both **Studies 3 and 4** to methodically assess relative contribution of ELS and MDD (and possible interactions) on two morphological indices (GMV and cortical thickness), using a novel study design aimed to better address the pertinent research question.

6. Conclusion

There is increasing evidence that both MDD and ELS have shared effects on cognitive and neurobiological measures, including affective cognition and brain structure. Despite the documented high prevalence of ELS in MDD and undisputed role of ELS as a key environmental risk factor for MDD, very few studies in depression have controlled for ELS. This has made it difficult to disentangle the relative effects of ELS and MDD (and possible interactions) on affective cognition and morphological measures. This PhD sought to

address this gap in the literature by designing and implementing highly controlled studies that allowed for a systematic and comprehensive assessment of affective cognition and analysis of GMV and cortical thickness in MDD and ELS.

REFERENCES

Aghamohammadi-Sereshki, A., Coupland, N. J., Silverstone, P. H., Huang, Y., Hegadoren, K. M., Carter, R., ... & Malykhin, N. V. (2021). Effects of childhood adversity on the volumes of the amygdala subnuclei and hippocampal subfields in individuals with major depressive disorder. *Journal of psychiatry and neuroscience*, *46*(1), E186-E195.

Anda, R. F., Whitfield, C. L., Felitti, V. J., Chapman, D., Edwards, V. J., Dube, S. R., & Williamson, D. F. (2002). Adverse childhood experiences, alcoholic parents, and later risk of alcoholism and depression. *Psychiatric services*, *53*(8), 1001-1009.

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of neuropsychiatry and clinical neurosciences*, *20*(3), 292-301.

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

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., ... & Thomas, S. M. (2008). The mental wealth of nations. *Nature*, *455*(7216), 1057-1060.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child abuse & neglect*, *27*(2), 169-190.

Bland, A. R., Roiser, J. P., Mehta, M. A., Schei, T., Boland, H., Campbell-Meiklejohn, D. K., ... & Viding, E. (2016). EMOTICOM: A Neuropsychological Test Battery to Evaluate Emotion, Motivation, Impulsivity, and Social Cognition. *Frontiers in behavioral neuroscience*, *10*.

Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., ... & O'Farrell, D. (2012). *The global economic burden of noncommunicable diseases* (No. 8712). Program on the Global Demography of Aging.

Bouhuys, A. L., Geerts, E., & Gordijn, M. C. (1999). Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study. The Journal of nervous and mental disease, 187(10), 595-602.

Bounoua, N., Miglin, R., Spielberg, J. M., & Sadeh, N. (2020). Childhood assaultive trauma and physical aggression: Links with cortical thickness in prefrontal and occipital cortices. *Neuroimage: clinical*, *27*, 102321.

Bremner, J. D., Narayan, M., Anderson, E. R., Staib, L. H., Miller, H. L., & Charney, D. S. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry*, *157*(1), 115-118.

Calem, M., Bromis, K., McGuire, P., Morgan, C., & Kempton, M. J. (2017). Meta-analysis of associations between childhood adversity and hippocampus and amygdala volume in non-clinical and general population samples. *NeuroImage: Clinical*, *14*, 471-479.

Campbell, S., Marriott, M., Nahmias, C., & MacQueen, G. M. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry*, *161*(4), 598-607.

Cardinale, F., Chinnici, G., Bramerio, M., Mai, R., Sartori, I., Cossu, M., ... & Ferrigno, G. (2014). Validation of FreeSurfer-estimated brain cortical thickness: comparison with histologic measurements. *Neuroinformatics*, *12*(4), 535-542.

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.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., ... & Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, *301*(5631), 386-389.

Cassiers, L. L., Sabbe, B. G., Schmaal, L., Veltman, D. J., Penninx, B. W., & Van Den Eede, F. (2018). Structural and functional brain abnormalities associated with exposure to different childhood trauma subtypes: A systematic review of neuroimaging findings. *Frontiers in psychiatry*, *9*, 329.

Chaney, A., Carballedo, A., Amico, F., Fagan, A., Skokauskas, N., Meaney, J., & Frodl, T. (2014). Effect of childhood maltreatment on brain structure in adult patients with major depressive disorder and healthy participants. *Journal of psychiatry & neuroscience: JPN*, *39*(1), 50.

Clarkson, M. J., Cardoso, M. J., Ridgway, G. R., Modat, M., Leung, K. K., Rohrer, J. D., ... & Ourselin, S. (2011). A comparison of voxel and surface based cortical thickness estimation methods. *Neuroimage*, *57*(3), 856-865.

Cohen, J. R., McNeil, S. L., Shorey, R. C., & Temple, J. R. (2019). Maltreatment subtypes, depressed mood, and anhedonia: A longitudinal study with adolescents. *Psychological trauma: theory, research, practice, and policy*, *11*(7), 704.

Colle, R., Segawa, T., Chupin, M., Dong, M. N. T. K. T., Hardy, P., Falissard, B., ... & Corruble, E. (2017). Early life adversity is associated with a smaller hippocampus in male but not female depressed in-patients: a case–control study. *BMC psychiatry*, *17*(1), 1-8.

Couture, S. M., Penn, D. L., & Roberts, D. L. (2006). The functional significance of social cognition in schizophrenia: a review. *Schizophrenia bulletin*, *32*(suppl 1), S44-S63.

Cross, D., Fani, N., Powers, A., & Bradley, B. (2017). Neurobiological development in the context of childhood trauma. *Clinical Psychology: Science and Practice*, *24*(2), 111.

Danese, A., Moffitt, T. E., Harrington, H., Milne, B. J., Polanczyk, G., Pariante, C. M., ... & Caspi, A. (2009). Adverse childhood experiences and adult risk factors for age-related disease: depression, inflammation, and clustering of metabolic risk markers. *Archives of pediatrics & adolescent medicine*, *163*(12), 1135-1143.

Dannlowski, U., Stuhrmann, A., Beutelmann, V., Zwanzger, P., Lenzen, T., Grotegerd, D., ... & Kugel, H. (2012). Limbic scars: long-term consequences of childhood maltreatment revealed by functional and structural magnetic resonance imaging. *Biological psychiatry*, *71*(4), 286-293.

Disner, S. G., Beevers, C. G., Haigh, E. A., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467-477.

Drevets, W. C., Price, J. L., & Furey, M. L. (2008). Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain structure and function*, *213*(1-2), 93-118.

Elliott, R., Zahn, R., Deakin, J. W., & Anderson, I. M. (2011). Affective cognition and its disruption in mood disorders. *Neuropsychopharmacology*, *36*(1), 153-182.

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological bulletin*, *139*(6), 1342.

Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*(2), 774-781.

Frodl, T., Janowitz, D., Schmaal, L., Tozzi, L., Dobrowolny, H., Stein, D. J., ... & Grabe, H. J. (2017). Childhood adversity impacts on brain subcortical structures relevant to depression. *Journal of psychiatric research*, *86*, 58-65.

Gerritsen, L., Van Velzen, L., Schmaal, L., van der Graaf, Y., van der Wee, N., Van Tol, M. J., ... & Geerlings, M. (2015). Childhood maltreatment modifies the relationship of depression with hippocampal volume. *Psychological medicine*, *45*(16), 3517.

Gold, A. L., Sheridan, M. A., Peverill, M., Busso, D. S., Lambert, H. K., Alves, S., ... & McLaughlin, K. A. (2016). Childhood abuse and reduced cortical thickness in brain regions involved in emotional processing. *Journal of Child Psychology and Psychiatry*, *57*(10), 1154-1164.

Green, M. F., Hellemann, G., Horan, W. P., Lee, J., & Wynn, J. K. (2012). From perception to functional outcome in schizophrenia: modeling the role of ability and motivation. *Archives of General Psychiatry*, *69*(12), 1216-1224.

Guyer, A. E., Kaufman, J., Hodgdon, H. B., Masten, C. L., Jazbec, S., Pine, D. S., & Ernst, M. (2006). Behavioral alterations in reward system function: the role of childhood maltreatment and psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1059-1067.

Hamilton, J. P., Siemer, M., & Gotlib, I. H. (2008). Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Molecular psychiatry*, *13*(11), 993-1000.

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, *23*(1), 56.

Han, K. M., Won, E., Sim, Y., & Tae, W. S. (2016). Hippocampal subfield analysis in medication-naive female patients with major depressive disorder. *Journal of affective disorders*, *194*, 21-29.

Harmer, C. J., Bhagwagar, Z., Perrett, D. I., Völlm, B. A., Cowen, P. J., & Goodwin, G. M. (2003). Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology*, *28*(1), 148-152.

Harmer, C. J., de Bodinat, C., Dawson, G. R., Dourish, C. T., Waldenmaier, L., Adams, S., ... & Goodwin, G. M. (2011). Agomelatine facilitates positive versus negative affective processing in healthy volunteer models. *Journal of Psychopharmacology*, *25*(9), 1159-1167.

Heim, C. M., Mayberg, H. S., Mletzko, T., Nemeroff, C. B., & Pruessner, J. C. (2013). Decreased cortical representation of genital somatosensory field after childhood sexual abuse. *American Journal of Psychiatry*, *170*(6), 616-623.

Heim, C., & Binder, E. B. (2012). Current research trends in early life stress and depression: Review of human studies on sensitive periods, gene–environment interactions, and epigenetics. *Experimental neurology*, *233*(1), 102-111.

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

Heim, C., Plotsky, P. M., & Nemeroff, C. B. (2004). Importance of studying the contributions of early adverse experience to neurobiological findings in depression. *Neuropsychopharmacology*, *29*(4), 641-648.

Hengartner, M. P. (2020). How effective are antidepressants for depression over the long term? A critical review of relapse prevention trials and the issue of withdrawal confounding. *Therapeutic Advances in Psychopharmacology*, *10*, 2045125320921694.

Huang, Y., Coupland, N. J., Lebel, R. M., Carter, R., Seres, P., Wilman, A. H., & Malykhin, N. V. (2013). Structural changes in hippocampal subfields in major depressive disorder: a high-field magnetic resonance imaging study. *Biological psychiatry*, *74*(1), 62-68.

Iglesias, J. E., Augustinack, J. C., Nguyen, K., Player, C. M., Player, A., Wright, M., ... & Alzheimer's Disease Neuroimaging Initiative. (2015). A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. *Neuroimage*, *115*, 117-137.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*(7), 748-751.

Jaworska-Andryszewska, P., & Rybakowski, J. K. (2019). Childhood trauma in mood disorders: neurobiological mechanisms and implications for treatment. *Pharmacological Reports*, *71*(1), 112-120.

Jaworska, N., MacMaster, F. P., Gaxiola, I., Cortese, F., Goodyear, B., & Ramasubbu, R. (2014). A preliminary study of the influence of age of onset and childhood trauma on cortical thickness in major depressive disorder. *BioMed research international*, *2014*.

Jiang, S., Postovit, L., Cattaneo, A., Binder, E. B., & Aitchison, K. J. (2019). Epigenetic modifications in stress response genes associated with childhood trauma. *Frontiers in psychiatry*, *10*, 808.

Keefe, R. S., & Harvey, P. D. (2012). Cognitive impairment in schizophrenia. In *Novel antischizophrenia treatments* (pp. 11-37). Springer Berlin Heidelberg.

Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). A Swedish national twin study of lifetime major depression. *American Journal of Psychiatry*, *163*(1), 109-114.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). A population-based twin study of major depression in women: the impact of varying definitions of illness. Archives of general psychiatry, 49(4), 257-266.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., … Wang, P. S. (2011). The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. *Epidemiologia E Psichiatria*

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., ... & Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British journal of psychiatry*, *197*(5), 378-385.

Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A., Weinberger, D. R., & Mattay, V. S. (2012). Normal age-related brain morphometric changes: nonuniformity across cortical thickness, surface area and gray matter volume?. *Neurobiology of aging*, *33*(3), 617-e1.

Li, M., Liu, S., D'Arcy, C., Gao, T., & Meng, X. (2020). Interactions of childhood maltreatment and genetic variations in adult depression: a systematic review. *Journal of affective disorders*.

Li, Q., Zhao, Y., Chen, Z., Long, J., Dai, J., Huang, X., ... & Sweeney, J. A. (2020). Meta-analysis of cortical thickness abnormalities in medication-free patients with major depressive disorder. *Neuropsychopharmacology*, *45*(4), 703-712.

Lim, L., Hart, H., Mehta, M., Worker, A., Simmons, A., Mirza, K., & Rubia, K. (2018). Grey matter volume and thickness abnormalities in young people with a history of childhood abuse. *Psychological Medicine*, *48*(6), 1034-1046.

Lin, E., & Tsai, S. J. (2019). Epigenetics and depression: an update. *Psychiatry investigation*, *16*(9), 654.

Lindert, J., von Ehrenstein, O. S., Grashow, R., Gal, G., Braehler, E., & Weisskopf, M. G. (2014). Sexual and physical abuse in childhood is associated with depression and anxiety over the life course: systematic review and meta-analysis. *International journal of public health*, *59*(2), 359-372.

Lindqvist, D., Mueller, S., Mellon, S. H., Su, Y., Epel, E. S., Reus, V. I., ... & Wolkowitz, O. M. (2014). Peripheral antioxidant markers are associated with total hippocampal and CA3/dentate gyrus volume in MDD and healthy controls–preliminary findings. *Psychiatry Research: Neuroimaging*, *224*(3), 168-174.

Lyons, M. J., Eisen, S. A., Goldberg, J., True, W., Lin, N., Meyer, J. M., ... & Tsuang, M. T. (1998). A registry-based twin study of depression in men. *Archives of general psychiatry*, *55*(5), 468-472.

Maksimovskiy, A. L., Fortier, C. B., Milberg, W. P., & McGlinchey, R. E. (2019a). A structural MRI study of differential neuromorphometric characteristics of binge and heavy drinking. *Addictive behaviors reports*, *9*, 100168.

Maksimovskiy, A. L., Oot, E. N., Seraikas, A. M., Rieselbach, M., Caine, C., Sneider, J. T., ... & Silveri, M. M. (2019b). Morphometric biomarkers of adolescents with familial risk for alcohol use disorder. *Alcoholism: Clinical and Experimental Research*, *43*(11), 2354-2366.

McCrory, E. J., Gerin, M. I., & Viding, E. (2017). Annual research review: childhood maltreatment, latent vulnerability and the shift to preventative psychiatry–the contribution of functional brain imaging. *Journal of child psychology and psychiatry*, *58*(4), 338-357.

McLaughlin, K. A., & Sheridan, M. A. (2016). Beyond cumulative risk: A dimensional approach to childhood adversity. *Current directions in psychological science*, *25*(4), 239-245.
Mikolas, P., Tozzi, L., Doolin, K., Farrell, C., O'Keane, V., & Frodl, T. (2019). Effects of early life adversity and FKBP5 genotype on hippocampal subfields volume in major depression. *Journal of affective disorders*, *252*, 152-159.

Miu, A. C., Bîlc, M. I., Bunea, I., & Szentágotai-Tătar, A. (2017). Childhood trauma and sensitivity to reward and punishment: Implications for depressive and anxiety symptoms. *Personality and Individual Differences*, *119*, 134-140.

Molnar, B. E., Buka, S. L., & Kessler, R. C. (2001). Child sexual abuse and subsequent psychopathology: results from the National Comorbidity Survey. *American journal of public health*, *91*(5), 753.

Murphy, S. E., Downham, C., Cowen, P. J., & Harmer, C. J. (2008). Direct effects of diazepam on emotional processing in healthy volunteers. *Psychopharmacology*, *199*(4), 503-513.

Nanni, V., Uher, R., & Danese, A. (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *American Journal of Psychiatry*, *169*(2), 141-151.

Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. *PLoS medicine*, *9*(11), e1001349.

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., ... & Marder, S. R. (2008). The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203-213.

Opel, N., Redlich, R., Zwanzger, P., Grotegerd, D., Arolt, V., Heindel, W., ... & Dannlowski, U. (2014). Hippocampal atrophy in major depression: a function of childhood maltreatment rather than diagnosis?. *Neuropsychopharmacology*, *39*(12), 2723-2731

Papmeyer, M., Giles, S., Sussmann, J. E., Kielty, S., Stewart, T., Lawrie, S. M., ... & McIntosh, A. M. (2015). Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder. *Biological psychiatry*, *78*(1), 58-66.

Paquola, C., Bennett, M. R., & Lagopoulos, J. (2016). Understanding heterogeneity in grey matter research of adults with childhood maltreatment—A meta-analysis and review. *Neuroscience & Biobehavioral Reviews*, *69*, 299-312.

Paquola, C., Bennett, M. R., & Lagopoulos, J. (2016). Understanding heterogeneity in grey matter research of adults with childhood maltreatment—A meta-analysis and review. *Neuroscience & Biobehavioral Reviews*, *69*, 299-312.

Paquola, C., Bennett, M. R., & Lagopoulos, J. (2016). Understanding heterogeneity in grey matter research of adults with childhood maltreatment—A meta-analysis and review. *Neuroscience & Biobehavioral Reviews*, *69*, 299-312.

Pardoe, H. R., Abbott, D. F., Jackson, G. D., & Alzheimer's Disease Neuroimaging Initiative. (2013). Sample size estimates for well-powered cross-sectional cortical thickness studies. *Human brain mapping*, *34*(11), 3000-3009.

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: an integrated review of human literature. *Psychopharmacology*, *214*(1), 55-70.

Pechtel, P., Lyons-Ruth, K., Anderson, C. M., & Teicher, M. H. (2014). Sensitive periods of amygdala development: the role of maltreatment in preadolescence. *Neuroimage*, *97*, 236-244.

Penner-Goeke, S., & Binder, E. B. (2019). Epigenetics and depression. *Dialogues in clinical neuroscience*, *21*(4), 397.

Rao, U., Chen, L. A., Bidesi, A. S., Shad, M. U., Thomas, M. A., & Hammen, C. L. (2010). Hippocampal changes associated with early-life adversity and vulnerability to depression. *Biological psychiatry*, *67*(4), 357-364.

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological medicine*, *44*(10), 2029-2040.

Roiser, J. P., & Sahakian, B. J. (2013). Hot and cold cognition in depression. *CNS spectrums*, *18*(3), 139-149.

Roiser, J. P., Cannon, D. M., Gandhi, S. K., Tavares, J. T., Erickson, K., Wood, S., ... & Drevets, W. C. (2009). Hot and cold cognition in unmedicated depressed subjects with bipolar disorder. *Bipolar disorders*, *11*(2), 178-189.

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*, *37*(1), 117-136.

Saleh, A., Potter, G. G., McQuoid, D. R., Boyd, B., Turner, R., MacFall, J. R., & Taylor, W. D. (2017). Effects of early life stress on depression, cognitive performance and brain morphology. *Psychological medicine*, *47*(1), 171-181.

Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinssen, R. K., ... & Cuthbert, B. N. (2010). Developing constructs for psychopathology research: research domain criteria. *Journal of abnormal psychology*, *119*(4), 631.

Saygin, Z. M., Kliemann, D., Iglesias, J. E., van der Kouwe, A. J., Boyd, E., Reuter, M., ... & Alzheimer's Disease Neuroimaging Initiative. (2017). High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: manual segmentation to automatic atlas. *Neuroimage*, *155*, 370-382.

Scott, K. M., Smith, D. R., & Ellis, P. M. (2010). Prospectively ascertained child maltreatment and its association with DSM-IV mental disorders in young adults. *Archives of general psychiatry*, *67*(7), 712-719.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. In *Journal of Clinical Psychiatry* (Vol. 59, pp. 22–33). http://doi.org/10.1016/S0924-9338(99)80239-9.

Smith, K. E., & Pollak, S. D. (2020). Early life stress and development: potential mechanisms for adverse outcomes. *Journal of Neurodevelopmental Disorders*, *12*(1), 1-15.

Suh, J. S., Schneider, M. A., Minuzzi, L., MacQueen, G. M., Strother, S. C., Kennedy, S. H., & Frey, B. N. (2019). Cortical thickness in major depressive disorder: a systematic review and meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *88*, 287-302.

Teicher, M. H., Anderson, C. M., & Polcari, A. (2012). Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proceedings of the National Academy of Sciences*, *109*(9), E563-E572.

Treadway, M. T., Waskom, M. L., Dillon, D. G., Holmes, A. J., Park, M. T. M., Chakravarty, M. M., ... & Pizzagalli, D. A. (2015). Illness progression, recent stress, and morphometry of hippocampal subfields and medial prefrontal cortex in major depression. *Biological psychiatry*, *77*(3), 285-294.

Videbech, P., & Ravnkilde, B. (2004). Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry*, *161*(11), 1957-1966.

Vythilingam, M., Heim, C., Newport, J., Miller, A. H., Anderson, E., Bronen, R., ... & Bremner, J. D. (2002). Childhood trauma associated with smaller hippocampal volume in women with major depression. *American Journal of Psychiatry*, *159*(12), 2072-2080.

Widom, C. S., DuMont, K., & Czaja, S. J. (2007). A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. *Archives of general psychiatry*, *64*(1), 49-56.

Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., & Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: data from the randomized international study to predict optimized treatment for depression. *Translational psychiatry*, *6*(5), e799-e799.

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., ... & Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage*, *53*(3), 1135-1146.

Xie, P., Wu, K., Zheng, Y., Guo, Y., Yang, Y., He, J., ... & Peng, H. (2018). Prevalence of childhood trauma and correlations between childhood trauma, suicidal ideation, and social support in patients with depression, bipolar disorder, and schizophrenia in southern China. *Journal of affective disorders*, *228*, 41-48.

Yang, S., Cheng, Y., Mo, Y., Bai, Y., Shen, Z., Liu, F., ... & Xu, X. (2017). Childhood maltreatment is associated with gray matter volume abnormalities in patients with first-episode depression. *Psychiatry Research: Neuroimaging*, *268*, 27-34.

Study 1: Comprehensive Assessment of Affective Cognition using the EMOTICOM Test Battery in Depression and Early Life Stress

Franziska Goer[1], Amy Bland[2], Richard Drake[1], Joanna Neill[1], Rebecca Elliott[1]
[1]University of Manchester, Oxford Road, Manchester M13 9PL, UK
[2] Manchester Metropolitan University, All Saints, Manchester, M15 6BH

Author Contributions:

Franziska Goer contributed to study design, submitted applications to the University of Manchester and Research Ethics Service, created all study materials, collected the data (including conducting all clinical interviews), conducted all analyses, wrote the manuscript and completed all revisions following review by the co-authors.

Professor Rebecca Elliott also contributed to study design, provided methodological guidance and supervisory input throughout the entire project, gave guidance and feedback on data collection and analyses, and critically reviewed and provided feedback on drafts of the manuscript prepared by FG.

Dr. Amy Bland contributed to study design relating to EMOTICOM tasks, provided the coding for tasks, and provided guidance on individual task analysis. Dr. Richard Drake provided supervisory input and guidance for recruitment and study design. Professor Joanna Neill provided supervisory input.

Funding Information:
This study was supported by funding from the Medical Research Council and University of Manchester as part of Franziska Goer's Medical Research Council (MRC) Doctoral Training Scholarship.

ABSTRACT

Introduction: Major depressive disorder (MDD) has been consistently associated with changes in affective cognition, characterised in particular by a negative affective bias in emotion recognition and categorisation, as well as reduced effort/motivation and reward sensitivity, as well as changes in social cognition (such as moral emotions). Emerging research in early life stress (ELS) has highlighted several similarities in changes in affective cognition. However, since the vast majority of studies have investigated either MDD or ELS in isolation, it is unclear how their contributions to affective cognition interact. Drawing conclusions from previous research has been further complicated by lack of a standardised measure of affective cognition. A large variety of individual tasks have been applied that measure different components of the multifaceted construct of affective cognition and consequently hinder comparisons between studies. The aim of the present investigation was hence to systematically and comprehensively investigate different components of affective cognition in MDD and ELS within a single study design.
Methods: 52 participants (aged 18 - 62), including 27 healthy controls (HC) and 25 participants with current MDD, completed 11 tasks of the standardised and validated EMOTICOM test battery comprising emotion processing, motivation and reward, and social cognition. All participants were screened for major psychiatric disorders using the Mini-International Neuropsychiatric Interview (MINI) version 7.0 (for DSM 5) and completed the Childhood Trauma Questionnaire (CTQ) to assess ELS.
Results: Results indicated that specific constructs of affective cognition were affected primarily by diagnosis (including emotion recognition and categorisation, motivation/effort, and emotional memory) while others revealed no group differences (MDD versus HC) and instead appeared more sensitive to ELS (including attentional bias, value-based choice, and moral emotions). Null results were recorded for the Monetary Incentive Reward (MIR) task, a measure of reward sensitivity, and three social cognition tasks including social economic exchange games and theory of mind.
Conclusion: Collectively, the results emphasise that affective cognition as measured by the EMOTICOM test battery, is comprised of multiple, distinct constructs, that appear to be differentially affected by MDD and ELS. The study highlights the importance of measuring and controlling for ELS in studies of depression and affective cognition to avoid misattribution of findings (given the high prevalence of ELS in MDD). This study was limited in sample size and consequently underpowered, and hence did not correct for multiple comparisons – results should hence be interpreted cautiously, and future studies are needed to replicate findings.

1. Introduction

1.1. Affective 'Hot' and Non-affective 'Cold' Cognition

Cognition may be divided into two general constructs - 'hot' (emotion-laden or affective) and 'cold' (emotion-independent or non-affective) cognition (Roiser et al., 2009; Roiser & Sahakian, 2013). While depression has been associated with deficits in cold cognition, including executive functioning, memory, and attention (Keefe and Harvey, 2012; Rock et al., 2014), research has indicated that specific deficits in affective cognition may not only lead to development of psychopathology and maintenance of symptoms, but also predict functional outcomes (Couture et al., 2006; Insel et al., 2010; Roiser et al., 2012; Sanislow et al., 2010). Changes in affective cognition in depression are various, including mood-congruent processing biases and deficits in motivation, reward and punishment processing/learning (Roiser & Sahakian, 2013). Importantly, several affective cognition changes observed in major depressive disorder (MDD) have been similarly reported in individuals with a history of significant early life stress (ELS; Elliott et al., 2011; Pechtel & Pizzagalli, 2011). Despite the greater than 2-fold increase of ELS in MDD as compared to the general population (estimated as up to 62.5% of MDD patients with significant levels of ELS), the majority of studies of cognition in MDD have not measured or controlled for ELS in statistical analyses. It is hence unclear whether ELS may contribute to changes in affective cognition observed in MDD.

1.2. Measuring affective cognition

A multitude of behavioural tests have been applied to investigate affective cognition, and this lack of standardisation across studies complicates drawing conclusions about changes observed in various psychiatric disorders, including MDD (Elliott et al., 2011). In order to

comprehensively assess affective cognition in MDD and ELS, this study applied a recently developed and validated test battery for affective cognition (EMOTICOM; Bland et al., 2016). While the original test battery incorporates 17 tasks spanning 4 proposed main domains of affective cognition, including emotional processing, motivation, social cognition, and impulsivity, the present study excluded the final category to make testing duration more feasible in a clinical population. The decision to eliminate impulsivity related tasks was based on both practical limitations (recommendations by the original task developers based on need for modifications to improve some of the tasks in this domain; Dam et al., 2019) and literature suggesting a complex picture of impulsivity in MDD, which may be mediated by comorbid anxiety (Bellani et al., 2012; Del Carlo et al., 2012) and dependent on age (Moustafa et al., 2017). However, a recently published systematic review has highlighted higher levels of impulsivity in MDD relative to healthy controls (HC; Fields et al., 2021), and hence future studies may consider including this domain in the analysis of affective cognition.

1.3. Affective cognition in MDD and ELS

1.3.1. Emotion processing domain

Studies have found effects of both ELS and MDD on emotion processing. The EMOTICOM test battery includes 4 tasks in the emotion processing domain, measuring diverse constructs including emotion recognition/categorisation, attentional bias, and emotional memory (see **Table 1** for a detailed list of tasks and constructs measured).

Previous studies in patients with depression have demonstrated a negative (mood-congruent) processing bias in which participants are either quicker, more accurate, or both,

in responding to negative stimuli than positive or neutral stimuli, in contrast to healthy control participants who show the opposite pattern of responding with a bias towards positive stimuli (Erickson et al., 2005; Murphy et al., 1999). Similarly, MDD patients show reduced accuracy in identifying positive facial expressions than healthy controls, an effect that is reversed in patients following antidepressant drug administration (Harmer et al., 2009). Furthermore, depressed patients have demonstrated a negative processing bias in the interpretation of neutral facial expressions - they are significantly more likely to interpret neural facial expression as sad than HC (Gollan et al., 2008). Fewer studies have investigated emotion processing in ELS, but studies in children exposed to ELS have indicated abnormal emotional processing and categorisation of facial expressions, with a heightened propensity to identify potential threats in the form of angry facial expressions (Pollak and Kistler, 2002) and a heightened neural response to negative facial expressions (particularly anger) compared to happy facial expressions (Pollak et al., 2000; Pollak et al., 2001).

Research has highlighted that depression may not merely be associated with general memory impairment but in fact show a negative memory bias, in which memory for negative stimuli is enhanced compared to healthy controls (Bradley et al., 1995; Hamilton & Gotlib, 2008; Ridout et al., 2003; Watkins et al., 1992). Furthermore, while healthy individuals exhibit a relative better memory for positive than negative or neutral material, this effect has been found to be reduced/absent in those with MDD (Gotlib et al., 2004; Hamilton & Gotlib, 2008). Deficits in recall of positive memories may in fact be even more pronounced and reliable than evidence of mood-congruent memory bias (improved accuracy for recall of negative stimuli) in MDD (Burt et al., 1995).

Some studies have suggested that memory impairments in depression, including enhanced memory for negative material and diminished memory for positive stimuli may be caused by biological sequelae of chronic stress (Dillon & Pizzagalli, 2018). It is hence possible that ELS may play a crucial role in the development of an affective bias in memory in MDD. Unfortunately, to date, no studies, to the authors knowledge, have investigated memory biases in depression while controlling for ELS. Several studies point to general memory impairments in ELS (Majer et al., 2010; Pizzagalli & Pechtel, 2010), however, these studies have used measures of non-affective memory, making it impossible to assess possible affective memory biases.

1.3.2. Motivation and reward

Deficits in motivation and reward have been consistently implicated in MDD, and emerging research in ELS has identified similar deficits. Three tasks in the EMOTICOM battery fall into this domain, measuring specific constructs of reward/punishment sensitivity, incentive motivation and value-based choice (see **Table 1** for a detailed list of tasks on corresponding constructs measured).

Anhedonia (characterised by both anticipatory/decisional anhedonia, decreased motivation for obtaining rewards, and consummatory anhedonia, decreased pleasure from reward) is one of the core diagnostic features of MDD (American Psychiatric Association, 1994). Decreased motivation for rewards has been observed extensively in behavioural studies in MDD, often using the Effort- Expenditure for Rewards Task (EEfRT; Treadway et al., 2009) in which participants decide whether to complete the 'easy task' (fewer button presses using

finger of the dominant hand for a smaller monetary reward) versus the 'hard task' (significantly more button presses using the finger of the non-dominant hand for a larger monetary reward). Studies using the EEfRT have demonstrated lower motivation (less willingness to expend effort to receive monetary reward) in individuals with MDD and those with subclinical depressive symptoms compared to healthy controls (Treadway et al., 2012; Yang et al., 2014). Evidence suggests that anticipatory anhedonia may be related to abnormalities in dopamine circuitry in depressed individuals (Treadway & Zald, 2013).

The Progressive Ratio Task (PRT) used in the EMOTICOM battery is a behavioural task that has been extensively applied to measure motivation to obtain a reward in both the animal literature and human studies. A simple task is performed to obtain reward, which gets progressively longer/harder (more trials needed to obtain reward) and progressively less highly rewarded (the monetary value of reward decreases with increasing trials). Motivation is typically measured as the 'breakpoint' – the number of trials participants complete before choosing to quit the task. The PRT has been successfully used to assess motivation for reward in a variety of clinical samples, ranging from gastric bypass surgery patients (Miras et al., 2012) to substance abuse (Stoops, 2008). Furthermore, studies have reported positive correlations between self-reported motivation and diligence and higher breakpoints on the PRT (Dam et al., 2019). However, applications of the PRT in depression research have been limited to date. The few existing studies appear to suggest lower breakpoints in individuals with depression that may be reversed following successful treatment. One study demonstrated increased breakpoints in depressed patients who showed improvements following antidepressant treatment while no change in breakpoints was observed for patients showing no improvements (Hughes et al., 1985). However, this study did not have a

control group and was limited by very small sample size (N=6). More recently, a study observed significantly lower breakpoints on a progressive ratio task in both unipolar and bipolar depression, relative to healthy controls (Hershenberg et al., 2016). In regard to ELS, studies on motivation and effort are limited to date. Several studies point to decreased sensitivity to reward in individuals with ELS which may even mediate the relationship between ELS and depressive symptoms (Guyer et al., 2006; Miu et al., 2017).

Value-based choice and risk adjustment in the EMOTICOM battery is assessed using the adapted Cambridge Gambling Task (CGT). Relative to healthy controls, individuals with MDD have been shown to be risk-averse in decision-making and show a heightened tendency to employ a conservative/risk-averse strategy even when the probability of winning a reward is high (Murphy et al., 2001), a pattern that has also been observed in healthy young adults with a first degree relative of MDD (Mannie et al., 2015). Studies in ELS have been limited to date – in a study of children (aged 8-14), those with ELS were found to be quicker to select high risk options compared to healthy controls, while children with both ELS and depressive disorders were more risk averse than healthy controls (Guyer et al., 2006). Interestingly, a study using two different types of monetary decision-making tasks (with varying probabilities of win/loss), found that ELS was the single biggest predictor of loss aversion in MDD patients (Huh et al., 20166), though it should be noted that no HC comparison group was studied, and as such these findings were limited to depressed patients.

The Monetary Incentive Reward (MIR) task employed in the EMOTICOM test battery is a frequently used measure of reward anticipation and sensitivity, which has been implicated in both ELS and MDD. Both behavioural and neuroimaging studies have shown dysfunctional

reinforcement anticipation and response in MDD (Murphy et al., 2003; Pizzagalli et al., 2008). While studies using the MIR tasks in MDD have often failed to identify any effects of diagnosis on behavioural outcome measures, functional magnetic resonance imaging (fMRI) studies have documented reduced activation in neural areas associated with reward in MDD relative to HC (Knutson et al., 2008; Pizzagalli et al., 2009). In regard to ELS, animal literature has demonstrated decreased anticipation for reward in rats following early maternal separation (Mathews & Robbins, 2003) and parental deprivation in marmoset monkeys (Pryce et al., 2004). Studies in humans suggest that timing of ELS may play a crucial role – while ELS during younger childhood appears to lead to decreased reward responsiveness and approach motivation, ELS occurring during later childhood (teenage years) may show the opposite pattern (Novick et al., 2018). This may be attributable to different stages of brain development for different brain regions, which evidence suggest may be sensitive to ELS during times of peak maturation (Andersen et al., 2008). Decreased responsiveness to reward anticipation in ELS has also been documented in corresponding neural pathways, as seen in fMRI studies of the MIR task which revealed decreased activity of structures of the basal ganglia during reward anticipation in ELS (decreased sensitivity to anticipation of reward; Boecker et al., 2014; Dillon et al., 2009; Mehta et al., 2010). However, it should be noted that these studies did not control for depression – moreover, self-reported depressive symptoms and anhedonia were heightened in ELS subjects relative to healthy controls in one study, which may also contribute to decreased reward sensitivity (Dillon et al., 2009).

1.3.3. Social cognition

The EMOTICOM test battery allows for a detailed assessment of moral emotions (using a novel task with cartoon scenarios), theory of mind (correctly identifying the mental states of other people; Frith & Frith, 2003) and social decision-making (using economic games including the Ultimatum Game and Prisoners' Dilemma; Bland et al., 2016). Moral emotions are responses to moral norms and shape humans' unique moral thoughts and behaviours (Tangney et al., 2007). Haidt (2003) argues that moral emotions are comprised of other-condemning emotions (including contempt and disgust), self-conscious emotions (including shame and guilt), other-suffering emotions (compassion) and other-praising emotions such as gratitude. Depression has been associated with abnormal processing of moral emotions such as guilt and shame, the latter of which may be particularly maladaptive in the maintenance of the disorder (Orth et al., 2006; Tangney et al., 1992). While studies of moral emotions in ELS are sparse to date, there is growing evidence that self-blame may play a critical role in mediating the relationship between ELS and later life psychiatric disorders including PTSD (Sharma-Patel & Brown, 2016) and depression (Dorresteijn et al., 2019). Furthermore, mirroring findings in MDD, ELS appears to be associated with heightened levels of shame and guilt which in turn are linked with increased depressive symptoms (Sekowski et al., 2020).

In regard to theory of mind (ToM), results in MDD are mixed and likely depend on the nature of the task purportedly measuring ToM (Wolkenstein et al., 2011). Studies in ELS suggest that ToM may be impaired, in particular in those with parental physical abuse (Germine et al., 2015). Some evidence suggests that ELS may be the driving factor of ToM impairments observed in MDD (in particular via emotional and physical neglect; Simon et al., 2019). However, the majority of cited studies employed the Reading the Mind in the

Eyes Test (RMET) task of ToM in which participants are asked to infer emotional states based solely on pictures of eyes. Some have argued that the RMET is less an assessment of the ability to infer the mental states of others and rather resembles a more basic emotion identification and classification task, which has been supported by research from autism spectrum disorder and alexithymia (Oakley et al., 2016). As such, more studies are needed in both MDD and ELS that employ alternative measures of ToM using more suitable tasks. The Social Information Preference Test in the EMOTICOM test battery aims to do so, by using cartoons of social scenarios and asking participants to choose information (thoughts, facial expressions, or facts) to help them make judgments about others' mental states in the cartoon (Bland et al., 2016).

The Ultimatum Game (UG) and Prisoners' Dilemma (PD) are social economic exchange games that assess sensitivity to fairness and cooperation with an opponent, respectively. While frequently studied in the field of economic decision-making in healthy controls, these games have rarely been assessed in depression to date. One study found that depressed individuals had a more aversive emotional reaction to unfair offers in the ultimatum game but were nevertheless less likely to reject these offers than healthy control participants (Harlé et al., 2010). Similarly, another study found higher levels of guilt in depressed compared to healthy individuals in response to receiving unfair offers in the ultimatum game (Pulcu et al., 2014). Interestingly, the opposite pattern was observed in healthy individuals who underwent a negative mood induction (by watching a sad movie clip) who subsequently had significantly lower acceptance rate to unfair offers (while a positive mood induction had no significant effects; Harlé & Sanfey, 2007). To the authors' knowledge,

neither UG nor PD style social economic exchange games have been studies in the context of ELS.

1.4. Study aims and hypotheses

While affective cognition in MDD has been extensively studied, previous research has not taken into account the potential confounding factor of ELS. The research reviewed above indicates that several constructs of affective cognition purportedly affected by depression may be similarly implicated in ELS. It is hence possible that ELS may be at least partially moderating effects of depression on affective cognition. However, since studies to date have tended to assess effects of ELS or MDD on affective cognition in isolation and using different tasks that may measure slightly different and hence not comparable constructs, it is not yet possible to disentangle individual effects. The aim of this study hence was twofold:

1. To systematically assess affective cognition using a standardised, validated test battery (EMOTICOM) in MDD while controlling for ELS (by matching MDD and HC groups for ELS exposure)
2. To assess the effect of ELS on affective cognition by examining correlations between CTQ scores (including subscales of different types of abuse and neglect) and affective cognition outcome variables, both across participants and within each diagnostic group.

Given the lack of previous studies simultaneously assessing MDD and ELS in affective cognition, we did not have specific hypotheses about the exact nature of MDD versus ELS contributions for individual EMOTICOM tasks. In line with previous literature, we expected

that MDD would be associated with negative biases during emotion processing and impaired motivational/reward function, though some of these findings may be, at least partially, accounted for by ELS. We expected ELS to be more sensitive to emotion processing biases toward threat related stimuli, such as angry and fearful faces. While previous literature on social cognition in MDD/ELS is sparser, we hypothesised that MDD/ELS would be associated with changes in moral emotions, in particular heightened feelings of guilt/shame and self-blame. Due to the inconsistent nature of previous studies for ToM and extremely limited data on social economic exchange games in MDD (and no studies to date in ELS), we had no specific *a priori* hypotheses for these three tasks in the social cognition domain.

2. Methods

2.1. Ethical approval, recruitment, testing location

The present investigation received ethical approval from both the University of Manchester Research Ethics Committee (2017-0370-2542) and the North West Research Ethics Committee (18/NW/0071). Participants were recruited from the community via posters, social media, research participant recruiting platforms, and University of Manchester staff and student email circulations. Specific advertisements for individuals with ELS were included to aim to recruit high ELS participants for both the MDD and HC groups. Interested individuals completed an online pre-screening tool to ascertain eligibility and were subsequently phone-screened if deemed likely to be eligible. All study sessions took place at the University of Manchester in designated behavioural testing rooms.

2.2. Participants

A total of 507 participants completed the online screening form, of whom approximately 25% were further assessed for eligibility via a brief phone screen. 60 participants came in for the first study session and provided informed consent, of whom 8 were ineligible (due to subthreshold major depressive episode (MDE), N= 3, current alcohol/substance abuse (N=3), and current anorexia nervosa (N=1)) and discontinued the study following the clinical interview. The remaining 52 participants (aged 18 - 62) completed the full study and all experimental measures. Participant groups (MDD versus HC) did not significantly differ on demographic factors, including age, gender, ethnicity, education, native English speakers, or current smokers (see **Table 3** for detailed results).

2.2.1. Inclusion/exclusion criteria

Inclusion criteria for all participant included age 18-65, capable of giving written informed consent, and English fluency. Exclusion criteria for all participants included history or presence of medical condition requiring centrally acting medications, history or presence of neurological disease (including but not limited to, stroke, epilepsy, space occupying lesions, multiple sclerosis, Parkinson's disease, vascular dementia, transient ischemic attack, that may influence the outcome of any cognitive testing), clinically significant head injury (e.g., requiring medical or surgical intervention), neuroendocrine disorder, including impaired thyroid function and steroid use, unwillingness or inability to follow the procedures outlined in the protocol, and female participants who are, or may be, pregnant. Further eligibility criteria for HC participants included no first degree relative (FDR) diagnoses with MDD or schizophrenia, no psychiatric illness (past or present). Additional inclusion criteria for the MDD group included meeting DSM-5 diagnostic criteria for current MDD and absence of any psychotropic medications for at least 2 weeks (6 weeks for fluoxetine; 6 months for

neuroleptics; 2 weeks for benzodiazepines; 2 weeks for any other antidepressants). While some comorbid diagnoses were not excluded in MDD participants (due to the high level of psychiatric comorbidity in MDD; Thaipisuttikul et al., 2014) exclusion criteria included meeting criteria for current or past (moderate or severe) alcohol or substance use, current or past psychotic disorder or manic/hypomanic episode, current bulimia nervosa or binge-eating disorder, current or past anorexia, and primary[2] diagnoses of current or past panic disorder, agoraphobia, social anxiety disorder, obsessive-compulsive disorder, posttraumatic stress disorder, bulimia nervosa or binge-eating disorder. The Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998) version 7.0 (for DSM 5) was used to screen all participants for major psychiatric disorders.

2.3. Procedures

Signed and dated consent sheets were obtained from each participant at the start of the first study session. Participants were then screened for eligibility using the MINI and, if eligible, continued with questionnaires and EMOTICOM tasks. All participants completed the study over two sessions to reduce fatigue from cognitive testing. Tasks order was randomised, with exception of the Emotional Memory Recognition task which was always performed at the start and end of the first session to ensure a 1-hour interval between encoding and retrieval. Participants were closely monitored for mood changes using a mood rating form before and after each study session to ensure participant well-being. The first study session took approximately 2- 3 hours (depending on length of the clinical interview), and the second study session lasted approximately 1.5 - 2 hours.

[2] These diagnoses were not exclusionary if secondary to MDD diagnosis (i.e., MDD was the primary diagnosis)

2.4. Measures

2.4.1. MDD and ELS

MDD was assessed using the MINI interview (Sheehan et al., 1998), Beck Depression Inventory-II (BDI-II; Beck et al., 1996) and the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). ELS was assessed using the short form of the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003), a 28-item self-report questionnaire of 5 types of ELS experienced before age 18: emotional abuse (EA), physical abuse (PA), sexual abuse (SA), emotional neglect (EN), and physical neglect (PN).

2.4.2. EMOTICOM test battery

11 tasks form the original EMOTICOM test battery were selected, representing 3 of the 4 domains included in the original study (Emotion Processing, Motivation and Reward, and Social Cognition; Bland et al., 2016). See **Table 1** for a list of each task, constructs assessed, and duration. The Emotional Memory Recognition Task was the only task altered from the original study (upon recommendation from the task developers) since it had been associated with high ceiling effects in 200 healthy controls. To increase difficulty of the task, rather than being asked to identify images seen at an earlier encoding stage (1 hour previously) relative to novel images, participants were asked to identify the original image from the mirror image.

One of the advantages of the EMOTICOM test battery is the nuanced outcomes for each task, allowing researchers to investigate specific research questions. For purposes of this paper, and to reduce the number of multiple comparisons, outcome variables were significantly reduced and mirrored those selected by the original publication (Bland et al.,

2016) and/or more recent Danish validation study (Dam et al., 2019). Some adjustments were made based on recommendations from these papers (such as using reaction time (RT), rather than *d*-prime for the Face Affective Go/ No-Go Task, since RT is less vulnerable to the ceiling effects observed in the task; Dam et al., 2019). Outcome variables for each task are given in **Table 2.** Detailed descriptions of each task and calculation of outcome variables is described below.

Affective cognition domain	**Name of Task**	**Main construct measured**	**Time to complete (approximate)**
Emotion Processing	Emotional Recognition Task (ERT)	Emotion recognition/categorisation	12 minutes
	Emotional Intensity Morphing Task	Emotion recognition	5 minutes
	Face Affective Go No-Go (FAGN) Task	Attentional bias (information processing bias)	6 minutes
	Emotional Memory Recognition Task	Emotional memory	10 minutes
Motivation and Reward	Monetary Incentive Reward (MIR) Task	Reward/punishment sensitivity and incentive motivation	10 minutes
	Progressive Ratio Task (PRT)	Incentive motivation / effort	20 minutes
	Adapted Cambridge Gambling Task (CGT)	Value-based choice	10 minutes
Social Cognition	Moral Emotions Task	Moral emotion	13 minutes
	Social Information Preference Test	Theory of mind	10 minutes
	Prisoners' Dilemma (PD)	Social economic exchange game	10 minutes
	Ultimatum Game (UG)	Social economic exchange game	12 minutes
		Total time	**118 minutes**

Table 1. EMOTICOM tasks and duration.

2.4.2.1. Emotional Recognition Task (ERT)

The ERT measures emotion recognition/categorisation. In this task, participants are shown a series of emotional faces, varying from low intensity (1 – neutral) to high intensity (10 – maximum intensity), that flash on the screen briefly after which they are asked to identify the emotion (happiness, sadness, fear, or anger). Participants complete 4 practice trials (disregarded in analyses) and are then shown a total of 80 faces in randomised order (each level of intensity for each emotion is shown twice). There is also a short (16 trial) control block (also disregarded in analyses) in which participants are asked to judge the age of the face shown (child, young adult, middle aged, elderly). A touchscreen is used for participants to make their selection (from the four choices) after each face is displayed.

The mean accuracy (%) for each emotion was calculated. Additionally, an affective bias score was calculated (mean accuracy for happy faces minus mean accuracy for sad faces).

2.4.2.2. Emotional Intensity Morphing Task

The Emotional Intensity Morphing Task is a measure of emotion recognition in which participants are asked to judge at which level of intensity they can either detect or no longer detect a given (pre-specified) facial expression of emotion. Both the increase condition (neutral to high intensity) and decrease condition (high intensity to neutral) consist of the same 15 levels of intensity for each of 5 emotions: happiness, sadness, anger, fear, and disgust. Half of the trials consist of female faces; the other half consist of male faces. A practice block is administered at the start of the task using the emotion surprise. Prior to each trial, participants are informed whether to press space when they SEE the emotion

(increase condition) or NO LONGER see the emotion (decrease condition). For each trial, participants are told in advance which emotion they will be asked to detect.

The average detection point (level at which participants either detected or no longer detected a given emotion) was calculated separately for increasing and decreasing conditions for each emotion. Disgust was excluded to decrease number of outcome variables and multiple comparisons, as it has not been as frequently studied in MDD/ELS and hence was less central to our research question and aims.

Variables of interest were limited to increasing trials, while decreasing trials served as a control measure for impulsive responding (I.e. to ensure participants did not merely impulsively/quickly press space on all trials, regardless of emotion/condition).

2.4.2.3. Face Affective Go No-Go (FAGN) Task

The FAGN Task measures attentional bias/ information processing bias of emotional faces (happy, sad, and neutral). On each trial, participants are told a target emotion to which to respond (by pressing the space bar). Participants then see a series of faces briefly flash up on the screen and are asked to press the spacebar only in response to the target emotion, and refrain from button presses for the distractor faces. The task consists of 6 blocks of 20 trials each (120 total trials) consisting of six different conditions: (1) happy targets with neutral distractors, (2) happy targets with sad distractors, (3) neutral targets with happy distractors, (4) neutral targets with sad distractors, (5) sad targets with happy distractors, and (6) sad targets with neutral distractors.

Mean reaction times (RT) of correct responses (hit rate) were calculated for each condition. The variable of interest was an affective bias score, calculated by subtracting the mean RT for sad target/happy distractor from mean RT for happy target/sad distractor. By computing

this affective bias score, individual differences in RT are taken into account and hence do not unduly affect possible group differences/correlations with CTQ.

2.4.2.4. Emotional Memory Recognition Task

This task measures emotional memory. Participants are shown 30 images during the encoding stage (10 each for positive, neutral, and negative). Participants are then asked how positive/negative the image made them feel and the intensity of this feeling. One hour later (after completing a series of other EMOTICOM tasks in the meantime), participants complete the recall portion of the task, in which they are shown all 30 images again next to the exact mirror image of the original presentation, and asked to identify the version previously seen during the encoding stage. See **Appendix, Figure 1**, for an example of an image and its mirror image counterpart used in the task.

The mean retrieval accuracy score (% of images correctly identified) was calculated for each valence (positive, negative, and neutral), in addition to an overall retrieval accuracy score (all three conditions combined).

2.4.2.5. Monetary Incentive Reward (MIR) Task

This task is a measure of reward/punishment sensitivity. Participants are shown two circles with either green or red coloured lines, indicating whether the following trial is a win or loss trial (association of colour and condition are randomised throughout blocks and participants). The spacing of the two lines in the circle further indicate the size of the loss/gain (whether participants will lose or gain smaller/larger amounts of money). On all trials, regardless of condition, participants are asked to respond as quickly as possible (by pressing the spacebar) when a black square appears in the centre of the screen. The quicker

the response, the more money is won/ less money is lost (this remains constant for the entire task). There are five different conditions: (1) high win, (2) low win, (3) neutral (no win/loss), (4) high loss, and (5) low loss. Participants initially complete a 30-trial neutral practice block (disregarded in final analyses), followed by 2 blocks of 50 trials each of the above 5 conditions (in randomised order).

To account for individual differences in reaction time (RT), a standardised RT was calculated for both win and loss conditions by subtracting RT from the neutral condition. Hence two outcome variables were used in subsequent analyses: (1) mean RT win minus mean RT neutral and (2) mean RT loss minus mean RT neutral.

2.4.2.6. Progressive Ratio Task (PRT)

The PRT assesses motivation/effort to attain reward. There are three blocks with progressively decreasing reward value (£1, 20p, and 4p). In each trial participants are presented with four red squares and asked to select the odd one out (one square is visibly different in size from the other three) and then press continue to complete the next trial. Initially, it takes 4 completed trials to receive reward, this subsequently doubles after each received reward (to 8, 16, 32, 64 etc.). Participants complete a total of 436 trials which take approximately 20 minutes to fully complete. If fully completed, participants receive 4 rewards of £1, 5 of 20p, and 6 of 4p. At any time participants can quit the task by pressing the quit button in the centre of the screen, however, they are told prior to the task that they will then be required to sit and wait for the remainder of the task (and cannot use their phones). If a participant quits, a screen appears with a timer showing the countdown of how long they must wait (until 20 minutes of task time has been reached).

The total number of trials completed signifies the *breakpoint* (i.e. after how many trials the participant decided to quit the task). Since only 6 participants (11.54%) quit the task prior to completion, breakpoints could not be calculated for the vast majority (88.46%) of participants who completed all trials of the PRT.

2.4.2.7. Adapted Cambridge Gambling Task (CGT)

The adapted CGT task is designed to assess risk-taking and decision making in both a loss and reward condition (importantly, unlike reinforcement learning tasks, the CGT task is void of a learning component). On each trial, participants are shown a roulette wheel (pie) shaded in different proportions of orange and purple. The participants are then asked to place a monetary bet on the outcome they predict (which of the two colours they believe the roulette spinner will land). Participants begin each block (loss/win) with 10 poker chips worth 5 points (5p), 10 worth 10 points (10p) and 10 worth 20 points (20p) - hence participants start with £3.5. At each trial (one spin per trial) participants are asked to place any 2 chips. Hence the minimum possible bet is 10 points (£0.1) and the maximum possible bet is 20 points (£0.4). In the win condition, bets are doubled for correct selection of the outcome and retained for wrong outcomes. In the loss condition, bets are lost if the selection is incorrect, and bets are retained for correct predictions. There are 5 different levels of probability (wheel proportions), ranging from 50% to 90% (unlikely to likely). Participants complete two practice trails (with different pie colours) prior to commencing the main task. Each block (win/loss) consists of 14 trials consisting of 2*50%, 4*60, 4 * 70%, 2*80%, and 2x90% probabilities.

The average bet placed for each of the five-wheel probabilities was calculated for both win and loss blocks. This was then used to calculate the risk adjustment (RA) score (for both win and loss blocks separately) using the following formula previously implemented by the creators of the EMOTICOM task (Bland et al., 2016) and subsequent studies (Dam et al., 2019; Savulich et al., 2021):

Risk adjustment (RA) = (((2 90% bet) + (1* 80% bet) + (0* 70% bet)) – ((1* 60% bet) + (2*50% bet))) / Average bet*

RA is a measure of adjustment of the bet placed depending on probability of reward/loss. A higher score (for both win and loss conditions) signifies better choices (more sensible and proportionate to wheel probability).

2.4.2.8. Moral Emotions Task

The Moral Emotion Task assesses the effect of intent (deliberate or accidental harm) upon moral judgments. Participants are shown 28 cartoons (in randomised order) depicting moral scenarios and are asked to try and identify with the person in the cartoon (victim or agent) and imagine how they would feel in that situation. The task has four conditions (victim/agent with intent/no intent) and participants are shown 7 different cartoons for each of these 4 conditions. Half of the cartoons depict deliberate harm while the other half depict accidental harm. After reading the cartoon, participants are shown which character (the agent or the victim) they are asked to identify with. They are then asked to rate their emotion in this hypothetical scenario on a 7-point Likert scale ranging from ‘not at all’ to ‘extremely’ for the emotions of guilt, shame and annoyance, as well as another 7-point Likert scale ranging from ‘bad’ to ‘good’. Following recommendations from the recent Danish validation study of EMOTICOM, ratings of ‘annoyance’ were disregarded as it

appeared this question was ambiguously and differentially interpreted across participants (Dam et al., 2019). Please see **Appendix, Figure 2**, for an example of a cartoon (victim intentional).

Outcome variables included in statistical analyses were mean ratings for agent and victim for shame, guilt, and bad/good. To reduce number of multiple comparisons, no additional analyses were conducted separating intent.

2.4.2.9. Social Information Preference Test

This task is a measure of Theory of Mind (ToM) which assesses participants choice of information source to decipher socially ambiguous situations. Participants are shown 18 different cartoons with 9 pieces of information hidden from view (3 facts, 3 faces, and 3 thoughts). Participants are then able to reveal 4 out of these 9 pieces of information to help decipher the scene (they are subsequently asked to choose one of 4 descriptions of what they believe is occurring in the cartoon).

The proportion (%) of selected faces, thoughts, and facts was calculated for each participant.

2.4.2.10. Prisoners' Dilemma (PD)

This task is a social economic exchange game involving interaction with different computerised opponents to assess social interaction. At the beginning of each trial, participants press the space bar as quickly as possible to fill a jar with coins. They then are shown how their opponent did at the same task and the sum of the two amounts is presented as the total money pot for that trial. Unbeknownst to the participant, this is pre-determined by the task design such that in a third of trials the participant contributes more

to the money pot, a third they contribute equally, and in a third the opponent contributes more. Participants are then asked to either split or steal the total sum, after which they see their opponent's choice (steal/split) and the outcome (amount of money, if any, received). Participants are informed of the rules prior to the task: if both players choose to split the sum, they both receive half; if one player chooses to split and the other chooses to steal, the player who selected steal receives the full sum; if both players decide to steal, neither receives any money. There are 27 total trials, through which participants play 3 different kind of opponents (9 trials) each: aggressive (begins with a steal, then mirrors participants choice), tit for two tats (begins with split, then changes choice after 2 consecutive steals by the participant), and cooperative (always splits).

The outcome variables selected for analyses was the proportion (%) of trials the participant chose to steal (across opponent type).

2.4.2.11. Ultimatum Game (UG)

This task is a social economic exchange game involving interaction with different computerised opponents to assess sensitivity to fairness. At the start of each trial, participants are asked to select 3 of 9 yellow balls which may be uncovered to be red or black – each black ball is worth £3, red balls are worth nothing. The computerised opponent then completes the same task, and the wins are combined as a shared pot. Similar to the PD, the task is pre-determined such that for a third of trials the participant contributes more money, for a third the opponent contributes more, and for a third they contribute equally. In 70% of trials the opponent makes the offer, for the other 30% the participant decides the offer. There are 7 levels of offers, ranging from fair (50/50 split) to increasingly unfair (10/90 split). The player not making the offer then gets to decide whether to accept or reject the

offer. Participants are informed at the beginning of the task that if they accept the offer they get to keep whatever amount was offered to them (and the opponent keeps their share), while if they reject the offer, neither player receives any money. Participants complete 51 trials, including 36 trials in which the opponent makes the offer, and 15 trials in which the participant makes the offer.

The proportion (%) of offers accepted by the participant was calculated as the variable of interest.

Task	**Variables of Interest Included in Analyses**
Emotional Recognition Task (ERT - face)	- Mean accuracy (%) for happy, sad, and fearful faces - Affective bias score (accuracy of happy minus sad)
Emotional Intensity Morphing Task	- Mean level of detection (range 1-15) for happy, sad, angry, and fearful faces. Variables of interest were confined to increasing trials, while decreasing trials were used as a control for impulsive responding. - Affective bias score (mean level of detection for happy minus sad)
Face Affective Go No-Go (FAGN) Task	- Affective bias score (RT happy target with sad distractor minus RT sad target with happy distractor)
Emotional Memory Recognition Task	- Retrieval accuracy (%) for negative, positive, and neutral pictures - Overall retrieval accuracy (%)
Monetary Incentive Reward (MIR) Task	- Mean RT win minus neutral - Mean RT loss minus neutral
Progressive Ratio Task (PRT)	- Breakpoint (number of trials completed before quitting task)
Adapted Cambridge Gambling Task (CGT)	- Risk adjustment (RA) win and loss
Moral Emotions Task	- Ratings for agent and victim perspectives for shame, guilt and bad/good
Social Information Preference Test	- Percentage of selected faces, thoughts, and facts
Prisoners' Dilemma	- Proportion (%) of trials participant chose to steal across opponent type

Ultimatum Game	- Proportion (%) of offers accepted
CTQ correlations	- Correlations were run for all above variables of interest with total CTQ and subtype (EA, PA, SA, EN, PN). These correlations were run across participants and also within MDD and HC groups only.

Table 2. Outcome measures for each EMOTICOM task
Abbreviations: RT, reaction time; CTQ, Childhood Trauma Questionnaire (Bernstein et al., 2003); EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect.

2.5. Power analysis

An a priori power analysis was conducted using the G*Power program (Faul et al., 2007) to determine required sample size. A sample size of 36 participants per group was suggested given power at 80% for an approximately medium between group effect size of Cohen's d = 0.6 (Cohen, 1988). Due to various constraints, the actual sample size was only 27 HC and 25 MDD which resulted in a reduced observed power of 68.7% (assuming a 0.6 effect size).

2.6. Statistical Analyses

All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) software, version 25. Separate univariate ANOVAs were performed for each task, comparing HC to MDD. If a significant group difference was detected (MDD versus HC), correlations between depression severity (using total BDI-II score and total HRSD score) and outcome measures were examined within the MDD group. These correlations were not examined in the HC group due to lack of HRSD scores and very low BDI scores which would likely skew results. For all tasks, correlations were examined between CTQ (CTQ total and the 5 subscales) and task outcome variables, both across participants and split by diagnostic group. Significance threshold was set at $p<0.05$. Additionally, results of $p<0.055$ were interpreted as trends in the results and discussion. No correction for multiple comparisons

was applied due to lack in statistical power, hence findings should be interpreted cautiously and seen primarily as exploratory in nature, aiming to inform future studies of affective cognition in MDD/ELS.

3. Results

3.1. Demographics and clinical characteristics

Demographics and clinical characteristics can be found in **Table 3**. Groups did not significantly differ on demographic characteristics (including age, gender, ethnicity, education, native language, or smoking status). ELS was also well matched across groups, both for total CTQ and subscales of CTQ, with the exception of emotional abuse (EA) which was significantly higher in MDD than HC (p=0.047). However, applying a previously validated and widely applied cut-off score for significant levels of emotional abuse (score ≥10 for EA; Bernstein & Fink, 1998; Bevilacqua et al., 2012; Walker et al., 1999), revealed no significant difference in proportion of participants meeting for significant EA in HC (N=11) versus MDD (N=14), X^2 (1, N = 52) = 1.211, p = 0.271. Furthermore, the proportion of participants meeting cut-off scores[3] for significant levels of abuse/neglect on at least 2 different subtypes did not differ between MDD (N=13/25) and HC (N=14/27), X^2 (1, N = 52) = 0.000, p = 0.991. Finally, it should be noted that all correlations between MDD symptom severity measures (BDI-II and HRSD) and ELS (CTQ total and all 5 subscales) were not significant (all p>0.05).

[3] Cut-off scores (Bernstein & Fink, 1998; Bevilacqua et al., 2012; Walker et al., 1999) were: 8 or higher for physical abuse, 8 for sexual abuse, 10 for emotional abuse, 8 for physical neglect, and 15 for emotional neglect.

The most frequent comorbidity present in the MDD sample was generalised anxiety disorder (GAD, N = 10); other comorbidities present are as followed: dysthymia (N=3 current), panic disorder (N=2 current, N=2 past), agoraphobia (N=2 current, N=1 past), social phobia (N=8 current, N=1 past), obsessive compulsive disorder (OCD; N=3 current, N=1 past), posttraumatic stress disorder (PTSD; N=2 past), alcohol use disorder (N=2 past), substance use (N=3 past), anorexia (N=2 past), bulimia (N=1 past), and binge eating disorder (N=2 past).

	HC	**MDD**	**Statistic**	**Significance**	**Effect Size**
Demographic					
Age (years)	31.592 (12.938)	27.801 (11.655)	*F*= 1.226	*p*= 0.273	ηp^2=0.024
Female (%)	59.32	80	χ^2= 2.621	*p*= 0.105	*V* = 0.225
Caucasian (%)	74.14	76	χ^2= 0.026	*p*= 0.873	*V* = 0.022
Education (years)	17.042 (4.192)	15.424 (2.125)	*F*= 2.915	*p*= 0.094	ηp^2= 0.056
English native language (%)	77	80	χ^2= 0.071	*p*= 0.789	*V* = 0.037
Current smokers	2	1	χ^2= 0.277	*p*= 0.599	*V* = 0.073
Clinical					
CTQ total	40.892 (14.516)	45.883 (13.630)	*F*= 1.627	*p*= 0.208	ηp^2=0.032
CTQ EA	9.564 (4.862)	12.043 (3.824)	*F*= 4.149	***p*= 0.047***	ηp^2=0.077
CTQ PA	6.963 (3.144)	7.000 (2.843)	*F*= 0.002	*p*= 0.965	ηp^2=0.000
CTQ SA	6.114 (2.665)	6.482 (3.513)	*F*= 0.184	*p*= 0.670	ηp^2=0.004
CTQ EN	10.221 (4.862)	12.440 (4.321)	*F*= 3.004	*p*= 0.089	ηp^2=0.057
CTQ PN	8.042 (3.684)	7.923 (2.999)	*F*= 0.016	*p*= 0.901	ηp^2=0.000
BDI-II	7.524 (7.066)	27.611 (5.31)	*F*= 110.558	***p*< 0.001***	ηp^2=0.734
HRSD total	NA	17.750 (4.131)			

Table 3. Demographics and clinical characteristics.
Abbreviations: BDI-II, Beck Depression Inventory-II (Beck et al., 1996); HRSD, Hamilton Rating Scale for Depression (Hamilton, 1960); MDD, major depressive disorder; HC, healthy control; NA, not applicable; CTQ, childhood trauma questionnaire CTQ; (Bernstein et al., 2003); EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect.
Mean values are displayed with standard deviations in parentheses where applicable.
*p<0.05 comparing MDD and HC

3.2. EMOTICOM Tasks

3.2.1 Emotional Recognition Task (ERT)

HC (M=14.231, SD=21.290) had significantly higher affective bias scores (accuracy of happy minus sad) than MDD (M=3.000, SD=18.431), $F(1,49)= 4.043$, $p= 0.050$, $\eta p^2 = 0.0076$. This appeared to be driven mainly by a lower accuracy (%) in identifying happy faces in MDD (M=83.000, SD =12.247) than in HC (M=89.038, SD=9.267), $F(1,49)= 3.960$, $p= 0.052$, $\eta p^2 = 0.075$. The opposite trend was observed for sad faces, though this did not approach statistical significance. Correlations of outcome measures with depression severity scores (in MDD only) revealed a significant association between total HRSD score and accuracy in identifying fearful faces, $r(23) = 0.506$, $p= 0.012$. This indicates that higher levels of HRSD scores (indicating greater severity of current depressive symptoms) were associated with greater accuracy for identifying fearful faces. No other significant correlations emerged with any other outcome variables or with BDI-II.

Correlations with CTQ scores across all participants revealed a significant positive correlation between EA and accuracy for fearful faces, $r(50) = 0.277$, $p=0.047$. Furthermore, within the HC group (but not within MDD), affective bias (happy minus sad) was significantly correlated with PA, $r(50) = -0.458$, $p=0.019$.

3.2.2 Emotional Intensity Morphing Task

Participants with MDD detected fearful faces significantly earlier than HC, $F(1,49)= 7.484$, $p= 0.009$, $\eta p^2 = 0.133$. There was also a trend approaching significance for MDD detecting sad faces more quickly than HC, $F(1,49)= 3.976$, $p= 0.052$, $\eta p^2 = 0.075$. It is important to note that the corresponding decreasing trials (decreasing fear and sadness) were not significantly different between groups (see **Table 4**) and as such the aforementioned findings are not thought to represent a mere heightened impulsiveness in MDD. There were no group differences in affective bias score ($p>0.05$). Correlations of outcome measures with depression severity scores (in MDD only) revealed a significant negative association between total BDI-II score and number of intensity levels needed to detect fearful faces, $r(23) = -0.422$, $p= 0.045$. This indicates that the greater the depression severity of MDD participants (as measured using BDI-II), the earlier they detected fearful faces.

	HC	MDD	Statistic	Significance
Increasing Fear	10.327 (1.721)	8.688 (2.525)	*F*= 7.484	***p*= 0.009***
Decreasing Fear	11.093 (2.484)	11.063 (2.212)	*F*=0.002	*p*=0.963
Increasing Sad	10.327 (2.304)	8.979 (2.524)	*F*=3.976	***p*=0.050***
Decreasing Sad	11.364 (1.382)	10.736 (1.785)	*F*=1.997	*p*=0.164

Table 4. Emotional Intensity morphing task mean values. Number of levels before emotion was detected (increasing) or was no longer detected (decreasing) are displayed with standard deviations in parentheses. *$p<0.05$ comparing MDD and HC

Correlations with CTQ revealed a significant positive correlation between identification of angry faces and SA across participants, $r(50)=0.299$, $p =0.033$, indicating that higher levels of SA are associated with needing more 'levels' (number of pictures of increasing emotional intensity) to detect anger. Similarly, within MDD (but not HC), a positive correlation between SA and fearful faces emerged, $r(50)=0.398$, $p =0.049$.

3.2.3 Face Affective Go No-Go (FAGN) Task

There were no significant group differences in reaction time (RT) for affective bias scores (RT happy target with sad distractor minus RT sad target with happy distractor), p=0.415.

There were, however, significant positive correlations between affective bias and CTQ, including, in participants overall, with physical abuse, $r(50)=0.303$, $p=0.030$, sexual abuse, $r(50)=0.301$, $p=0.032$, and CTQ total, $r(50)=0.284$, $p=0.044$. In addition, in MDD only, affective bias score was significantly correlated with emotional abuse, $r(50)=0.447$, $p=0.025$, and sexual abuse $r(50)=0.454$, $p=0.023$. Higher affective bias scores indicate slower reaction times for targets of happy faces (with sad face distractors) than for targets of negative faces (with happy distractors).

3.2.4. Emotional Memory Recognition Task

MDD participants had significantly lower overall (negative, positive, and neutral images combined) retrieval accuracy than HC, $F(1,49)= 6.878$, $p= 0.012$, $\eta p^2 = 0.123$. While there were no significant group differences in neutral or positive image memory, MDD participants had significantly lower retrieval accuracy for negative images than HC, $F(1,49)= 8.171$, $p= 0.006$, $\eta p^2 = 0.143$ (see Table 5 for means). No significant correlations emerged between total BDI-II or total HRSD scores and any outcome variables.

	HC	MDD	Significance
Retrieval accuracy negative (%)	73.846 (17.453)	59.600 (18.130)	***p*= 0.006***
Retrieval accuracy overall (%)	71.538 (12.229)	62.933 (11.152)	***p*= 0.012***

Table 5. Emotional memory mean retrieval accuracy. Standard deviations are given in parentheses. *p<0.05 comparing MDD and HC

Correlations with CTQ revealed a significant negative correlation between EA and accuracy for negative images in participants overall, $r(50)= -0.282$, $p = 0.045$. This correlation indicates that the higher the level of EA in participants (both HC and MDD), the lower their retrieval accuracy for negative images was.

3.2.5. Monetary Incentive Reward (MIR) Task

There were neither any significant group differences, nor any significant correlations with CTQ, for mean RT win minus neutral or mean RT loss minus neutral (all $p>0.05$).

3.2.6. Progressive Ratio Task (PRT)

Only 1 HC and 5 MDD had breakpoints, all other participants (88.462%) fully completed the task. Due to this very small sample of participants with breakpoints, a group analysis comparing breakpoint level between MDD and HC was not conducted. However, a Pearson Chi square was conducted to compare the presence of a breakpoint between groups and revealed a significantly higher frequency of breakpoint in MDD (N= 5, 20%) than HC (N=1, 3.7%), X^2 (2, N = 52) = 9.630, p = 0.008. Given the small subsample of participants with breakpoints, no correlations with symptom measures were conducted.

While this test did not allow for exploration of CTQ scores, there was an even distribution of participants meeting for cut-offs for significant ELS (defined as meeting criteria for significant levels of abuse for at least two subtypes of the CTQ according to previously validated and widely applied cut-off scores; Bernstein & Fink, 1998; Walker et al., 1999) and those who did not - of the 5 participants with a breakpoint, 1 HC and 2 MDD did not meet for ELS using the aforementioned criteria, while 3 MDD met for significant ELS.

3.2.7. Adapted Cambridge Gambling Task (CGT)

There were no significant group differences between MDD and HC for risk adjustment (RA), neither for win nor loss.

Correlations between RA and CTQ, however, revealed several significant findings, both across participants and within just HC and MDD (see **Table 6**). All significant findings constituted negative correlations between both RA win and loss and CTQ measures, indicating reduced optimisation in both conditions was associated with higher scores on various CTQ subscales/total score.

	Overall	HC only	MDD only
Risk Adjustment Win			
- EA	*r*= -0.317 ***p*= 0.024***		
- PA	*r*= -0.311 ***p*= 0.026***		*r*= -0.480 ***p*= 0.015***
- SA	*r*= -0.289 ***p*= 0.04***		
- Total CTQ	*r*= -0.360 ***p*= 0.010***		
Risk Adjustment Loss			
- PA	*r*= -0.322 ***p*= 0.021***	*r*= -0.509 ***p*= 0.008***	
- PN	*r*= -0.389 ***p*= 0.005***	*r*= -0.488 ***p*= 0.012***	
- EN		*r*= -0.414 ***p*= 0.035***	
- Total CTQ		*r*= -0.400 ***p*= 0.043***	

Table 6. Correlations between win and loss condition risk adjustment and CTQ scores. Both test statistic (pearson's *r*) and significance (*p* value) are given. Only significant correlations ($p<0.05$) are listed in the table.
Abbreviations: EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect; MDD, major depressive disorder; HC, healthy control; CTQ, childhood trauma questionnaire CTQ; (Bernstein et al., 2003).

3.2.8. Moral Emotions Task

There were no significant group differences for any of the 6 variables (agent/victim ratings for guilt, shame, and bad/good), all $p > 0.05$.

There were, however, several significant correlations with CTQ, most prevalent within MDD participants (see **Table 7** for correlations within MDD). In addition, in participants overall, ratings of shame for victims, $r(50) = 0.422$, $p = 0.004$, as well as guilt for victims, $r(50) = 0.429$, $p = 0.003$, correlated significantly with PN. In HC only, the only significant correlation was similarly with PN and shame ratings for the victim, $r(50) = 0.443$, $p = 0.027$. These correlations indicate that higher levels of physical neglect were associated with higher ratings of shame for the victim in HC and participants overall, as well as higher ratings of guilt for the victim (across participants). The findings within MDD and CTQ correlations were particularly widespread, indicating a decreased attribution of guilt to the agent the higher the history of EN ($p = 0.003$), increased ratings of guilt and shame for the victim related to increasing levels of PN ($p = 0.005$,) and EA ($p = 0.039$), respectively, and several significant correlations indicating higher ratings of bad (versus good) toward the victim associated with higher levels of EA ($p = 0.026$), PA ($p = 0.043$), EN ($p = 0.027$), and total CTQ score ($p = 0.013$). Please see **Table 7** for complete test statistics within MDD.

	Statistic	**Significance**
Guilt agent & EN	*r*= -0.484	***p*= 0.030**
Guilt victim & PN	*r*= 0.601	***p*= 0.005***
Shame victim & EA	*r*= 0.464	***p*= 0.039***
Bad/Good victim &		
- EA	*r*= -0.497	***p*= 0.026***

- PA	*r*= -0.456	***p*= 0.043***
- EN	*r*= -0.493	***p*= 0.027***
- Total CTQ	*r*= -0.543	***p*= 0.013***

Table 7. Correlations between ratings of moral emotions for victim and agent with CTQ scores within MDD participants. Only significant correlations (p<0.05) are shown.

3.2.9. Social Information Preference Test

There were no significant group differences between MDD and HC in proportion (%) of selected faces (*p*=0.388), thoughts (*p*=0.809), or facts (*p*=0.255). Similarly, there were no significant correlations between these variables and CTQ scores, neither across participants nor within solely MDD/HC (all *p*>0.05).

3.2.10. Prisoners' Dilemma

There were no significant group differences between MDD and HC in proportion (%) of steals (for aggressive, tit for two tats, cooperative, or overall steal), all *p*>0.05. There were similarly no significant correlations between proportion of steals and CTQ scores.

3.2.11. Ultimatum Game

There were no significant group differences between MDD and HC in proportion (%) of acceptance of offers (*p*=0.189). Similarly, no significant correlations emerged between acceptance rate and CTQ scores, all *p*>0.05.

Tasks	**Main construct measured**	**MDD** (versus HC)	**ELS** (correlations)
Emotional Recognition Task (ERT)	Emotion recognition/ categorisation	✓	✓ EA, PA
Emotional Intensity Morphing Task	Emotion recognition	✓	✓ SA
Face Affective Go No-Go (FAGN) Task	Attentional bias	✗	✓ PA, SA, EA, CTQ total
Emotional Memory Recognition Task	Emotional memory	✓	✓ EA
Monetary Incentive Reward (MIR) Task	Reward/ punishment sensitivity	✗	✗
Progressive Ratio Task (PRT)	Incentive motivation / effort	✓	✗
Adapted Cambridge Gambling Task (CGT)	Value-based choice	✗	✓ All subscales & CTQ total
Moral Emotions Task	Moral emotion	✗	✓ EA, PA, EN, PN, CTQ Total
Social Information Preference Test	Theory of mind	✗	✗
Prisoners' Dilemma (PD)	Social economic exchange game	✗	✗
Ultimatum Game (UG)	Social economic exchange game	✗	✗

Table 8. General overview of main findings. Green checkmarks indicate significant findings (for diagnosis [MDD versus HC], and/or ELS [correlations with CTQ scores]). Red crosses indicate null results. For ELS, all significant correlations are indicated (across participants and/or within only HC/MDD groups).
Abbreviations: EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect; MDD, major depressive disorder; HC, healthy control; CTQ, childhood trauma questionnaire CTQ; (Bernstein et al., 2003).

4. Discussion

Analyses of EMOTICOM tasks revealed that certain affective cognition constructs appeared to be sensitive to current depressive symptoms, including emotion recognition and incentive motivation/effort (as seen by significant differences between MDD and HC groups) while other constructs revealed no group differences but instead significant associations with various types of ELS (including attentional bias, value-based choice, and moral emotions). Unexpectedly, in our sample, contradictory results to the literature were found for emotional memory (with decreased memory for negative images in MDD versus HC) and null results emerged for the Monetary Incentive Reward (MIR) task and all but one (moral emotions) social cognition tasks, including two social economic exchange games (PD and UG) and theory of mind (Social Information Preference Test). See **Table 8** above for a summary of overall findings.

Four tasks revealed an effect of diagnosis on outcome variables, including two emotion recognition tasks, one task measuring incentive motivation/effort, and one measuring emotional memory. Specifically, the Emotional Recognition Task (ERT) revealed significantly higher affective bias (accuracy of identifying happy minus sad faces) in HC than MDD (p=0.050), which appeared to be driven mainly by a lower accuracy for happy faces in MDD than HC (p=0.052). Another task mainly driven by current symptoms was the Emotional Intensity Morphing task, which indicated an earlier detection (fewer 'levels' of emotional faces needed to detect presence of emotion) of fearful faces (p =0.009) and approaching significance for sad faces in MDD (p =0.052) compared to HC. Correlations with symptom severity measures further revealed that the observed effect for fearful faces in MDD was associated with BDI-II scores (the higher total BDI-II score, the earlier fearful face detection occurred), however this was not replicated in HRSD scores. The only CTQ correlation that

emerged was with SA and indicated the opposite trend observed for MDD, suggesting that observed effects appeared to be driven by MDD, independent of ELS. The fact that current MDD appears to influence affective bias in these two emotion recognition tasks, independently of ELS, is further supported by findings that antidepressant interventions can successfully reverse such affective biases in patients with MDD, even prior to any changes in mood or symptoms are observed (Harmer et al., 2009). If affective bias in emotion recognition tasks were more strongly affected by ELS history than MDD, psychopharmacological intervention would perhaps not be expected to reverse the observed changes in emotion processing.

The Progressive Ratio Task resulted in significantly (p=0.008) greater number of breakpoints in MDD (20%) than HC (3.7%). These results are in line with previous findings showing significantly reduced effort/incentive motivation in MDD (Treadway et al., 2012; Yang et al., 2014). Given the even spread of ELS across groups and among participants with breakpoints (N=3 with and N=3 without significant levels of ELS), it may be deduced that reduced effort to receive reward may be more associated with MDD itself than a history of ELS. However, the results from the PRT must be tentatively interpreted since statistical analyses were based on a small subset of participants (N=6, 11.54%) as all other participants fully completed the task. Interestingly, the completion rate in our sample is significantly higher than the proportion of participants who completed the task in the original paper that administered the EMOTICOM test battery to 200 healthy volunteers (in which 57% completed the task; Bland et al., 2016). Though the authors do not report an exact percentage, a Danish validation study of the EMOTICOM test battery mentions a “large proportion of participants who met criteria for ceiling effects” in the PRT, presumably also

indicating a high number of participants who fully completed the task (Dam et al., 2019). Future studies using the PRT should consider this observed ceiling effect and potentially consider altering task instructions or the testing environment to reduce this effect and enable for more variation between participants.

The fourth, and final, task revealing an effect of diagnosis was the Emotional Memory Recognition Task. While no specific a priori predictions had been made regarding overall recall accuracy (positive, negative, and neutral combined), our findings of reduced overall recall accuracy in MDD versus HC is in line with some studies demonstrating general (not affective dependent) memory impairments in depression (Cale, 1996; Nikolin et al., 2021; Rock et al., 2014). In regard to affective bias, the primary variable of interest, findings were divergent from our priori hypothesis. The vast majority of existing literature has consistently identified an affective bias of memory in depression, whereby MDD is characterised by increased recall accuracy for negative stimuli and decreased recall for positive stimuli compared to HC (Gotlib et al., 2004; Hamilton & Gotlib, 2008). Instead, the opposite effect was found, in which MDD participants had significantly lower retrieval accuracy for negative images than HC, but no significant differences emerged between groups for neutral or positive images. Interestingly, a few other studies have also failed to replicate an affective memory bias in depression (Ellwart et al., 2003; Danion et al., 1995) and some have even observed improved memory for positive over negative words in MDD (Calev, 1996). It should be noted that the only significant correlation with CTQ that emerged indicated lower accuracy for negative images in participants with higher levels of EA. Emotional abuse was the only CTQ subscale that was significantly different between diagnostic groups (higher in

MDD) and hence it is possible that either MDD or EA (or a combination of the two) had an effect on the observed results.

Furthermore, though the ERT and Emotional Intensity Morphing Task appeared to be primarily driven by group differences (MDD versus HC), there were a few significant correlations with specific CTQ subscale scores. Within the Morphing Task, SA was found to increase the levels (number of faces with increasing intensity of emotion) needed for detection of both angry and fearful faces. This finding was unexpected, given previous studies that have documented attentional biases specifically toward angry faces (not for happy or sad faces) in children with a history of ELS (Pollak & Tolly-Schell, 2003; Pollak & Kistler, 2002). This finding has been replicated in young adults in a study that reported preferential attention to and increased accuracy in identification at lower levels of emotional intensity for angry faces (and not for happy or sad faces) in those reporting moderate levels of ELS compared to those with no ELS, independent of depressive symptoms (Gibb et al., 2008). Interestingly, the direction of attentional bias (either towards threat or avoidance), may depend on severity and type of abuse. One study reported a significant correlation between severity of physical abuse and avoidance of threat (attentional bias away from angry or threatening faces; Pine et al., 2005). This may be related to emotional blunting, which has been observed in individuals with ELS (Locher et al., 2014). Interestingly, emotional numbing is also a hallmark of post-traumatic stress disorder (PTSD), and one prospective study found that early numbing following experience of sexual or non-sexual assault in women predicted development of PTSD (Feeny et al., 2000). However, it should be noted that this abuse occurred during adulthood and emotional numbing was measured shortly afterward, and as such may not be comparable to

ELS. More studies are needed to investigate whether attentional biases away from fearful and angry faces may represent emotional blunting and may depend on the type and severity of abuse. It is possible that the slower detection of angry and fearful faces in participants with higher levels of SA observed in the Emotional Intensity Morphing Task of the present investigation indicates a threat avoidance potentially specific to a significant history of childhood SA, but this is a purely speculative hypothesis and requires further study.

In the ERT on the other hand, the two correlations that emerged with CTQ indicated increased attentional bias toward fearful and sad (relative to happy) faces, more in line with previous research. Specifically, EA (across participants) was correlated with increased accuracy in recognizing fearful faces, while PA (in HC only) was correlated with lower affective bias scores (indicate greater accuracy for sad versus happy faces). As aforementioned, increased accuracy or preferential responding to threatening stimuli has been reported following ELS (Pollak, 2003; Pollak & Tolly-Schell, 2003; Pollak & Kistler, 2002), however these studies have generally highlighted the specificity toward angry/fearful faces, and an absence of an effect for happy/sad. It is hence unclear why the correlation between decreased affective bias accuracy and PA emerged, and why this was limited to HC – future studies are needed to investigate whether this effect is replicated.

Three tasks, measuring attentional bias (FAGN task), value-based choice (Adapted CGT), and moral emotions (Moral Emotions Task) revealed no group differences (MDD versus HC) but significant correlations with CTQ, suggesting that these constructs of affective cognition may be more influenced by history of ELS rather than current depression. Specifically, the FAGN

task revealed significant positive correlations between affective bias (RT of happy target with sad distract minus RT sad target with happy distract) and PA, SA, and CTQ total (across participants), and EA and SA (in MDD), indicating that higher levels of ELS in these subtypes/total score are associated with quicker response to targets of sad faces relative to happy faces. The adapted CGT also indicated several significant correlations between both risk adjustment (RA) for win and loss and various CTQ scores, including PA, SA, CTQ total (for RA win) and PA, PN, EN and CTQ total (for RA loss). The lack of group effect but significant correlations with total CTQ and various subtypes of ELS indicate that findings have been previously misattributed to MDD when perhaps ELS (which is highly prevalent within MDD patients Williams et al., 2016; Xie et al., 20 18) may instead be primarily driving this effect. This is supported by findings that report impaired optimisation of bets (lower rates of RA) in medicated, unmedicated and remitted MDD (Murphy et al., 2001; Rawal et al., 2013; Roiser & Sahakian, 2013), further indicating that current depressive symptoms are unlikely to be the driving factor behind the observed behavioural changes.

Similarly, while the Moral Emotions Task revealed no group differences between MDD and HC, a host of significant correlations emerged with CTQ scores, in particular relating to heightened levels of shame and guilt perceived for the victim of harm. While a few correlations (with PN) emerged across participants overall and in HC, all further significant correlations were exclusive to MDD, suggesting that potentially ELS within MDD participants has a heightened effect on moral judgment. Higher levels of ELS (including EA, PA, EN, and total CTQ) were associated with increased ratings of shame, guilt and bad (versus good) for the victim of harm in MDD participants (see **Table 7** for details). An extensive literature has reported heightened levels of negative moral emotions, including self-blame, shame, and

guilt in MDD (Zahn et al., 2015). Interestingly, these same emotions have also been found to be heightened in individuals with ELS exposure and may in fact represent a mediating factor between ELS exposure (including PN, SA, PA, and EN) and later life depression (Sekowski et al., 2020; Tanzer et al., 2021) and more generally psychological distress in adulthood (Coffey et al., 1996). Similar findings have been reported in studies of military personnel with ELS, in which ELS predicted subsequent anxiety and depression via self-blame (Dorresteijn et al., 2019). The results from the Moral Emotions Task complement these findings and suggest that heightened levels of self-blame, guilt, and shame for victims of harm (whether intentional or unintentional) appear to be driven especially by ELS exposure, apparent to some degree across all participants including HC (for PN) but possibly especially heightened in those with MDD. Heightened negative moral emotions may hence potentially represent a key mechanism linking history of ELS and development of later life depression, though more studies are necessary to more directly investigate this possible relationship.

Null results (for both group differences and correlations with ELS) emerged for the remaining three social cognition tasks as well as the MIR task. A lack in group difference between MDD and HC in the MIR task is in line with fMRI studies in depression that have effectively used the MIR task to demonstrate diminished activation of both the dorsal and ventral striatum in response to rewards (including the caudate and nucleus accumbens) but have not yielded significant behavioural findings (Pizzagalli et al., 2009; Knutson et al., 2008). While the EMOTICOM MIR task was amended to attempt to increase sensitivity for potential group differences (Bland et al., 2016), it is possible that either these differences are too subtle to be picked up by the MIR task design, or potentially there are indeed no

behavioural differences and other methods, such as the aforementioned fMRI studies, may provide a more fruitful avenue to exploring differences in reward anticipation and sensitivity.

Social economic exchange games (UG and PD) and Theory of Mind also appeared less sensitive to MDD and ELS. Interestingly, while several studies have reported decreased theory of mind in depression, particularly in chronic depression (Mattern et al., 2015; Zobel et al., 2010), this may be closely linked to more general deficits in cold cognition, as effects may be fully accounted for by deficits in logical and working memory (Zobel et al., 2010). Moreover, the exact nature of the task putatively designed to measure Theory of Mind may also determine whether effects are seen in MDD, and suggest that different tasks may in fact measure different constructs. For instance, one study reported deficits in MDD on the frequently used ”Reading the Mind in the Eyes Test” (RMET) compared to HC, but no group differences on a task that required using contextual information to make judgments about other people (“Movie for the Assessment of Social Cognition”, MASC; Wolkenstein et al., 2011). The RMET is frequently used to study Theory of Mind, however, this task is identical in design (apart from using eyes only versus full faces) to the ERT in the EMOTICOM test battery, and as such may perhaps be more accurately described as an emotion recognition and categorisation task. It is interesting to note that no group differences were observed between MDD and HC in the MASC in the aforementioned study (Wolkenstein et al., 2011), which, similar to the Social Information Preference Test used in EMOTICOM, requires integrating more complex social information and reasoning about others’ state of mind.

Social economic exchange games, such as the UG and PD, have been primarily studied in healthy individuals to investigate economic decision-making, and the few studies in depression have yielded inconsistent findings (Harlé et al., 2010; Pulcu et al., 2014; Harlé & Sanfey, 2007). To the authors' knowledge, neither of these tasks have been previously studied in the context of ELS. A recently published study investigating performance on a monetary exchange game designed to measure trustworthiness of opponents, showed that ELS (as measured using the CTQ) was significantly correlated with heightened ratings of distrust (Hepp et al., 2021). Given these initial findings of social decision-making in ELS and conflicting findings in a small number of studies of MDD, future studies may benefit from investigating more nuanced outcome measures, to potentially enable revealing more complex patterns of aberrant social decision-making possibly present in ELS and/or MDD. Additional variables potentially of interest include breaking down proportion of steals by opponent type (aggressive, tit-for-two-tats, cooperative) in the PD and calculating the proportion of accepted offers for each level of offer (ranging from 50% - 90%) and each level of contribution (whether the opponent or participant contribute more to the pot) in the UG. The present investigation limited PD and UG analyses to proportion of steals across opponent type and proportion of offers accepted overall, respectively, in line with previous studies using these EMOTICOM tasks (Bland et al., 2016; Dam et al., 2019), to avoid excessive multiple comparisons as the nature of multiple tasks in the EMOTICOM test battery already resulted in a large number of outcome variables analysed. However, reducing the number of variables considerably also decreases sensitivity of these two tasks which allow for more detailed & graded analyses and subsequent interpretation of social exchange behaviour.

Overall, our findings from the EMOTICOM test battery suggest that different constructs of affective emotion are predominantly affected by MDD or ELS and others appear insusceptible to either (see **Table 8** for a summary). The spread of results across initially proposed affective cognition domains (emotion processing, motivation and reward, and social cognition, Bland et al., 2016) is in line with previous studies that have rejected the domain approach of the EMOTICOM tasks and instead suggest individual tasks measure specific, heterogenous, constructs (Bland et al., 2016; Dam et al., 2019). While domains may be useful in general classification of tasks, individual tasks clearly map onto distinct constructs of affective cognition, which, as seen in the present investigation, appear to be differentially affected by MDD and ELS. This further highlights the need for comprehensive, standardised measures, such as the EMOTICOM test battery, for the study of affective cognition, since evidently individual tasks measure distinct constructs and as such cannot be used as a singular proxy for affective cognition.

4.1. Strengths, limitations, and future directions

The present investigation had several strengths, including the thorough assessment of affective cognition using a recently developed and validated test battery, assessment of both depressive diagnosis and symptoms and ELS, and ensuring that groups were well matched for history of ELS to allow for a controlled analysis of the effect of MDD on affective cognition without ELS as a potential confounding factor. It should, however, be noted that while groups were generally well matched for ELS exposure, the MDD group reported significantly higher rates of EA than HC. Importantly, the proportion of participants meeting for significant levels of EA did not differ between groups, however, correlations with EA should nevertheless be cautiously interpreted. The use of CTQ scores allowed for a

nuanced analysis of ELS subtypes and correlations with outcome variables which is another strength of the current investigation. This study design allowed us to tease apart potential effects of MDD and different types of ELS on affective cognition. Furthermore, MDD participants were all unmedicated, ruling out any potential confounding factors of antidepressant medication, which has been found to reverse negative emotional cognitive biases observed in MDD (Harmer & Cowen, 2013).

Several limitations should be highlighted. The current investigation was significantly underpowered, and our current sample size yielded only 68.7% observed power (assuming a $d = 0.6$ effect size) and as a result, due to already diminished power to detect significant results, we did not correct for multiple comparisons. Similar to previous studies using EMOTICOM tasks, we attempted to mitigate this by a priori selection of a reduced number of outcome variables (Bland et al., 2016; Dam et al., 2019). Nevertheless, the findings presented here should be interpreted cautiously, and additional studies are needed to confirm and replicate reported results. It should also be noted, that while the EMOTICOM tasks are purposefully designed in a way to be largely independent of cold cognition (and in tasks in which reaction time was used as an outcome variable it was standardised to the neutral RT to account for individual differences), future studies would benefit from including a measure of cold cognition, such as the Cambridge Neuropsychological Test Automated Battery (CANTAB) which has been extensively studied in MDD and revealed moderate deficits in various cold cognition domains (including executive functioning, memory, and attention) in MDD relative to HC (for a systematic review and meta-analysis, see Rock et al., 2014). In addition to outcome variables based on RT (which were controlled for, as discussed), additional variables potentially susceptible to changes in cold cognition may be

those in which accuracy was measured – in the ERT and Emotional Memory Recognition Task. While the latter did find an overall deficit of accuracy in MDD vs HC, and hence possibly could be partially accounted for by cold cognition deficits in memory in MDD, the ERT did not reveal any overall differences in accuracy in MDD versus HC (instead this was specific to sad relative to happy faces). While it is hence unlikely that differences in cold cognition was influencing results of the EMOTICOM test battery reported here (bar potentially the Emotional Memory Recognition Task), future studies may consider adjunctly collecting cold cognition data to systematically control for any potential confounders of cold cognition in statistical analyses of the EMOTICOM tasks.

A general limitation of cross-sectional studies of ELS is the retroactive assessment of childhood exposure to abuse/neglect. While the CTQ is the most commonly used instrument to assess ELS, it is by design subject to possible recall bias. In addition, while the CTQ collects information on severity of 5 different forms of abuse/neglect and hence allowed us to differentiate our analyses by ELS subtype, it does not provide any information on timing or duration of abuse. It is becoming increasingly apparent that there seem to be sensitive periods during which ELS has particular effects, which may differentially affect brain development (Andersen et al., 2008) and risk for later life depression (Khan et al., 2015). Future studies should hence consider applying an ELS measure that not only measures type but also timing of abuse, such as the Maltreatment and Abuse Chronology of Exposure (MACE) Scale (Teicher & Parigger, 2015), though additional considerations in study design will need to be made, such as larger sample sizes to allow for meaningful statistical analyses of timing of abuse, and adequate experimental time to complete the lengthy and detailed MACE questionnaire.

5. Conclusion

The present investigation assessed multiple constructs of affective cognition in both HC and MDD with matching levels of ELS history to determine effects of MDD and ELS. Collectively, the results emphasise that affective cognition as measured by the EMOTICOM test battery, is comprised of multiple, distinct constructs, that appear to be differentially affected by MDD and ELS. Specifically, an MDD diagnosis was primarily related to changes in emotion recognition/categorisation, emotional memory, and motivation/effort, independent of ELS. On the other hand, tasks measuring attentional bias, value-based choice, and moral emotion were affected solely by ELS (and not depression). This study highlights the need for measurement of ELS in studies of cognition in depression, which is frequently omitted despite significantly elevated levels of ELS in this population, and may lead to misinterpretation of findings given the effect of ELS on certain affective cognition constructs. The present investigation was underpowered and limited by multiple comparisons, and future studies assessing affective cognition are needed to ascertain whether reported results replicate. Future studies may also want to consider including additional information about ELS timing (to assess for sensitive periods), and possibly investigate relationships between affective cognition and functional (e.g. well-being and social functioning) and clinical outcomes (e.g. number of MDEs, remission) in prospective studies.

6. Appendix

Figure 1. Example of Emotional Memory Recognition Task.

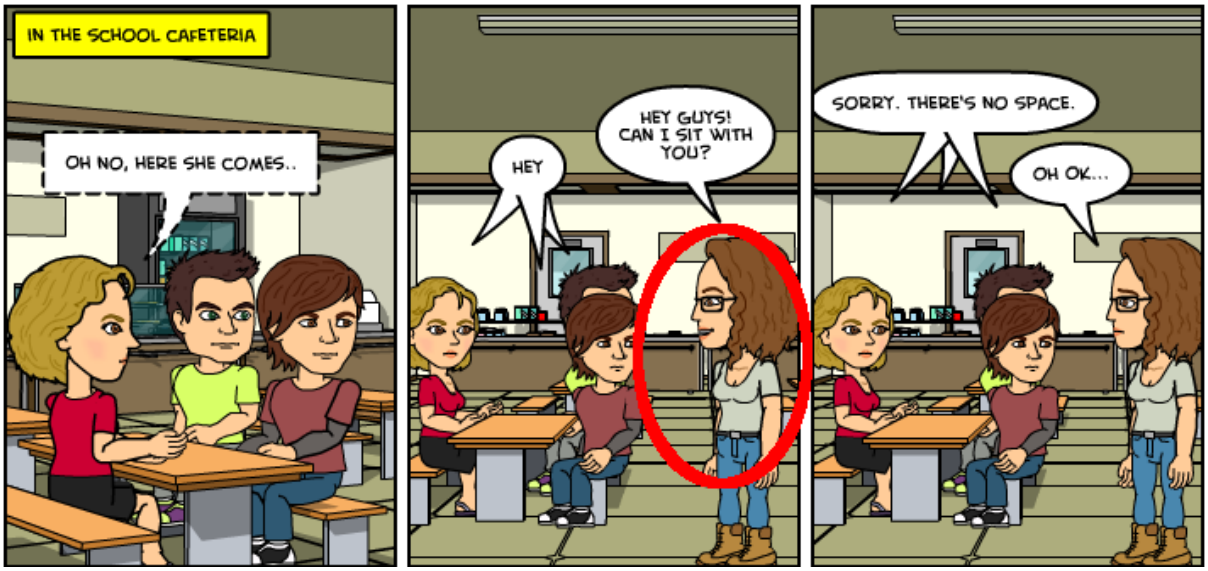


Figure 2. Example of Moral Emotions Task, depicting the condition victim intentional.

REFERENCES

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of neuropsychiatry and clinical neurosciences*, *20*(3), 292-301.

Barsky, A. J., Ahern, D. K., Bailey, E. D., Saintfort, R. L., Liu, E. B., & Peekna, H. M. American Psychiatric Association.(1994). Diagnostic and statistical manual of mental disorders . Washington, DC: Author. Andrews (1989). Psychotherapy of depression: A self-confirmation model. Psychological Review, 96, 576-607. Bargh, JA, & Chartrand, TL (1999). The unbearable automaticity of being. American. *Psychological Review*, *96*, 576-607.

Bazin, N., Perruchet, P., & Féline, A. (1996). Mood congruence effect in explicit and implicit memory tasks: a comparison between depressed patients, schizophrenic patients and controls. *European Psychiatry*, *11*(8), 390-395.

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

Bellani, M., Hatch, J. P., Nicoletti, M. A., Ertola, A. E., Zunta-Soares, G., Swann, A. C., ... & Soares, J. C. (2012). Does anxiety increase impulsivity in patients with bipolar disorder or major depressive disorder?. *Journal of psychiatric research*, *46*(5), 616-621.

Bernstein, D. P., & Fink, L. (1998). CTQ: Childhood trauma questionnaire manual.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child abuse & neglect*, *27*(2), 169-190.

Bevilacqua, L., Carli, V., Sarchiapone, M., George, D. K., Goldman, D., Roy, A., & Enoch, M. A. (2012). Interaction between FKBP5 and childhood trauma and risk of aggressive behavior. *Archives of general psychiatry*, *69*(1), 62-70.

Bland, A. R., Roiser, J. P., Mehta, M. A., Schei, T., Boland, H., Campbell-Meiklejohn, D. K., ... & Viding, E. (2016). EMOTICOM: A Neuropsychological Test Battery to Evaluate Emotion, Motivation, Impulsivity, and Social Cognition. *Frontiers in behavioral neuroscience*, *10*.

Boecker, R., Holz, N. E., Buchmann, A. F., Blomeyer, D., Plichta, M. M., Wolf, I., ... & Laucht, M. (2014). Impact of early life adversity on reward processing in young adults: EEG-fMRI results from a prospective study over 25 years. *PloS one*, *9*(8), e104185.

Bradley, B. P., Mogg, K., & Williams, R. (1995). Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety. *Behaviour research and therapy*, *33*(7), 755-770.

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychological bulletin*, *117*(2), 285.

Calev, A. (1996). Affect and memory in depression: Evidence of better delayed recall of positive than negative affect words. *Psychopathology*, *29*(2), 71-76.

Coffey, P., Leitenberg, H., Henning, K., Turner, T., & Bennett, R. T. (1996). Mediators of the long-term impact of child sexual abuse: Perceived stigma, betrayal, powerlessness, and self-blame. *Child Abuse & Neglect*, *20*(5), 447-455.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Academic press.
Couture, S. M., Penn, D. L., & Roberts, D. L. (2006). The functional significance of social cognition in schizophrenia: a review. *Schizophrenia bulletin*, *32*(suppl 1), S44-S63.

Dam, V. H., Thystrup, C. K., Jensen, P. S., Bland, A. R., Mortensen, E. L., Elliott, R., ... & Stenbæk, D. S. (2019). Psychometric properties and validation of the EMOTICOM test battery in a healthy Danish population. Frontiers in psychology, 10, 2660.

Danion, J. M., Kauffmann-Muller, F., Grangé, D., Zimmermann, M. A., & Greth, P. (1995). Affective valence of words, explicit and implicit memory in clinical depression. *Journal of affective disorders*, *34*(3), 227-234.

Del Carlo, A., Benvenuti, M., Fornaro, M., Toni, C., Rizzato, S., Swann, A. C., ... & Perugi, G. (2012). Different measures of impulsivity in patients with anxiety disorders: a case control study. *Psychiatry research*, *197*(3), 231-236.

Dillon, D. G., & Pizzagalli, D. A. (2018). Mechanisms of memory disruption in depression. *Trends in neurosciences*, *41*(3), 137-149.

Dillon, D. G., Holmes, A. J., Birk, J. L., Brooks, N., Lyons-Ruth, K., & Pizzagalli, D. A. (2009). Childhood adversity is associated with left basal ganglia dysfunction during reward anticipation in adulthood. *Biological psychiatry*, *66*(3), 206-213.

Dorresteijn, S., Gladwin, T. E., Eekhout, I., Vermetten, E., & Geuze, E. (2019). Childhood trauma and the role of self-blame on psychological well-being after deployment in male veterans. *European journal of psychotraumatology*, *10*(1), 1558705.

Elliott, R., Zahn, R., Deakin, J. W., & Anderson, I. M. (2011). Affective cognition and its disruption in mood disorders. *Neuropsychopharmacology*, *36*(1), 153-182.

Ellwart, T., Rinck, M., & Becker, E. S. (2003). Selective memory and memory deficits in depressed inpatients. *Depression and Anxiety*, *17*(4), 197-206.

Erickson, K., Drevets, W. C., Clark, L., Cannon, D. M., Bain, E. E., Zarate Jr, C. A., ... & Sahakian, B. J. (2005). Mood-congruent bias in affective go/no-go performance of

unmedicated patients with major depressive disorder. *American Journal of Psychiatry*, *162*(11), 2171-2173.

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*(2), 175–91. http://doi.org/10.3758/BF03193146

Feeny, N. C., Zoellner, L. A., Fitzgibbons, L. A., & Foa, E. B. (2000). Exploring the roles of emotional numbing, depression, and dissociation in PTSD. *Journal of traumatic stress*, *13*(3), 489-498.

Fields, S. A., Schueler, J., Arthur, K. M., & Harris, B. (2021). The Role of Impulsivity in Major Depression: A Systematic Review. Current Behavioral Neuroscience Reports, 1-13.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *358*(1431), 459-473.

Germine, L., Dunn, E. C., McLaughlin, K. A., & Smoller, J. W. (2015). Childhood adversity is associated with adult theory of mind and social affiliation, but not face processing. *PloS one*, *10*(6), e0129612.

Gibb, B. E., Schofield, C. A., & Coles, M. E. (2009). Reported history of childhood abuse and young adults' information-processing biases for facial displays of emotion. *Child maltreatment*, *14*(2), 148-156.

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

Gotlib, I. H., Kasch, K. L., Traill, S., Joormann, J., Arnow, B. A., & Johnson, S. L. (2004). Coherence and specificity of information-processing biases in depression and social phobia. *Journal of abnormal psychology*, *113*(3), 386.

Guyer, A. E., Kaufman, J., Hodgdon, H. B., Masten, C. L., Jazbec, S., Pine, D. S., & Ernst, M. (2006). Behavioral alterations in reward system function: the role of childhood maltreatment and psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1059-1067.

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.

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, *23*(1), 56.

Harlé, K. M., & Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion*, *7*(4), 876.

Harlé, K. M., Allen, J. J., & Sanfey, A. G. (2010). The impact of depression on social economic decision making. *Journal of abnormal psychology*, *119*(2), 440.

Harmer, C. J., & Cowen, P. J. (2013). 'It's the way that you look at it'—a cognitive neuropsychological account of SSRI action in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1615), 20120407.

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., ... & Cowen, P. J. (2009). Effect of acute antidepressant administration on negative affective bias in depressed patients. *American Journal of Psychiatry*, *166*(10), 1178-1184.

Hepp, J., Schmitz, S. E., Urbild, J., Zauner, K., & Niedtfeld, I. (2021). Childhood maltreatment is associated with distrust and negatively biased emotion processing. *Borderline personality disorder and emotion dysregulation*, *8*(1), 1-14.

Hershenberg, R., Satterthwaite, T. D., Daldal, A., Katchmar, N., Moore, T. M., Kable, J. W., & Wolf, D. H. (2016). Diminished effort on a progressive ratio task in both unipolar and bipolar depression. Journal of affective disorders, 196, 97-100.

Hughes, J. R., Pleasants, C. N., & Pickens, R. W. (1985). Measurement of reinforcement in depression: a pilot study. *Journal of behavior therapy and experimental psychiatry*, *16*(3), 231-236.

Huh, H. J., Baek, K., Kwon, J. H., Jeong, J., & Chae, J. H. (2016). Impact of childhood trauma and cognitive emotion regulation strategies on risk-aversive and loss-aversive patterns of decision-making in patients with depression. *Cognitive neuropsychiatry*, *21*(6), 447-461.
Keefe, R. S., & Harvey, P. D. (2012). Cognitive impairment in schizophrenia. In *Novel antischizophrenia treatments* (pp. 11-37). Springer Berlin Heidelberg.

Khan, A., McCormack, H. C., Bolger, E. A., McGreenery, C. E., Vitaliano, G., Polcari, A., & Teicher, M. H. (2015). Childhood maltreatment, depression, and suicidal ideation: critical importance of parental and peer emotional abuse during developmental sensitive periods in males and females. Frontiers in psychiatry, 6, 42.

Knutson, B., Bhanji, J. P., Cooney, R. E., Atlas, L. Y., & Gotlib, I. H. (2008). Neural responses to monetary incentives in major depression. *Biological psychiatry*, *63*(7), 686-692.
Locher, S. C., Barenblatt, L., Fourie, M. M., Stein, D. J., & Gobodo-Madikizela, P. (2014). Empathy and childhood maltreatment: A mixed-methods investigation. *Annals of Clinical Psychiatry*, *26*(2), 97-110.

Majer, M., Nater, U. M., Lin, J. M. S., Capuron, L., & Reeves, W. C. (2010). Association of childhood trauma with cognitive function in healthy adults: a pilot study. *BMC neurology*, *10*(1), 1-10.

Mannie, Z. N., Williams, C., Browning, M., & Cowen, P. J. (2015). Decision making in young people at familial risk of depression. *Psychological medicine*, *45*(2), 375-380.

Mattern, M., Walter, H., Hentze, C., Schramm, E., Drost, S., Schoepf, D., ... & Schnell, K. (2015). Behavioral evidence for an impairment of affective theory of mind capabilities in chronic depression. *Psychopathology*, *48*(4), 240-250.

Matthews, K., & Robbins, T. W. (2003). Early experience as a determinant of adult behavioural responses to reward: the effects of repeated maternal separation in the rat. *Neuroscience & Biobehavioral Reviews*, *27*(1-2), 45-55.

Mehta, M. A., Gore-Langton, E., Golembo, N., Colvert, E., Williams, S. C., & Sonuga-Barke, E. (2010). Hyporesponsive reward anticipation in the basal ganglia following severe institutional deprivation early in life. *Journal of cognitive neuroscience*, *22*(10), 2316-2325.

Miu, A. C., Bîlc, M. I., Bunea, I., & Szentágotai-Tătar, A. (2017). Childhood trauma and sensitivity to reward and punishment: Implications for depressive and anxiety symptoms. *Personality and Individual Differences*, *119*, 134-140.

Moustafa, A. A., Tindle, R., Frydecka, D., & Misiak, B. (2017). Impulsivity and its relationship with anxiety, depression and stress. *Comprehensive psychiatry*, *74*, 173-179.

Murphy, F. C., Michael, A., Robbins, T. W., & Sahakian, B. J. (2003). Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. *Psychological medicine*, *33*(03), 455-467.

Murphy, F. C., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., Paykel, E. S., & Sahakian, B. J. (2001). Decision-making cognition in mania and depression. *Psychological medicine*, *31*(4), 679-693.

Murphy, F. C., Sahakian, B. J., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., & Paykel, E. S. (1999). Emotional bias and inhibitory control processes in mania and depression. *Psychological medicine*, *29*(06), 1307-1321.

Nikolin, S., Tan, Y. Y., Schwaab, A., Moffa, A., Loo, C. K., & Martin, D. (2021). An investigation of working memory deficits in depression using the n-back task: A systematic review and meta-analysis. *Journal of Affective Disorders*.

Novick, A. M., Levandowski, M. L., Laumann, L. E., Philip, N. S., Price, L. H., & Tyrka, A. R. (2018). The effects of early life stress on reward processing. *Journal of psychiatric research*, *101*, 80-103.

Oakley, B. F., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of abnormal psychology*, *125*(6), 818.

Orth, U., Berking, M., & Burkhardt, S. (2006). Self-conscious emotions and depression: Rumination explains why shame but not guilt is maladaptive. *Personality and social psychology bulletin*, *32*(12), 1608-1619.

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: an integrated review of human literature. *Psychopharmacology*, *214*(1), 55-70.

Pine, D. S., Mogg, K., Bradley, B. P., Montgomery, L., Monk, C. S., McClure, E., ... & Kaufman, J. (2005). Attention bias to threat in maltreated children: Implications for vulnerability to stress-related psychopathology. *American Journal of Psychiatry*, *162*(2), 291-296.

Pizzagalli, D. A., Evins, A. E., Schetter, E. C., Frank, M. J., Pajtas, P. E., Santesso, D. L., & Culhane, M. (2008). Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology*, *196*(2), 221-232.

Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., ... & Fava, M. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *American Journal of Psychiatry*, *166*(6), 702-710.

Pollak SD, Cicchetti D, Hornung K, Reed A (2000) Recognizing emotion in faces: developmental effects of child abuse and neglect. Dev Psychol 36:679–688.

Pollak SD, Klorman R, Thatcher JE, Cicchetti D (2001) P3b reflects maltreated children's reactions to facial displays of emotion. Psychophysiology 38:267–274.

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences*, *99*(13), 9072-9076.

Pollak, S. D., & Tolley-Schell, S. A. (2003). Selective attention to facial emotion in physically abused children. *Journal of abnormal psychology*, *112*(3), 323.

Pryce, C. R., Dettling, A. C., Spengler, M., Schnell, C. R., & Feldon, J. (2004). Deprivation of parenting disrupts development of homeostatic and reward systems in marmoset monkey offspring. *Biological psychiatry*, *56*(2), 72-79.

Pulcu, E., Thomas, E. J., Trotter, P. D., McFarquhar, M., Juhász, G., Sahakian, B. J., ... & Elliott, R. (2015). Social-economical decision making in current and remitted major depression. *Psychological medicine*, *45*(06), 1301-1313.

Rawal, A., Collishaw, S., Thapar, A., & Rice, F. (2013). 'The risks of playing it safe': a prospective longitudinal study of response to reward in the adolescent offspring of depressed parents. *Psychological medicine*, *43*(1), 27-38.

Ridout, N., Astell, A., Reid, I., Glen, T., & O'Carroll, R. (2003). Memory bias for emotional facial expressions in major depression. *Cognition and Emotion*, *17*(1), 101-122.

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological medicine*, *44*(10), 2029-2040.

Roiser, J. P., & Sahakian, B. J. (2013). Hot and cold cognition in depression. *CNS spectrums*, *18*(3), 139-149.

Roiser, J. P., Cannon, D. M., Gandhi, S. K., Tavares, J. T., Erickson, K., Wood, S., ... & Drevets, W. C. (2009). Hot and cold cognition in unmedicated depressed subjects with bipolar disorder. *Bipolar disorders*, *11*(2), 178-189.

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*, *37*(1), 117-136.

Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinssen, R. K., ... & Cuthbert, B. N. (2010). Developing constructs for psychopathology research: research domain criteria. *Journal of abnormal psychology*, *119*(4), 631.

Savulich, G., Bowden-Jones, O., Stephenson, R., Brühl, A. B., Ersche, K. D., Robbins, T. W., & Sahakian, B. J. (2021). "Hot" and "Cold" Cognition in Users of Club Drugs/Novel Psychoactive Substances. Frontiers in psychiatry, 12, 333.

Sekowski, M., Gambin, M., Cudo, A., Wozniak-Prus, M., Penner, F., Fonagy, P., & Sharp, C. (2020). The relations between childhood maltreatment, shame, guilt, depression and suicidal ideation in inpatient adolescents. *Journal of affective disorders*, *276*, 667-677.

Sharma-Patel, K., & Brown, E. J. (2016). Emotion regulation and self blame as mediators and moderators of trauma-specific treatment. *Psychology of Violence*, *6*(3), 400.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. In *Journal of Clinical Psychiatry* (Vol. 59, pp. 22–33). http://doi.org/10.1016/S0924-9338(99)80239-9

Simon, M., Németh, N., Gálber, M., Lakner, E., Csernela, E., Tényi, T., & Czéh, B. (2019). Childhood adversity impairs theory of mind abilities in adult patients with major depressive disorder. *Frontiers in psychiatry*, *10*, 867.

Stoops, W. W. (2008). Reinforcing effects of stimulants in humans: sensitivity of progressive-ratio schedules. *Experimental and clinical psychopharmacology*, *16*(6), 503.
Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual review of psychology*, *58*, 345.

Tangney, J.P., Wagner, P.E. & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology 103*, 469-478.

Tanzer, M., Salaminios, G., Morosan, L., Campbell, C., & Debbané, M. (2021). Self-blame mediates the link between childhood neglect experiences and internalizing symptoms in low-risk adolescents. *Journal of Child & Adolescent Trauma*, *14*(1), 73-83.

Teicher, M. H., & Parigger, A. (2015). The 'Maltreatment and Abuse Chronology of Exposure'(MACE) scale for the retrospective assessment of abuse and neglect during development. *PLoS one*, *10*(2), e0117423.

Thaipisuttikul, P., Ittasakul, P., Waleeprakhon, P., Wisajun, P., & Jullagate, S. (2014). Psychiatric comorbidities in patients with major depressive disorder. *Neuropsychiatric disease and treatment*, *10*, 2097.

Treadway, M. T., & Zald, D. H. (2013). Parsing anhedonia: translational models of reward-processing deficits in psychopathology. Current Directions in Psychological Science, 22(3), 244-249.

Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *Journal of abnormal psychology*, *121*(3), 553.

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PloS one*, *4*(8), e6598.

Walker, E. A., Unutzer, J., Rutter, C., Gelfand, A., Saunders, K., VonKorff, M., ... & Katon, W. (1999). Costs of health care use by women HMO members with a history of childhood abuse and neglect. *Archives of general psychiatry*, *56*(7), 609-613.

Watkins, P. C., Mathews, A., Williamson, D. A., & Fuller, R. D. (1992). Mood-congruent memory in depression: Emotional priming or elaboration?. *Journal of Abnormal Psychology*, *101*(3), 581.

Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., & Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: data from the randomized international study to predict optimized treatment for depression. *Translational psychiatry*, *6*(5), e799-e799.

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but I can't deal with it: Impaired theory of mind in depression. *Journal of affective disorders*, *132*(1-2), 104-111.

Xie, P., Wu, K., Zheng, Y., Guo, Y., Yang, Y., He, J., ... & Peng, H. (2018). Prevalence of childhood trauma and correlations between childhood trauma, suicidal ideation, and social support in patients with depression, bipolar disorder, and schizophrenia in southern China. *Journal of affective disorders*, *228*, 41-48.

Yang, X. H., Huang, J., Zhu, C. Y., Wang, Y. F., Cheung, E. F., Chan, R. C., & Xie, G. R. (2014). Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry research*, *220*(3), 874-882.

Zahn, R., Lythe, K. E., Gethin, J. A., Green, S., Deakin, J. F. W., Young, A. H., & Moll, J. (2015). The role of self-blame and worthlessness in the psychopathology of major depressive disorder. *Journal of affective disorders*, *186*, 337-341.

Zobel, I., Werden, D., Linster, H., Dykierek, P., Drieling, T., Berger, M., & Schramm, E. (2010). Theory of mind deficits in chronically depressed patients. *Depression and anxiety*, *27*(9), 821-828.

Study 2: Systematic Review of Grey Matter Volume Changes in Depression and Early Life Stress

Franziska Goer[1], Kieran Lyon[1], Richard Drake[1], Joanna Neill[1], Rebecca Elliott[1]
[1]University of Manchester, Oxford Road, Manchester M13 9PL, UK

Author Contributions:

Franziska Goer conceived the original idea for the systematic review, decided on eligibility criteria and databases to be searched, conducted the literature search, screened all articles and assessed eligibility, conducted data extraction, interpreted findings, wrote the manuscript and completed all revisions following review by the co-authors.

Kieran Lyon was the second screener for eligibility assessment of articles and second rater for assessment of bias. Professor Rebecca Elliott also contributed to design of the review, provided guidance and supervisory input throughout the entire project, gave guidance and feedback on the literature search and data extraction, and critically reviewed and provided feedback on drafts of the manuscript prepared by FG. Dr. Richard Drake and Professor Joanna Neill provided supervisory input.

Funding Information:
This study was supported by funding from the Medical Research Council and University of Manchester as part of Franziska Goer's Medical Research Council (MRC) Doctoral Training Scholarship.

ABSTRACT

Background: Both Major Depressive Disorder (MDD) and Early Life Stress (ELS) have been associated with various structural brain changes, including a consistent finding of decreased hippocampal volume for both. However, most studies have investigated either ELS or MDD in isolation, making it difficult to ascertain the effect of each on grey matter volume (GMV) changes. Given the high levels of ELS present in individuals with MDD, ELS may present a confounding factor which may potentially be contributing to, or even driving, several of the GMV effects often attributed to MDD alone.
Methods: The aim of this systematic review was to provide a comprehensive overview of all empirical studies that measured GMV, MDD and ELS. PRISMA guidelines were followed in the planning and execution of the systematic review, and a detailed protocol was registered with PROSPERO prior to commencement. An electronic literature search was conducted in five databases (PsychInfo, Ovid MEDLINE, EMBASE, OpenGrey & medRxiv/bioRxiv) and yielded 5129 records, including 4644 published papers and 548 unpublished papers (grey literature). Following screening of titles/abstract and subsequent full-text review of remaining records by two independent reviewers, 20 studies met the full eligibility criteria and were included in the systematic review. All 20 included studies met criteria for low risk of bias (using the Joanna Briggs Institute Critical Appraisal Checklist for Analytical Cross-Sectional Studies).
Results: Analysis of included studies revealed that ELS is associated with hippocampal GMV reductions (both for the whole hippocampus and its subfields, particularly the cornu ammonis, dentate gyrus, and subiculum) independent of MDD diagnosis. In studies without a healthy control (HC) group, hippocampal volume reductions were more pronounced in depressed participants with ELS than those without. Additional regions, including the caudate, orbitofrontal cortex, and posterior cingulate cortex, were found to have lower GMV in ELS, independent of MDD diagnosis.
Conclusion: This systematic review found that ELS appears to be the driving factor in GMV reductions, independent of MDD, most consistently reported in the hippocampus but also in several other key brain areas frequently implicated in depression. However, due to low levels of ELS in the HC group of several studies, these findings should be cautiously interpreted, and more studies, ideally including a high ELS HC group, are needed to ascertain the effect of ELS independent of depressive symptoms. Overall, the findings highlight the need for measures of ELS in studies of brain structure in depression.

1. Introduction

Early life stress (ELS) has been implicated in both animal models and clinical studies as a significant risk factor in the aetiology of mental health disorders, including major depressive disorder (MDD; Lindert et al., 2014; McCrory et al., 2017). However, the precise mechanism by which ELS increases the risk for psychiatric disorders remains unknown. One proposed theory is that ELS may lead to stress-induced structural brain changes that in turn may increase vulnerability to disorders such as MDD (Hammen et al., 2000; Lupien et al., 2008; Magarin & McEwen, 1995). Both MDD and ELS have been associated with various structural brain changes, with reductions in hippocampal volume (both hippocampus proper and various subfields) being the most frequently studied and reported for both; Bremner et al., 2000; Calem et al., 2017; Campbell et al., 2004; Paquola et al., 2016; Videbech & Ravnkilde, 2004). However, the vast majority of these studies have analysed MDD or ELS in isolation, making it very difficult to disentangle potential effects of one versus the other on changes in grey matter volume (GMV). The issue of not controlling for ELS in studies of depression is particularly problematic given the high prevalence of ELS exposure in individuals with MDD: estimates of significant ELS in depression have been reported between 50.5% to 62.5% in individuals with MDD diagnosis; Williams et al., 2016; Xie et al., 2018), raising the possibility that unmeasured ELS in participants may be contributing to (or perhaps even driving) findings in depression.

A few studies (which did control for both MDD and ELS in their analyses) have reported that ELS may in fact be the driving factor behind the frequently reported hippocampal GMV reduction in MDD (Chaney et al., 2014; Opel et al., 2014; Vythilingam et al., 2002). However, contradictory findings have also been published (Mikolas et al., 2019) and further

investigation is needed to ascertain whether ELS may indeed underlie GMV reductions in MDD and if so, whether these are limited to hippocampal GMV or perhaps extends to other brain areas implicated in MDD. Whole brain analyses and/or studies including multiple regions of interest (ROIs) may help address this latter question. For instance, a large multisite study of the ENIGMA framework with over 3000 MDD and HC participants reported a significant negative correlation between ELS and caudate volume, independent of ELS, but did not identify any additional significant brain areas in their whole brain volumetric analyses.

There is a significant need for an improved understanding of the role of ELS and MDD in GMV changes as it may help develop more targeted and well-timed treatments and interventions, with the hope of possibly avoiding or reducing structural brain changes, which in turn have been independently implicated in worse treatment outcomes in depression (Frodl et al., 2010) and other disorders such as substance use disorder (van Dam et al., 2014). Promisingly, both antidepressant treatment (Boldrini et al., 2009; Frodl et al., 2008) and cognitive behavioural therapy (Levy-Gigi et al., 2013) have shown promise in reversing hippocampal volume reduction in MDD and posttraumatic stress disorder (PTSD), respectively.

In order to attempt to ascertain whether ELS may primarily account for GMV changes observed in MDD, or whether there perhaps is an additive effect or mediation involving the two variables, a systematic review of the literature was conducted.

1.1. Aims and hypothesis

The aim of this systematic review is to provide a comprehensive overview of all empirical studies that measured GMV, MDD and ELS. First and foremost, the systematic review seeks to identify whether MDD and/or ELS are associated with changes in GMV (using both ROI and whole brain approaches) and importantly, aims to differentiate effects of depression and ELS on brain structure to determine whether one of these factors is driving GMV changes or whether there is perhaps a mediating or additive effect of these factors. Furthermore, this systematic review will highlight which brain regions may be implicated in potential GMV changes in ELS and/or MDD (though this may likely be biased and not fully comprehensive since we do not exclude studies that take an ROI rather than a whole brain approach).

Based on previous literature we hypothesised that ELS partially or fully mediates the reduction of GMV often observed in MDD, in particular in the hippocampus.

2. Methods

The PRISMA guidelines were adopted in the methodology and reporting of results of this systematic review (Moher et al., 2009). A detailed protocol was submitted and approved by PROSPERO (https://www.crd.york.ac.uk/prospero), an international prospective register of systematic reviews, prior to conducting literature searches (registration reference: CRD42020223533).

It should be noted that for purposes of this manuscript, the term ELS is used for all references to significant early life stress occurring during childhood (generally before the age of 18). The ELS literature uses many different terms and definitions, including early life adversity, childhood trauma, maltreatment, and others; for consistency, all of these

constructs are referred to as ELS henceforth. Specific details of how each included study defined and measured ELS are given in the results section.

2.1. Search strategy

An electronic literature search was conducted on 25th November 2020 in five databases: PsychInfo, Ovid MEDLINE, EMBASE, OpenGrey & medRxiv/bioRxiv. A manual search of the references of included studies was also conducted to identify any additional relevant studies. OpenGrey and medRxiv were included to decrease bias by searching unpublished studies (such as research reports, dissertations, conference papers, etc) and preprints in health sciences, respectively. The exact search phrase implemented for each database was: ((early AND life AND stress) OR (early AND life AND trauma) OR (childhood AND abuse) OR (childhood AND trauma) OR (childhood AND adversity) OR (childhood AND maltreatment) OR (childhood AND neglect)) AND (depress* OR (beck AND depression AND inventory) OR (low AND mood)) AND ((brain AND volume) OR (brain AND structural) OR (brain AND abnormalit*) OR (brain AND change*) OR neuroimaging OR (grey AND matter) OR (grey AND matter) OR mri OR brain)). All identified references (including abstracts) were exported into an EndNote library.

2.2. Study identification

PRISMA guidelines were followed for study identification: 2 reviewers (FG and KL) independently screened titles and abstracts (for all identified articles as part of the first screen) and full texts (during the second, in-depth screen for a subset of articles identified as potentially eligible after the initial screen) in random order to determine eligibility

according to the selection criteria. Any disagreements were discussed between authors and resolved through consensus.

2.3. Selection criteria

Inclusion criteria for studies were: (i) original data (no reviews) of human subjects (no animal studies), (ii) the patient population studied were adults (aged 18+), (iii) studies included at least one measure of ELS, (iv) studies included a validated measure of depression or MDD diagnosis, (v) studies included measures of GMV obtained from structural MRI scans performed on participants. Since different types of control groups could be appropriate (e.g. healthy controls with no psychiatric history, depressed individuals with and without ELS exposure, or healthy individuals with and without ELS exposure) we did not define a specific type of control group in our study criteria. Since ELS is a broad construct with many different definitions and measurements we did not confine our eligibility criteria to any specific type of ELS measurement and rather included any measure that established childhood maltreatment (defined as abuse and/or neglect of various types) as has been done in previous systematic reviews involving ELS (Cassiers et al., 2008).

2.4. Data extraction and risk of bias assessment

Data was extracted from the included papers by the first author. Data extraction included information on study site and population, demographics, inclusion/exclusion criteria, definitions of MDD and control group(s), measures of main variables (ELS, MDD, GMV), statistics, limitations/bias, funding/conflict of interest, and main findings (including null results). A single included study (Hassel et al., 2017) consisted of a published abstract only rather than a full manuscript. Since information provided about study design and inclusion

criteria was limited in the abstract, this information was complemented by a recently published manuscript of the same study (CAN-BIND study; Nogovitsyn et al., 2020). Quality assessment to identify risk of bias in individual studies was conducted independently by two authors (FG and KL) and any discrepancies were later discussed and resolved by consensus. Since all included studies were of cross-sectional design the Joanna Briggs Institute Critical Appraisal Checklist for Analytical Cross-Sectional Studies (Joanna Briggs Institute, 2020) was applied. The checklist includes 8 questions covering inclusion criteria; description of study subjects and setting; valid and reliable exposure measure; objective, standard criteria used for measurement of condition; identification of confounding factors; strategies to deal with confounding factors; reliable measurement of outcome variables; and appropriate statistical analysis. We rated each individual study on these 8 questions, giving them a score of 0 (absent/insufficient), 0.5 (partial), and 1 (satisfactory/complete), for each question, resulting in a total score with a possible range of 0-8 for each study. We applied previously used cut-offs (Melo et al., 2018) to determine low risk of bias (total score >6), moderate risk of bias (total score 4-6), and high risk of bias (total score <4).

3. Results

Due to the heterogenous design and measures of studies included (in regard to brain area(s) analysed, inclusion/exclusion criteria, medication status, ELS measure, scanner parameters and imaging analysis) a meta-analysis was not possible; hence a narrative summary and qualitative analysis of the results was conducted.

3.1. Study selection

Our search yielded a total of 4644 published papers broken down as follows by database: EMBASE (2406), APA PsycInfo (1188), Medline (1050). Once identical duplicates were removed a total of 3082 papers remained. Search of grey (unpublished) literature yielded 548 papers, 266 in openGrey and 282 in medRxiv. An additional 3 articles were included through manual screening of references of included articles. Following identification, screening, and assessment of eligibility, a final total of 20 studies were included in the systematic review. Details of identification, screening, eligibility, and inclusion can be seen in the PRISMA flowchart in **Figure 1**.

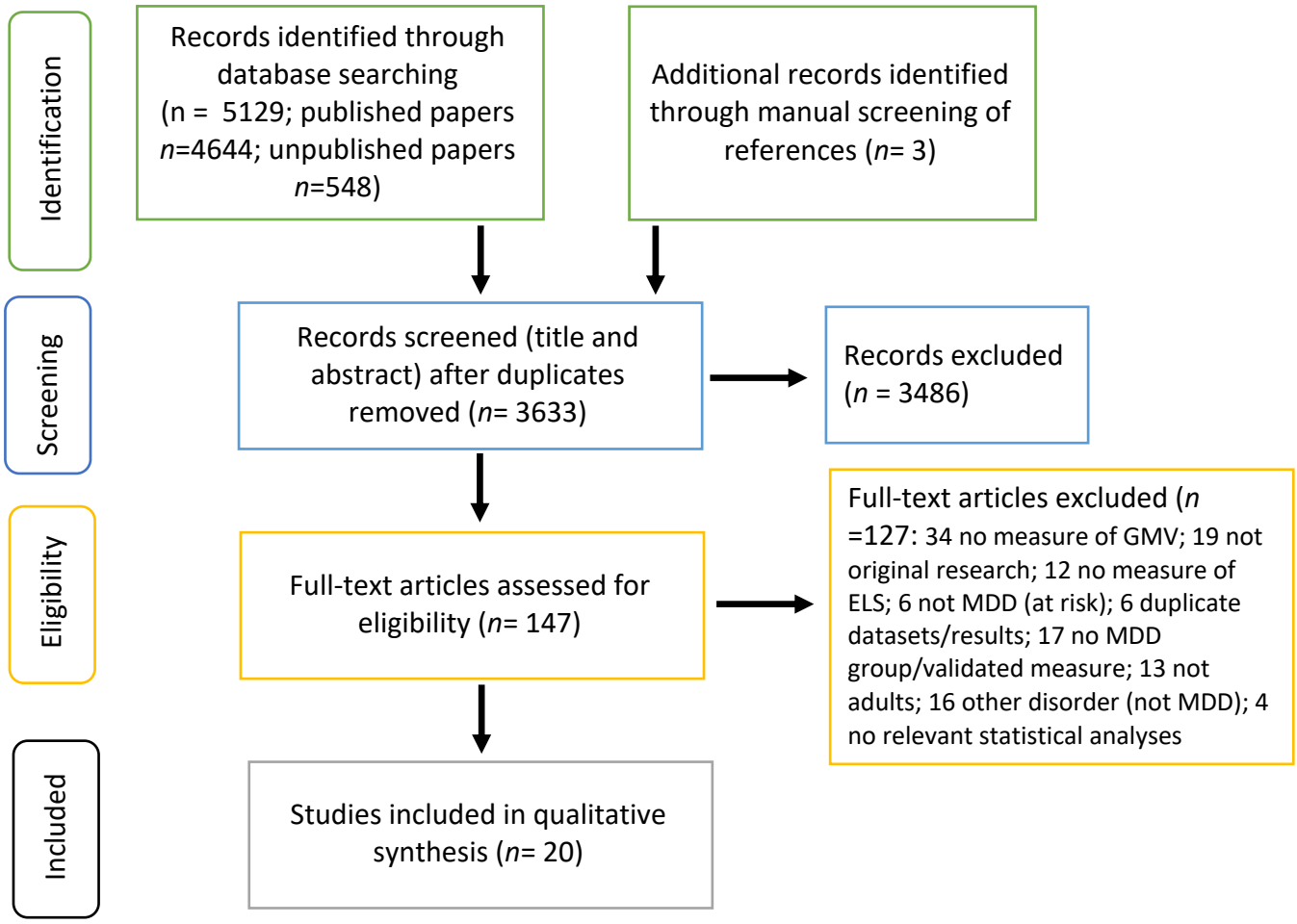


Figure 1. PRISMA flowchart of study identification and selection

3.2. Risk of bias assessment

The quality assessment scores using the Joanna Briggs Institute Critical Appraisal Checklist for Analytical Cross-Sectional Studies (Joanna Briggs Institute, 2020) are listed in **Table 1**. Applying cut-offs used in previous systematic reviews (Melo et al., 2018) revealed that all 20 included studies met criteria for low risk of bias (total score >6). Five studies were given partial scores for questions 5 and 6 (identification of and controlling for confounding factors) as they did not control for total intracranial volume (ICV) in their analyses (though they did control for other factors, such as age and gender). The only score of zero on any criteria was given for a lack of validated ELS measure in one study (Yuan et al., 2020) in which participants were merely asked a yes/no question about previous trauma experience.

Table 1. Risk of bias assessment

Study	Clearly defined inclusion criteria	Descript-ions of subjects and setting	Valid & reliable meas-ure ELS	Object-ive, standard criteria for MDD	Confound-ing factors identified	Control for confound-ing factors	Reliab-le & valid meas-ure of out-come	Statist-ical analyses	Total score
Bermingham et al., 2012	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🟢	**7.5**
Carballedo et al., 2013	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟢	**7**
Chaney et al., 2014	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟢	**7**
Frodl et al., 2017	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Colle et al., 2017	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Croy et al., 2013	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟢	**7**
Frodl et al., 2010	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Gerritsen et al., 2015	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🟢	**7.5**
Hassel et al., 2017	🟢	🟡	🟢	🟢	🟢	🟢	🟢	🟡	**7**
Lu et al., 2019	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟢	**7**
Lu et al., 2018	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟢	**7**
Opel et al., 2014	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**

Saleh et al., 2017	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Aghamohammadi-Sereshki et al., 2021	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Vythilingam et al., 2002	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Yang et al., 2017	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Yuan et al., 2020	🟢	🟢	🔴	🟢	🟢	🟢	🟢	🟢	**7**
Mikolas et al., 2019	🟡	🟡	🟢	🟢	🟢	🟢	🟢	🟢	**7**
Teicher et al., 2012	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**
Tannous et al., 2020	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	**8**

Table 1 key: Green circle = score of 1 (criteria satisfactorily fulfilled), orange circle = score of 0.5 (criteria partially fulfilled), red circle = score of 0 (criteria not fulfilled)

3.3. Study characteristics

All 20 included studies were cross-sectional studies published in peer-reviewed scientific journals and none of the authors reported any conflicts of interest. Please see **Table 2a-d** for detailed study characteristics of each study, including study setting, participant characteristics and demographics, ELS and MDD measures, imaging parameters and analysis software, and statistical analyses. For more detailed inclusion/exclusion criteria for each study, please see **Appendix Table 1**. Across all 20 studies, conducted in 10 countries (Ireland, Germany, Netherlands, Australia, South Africa, France, Canada, China, USA, and the UK) a grand total of 5994 adult (aged 18+) participants were included. 17 studies compared MDD participants to healthy controls (HC), of which 4 studies divided these groups further into four subgroups depending on presence/absence of ELS (MDD without ELS: MDD; MDD with ELS: MDDELS; HC without ELS: HC; and HC with ELS: HCELS), and one study divided them into three subgroups (as above but without a HCELS group due to limited numbers). Two studies compared MDDELS to MDD (without a healthy control comparison group) and

one study used a community sample with a range of participants including HC, MDD, and other disorders. See **Table 2a** for detailed demographics of study participants in each study.

Table 2a. Study details and demographics

Study (Authors)	Country	Total participants	Population studied	Age participants	Gender (% female)	Education	Ethnicity
Bermingham et al., 2012	Ireland	88	MDD vs. HC	T gene carriers/ no carrier: MDD: 45.8 (9.3)/ 40 (10.3) ; HC: 36.6 (11.8)/ 35.3 (13.4)	T gene carriers/no carrier: MDD: 46.2% / 71% ; HC: 50% / 69.2%	Not reported	Not reported
Carballedo et al., 2013	81 from Ireland, 52 from Germany	133	MDD vs. HC	MDD: 41.8 (11.1); HC: 38.4 (13.5)	MDD: 61.29%; HC: 61.97%	Not reported	Not reported
Chaney et al., 2014	Ireland	83	MDD vs. HC	MDD: 40.6 (10.4), MDDELS: 39.9 (9.7), HC: 34.2 (10.8), HCELS: 45.3 (15.8)	MDD: 70.59%, MDDELS: 45%, HC: 66.66%, HCELS: 40%	Not reported	Not reported
Frodl et al., 2017	9 study sites in 5 countries: Netherlands, Ireland, Germany, Australia, South Africa	3036	MDD vs. HC	Varied by sample, ranges: 25.5 (5.0) - 55.3 (12.8) for HC and 18.9 (3.0) - 53.6 (11.9) for MDD.	Range differed by study site, from 33.3-71.8% in MDD and 25-64.5% in HC	Not reported	Not reported
Colle et al., 2017	France	63	MDD with ELS vs. MDD without ELS (no healthy control group)	46.4 (±12.4)	58.70%	7 (11.1%) low educational level, 31 (49.2%) middle educational level, 25 (39.7%) high educational level.	Not reported
Croy et al., 2013	Germany	23	MDD with ELS vs. MDD without ELS (no healthy control group)	21-49, mean: 37.7 (9.6)	100% female	Not reported	Not reported

Frodl et al., 2010	Germany	87	MDD vs. HC	MDD: 44.2 (12.2), HC: 41.1 (12.5)	MDD: 60.46%, HC: 54.55%	Not reported	Not reported
Gerritsen et al., 2015	Netherlands	898 (2 samples: 262 NESDA + 636 SMART)	MDD (past year) vs. HC	NESDA: 37.8 (10.2), SMART: 61.5 (9.6)	NESDA: 68%, SMART 19%	Not reported	Not reported
Hassel et al., 2017	Canada	139	MDD vs. HC	mean age not reported, range: 18-60	MDD: 61.25%, HC: 67.78%	Not reported	Not reported
Lu et al., 2019	China	78	MDDELS (n=16), MDD (n=14), HCELS (n=24), HC (n=24)	MDDELS: 24.4 (4.79), MDD: 23.5 (5.77), HCELS: 21.5 (3.98), HC: 21.5 (3.69)	Overall: 55.13%, MDDELS: 31.25%, MDD: 57.14%, HCELS: 62.5%, HC: 62.5%	In years: MDDELS: 14.2 (2.26), MDD: 14.5 (3.30), HCELS: 14.0 (1.3), HC: 14.7 (1.92)	100% Han ethnicity (inclusion criteria)
Lu et al., 2018	China	168	MDDELS (51), MDD (31), HCELS (48), HC (38).	MDDELS: 33.2 (8.7), MDD: 36.1 (7.8), HCELS: 33.0 (7.8), HC: 33.8 (7.4)	MDDELS: 58.33, MDD: 57.14%, HCELS: 50%, HC: 50%	In years: MDDELS: 10.7 (3.8), MDD: 10.8 (4.2), HCELS: 11.0 (3.2), HC: 12.8 (3.7)	Not reported
Opel et al., 2014	Germany	170	MDD vs. HC	MDD: 37.6 (12.0), HC: 37.2 (11.8)	MDD: 63.53%, HC: 60%	In years: MDD: 3.1 (1.3), HC: 3.1 (1.7)	Not reported
Saleh et al., 2017	USA	104	MDD vs. HC	MDD: 35.1 (8.9), HC: 29.7 (9.2)	MDD: 60.9%, HC: 66.2%	In years: MDD: 15.2 (2.4), HC: 16.0 (1.9)	% White: MDD: 67.2%, HC: 55.4%
Aghamohammadi-Sereshki et al., 2021	Canada	70	MDD vs. HC	HC: 32.3 (10.0), MDD: 34.9 (8.7)	65.71% in both MDD and HC	In years: HC: 15.7 (1.7), MDD: 15.4 (1.8)	Not reported
Vythilingam et al., 2002	USA	46	MDD (with and without ELS) vs. HC	MDDELS: 33 (6), MDD: 34 (8), HC: 27 (5)	100%	Not reported	% white: MDDELS: 81%, MDD: 72.7%, HC: 64.3%.
Yang et al., 2017	China	168	MDD vs. HC	MDDELS: 31.7 (8.0), MDD: 30.1 (7.5), HCELS: 32.9 (7.6), HC: 29.8 (6.5)	MDDELS: 73.2%, MDD: 72.09%, HCELS: 75%, HC: 72.1%	In years: MDDELS: 11.5 (4.5), MDD: 11.9 (5.1), HCELS: 14.3 (5.2),	Not reported

						HC: 16.0 (4.2)	
Yuan et al., 2020	USA	87	MDD vs. HC	MDD: 34.8 (10.8), HC: 33.3 (11.5)	MDD: 63.41%, HC: 58.54%	In years: MDD:15.1 (2.4), HC: 16.2 (1.5)	% White MDD: 48.8, HC: 45.5, % African American: MDD: 31.7%, HC: 36.4%
Mikolas et al., 2019	Ireland	152	MDD vs. HC	MDD: 39.926 (8.979), HC: 36.841 (13.15)	MDD: 67.06%, HC: 64.29%	Not reported	Not reported
Teicher et al., 2012	USA	193	Community sample (including MDD and other disorders and HC)	21.9 (2.1)	62%	Not reported	Not reported
Tannous et al., 2020	UK	117	MDD vs. HC	MDD: 31.6 (10.2), HC: 31.5 (10.5)	MDD: 54.9%, HC: 54.3%	Not reported	Not reported

Table 2a. Abbreviations: MDD: Major Depressive Disorder; MDE: Major Depressive Episode; HC: healthy control; ELS: early life stress; MDDELS: participant with MDD and ELS; HCELS: healthy control with ELS

14 (70%) of included studies used the short form (28 questions) of the childhood trauma questionnaire (CTQ, Bernstein et al., 2003) to measure ELS, while 6 studies used another measure. The vast majority of studies used a global composite score of ELS in their analyses (such as total CTQ score) or presence/absence of ELS determined by meeting at least one subscale cut-off of the CTQ (sexual abuse (SA), physical abuse (PA), emotional abuse (EA), physical neglect (PN), or emotional neglect (EN); see **Table 2b** for details). Only three studies conducted analyses on ELS subtypes (Aghamohammadi-Sereshki et al., 2021; Frodl et al., 2010; Frodl et al., 2017), two studies limited ELS to sexual and/or physical abuse (Vythilingam et al., 2002; Yuan et al., 2020), 1 limited analyses to sexual abuse, emotional trauma and severe family conflict (Saleh et al., 2017), 3 did not report any subtypes (Colle et al., 2017; Croy et al., 2013; Hassel et al., 2017) while the remaining 11 studies reported

subtypes of abuse present in the sample but did not use these in analyses. The precise timing (age) of abuse was not reported in any of the included studies, though several specified a maximum cut-off including puberty, 15, 16, or 18 years of age.

Table 2b. Controls and ELS

Study (Authors)	Control group (#)	Control definition	ELS evaluation	ELS subtypes	ELS age
Bermingham et al., 2012	44 HC (N=18 T carriers)	Community sample HC (not mentioned if psychiatric history assessed)	Childhood Trauma Questionnaire (CTQ; Bernstein & Fink, 1998) - (according to accepted cut-off values, Bernstein et al., 1994). But do not clarify how many subtypes cut-offs must be met for	CTQ subtypes listed but not used in analyses	Not reported
Carballedo et al., 2013	71 HC	Community sample HC (not mentioned if psychiatric history assessed)	CTQ rated as present if one of the subcategories: SA, PA, EA, PN or EN met cut-offs (Bernstein et al., 1994).	CTQ subtypes listed but not used in analyses	Not reported
Chaney et al., 2014	46 HC (10 with ELS)	No current/past psychiatric disorders (Axis I or II)	CTQ (rated as present if at least one of the subcategories: SA, PA, EA, PN or EN met cut-offs). 20 MDD & 10 HC met for ELS cut-offs	CTQ subtypes listed but not used in analyses	Not reported
Frodl et al., 2017	2078 HC	2 excluded any lifetime diagnosis of depression or antidepressant use, 7 excluded any current or past axis-I disorder	CTQ (rated as present if at least one of the subcategories: SA, PA, EA, PN or EN met cut-offs, define these cut-offs in paper). Also used total CTQ score as dimensional variable	Yes, separate analyses for subtypes.	Not reported
Colle et al., 2017	No healthy control group	NA	Death of caregiver or child abuse/maltreatment. Assessed by 2 independent psychiatrists using patient health records, patients, family members, psychiatrists etc. ELS present if at least one instance of death of caregiver OR abuse/maltreatment.	Not reported	Not reported
Croy et al., 2013	No healthy control group	NA	Interview by trained psychotherapists and CTQ (minimum average CTQ score of 11 representing 95th percentile)	Not reported	Not reported

Frodl et al., 2010	44 HC	HC (nor their first-degree relatives) had any history of neurological or mental illness.	CTQ	Yes, did separate analyses for physical and emotional neglect (but not other subtypes as levels for those very low and did not differ between MDD and HC)	≤18 years old
Gerritsen et al., 2015	715 HC (123 NESDA, 592 SMART)	Anyone not meeting for MDD in past year considered a control	Nemesis Trauma Interview (Spijker et al., 2002)	Yes, listed (emotional neglect, psychological abuse, physical abuse, sexual abuse) but not analysed separately	<16 years old
Hassel et al., 2017	59 HC	No history of Axis I or Axis II disorders as determined by MINI	Childhood Experience of Care and Abuse (CECA; Bifulco et al., 1994)	Not reported	≤18 years old
Lu et al., 2019	48 HC (HCELS, n=24; HC, n=24)	No current/past psychiatric disorders (Axis I or II) as determined by SCID	CTQ (meeting validated cut-offs for moderate-severe ELS exposure for at least one subscale, Bernstein & Fink, 1998).	CTQ subtypes listed but not used in analyses. Most common form of ELS was emotional neglect.	<16 years old
Lu et al., 2018	86 HC (HCELS, n=48; HC, n=38)	No current/past psychiatric disorders (Axis I or II) as determined by SCID, HAMD score <7	CTQ (meeting validated cut-offs for moderate-severe ELS exposure for at least one subscale, Bernstein & Fink, 1998)	CTQ subtypes listed but not used in analyses.	<16 years old
Opel et al., 2014	85 HC	No current/past psychiatric disorders (assessed by SCID) and <10 on BDI. No psychotropic medication	CTQ	CTQ subtypes listed but not used in analyses.	Not reported

Saleh et al., 2017	53 HC	No history of psychiatric disorder or psychotropic medication use	Modified Early Life Stress Questionnaire (ELSQ Sanders & Becker-Lausen, 1995)	19 subtypes of ELS assessed. Only 3 significantly and independently predicted MDD diagnosis: sexual abuse, emotional trauma and severe family conflict. These 3 subtypes were focused on in analyses (labelled 'predictive ELS').	<18 years old
Aghamohammadi-Sereshki et al., 2021	35 HC	No lifetime psychiatric disorders and no psychosis or mood disorders in first degree relatives	CTQ (only measured in MDD)	CTQ subtypes listed and exploratory analyses of subtypes.	≤18 years old
Vythilingam et al., 2002	14 HC (without ELS)	No lifetime psychiatric disorders	Early Trauma Inventory (Bremner et al., 2000) and independent validation (e.g. court, social services, medical records, family/friends) possible for 52% of MDD participants	ELS limited to physical and/or sexual abuse	Prepubertal (before first menstrual period)
Yang et al., 2017	84 HC (HCELS: 16, HC: 68)	Absence of lifetime (past and present) psychiatric disorders (assessed by SCID)	Chinese version of the CTQ (Zhao et al., 2005). Validated cut-off scores for moderate-severe maltreatment for each subscale was applied, meeting criteria for at least one subscale led to a rating of ELS.	CTQ subtypes listed but not used in analyses.	≤18 years old
Yuan et al., 2020	44 HC (2 additional HCELS were excluded from analyses)	Absence of lifetime (past and present) psychiatric disorders (assessed by SCID)	No validated measure, participants asked 2 yes/no questions: "Any history of physical and/or sexual abuse over your lifetime?", if yes, this was followed up with: "Did the abuse take place before 15 years of age?"	ELS limited to physical and/or sexual abuse	Before age 15
Mikolas et al., 2019	67 HC	Not specified	CTQ - categorised as ELS participant if met cut-offs for at least 1 subscale of CTQ	CTQ subtypes listed but not used in analyses.	≤18 years old
Teicher et al., 2012	NA	NA	CTQ and 100-item semi structured Traumatic Antecedents Questionnaire (TAQ; Vanderkolk et al., 1991)	CTQ subtypes listed but not used in analyses.	≤18 years old

Tannous et al., 2020	46 HC	No current or past history of Axis I disorders as determined by SCID	CTQ	CTQ subtypes listed but not used in analyses.	≤18 years old

Table 2b. Abbreviations: HC: healthy control; ELS: early life stress; HCELS: healthy control with ELS; CTQ: Childhood Trauma Questionnaire; SA: sexual abuse; PA: physical abuse; EA: emotional abuse; PN: physical neglect; EN: emotional neglect.

All but one study (Croy et al., 2013) used Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, American Psychiatric Association, 1994) criteria to diagnose MDD, 13 of these using the Structured Clinical Interview for DSM-IV (SCID, First et al., 2002), 4 using the Mini International Neuropsychiatric Interview (MINI; Lecrubier et al., 1997), one using the Composite International Diagnostic Interview (CIDI; World Health Organisation, 1994), and one multisite study (Frodl et al., 2017) using either the MINI or CIDI depending on study site. All but two studies included only participants with a current major depressive episode (MDE) at the time of the study in the MDD group, while one study limited MDD to within the past year (but not necessarily current MDE at time of data collection, Gerritsen et al., 2015) and another did not distinguish between past or present MDD (Teicher et al., 2012). 10 studies included MDD participants on psychotropic medication, 7 studies included exclusively unmedicated participants, and 3 studies did not report medication status. Of the 10 studies with medicated MDD participants, 6 controlled for antidepressant use in their analyses (Aghamohammadi-Sereshki et al., 2021; Bermingham et al., 2012; Colle et al., 2017; Frodl et al., 2017; Gerritsen et al., 2015; Opel et al., 2014) while four did not (Chaney et al., 2014; Frodl et al., 2010; Mikolas et al., 2019; Tannous et al., 2020). For additional details on comorbidities and MDD age of onset, see **Table 2c**.

Table 2c. MDD

Study (Authors)	MDD group (#)	MDD evaluation	MDD on anti-depressants (%)	MDD comorbidities	MDD age of onset

Bermingham et al., 2012	44 MDD (N=13 T-carriers)	Clinical diagnosis based on DSM-IV. Confirmed by independent psychiatrist using SCID interview	70.45%	None (all exclusionary)	28.2 (15.4) / 24.1 (12.5) T-carriers/ not carriers
Carballedo et al., 2013	62 MDD	Clinical diagnosis based on DSM-IV. Confirmed by independent psychiatrist using SCID interview	Not reported	None (all exclusionary)	Not reported
Chaney et al., 2014	37 MDD (20 with ELS)	SCID for DSM-IV for psychiatric diseases (SCID-I) and for personality assessment (SCID-II).	64.86%	None (all exclusionary)	MDD: 26.3 (10.9), MDDELS: 22 (12.1)
Frodl et al., 2017	958 MDD	Differed by sample, 3x CIDI interview, 4x SCID interview, 1 x not reported, 1 study only recruited HC.	All but one sample included participants on antidepressants, which ranged from 17.3% - 86.7%	Differed by sample: 4 excluded any other comorbid Axis-I disorder, 3 permitted anxiety disorders, 1 had no exclusion criteria.	Differed by sample, ranged from 14.4 (2.9) - 38.3 (13.2)
Colle et al., 2017	63 MDD (MDDELS: 28, MDD only: 35)	MINI (DSM-IV) and ≥18 on HDRS	54%	Anxiety & personality disorders not exclusionary, all others exclusionary.	37.5 (±15.6)
Croy et al., 2013	23 MDD (15 with ELS, 8 without)	In-patient MDD	Not reported	None exclusionary	Not reported
Frodl et al., 2010	43 MDD	SCID for DSM-IV	86%	None (all exclusionary)	40.1 (11.2)
Gerritsen et al., 2015	183 MDD (139 NESDA, 44 SMART)	CIDI (DSM-IV), had to meet criteria in past year	26% NESDA, 7% SMART	None exclusionary	Not reported
Hassel et al., 2017	80 MDD	MINI to assess current MDE, and <23 on Montgomery-Åsberg Depression Rating Scale (MADRS).	Not reported	None (all exclusionary)	Not reported
Lu et al., 2019	30 MDD (MDD ELS, n=16; MDD, n=14)	SCID for DSM-IV (had to meet for current MDD and current MDE), evaluated by 2 psychiatrists	Unmedicated for at least 2 weeks	None (all exclusionary)	Not reported
Lu et al., 2018	82 MDD (MDDELS, n=52; MDD, n=31)	SCID for DSM-IV (had to meet for current MDD and current MDE), evaluated by 2 psychiatrists, and HAMD score of at least 20.	Unmedicated for at least 2 weeks	None (all exclusionary)	MDDELS: 30.1 (8.5), MDD: 31.8 (7.5)
Opel et al., 2014	85 MDD	SCID for DSM-IV (current MDE) and >18 score on BDI	95.29%	Not reported	Not reported
Saleh et al., 2017	51 MDD	MINI (version 5.0) diagnosis of current MDD (based on DSM-IV) & psychiatric interview, age of	Unmedicated for at least 4 weeks	Only anxiety disorders occurring within MDE were	before age 35

		onset before age 35, MADRS score of >14		permitted, all other disorders were exclusionary.	
Aghamohammadi-Sereshki et al., 2021	35 MDD	Anxiety Disorders Interview Schedule for DSM-IV-Lifetime Version	71.40%	Majority of comorbidities exclusionary, details not reported.	Not reported
Vythilingam et al., 2002	32 MDD (21 MDD ELS, 11 MDD)	Met for current MDD as assessed by structured clinical interview for DSM-IV	Unmedicated	Bipolar disorder, schizophrenia and schizoaffective disorder were exclusionary, other comorbidities (including anxiety disorders, PTSD, eating disorders and substance & alcohol abuse/dependence) were permitted.	Not reported
Yang et al., 2017	84 MDD (MDDELS: 41, MDD: 43)	SCID for DSM-IV (current MDE) and HRSD score >17	Unmedicated	Only secondary anxiety disorders permitted, all other disorders were exclusionary.	Not reported
Yuan et al., 2020	41 MDD (17 MDDELS, 24 MDD)	SCID for DMS-IV (current MDE) AND HRSD score >15	Unmedicated	7 comorbid social anxiety disorder, 3 comorbid panic disorder, 1 specific phobia, 1 agoraphobia.	Not reported
Mikolas et al., 2019	85 MDD	DSM-IV criteria applied by consultant, then confirmed by independent researcher using SCID.	76.47%	None (all exclusionary)	Not reported
Teicher et al., 2012	49 MDD (25% of sample)	SCID for DSM-IV axis I and II psychiatric disorders	Unmedicated	Not reported	Not reported
Tannous et al., 2020	71	SCID for DSM-IV axis I psychiatric disorders	22.50%	None (all exclusionary)	Not reported

Table 2c. Abbreviations: MDD: Major Depressive Disorder; MDE: Major Depressive Episode; MDDELS: participant with MDD and ELS; SCID: Structured Clinical Interview for DSM-IV; MINI: Mini International Neuropsychiatric Interview; CIDI: Composite International Diagnostic Interview; HRSD: Hamilton Depression Rating Scale.

In terms of neuroimaging acquisition, 3 studies used a 1.5T scanner, 9 used 3T, 1 used 4.7T, 4 used both 1.5T and 3T (due to multiple study sites), 1 used 7T, and 2 did not report MRI field strength. 6 studies used manual tracing for cortical segmentation and calculation of GMV, 7 used some version of FreeSurfer, 4 used the VBM8 toolbox in SPM8, 1 used a combination of Freesurfer and manual tracing (depending on study site), 1 used both FSL

FIRST and VBM8, and one used automated segmentation using SACHA software. 3 studies took a whole brain approach, 2 studies conducted both whole brain and ROI analyses, and 15 conducted exclusively ROI based analyses (of which 14 included the hippocampus and/or hippocampal subfields; see **Table 2d** for details). Studies differed in statistical approaches and whether confounding variables were identified and controlled for (such as total intracranial volume (ICV), and, depending on sample characteristics, medication status, age, gender, and others; see **Table 2d** for details).

Table 2d. MRI acquisition and data analysis/statistics

Study (Authors)	MRI field strength	Software used for analysis	ROI or whole brain	Statistics: confounding variables controlled for in analyses
Bermingham et al., 2012	3T	Manual tracing of bilateral hippocampus (whole, head, body, tail) using software BRAINS2. Tracers were blind to participant group assignment.	ROI (hippocampus whole, body, head, tail)	Controlled for age, gender, medication status, total ICV (all covariates in ANCOVA). Post hoc ANCOVA carried out for any significant interactions, using Bonferroni corrections.
Carballedo et al., 2013	1.5T (Munich), 3T (Dublin)	Software package ANALYZE was used for image processing. Manual tracing of bilateral hippocampus (whole, head, body, tail) using software BRAINS2. Tracers were blind to participant group assignment.	ROI (hippocampus whole, body, head, tail)	Controlled for age and gender (covariates in ANCOVA)
Chaney et al., 2014	3T	SPM8 (VBM8 toolbox used for segmentation)	Whole brain & ROI (hippocampus, anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), dorsomedial prefrontal cortex (DMPFC) and orbitofrontal cortex (OFC)	Controlled for age and gender (covariates in ANCOVA). Corrected for multiple comparisons. Did not correct for ICV or medication status.
Frodl et al., 2017	6 x 3T, 3 x 1.5T	FreeSurfer (version 5.0 and higher) semi-automated processing and segmentation pipeline	7 ROI (nucleus accumbens, amygdala, caudate nucleus, hippocampus,	Controlled for total ICV, age, gender & imaging site. Also did additional analysis comparing with and without antidepressant medication. Corrected for multiple comparisons.

			pallidum, putamen, thalamus)	
Colle et al., 2017	1.5T and 3T	SACHA software used for automated segmentation of hippocampus . SPM5 used to assess total brain volume	ROI: hippocampus	Controlled for total brain volume, age, gender, medication and other factors)
Croy et al., 2013	1.5T	AMIRA 3D visualisation and modelling system (Visage Imaging) used to outline left and right OB, manual tracing used to ascertain volume.	ROI: olfactory bulb	Controlled for depression (BDI) but not ICV or any other variable
Frodl et al., 2010	1.5T	Software package ANALYZE was used for image processing. Manual tracing (blind to group status)of bilateral hippocampus using software BRAINS2. Whole brain VBM analysis performed using VBM5 toolbox in SPM5.	ROI: (hippocampus) and whole brain	Controlled for age, gender and total brain volume as covariates
Gerritsen et al., 2015	1.5T (SMART), 3T (NESDA)	SMART: manual outlining of hippocampus (2 investigators, blind to group status); NESDA: FreeSurfer to calculate ICV and hippocampal subfield volumes	ROI: hippocampus (proper, subiculum, fimbria, alveus, dentate gyrus)	Controlled for age, sex, education, blood pressure, smoking status, alcohol intake, diabetes, antidepressant use, white matter volume, MMSE scores (Mini Mental State Examination)
Hassel et al., 2017	3T	Freesurfer used for processing, segmentation of hippocampus, and calculation of total intracranial volume(ICV)	ROI: hippocampus	Controlled for variables including age and total ICV according to standardised ENIGMA protocol
Lu et al., 2019	3.0T	VBM8 toolbox in SPM8	Whole brain	Controlled for age, gender, educational level but not total ICV
Lu et al., 2018	3.0T	VBM8 toolbox in SPM8	Whole brain	Controlled for age, gender, educational level but not total ICV
Opel et al., 2014	not reported	FSL-FIRST (automated segmentation) and VBM8 (SPM8)	ROI: hippocampus	All analyses controlled for ICV (did not control for medication but no effect of medication on GMV or interaction with ELS)
Saleh et al., 2017	not reported	FreeSurfer (version 5.1) used to calculate GMV of ROIs and ICV	ROIs: ACC, OFC, amygdala, hippocampus, caudate	Controlled for diagnosis, age, sex, and ICV
Aghamohammadi-Sereshki et al., 2021	4.7T	Manual delineation of amygdala subnuclei and hippocampal subfields (cornu ammonis, subiculum, dentate gyrus)	ROI: amygdala and hippocampus subfields	Normalised raw volumetric measurements to ICV; compared medicated to unmedicated MDD; controlled for ICV in all analyses (included as covariate in ANCOVA)
Vythilingam et al., 2002	1.5T	Software package ANALYZE was used for image processing. Manual tracing (blind to group status)of hippocampus. Hippocampus was further	ROI: hippocampus (whole hippocampus, body, and head).	Controlled for age, race, PTSD, alcohol/substance dependence, and whole brain volume

		segmented into head, body, and tail.		
Yang et al., 2017	3T	VBM8 toolbox in SPM8	Whole brain	Controlled for age, total GMV, educational level
Yuan et al., 2020	3T	Freesurfer (version 6.0) used for hippocampal subfield segmentation	ROI: hippocampus whole & subfields	Controlled for age, sex, sequence, total intracranial volume (TIV), abuse group or diagnostic group
Mikolas et al., 2019	3T	Freesurfer: version 5.3 used for whole brain segmentation and version 6 dev. used for hippocampal subfield segmentation	ROI: hippocampus whole & subfields	Controlled for study site, sex, brain volume
Teicher et al., 2012	3T	Freesurfer (version not specified) used for cortical reconstruction and volumetric segmentation, including hippocampal subfields	ROI: hippocampus whole & subfields	Controlled for subcortical GMV, age, gender and socioeconomic status
Tannous et al., 2020	7T	Freesurfer (version 6.0) used for hippocampal subfield segmentation	ROI: hippocampus whole & subfields (selected 8/12 subfields a prior)	Controlled for age, gender, total intracranial volume

Table 2d. Abbreviations: ICV: intracranial volume; ROI: region of interest; GMV: grey matter volume; ANCOVA: analysis of covariance; SPM: Statistical Parametric Mapping.

3.4. Main findings

Table 3 summarizes the main findings and conclusions of each study and also lists possible limitations/risks of bias for each study.

3.4.1 Hippocampus

The most commonly investigated region was the hippocampus (14/20 studies included the hippocampus as an ROI). Combined with 5 further studies that conducted whole brain analyses, this results in 19 studies (all but Croy et al., 2013 which focused exclusively on the olfactory bulb) that can address whether hippocampal GMV changes are associated with ELS and/or MDD. While study design and variables slightly vary for each study, a general trend

of ELS driving hippocampal volume reduction irrespective of diagnosis can be clearly observed, and in studies within MDD alone, individuals with MDD and ELS appear to have lower hippocampal volume than those without ELS exposure. Specifically, eight studies found that ELS was driving hippocampal volume reductions, irrespective of MDD diagnosis (Bermingham et al., 2012; Carballedo et al., 2013; Chaney et al., 2014; Lu et al., 2018; Opel et al., 2014; Teicher et al., 2012; Vythilingam et al., 2002; Yuan et al., 2020). Furthermore, six studies found that ELS was associated with decreased hippocampal volume in MDD (i.e. MDDELS had significantly smaller hippocampal GMV than MDD only; Aghamohammadi-Sereshki et al., 2021; Colle et al., 2017; Frodl et al., 2010; Gerritsen et al., 2015; Hassel et al., 2017; Saleh et al., 2017). Of the remaining 5 studies, 4 reported null results – 2 were whole brain analyses that did not identify hippocampal GMV as an area affected by MDD or ELS (Lu et al.. 2019; Yang et al., 2017) and two were ROI studies that similarly found no effect of MDD or ELS on hippocampal GMV (Frodl; et al., 2017; Tannous et al., 2020). Interestingly, one study that reported significant effects of ELS on hippocampal GMV (irrespective of MDD) in their ROI analyses, did not find any significant areas (including the hippocampus) in their whole brain analyses after correcting for family wise error (FWE; Chaney et al., 2014). Finally, only a single study reported increased hippocampal GMV in MDDELS compared to MDD (Mikolas et al., 2019).

3.4.2 Hippocampal subfields

Five included studies investigated hippocampal subfields as part of their ROI analyses. The most replicated finding was significantly reduced cornu ammonis (CA1-3) GMV associated with ELS – due to study design, this was limited to MDD in two studies as they did not include a HCELS group or did not measure ELS in HC (Aghamohammadi-Sereshki et al., 2021;

Yuan et al., 2020), while the third study reported decreased CA2-3 GMV in ELS, irrespective of MDD (Teicher et al., 2012). The dentate gyrus and subiculum was further implicated by two studies, again showing decreased GMV associated with ELS (in MDD: right DG and left subiculum, Aghamohammadi-Sereshki et al., 2021; and irrespective of diagnosis: bilateral CA4-DG, left presubiculum, and subiculum, Teicher et al., 2012). Echoing their findings from the whole hippocampus reported above, a single study found the opposite effect of ELS increasing GMV in cornu ammonis and DG subfields in depressed participants while MDD with no history of ELS showed decreased GMV in these areas compared to HC (Mikolas et al., 2019). Finally, one study reported no effect of MDD or ELS on hippocampal subfields, just as they had reported for the hippocampus whole (Tannous et al., 2020).

3.4.3 Other brain regions

Eight studies revealed additional brain regions associated with ELS and/or MDD in either ROI (5 studies) or whole brain analyses (3 studies). Of these, only three described areas that were associated with increased GMV in ELS, including the dorsomedial prefrontal cortex (DMFC; Chaney et al., 2014; Lu et al., 2018), dorsolateral prefrontal cortex (DLPFC; Yang et al., 2017), left cerebellum anterior lobe (Yang et al., 2017) and the supplemental motor area (SMA; Lu et al., 2018). However, it should be noted that others found decreased left DLPFC in ELS, irrespective of diagnosis (Lu et al., 2018). Other brain areas identified were all associated with a decrease in GMV in ELS, including the caudate (Frodl et al., 2017; Saleh et al., 2017; both irrespective of MDD diagnosis), olfactory bulb (Croy et al., 2013; MDDELS < MDD, no HC group), orbitofrontal cortex (OFC; Saleh et al., 2017; Yang et al., 2017; both irrespective of MDD diagnosis), left posterior cingulate cortex (PCC) and left inferior occipital gyrus (Yang et al., 2017; both irrespective of MDD diagnosis), and the right

amygdala (Aghamohammadi-Sereshki et al., 2021; MDDELS<MDD, HC group without ELS measure). Finally, only one study reported any main effects of MDD that were independent of ELS, namely a decrease in GMV in the left medial prefrontal cortex (MPFC) and left superior parietal lobule (Yang et al., 2017). Finally, it is important to note that two whole brain analysis failed to identify a single brain region associated with GMV changes in MDD and/or ELS (Chaney et al., 2014; Tannous et al., 2020).

Table 3. Main findings and limitations/risks of bias

Study (Authors)	Main findings	Conclusion	Limitations, risks of bias
Bermingham et al., 2012	No main effect of diagnosis, ELS or BICC1 (bicaudal C homolog 1) gene, on hippocampal volume. Right hippocampal bodies of MDD and HC without ELS but with T-allele of BICC1 were significantly larger than those without T-allele. But, MDD with T-allele & history of ELS had significantly smaller hippocampal head volumes compared to those without ELS.	The minor T-allele of BICCI may have a protective role against hippocampal GMV reduction in MDD but this effect is removed in ELS: ELS associated with hippocampal GMV reduction in MDD & HC, even when the otherwise protective T-allele is present.	Cut-off for ELS not well-defined. No age range given for ELS. Female/male ratios differed significantly between T carrier and not carrier group (but gender included as covariate). Majority on medication (though included as covariate in analyses). HC not assessed for psychiatric disorders. Low levels of ELS in MDD and even lower in HC as not specifically recruited any ELS groups.
Carballedo et al., 2013	Main effect of MDD (smaller bilateral hippocampal volume) compared to HC, but not when controlling for ELS. Significant effect of ELS: MDD and HC with ELS & met allele showed significantly smaller hippocampal volume than those without ELS.	In individuals with met allele, ELS leads to decreased hippocampal volume, irrespective of diagnosis.	HC not assessed for psychiatric disorders. Statistics did not control for total ICV. ELS participants were significantly older than non ELS. Low levels of ELS in MDD and even lower in HC as not specifically recruited any ELS groups. Do not report whether medicated or not.
Chaney et al., 2014	Whole brain: no areas remained significant after corrections. ROI: significant main effect of ELS - smaller bilateral hippocampal volumes in MDD and HC with ELS than without. Also significantly larger right DMPFC in ELS, irrespective of diagnosis. MDDELS had significantly smaller hippocampal volume than MDD only. No significant difference in hippocampal volume in MDD vs. HC (when not controlling for ELS).	ELS is associated with GMV changes in bilateral hippocampus (decrease) and right DMPFC (increase) independent of sex, age, and depression.	MDDELS had significantly higher levels of CTQ scores than HCELS. Statistics did not control for total ICV or medication status (despite majority medicated). Small sample size once broken down into four groups, especially HCELS (n=10).

Frodl et al., 2017	Increased ELS exposure (total CTQ score) was associated with significantly smaller bilateral caudate volumes in females only, irrespective of MDD (not significant in males). This finding held even when distinguishing between MDD with versus without antidepressant medication. All subcategories of ELS (EA, PA, SA, EN, PN) had a significant negative correlation with caudate volume, in particular EN & PN, again independent from MDD, age, ICV and imaging site. There was no significant interaction between MDD & ELS or main effect of MDD alone on any ROIs.	ELS associated with volumetric reduction of caudate in females, irrespective of diagnosis.	Different MRI strengths used, low levels of childhood trauma (as not specifically recruited high ELS groups), no information on age of trauma exposure.
Colle et al., 2017	MDDELS had significantly smaller bilateral hippocampal GMV than MDD only, even when controlling for age, total brain volume, age of MDD onset, antidepressant medication, history of suicide attempt, MRI acquisition method, and severity of MDD (HDRS score). This association was only significant for men, not women.	ELS is associated with decrease hippocampal volume in males (but not females) with MDD, even when controlling for various other potential confounding factors.	No healthy control group (i.e. without MDD, neither with/without ELS), MDD were in-patients in psychiatric setting (may not be generalizable to whole MDD population).
Croy et al., 2013	Olfactory bulb (OB) volume in MDDELS was significantly decreased compared to MDD only (20% reduction in volume).		Very small sample size, in-patient MDD, did not control for ICV, no HC group, manual tracing of OB only rated reliable for 6 MDD and 14 MDDELS (20 participants/23 included)
Frodl et al., 2010	No significant difference between MDD and HC in hippocampal GMV. MDD with higher levels of emotional neglect had significantly smaller left hippocampal GMV than those with lower levels of emotional neglect (also significant in right hippocampus for male MDD). There was also a significant main effect of physical neglect: male MDD with physical neglect had significantly smaller hippocampal GMV compared to those without physical neglect. In HC only, physical neglect was significantly associated with decreased prefrontal GMV (but not in MDD). Both emotional neglect and brain structural abnormalities independently predicted cumulative illness duration.	PN and EN associated with reductions in hippocampal volume in MDD	Emotional and physical neglect significantly higher in MDD versus HC, other ELS subtypes not significantly different (all rather low levels in both groups), did not control for medication use.
Gerritsen et al., 2015	MDDELS (but not MDD only) associated with smaller hippocampal volume compared to HC. MDDELS associated with a 3% (NESDA) and 6% (SMART) smaller volume compared to both MDD alone and HC. But no main effect of ELS alone in either sample.	ELS appears to moderate the relationship between MDD and hippocampal volume - specifically individuals with both MDD and ELS have decreased hippocampal volume	Nemesis trauma interview asks only 1 yes/no question for each type of trauma; MDD defined as present in past 12months (so possible not all in current MDE), different field strength and different MRI processing/analysis methods between two samples. Much older sample, especially in the

		compared to those with only MDD or HC.	larger SMART study (mean age 61.5)
Hassel et al., 2017	MDD patients overall did not significantly differ from HC in hippocampal GMV. MDDELS had significantly decreased left and total hippocampal volume compared to MDD only.	MDD did not differ from HC in hippocampal GMV, but MDD with ELS had significantly smaller left and total hippocampal volume than MDD without ELS.	MDD had significantly more ELS than HC. No information given about medication status, no information about age/types/severity of ELS, no mean ages given for either group.
Lu et al., 2019	Significant main effect of ELS: smaller left dorsolateral prefrontal cortex (DLPFC) in ELS, irrespective of diagnosis.	Authors conclude that decreased left DLPFC GMV may be a function of ELS rather than MDD.	MDDELS had significantly higher levels of ELS (CTQ scores) than HCELS. Relatively small sample size once broken down into 4 groups, especially in the two MDD groups.
Lu et al., 2018	ELS associated with decreased parahippocampal gyrus/hippocampus GMV and left cuneus GMV irrespective of diagnosis. Opposite finding of increased GMV in ELS was found in the DMPFC and SMA. Main effects of diagnosis: decreased GMV in MDD in some areas (right medial temporal gyrus) but increased GMV in several other areas (left caudate, left posterior lobe of the cerebellum, right inferior temporal gyrus, and left lingual gyrus).	Mixed findings: ELS associated with decreased GMV in hippocampus and left cuneus GMV, irrespective of diagnosis, but also increases in GMV in other areas. Similarly MDD associated with increased/decreased volume in various areas compared to HC.	Relatively low levels of ELS across all groups. Very low levels of emotional, physical and sexual abuse, most common form of ELS was emotional neglect. Despite larger group numbers, no analyses were performed for different types of abuse. MDDELS participants had significantly greater severity of current depression (as measured by HRSD scores) than MDD only group.
Opel et al., 2014	MDD had significantly lower hippocampal volume than ELS, however, this association did not hold when ELS was regressed out. ELS was significantly associated with decreased hippocampal volume in both MDD and HC.	Results suggest that hippocampal GMV reductions in MDD may be explained by ELS.	Vast majority of MDD on current antidepressant medication (although there was no effect of medication on GMV); all MDD were inpatients so possibly not generalizable to general MDD population.
Saleh et al., 2017	No significant differences in GMV in any a priori defined ROIs between MDD and HC. Also no main effect of total ELSQ score on these ROIs. However, 'predictive' ELS was significantly associated with decreased GMV in the left lateral OFC and right caudate, irrespective of MDD. Furthermore, there was a significant interaction between diagnosis and predictive ELS in the hippocampus: predictive ELS was significantly associated with decreased left (and trend for right) hippocampal GMV in MDD but not HC. Additional exploratory analyses of duration of exposure revealed significant findings only for sexual abuse, in which longer exposure to CSA was	Exposure to at least one of 3 subtypes of ELS (EA, SA, & severe family conflict) were significantly associated with decreased GMV of the left lateral OFC and right caudate, irrespective of MDD diagnosis. Predictive ELS was also significantly associated with decreased GMV in the left hippocampus in MDD only.	Significant differences in demographics/clinical variables between MDD and HC (MDD group older, more medical illnesses, higher ELSQ scores). Only 10 HC (18.87%) but 62.7% of MDD reported at least one form of 'predictive' ELS (emotional trauma, sexual abuse, severe family conflict).

	significantly associated with decreased caudate GMV.		
Aghamohammadi-Sereshki et al., 2021	In MDD, total CTQ score significantly correlated with reductions in GMV of the whole hippocampus, hippocampal head (but not body or tail). Within the hippocampal head, the bilateral cornu ammonis (CA1-3), right dentate gyrus (DG), and left subiculum were significantly negatively correlated with total CTQ score. Exploratory ELS subtype analyses found significant negative correlations between emotional and physical abuse and emotional and physical neglect with hippocampal head and cornu ammonis GMV. There were no MDD vs HC group differences in total or amygdala subnuclei volumes. Total CTQ was negatively correlated with total right amygdala GMV in MDD.	Findings suggest ELS exposure in MDD may lead to GMV reductions in subfields of the hippocampal head and right amygdala.	CTQ only assessed in MDD (not in HC) so analyses on ELS limited to MDD group. Main effects of MDD analysed only for amygdala subnuclei and not for hippocampus (for which only childhood adversity in MDD was analysed). Medicated versus unmedicated analyses only performed for amygdala and not hippocampus.
Vythilingam et al., 2002	MDDELS had significantly smaller left hippocampal GMV than MDD only (18% reduction) and than HC (15% reduction). Right and left hippocampal volumes were similar in MDD only and HC groups (no significant differences). Right hippocampal volume was not significantly different between all groups. After correcting for multiple comparisons, there was no significant correlation between hippocampal GMV and total score on the Early Trauma Inventory or physical, sexual, and emotional subscale scores.	Reductions in hippocampal volume were exclusive to women with MDD and a history of childhood physical and/or sexual abuse (and were not observed in individuals with MDD without a history of ELS).	Do not include HCELS group as a control group; MDD groups significantly older than HC group, high percentage of comorbidities in MDD groups
Yang et al., 2017	Main effect of MDD (decreased GMV in left MPFC and left superior parietal lobule), independent of age, total GMV, education level and total CTQ score. Main effect of ELS on left posterior cingulate cortex, right medial OFC, & left inferior occipital gyrus (decrease GMV), right DLPFC and left cerebellum anterior lobe (increased GMV), independent of MDD diagnosis.	Authors conclude that ELS is associated with structural abnormalities in frontolimbic regions that are often exclusively attributed to MDD diagnosis.	Much smaller group of HCELS compared to other three groups. The most common types of ELS were physical and emotional neglect, there were significantly fewer instances of sexual, physical and emotional abuse.
Yuan et al., 2020	No main effect of diagnosis on hippocampal total volume or subfields. When divided into MDDELS and MDD only, found no differences between MDD only and HC in whole/subfield hippocampal GMV, but there was a significant diagnosis x region interaction for MDDELS versus HC. Comparison between MDDELS and MDD showed significantly smaller left CA1 volume in MDDELS versus MDD.	Authors conclude that the left CA1 may be particularly sensitive to ELS, independent of MDD.	No HCELS group (only 2 HC met for ELS so were excluded from analyses). No validated measure of ELS (simple yes/no question, risk of bias, no information on other subtypes, severity/age of exposure). MDD and HC were mostly matched on demographics, but the MDD group was significantly lower educated than HC (p=0.017).

Mikolas et al., 2019	Overall reductions in hippocampal volumes in MDD versus HC in the cornu ammonis (CA1, CA3, CA4) and DG subfields while, contrary to predictions, MDDELS showed increased volumes in CA1, CA3 and DG.	MDD leads to volume reductions in several hippocampal subfields compared to HC; MDDELS showed the opposite pattern.	Majority of participants on antidepressant treatment (and not controlled for), very low levels of ELS, especially in HC group. Pooled data from 2 study sites with different scanning parameters and found that results differed slightly between two sites.
Teicher et al., 2012	Significant reductions in hippocampal subfield GMV (most pronounced in left CA2-3 and CA4-DG subfields, but also in the left presubiculum and subiculum) were driven by ELS and were not mediated by lifetime histories of MDD or PTSD	ELS appears to significantly decrease GMV in various hippocampal subfields, independent of MDD.	Relatively low levels of ELS overall (since it was a community sample), no information of timing of MDD (lifetime history only), as expected, ELS was strongly associated with MDD and significantly lower levels of ELS were experienced by individuals without psychiatric diagnoses
Tannous et al., 2020	No differences in global hippocampal volume or subfields in individuals with MDD or ELS compared to HC.	No effect of ELS or MDD on global hippocampal volume and subfields	Significant proportion of MDD participants (23%) on antidepressant medication but not controlled for in analyses. Relatively low levels of ELS

Table 3. Abbreviations: ICV: intracranial volume; ROI: region of interest; GMV: grey matter volume; ANCOVA: analysis of covariance; HRSD: Hamilton Depression Rating Scale; CTQ: Childhood Trauma Questionnaire; SA: sexual abuse; PA: physical abuse; EA: emotional abuse; PN: physical neglect; EN: emotional neglect; MDD: Major Depressive Disorder; HC: healthy control; ELS: early life stress; MDDELS: participant with MDD and ELS; HCELS: healthy control with ELS; DLPFC: dorsolateral prefrontal cortex; DMPFC: dorsomedial prefrontal cortex; MPFC: medial prefrontal cortex; SMA: supplemental motor area; OFC: orbitofrontal cortex; CA1-3: cornu ammonis; DG: dentate gyrus.

4. Discussion

4.1. Summary and discussion of main findings

The systematic review identified 5129 records, of which 20 met predefined eligibility criteria and were included in the final analyses. All studies were assessed for risk of bias by two independent reviewers (FG and KL) and met criteria for low risk of bias overall (Melo et al., 2018). While there was considerable heterogeneity in study design and method that prevented a meta-analysis, several key findings emerged in the systematic review. All but one study included the hippocampus in their volumetric analyses (either as an ROI [14 studies] and/or as part of whole brain analyses [5 studies]). The vast majority of studies confirmed our hypothesis that ELS appears to drive hippocampal volume reductions,

independent of MDD diagnosis (reported by 8 studies; Bermingham et al., 2012; Carballedo et al., 2013; Chaney et al., 2014; Lu et al., 2018; Opel et al., 2014; Teicher et al., 2012; Vythilingam et al., 2002; Yuan et al., 2020), and an additional 6 studies reported significantly smaller hippocampal GMV in depressed individuals (but not HC) with a history of ELS than those without ELS (Aghamohammadi-Sereshki et al., 2021; Colle et al., 2017; Frodl et al., 2010; Gerritsen et al., 2015; Hassel et al., 2017; Saleh et al., 2017). Notably, two of these studies did not make comparisons to healthy controls since they either did not have a HC group (Colle et al., 2017) or did not measure ELS in their HC group (Aghamohammadi-Sereshki et al., 2021), making it impossible to ascertain whether ELS drives hippocampal GMV reductions irrespective of MDD. The other 4 studies that found ELS was associated with decreased hippocampal GMV in MDD did not find this association in HC (Frodl et al., 2010; Gerritsen et al., 2015; Hassel et al., 2017; Saleh et al., 2017), however, ELS levels in the healthy control groups were significantly lower than the MDD groups (a common problem in the vast majority of studies). Furthermore, Hassel et al., 2017 found no difference in hippocampal GMV between MDD and HC overall, only MDDELS had significantly smaller GMV compared to both MDD and HC, suggesting that ELS may indeed be the driving factor of volumetric reductions.

Of the 5 remaining studies that investigated the hippocampus, 4 reported null results (i.e. no effect of MDD or ELS on hippocampal GMV (Frodl; et al., 2017; Lu et al.. 2019; Tannous et al., 2020 Yang et al., 2017) while only a single study reported increased hippocampal GMV in MDDELS versus MDD only (Mikolas et al., 2019). Interestingly, only 1 whole brain analyses picked up hippocampus GMV differences in MDD and/or ELS (Frodl et al., 2010) while the other whole brain analyses reported either no significant regions (null results) or

identified other brain areas that survived family wise error (FWE) correction. One explanation may be that hippocampal GMV changes are too subtle to be picked out at whole brain level and survive stringent FWE corrections and are better identified in ROI analyses (e.g. Chaney et al., 2014 found evidence of significantly decreased hippocampal volume in ELS, irrespective of MDD, in their ROI analyses, but the hippocampus did not survive FWE correction in their whole brain analyses). Conversely, only 3 ROI studies did not report ELS driving volumetric reductions in the hippocampus (either irrespective or MDD or within MDD), including two with null results (Frodl et al., 2017; Tannous et al., 2020) and one with opposite findings (MDDELS>MDD hippocampal GMV, Mikolas et al., 2019). While one of these studies controlled for medication use (and found this a significant factor affecting hippocampal GMV when not controlled for, Frodl et al., 2017), the other two included high levels of antidepressant use within their sample and did not control for this in their statistical analyses (Mikolas et al., 2019; Tannous et al., 2020). Given that antidepressant medication has been shown to change GMV in various regions including the hippocampus (Boldrini et al., 2009; Frodl et al., 2008), this may affect overall results.

Another key replicated finding that emerged in the systematic review was that ELS was associated with a significant GMV reduction of the cornu ammonis (CA1-3), dentate gyrus (DG) and subiculum subfields of the hippocampus (Aghamohammadi-Sereshki et al., 2021; Teicher et al., 2012; Yuan et al., 2020). While two of these studies did not include a HCELS group or measure ELS in their HC group and hence were limited to making conclusions within depressed groups, Teicher and colleagues (2012) reported reductions in GMV in ELS irrespective of MDD diagnosis. These results are particularly of note given that studies of depression have consistently implicated GMV reduction in the CA1-3 (Huang et al., 2013;

Lim et al., 2012; Lindqvist et al., 2014) and DG (Huang et al., 2013; Treadway et al., 2015; Lindqvist et al., 2014; Han et al., 2016), however, these studies did not measure or control for ELS in their samples. It is hence possible, given the high levels of ELS commonly found in MDD (Williams et al., 2016; Xie et al., 2018), that unmeasured ELS in these studies may at least partially, if not perhaps fully, account for GMV reductions observed in CA1-3 and DG subfields of the hippocampus.

Eight of the 20 reviewed studies included other brain areas in their analysis (either ROI or in whole brain analyses). Given the heterogeneity of brain regions identified or selected as ROIs, overall conclusions are more difficult to ascertain. However, a general trend emerged that ELS was associated with GMV decrease in several areas, including in the caudate (Frodl et al., 2017; Saleh et al., 2017) orbitofrontal cortex (OFC; Saleh et al., 2017; Yang et al., 2017), left posterior cingulate cortex (PCC) and left inferior occipital gyrus (Yang et al., 2017), which were all irrespective of MDD diagnosis. Only a single study reported any main effects of MDD independent of ELS (decreased MPFC GMV; Yang et al., 2017), suggesting that ELS may again be the driving factor of GMV changes in other brain areas in addition to the hippocampus, rather than MDD.

4.2. Limitations

This systematic review sought to limit bias by following PRISMA guidelines, including a thorough and systematic search of two grey literature databases (including unpublished manuscripts, such as dissertations and abstracts), using well defined eligibility criteria, pre-registering the systematic review protocol with PROSPERO, and applying a validated risk of bias assessment tool for each included study. One of the key limitations in the conclusions

drawn is that these are based on a systematic review and narrative synthesis, as a more quantitative approach in the form of a meta-analysis was not possible in this case due to heterogeneity of measures and study design of included studies.

4.2.1 ELS measure and potential memory bias

In terms of included studies, several limitations must be addressed. Firstly, all studies assessed ELS retroactively, a common issue in ELS research, meaning there is potential for memory bias. A prospective study analysing memory bias of traumatic events found that people's memories were very accurate for potentially traumatic events (in particular to other, non-traumatic life events), however, this was measured in college students over a limited 4-year period (Lalande & Bonanno, 2011). Another study comparing retroactively reported childhood traumatic events in adults at two time points (over 10 years apart) and found that 39% of individuals were inconsistent in reporting ELS, an effect that appeared heightened in individuals who went on to develop mental health issues (including depression) and work and chronic stress (Colman et al., 2015). Furthermore, a recent systematic review found relatively poor agreement between retroactive and prospective measures of ELS (Baldwin et al., 2019). Interestingly, this agreement appears to improve significantly when interviews rather than questionnaires are used for retrospective assessment (Baldwin et al., 2019). The majority (70%) of included studies in this systematic review used the frequently applied CTQ which may hence be slightly more prone to risk of memory bias. Only one study used additional outside sources to confirm ELS (such as medical and court documents and statements from friends/relatives), which was possible for 53% of participants (Vythilingham et al., 2002). Ideally, future studies would adopt a similar external validation where possible or employ a prospective design. However, the

limitations of access to such records, the fact that a lot of childhood abuse goes unreported, and the difficult and expensive nature of conducting prospective studies, makes this a challenge, and hence using well-validated measures, such as the CTQ, appears to still be one of the more reliable, feasible methods for cross-sectional research studies.

Another issue often reported in ELS research including in the studies of this systematic review is the fact that MDD tend to have greater levels of ELS than HC. This was true for the vast majority of studies (Bermingham et al., 2012; Carballedo et al., 2013; Chaney et al., 2014; Frodl et al., 2010; Lu et al., 2019; Lu et al., 2018; Vythilingham et al., 2002; Yang et al., 2017), and even resulted in studies not including a HCELS group or not measuring ELS in their HC participants (Mikolas et al., 2019; Yuan et al., 2020). In order to best address whether ELS may be a driving factor of structural brain changes irrespective of MDD, ideally both high and low levels of ELS in HC and MDD groups should be recruited, with comparable levels between groups to avoid potential skewing of results. None of the included studies in this review specifically recruited for ELS, rather they recruited MDD and HC groups and assessed ELS subsequently and divided their participants into groups accordingly. In order to avoid low levels of ELS in HC groups, more efforts in future studies should be made to specifically recruit for HC with high levels of ELS.

Finally, in regard to ELS, recent research has highlighted the importance of both timing (age at which ELS occurred, Andersen et al., 2008; Pechtel et al., 2014) and type of ELS (Cassiers et al., 2018; Heim et al., 2013; Teicher et al., 2016) on structural brain changes, including GMV. However, of the included studies, only three conducted analyses on ELS subtypes (Aghamohammadi-Sereshki et al., 2021; Frodl et al., 2010; Frodl et al., 2017;), and three

limited their analyses to a specific subtype(s) (Saleh et al., 2017; Vythilingam et al., 2002; Yuan et al., 2020), whilst all others either did not report or conduct analyses on subtypes. It is important to note that frequently the limiting factor for ELS subtype analysis is sample size, as once broken down by subtype the groups are often underpowered for any meaningful statistical analyses. Even for the three studies that conducted subtype analyses, this was limited this to one or two due to low levels of exposure to certain types of ELS (Frodl et a., 2010) or conducted as exploratory post-hoc analyses due to small sample size (Aghamohammadi-Sereshki et al., 2021). Similarly, the precise timing of abuse was not specified by any study apart from a maximum cut-off (varying between onset of puberty to 18 years of age). As such, this review could not identify any results concerning the potential differing effect of certain ELS subtypes or age of exposure on GMV changes.

4.2.2 Controlling for confounding variables

Five studies did not control for total intracranial volume (ICV), which may potentially skew results if underlying differences in total GMV exist between groups (Carballedo et al., 2013; Chaney et al., 2014; Croy et al., 2013; Lu et al., 2018; Lu et al., 2019). Furthermore, while 6 of the 10 studies that included medicated MDD participants controlled for this in their analyses, four did not (Chaney et al., 2014; Frodl et al., 2010; Mikolas et al., 2019; Tannous et al., 2020), which may affect results given the potential of antidepressant medication to reverse GMV reductions (Boldrini et al., 2009; Frodl et al., 2008). Ideally future studies may seek to include only unmedicated MDD, or apply stringent statistical controls to ensure any GMV changes observed are not potentially mediated by antidepressant use in MDD groups.

5. Conclusion

This systematic review of 20 studies found that ELS is associated with hippocampal GMV reductions (including the hippocampus proper and various hippocampal subfields) that appear to be independent of MDD diagnosis. A limiting factor in firmly confirming this finding is that many studies do not recruit a HCELS group or had low levels of ELS in their HC group, making it more difficult to ascertain whether ELS may be an independent driving factor of hippocampal volume reductions, irrespective of MDD. In those studies that did not include a HCELS group, there was still a clear association between ELS and hippocampal volume reduction (MDDELS<MDD) in the vast majority of included studies. In fact, out of 19 studies that included hippocampal volume analyses (either as an ROI or whole brain analyses) only a single study reported contrary results (MDDELS> MDD; Mikolas et al., 2019), while four studies reported null results (for either MDD or ELS) on hippocampal GMV (Frodl; et al., 2017; Lu et al.. 2019; Tannous et al., 2020 Yang et al., 2017). Similar findings emerged for hippocampal subfields, with an emerging trend towards GMV reductions of the cornu ammonis, DG, and subiculum in ELS, independent of MDD, though more studies are needed as only 5 studies included subfield analyses.

While research has largely focused on the GMV changes in the hippocampus to date, several other key regions associated with ELS were identified in this systematic review (both from whole brain and ROI approaches) and highlight the importance of broadening the scope of future research to perhaps include these regions. Importantly, some of the key areas identified, including the caudate, OFC, and PCC, were all found to have lower GMV in ELS, independent of MDD diagnosis. In conclusion, this systematic review found that of studies including MDD and ELS measures in their analyses of GMV, ELS appears to be the driving factor in GMV reductions, most consistently reported in the hippocampus but also in several

other key brain areas frequently implicated in depression. Previous research in MDD that has cited reductions in GMV in these areas, particularly the hippocampus, may have been, at least partially, misattributing these findings to MDD. The findings of this systematic reviews highlight the need for future studies to include both measures of MDD and ELS in their analyses.

5. Appendix

Study (Authors)	Inclusion /exclusion criteria
Bermingham et al., 2012	Inclusion criteria: ages 18-65; exclusion criteria: history of neurological or comorbid psychiatric disorders (Axis I or Axis II), other severe medical illness, head injury, or substance abuse
Carballedo et al., 2013	Inclusion criteria: ages 18-65; Exclusion criteria: history of neurological or comorbid psychiatric disorders (Axis I or Axis II), other severe medical illness, head injury, or substance abuse
Chaney et al., 2014	Inclusion criteria: 18-65. Exclusion criteria: history of neurologic or severe internal disorders, head injury or substance abuse.
Frodl et al., 2017	Differ slightly for each of the 9 samples (see MDD and HC definitions for details)
Colle et al., 2017	Exclusion criteria: bipolar disorders, psychotic disorders, current substance abuse/dependence, organic brain syndromes, unstable medical conditions, and contra-indications to MRI,
Croy et al., 2013	No olfactory disorders (as assessed using validated Sniffin' Sticks test)
Frodl et al., 2010	18-65, Exclusion criteria for all subjects were previous head injury with loss of consciousness, earlier treatment with hydrocortisone, a history of alcohol or substance abuse, and neurological diseases. Additional exclusion criteria for MDD: comorbidity with other mental illnesses, (including bipolar disorders, personality disorders or psychotic symptoms) and past electroconvulsive therapy.
Gerritsen et al., 2015	SMART: patients with coronary artery disease, cerebrovascular disease, peripheral arterial disease or abdominal aortic aneurysm, and without magnetic resonance imaging (MRI) contraindications. NESDA: current MDD in the past 6 months according to DSM-IV criteria. Exclusion criteria for both patients and controls were the presence or history of major internal or neurological disorder, dependency or recent abuse (past year) of alcohol or drugs; hypertension (5180/130 mmHg), heavy smoker (55 cigarettes/day), and general MRI contraindications.
Hassel et al., 2017	MDD exclusion criteria: Axis I diagnosis other than MDD as a primary diagnosis, Axis II diagnosis, substance abuse within past 6 months, history of neurologic diseases, head trauma.
Lu et al., 2019	Exclusion criteria: significant medical illness; psychiatric axis-I or axis-II disorders (except MDD in patients); alcohol or substance abuse; family history of bipolar disorder; history of loss of consciousness; pregnant or breastfeeding women; history of seizures or family history of epilepsy; any psychotropic medication or hormone; MRI contraindications; left-handedness
Lu et al., 2018	General inclusion criteria: 18-55 years, general exclusion criteria: substance abuse or dependence, neurological or internal illness, MRI contraindications. Additional exclusion criteria for MDD: family member with bipolar disorder.
Opel et al., 2014	Exclusion criteria: any history of severe neurological (e.g., concussion, stroke, tumour, neuro-inflammatory diseases) and medical (e.g., cancer, chronic inflammatory or autoimmune diseases, infections) conditions.
Saleh et al., 2017	Inclusion criteria: 20-50 years. Exclusion criteria: other lifetime DSM-IV Axis I disorders (except comorbid anxiety symptoms occurring in context of depressive episodes), Axis II disorders, history of psychosis, acute suicidality, use of illicit substances in the last month, ECT in the last 6 months, a family history of bipolar disorder, any unstable medical condition, any history of neurological illness or head injury, or MRI contraindications.

Aghamohammadi-Sereshki et al., 2021	Age range 18-49; Exclusion criteria in MDD group: mild depressive episodes; psychotic or atypical features; seasonal affective disorder; lifetime schizophrenia, bipolar disorder, alcohol or substance dependence, anorexia nervosa, or predominant personality or anxiety disorders; antipsychotic or mood stabilizer treatment; corticosteroid use; or significant medical or neurologic diseases.
Vythilingam et al., 2002	Exclusion criteria: major medical illness, significant head trauma, irregular menses, MRI contraindication, hormonal (except contraceptives) or psychotropic medication.
Yang et al., 2017	Inclusion criteria (MDD): first episode depression; ages 18-45; right-handed; HAMD score >17; MDD illness duration between 2 months - 2 years; no comorbid Axis I disorder (except anxiety disorders); no psychotropic medication (except infrequent benzodiazepine use). HC: No psychiatric disorders. Exclusion criteria all participants: significant medical illness, neurological disease, head trauma, loss of consciousness, alcohol/substance abuse.
Yuan et al., 2020	Exclusion criteria (MDD): unstable medical conditions, alcohol/substance use disorder unless in remission for >6 months, other lifetime (past/present) psychiatric disorder (except comorbid anxiety disorders), pregnancy/breastfeeding, dementia, neurological disease or head trauma with cognitive impairment, first-degree family history of schizophrenia, MRI contraindications including claustrophobia. Exclusion criteria (HC): as above, additionally any lifetime (past or present) DSM-IV axis I diagnosis (as assessed using SCID) or family history of a mood or psychotic disorder.
Mikolas et al., 2019	Exclusion criteria: <18 or >65 years of age, history of neurological or comorbid Axis I or II psychiatric disorders, head injury, other severe medical illness or substance abuse.
Teicher et al., 2012	Inclusion criteria: unmedicated, right-handed, 18-25 years old, free from neurological disease or head trauma resulting in loss of consciousness for more than a few seconds. Exclusion criteria: high levels of drug or alcohol use, premature birth/birth complications, maternal substance abuse during pregnancy, experience of multiple unrelated forms of adversity (including natural disaster, motor vehicle accidents, animal attack, near drowning, house fire, mugging, witnessing or experiencing war, gang violence or murder, riot, or assault with a weapon.
Tannous et al., 2020	General inclusion criteria : 18-65 years old. General exclusion criteria: MRI contraindications, pregnancy/breastfeeding. Additional inclusion criteria MDD: DSM-IV diagnosis of current MDD without other Axis I comorbidities, no clinically significant risk of suicidal behaviour or need for urgent treatment. Additional inclusion criteria HC: no current or past history of Axis I disorders.

Appendix Table 1.0 Study inclusion/exclusion criteria

REFERENCES

Aghamohammadi-Sereshki, A., Coupland, N. J., Silverstone, P. H., Huang, Y., Hegadoren, K. M., Carter, R., ... & Malykhin, N. V. (2021). Effects of childhood adversity on the volumes of the amygdala subnuclei and hippocampal subfields in individuals with major depressive disorder. *Journal of psychiatry and neuroscience*, *46*(1), E186-E195.

American Psychiatric Association, A. P. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV)* (Vol. 886). Washington: Washington, DC: American psychiatric association.

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of neuropsychiatry and clinical neurosciences*, *20*(3), 292-301.

Bermingham, R., Carballedo, A., Lisiecka, D., Fagan, A., Morris, D., Fahey, C., ... & Frodl, T. (2012). Effect of genetic variant in BICC1 on functional and structural brain changes in depression. *Neuropsychopharmacology*, *37*(13), 2855-2862.

Bernstein, D. P., & Fink, L. (1998). CTQ: Childhood trauma questionnaire manual.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child abuse & neglect*, *27*(2), 169-190.

Bifulco, A., Brown, G. W., & Harris, T. O. (1994). Childhood Experience of Care and Abuse (CECA): a retrospective interview measure. Journal of Child Psychology and Psychiatry, 35(8), 1419-1435.

Boldrini, M., Underwood, M. D., Hen, R., Rosoklija, G. B., Dwork, A. J., Mann, J. J., & Arango, V. (2009). Antidepressants increase neural progenitor cells in the human hippocampus. *Neuropsychopharmacology*, *34*(11), 2376-2389.

Bremner, J. D., Narayan, M., Anderson, E. R., Staib, L. H., Miller, H. L., & Charney, D. S. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry*, *157*(1), 115-118.

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.

Calem, M., Bromis, K., McGuire, P., Morgan, C., & Kempton, M. J. (2017). Meta-analysis of associations between childhood adversity and hippocampus and amygdala volume in non-clinical and general population samples. *NeuroImage: Clinical*, *14*, 471-479.

Campbell, S., Marriott, M., Nahmias, C., & MacQueen, G. M. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry*, *161*(4), 598-607.

Carballedo, A., Morris, D., Zill, P., Fahey, C., Reinhold, E., Meisenzahl, E., ... & Frodl, T. (2013). Brain-derived neurotrophic factor Val66Met polymorphism and early life adversity affect hippocampal volume. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *162*(2), 183-190.

Cassiers, L. L., Sabbe, B. G., Schmaal, L., Veltman, D. J., Penninx, B. W., & Van Den Eede, F. (2018). Structural and functional brain abnormalities associated with exposure to different childhood trauma subtypes: A systematic review of neuroimaging findings. *Frontiers in psychiatry*, *9*, 329.

Chaney, A., Carballedo, A., Amico, F., Fagan, A., Skokauskas, N., Meaney, J., & Frodl, T. (2014). Effect of childhood maltreatment on brain structure in adult patients with major depressive disorder and healthy participants. *Journal of psychiatry & neuroscience: JPN*, *39*(1), 50.

Colle, R., Segawa, T., Chupin, M., Dong, M. N. T. K. T., Hardy, P., Falissard, B., ... & Corruble, E. (2017). Early life adversity is associated with a smaller hippocampus in male but not female depressed in-patients: a case–control study. *BMC psychiatry*, *17*(1), 1-8.

Colman, I., Kingsbury, M., Garad, Y., Zeng, Y., Naicker, K., Patten, S., ... & Thompson, A. H. (2016). Consistency in adult reporting of adverse childhood experiences. *Psychological medicine*, *46*(3), 543.

Croy, I., Negoias, S., Symmank, A., Schellong, J., Joraschky, P., & Hummel, T. (2013). Reduced olfactory bulb volume in adults with a history of childhood maltreatment. *Chemical senses*, *38*(8), 679-684.

First M, Spitzer R, Gibbon M and Williams J (2002) Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-Patient Edition. (SCID-I/NP). New York: Biometrics Research, New York State Psychiatric Institute.

Frodl, T., Jäger, M., Smajstrlova, I., Born, C., Bottlender, R., Palladino, T., ... & Meisenzahl, E. M. (2008). Effect of hippocampal and amygdala volumes on clinical outcomes in major depression: a 3-year prospective magnetic resonance imaging study. *Journal of psychiatry & neuroscience: JPN*, *33*(5), 423.

Frodl, T., Janowitz, D., Schmaal, L., Tozzi, L., Dobrowolny, H., Stein, D. J., ... & Grabe, H. J. (2017). Childhood adversity impacts on brain subcortical structures relevant to depression. *Journal of psychiatric research*, *86*, 58-65.

Frodl, T., Reinhold, E., Koutsouleris, N., Reiser, M., & Meisenzahl, E. M. (2010). Interaction of childhood stress with hippocampus and prefrontal cortex volume reduction in major depression. *Journal of psychiatric research*, *44*(13), 799-807.

Gerritsen, L., Van Velzen, L., Schmaal, L., van der Graaf, Y., van der Wee, N., van Tol, M. J., ... & Geerlings, M. (2015). Childhood maltreatment modifies the relationship of depression with hippocampal volume. *Psychological medicine*, *45*(16), 3517-3526.

Hammen, C., Henry, R., & Daley, S. E. (2000). Depression and sensitization to stressors among young women as a function of childhood adversity. *Journal of consulting and clinical psychology*, *68*(5), 782.

Han, K. M., Won, E., Sim, Y., & Tae, W. S. (2016). Hippocampal subfield analysis in medication-naive female patients with major depressive disorder. *Journal of affective disorders*, *194*, 21-29.

Hassel, S., Davis, A., Hall, G., Foster, J., Harkness, K., Harris, J., ... & MacQueen, G. (2017). 754. Inflammatory Markers as Mediators of the Association between Early Life Adversity and Hippocampal Volume in Adult Patients with depression: Findings from Phase 1 of the CAN-BIND Study. *Biological Psychiatry*, *81*(10), S306.

Heim, C. M., Mayberg, H. S., Mletzko, T., Nemeroff, C. B., & Pruessner, J. C. (2013). Decreased cortical representation of genital somatosensory field after childhood sexual abuse. *American Journal of Psychiatry*, *170*(6), 616-623.

Huang, Y., Coupland, N. J., Lebel, R. M., Carter, R., Seres, P., Wilman, A. H., & Malykhin, N. V. (2013). Structural changes in hippocampal subfields in major depressive disorder: a high-field magnetic resonance imaging study. *Biological psychiatry*, *74*(1), 62-68.

Joanna Briggs Institute. (2020). *Checklist for analytical cross-sectional studies.* https://jbi.global/critical-appraisal-tools

Lalande, K. M., & Bonanno, G. A. (2011). Retrospective memory bias for the frequency of potentially traumatic events: A prospective study. *Psychological Trauma: Theory, Research, Practice, and Policy*, *3*(2), 165.

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., ... & Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European psychiatry*, *12*(5), 224-231.

Levy-Gigi, E., Szabó, C., Kelemen, O., & Kéri, S. (2013). Association among clinical response, hippocampal volume, and FKBP5 gene expression in individuals with posttraumatic stress disorder receiving cognitive behavioral therapy. *Biological psychiatry*, *74*(11), 793-800.

Lim, H. K., Hong, S. C., Jung, W. S., Ahn, K. J., Won, W. Y., Hahn, C., ... & Lee, C. U. (2012). Automated hippocampal subfields segmentation in late life depression. *Journal of affective disorders*, *143*(1-3), 253-256.

Lindert, J., von Ehrenstein, O. S., Grashow, R., Gal, G., Braehler, E., & Weisskopf, M. G. (2014). Sexual and physical abuse in childhood is associated with depression and anxiety

over the life course: systematic review and meta-analysis. *International journal of public health*, *59*(2), 359-372.

Lindqvist, D., Mueller, S., Mellon, S. H., Su, Y., Epel, E. S., Reus, V. I., ... & Wolkowitz, O. M. (2014). Peripheral antioxidant markers are associated with total hippocampal and CA3/dentate gyrus volume in MDD and healthy controls–preliminary findings. *Psychiatry Research: Neuroimaging*, *224*(3), 168-174.

Lu, X. W., Guo, H., Sun, J. R., Dong, Q. L., Zhao, F. T., Liao, X. H., ... & Li, L. J. (2018). A shared effect of paroxetine treatment on gray matter volume in depressive patients with and without childhood maltreatment: A voxel-based morphometry study. *CNS neuroscience & therapeutics*, *24*(11), 1073-1083.

Lu, S., Xu, R., Cao, J., Yin, Y., Gao, W., Wang, D., ... & Xu, Y. (2019). The left dorsolateral prefrontal cortex volume is reduced in adults reporting childhood trauma independent of depression diagnosis. *Journal of psychiatric research*, *112*, 12-17.

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.

Magarin, A. M., & McEwen, B. S. (1995). Stress-induced atrophy of apical dendrites of hippocampal CA3c neurons: involvement of glucocorticoid secretion and excitatory amino acid receptors. *Neuroscience*, *69*(1), 89-98.

McCrory, E. J., Gerin, M. I., & Viding, E. (2017). Annual research review: childhood maltreatment, latent vulnerability and the shift to preventative psychiatry–the contribution of functional brain imaging. *Journal of child psychology and psychiatry*, *58*(4), 338-357.

Melo, G., Dutra, K. L., Rodrigues Filho, R., Ortega, A. D. O. L., Porporatti, A. L., Dick, B., ... & De Luca Canto, G. (2018). Association between psychotropic medications and presence of sleep bruxism: A systematic review. *Journal of oral rehabilitation*, *45*(7), 545-554.

Mikolas, P., Tozzi, L., Doolin, K., Farrell, C., O'Keane, V., & Frodl, T. (2019). Effects of early life adversity and FKBP5 genotype on hippocampal subfields volume in major depression. *Journal of affective disorders*, *252*, 152-159.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097.

Nogovitsyn, N., Muller, M., Souza, R., Hassel, S., Arnott, S. R., Davis, A. D., ... & MacQueen, G. M. (2020). Hippocampal tail volume as a predictive biomarker of antidepressant treatment outcomes in patients with major depressive disorder: a CAN-BIND report. *Neuropsychopharmacology*, *45*(2), 283-291.

Opel, N., Redlich, R., Zwanzger, P., Grotegerd, D., Arolt, V., Heindel, W., ... & Dannlowski, U. (2014). Hippocampal atrophy in major depression: a function of childhood maltreatment rather than diagnosis?. *Neuropsychopharmacology*, *39*(12), 2723-2731

Paquola, C., Bennett, M. R., & Lagopoulos, J. (2016). Understanding heterogeneity in grey matter research of adults with childhood maltreatment—A meta-analysis and review. *Neuroscience & Biobehavioral Reviews*, *69*, 299-312.

Pechtel, P., Lyons-Ruth, K., Anderson, C. M., & Teicher, M. H. (2014). Sensitive periods of amygdala development: the role of maltreatment in preadolescence. *Neuroimage*, *97*, 236-244.

Saleh, A., Potter, G. G., McQuoid, D. R., Boyd, B., Turner, R., MacFall, J. R., & Taylor, W. D. (2017). Effects of early life stress on depression, cognitive performance and brain morphology. *Psychological medicine*, *47*(1), 171-181.

Sanders, B., & Becker-Lausen, E. (1995). The measurement of psychological maltreatment: Early data on the child abuse and trauma scale. Child abuse & neglect, 19(3), 315-323.

Spijker J, de GR, Bijl RV, Beekman AT, Ormel J, Nolen WA (2002). Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). British Journal of Psychiatry 181, 208–213.

Tannous, J., Godlewska, B. R., Tirumalaraju, V., Soares, J. C., Cowen, P. J., & Selvaraj, S. (2020). Stress, inflammation and hippocampal subfields in depression: a 7 Tesla MRI Study. *Translational psychiatry*, *10*(1), 1-7.

Teicher, M. H., Anderson, C. M., & Polcari, A. (2012). Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proceedings of the National Academy of Sciences*, *109*(9), E563-E572.

Teicher, M. H., Samson, J. A., Anderson, C. M., & Ohashi, K. (2016). The effects of childhood maltreatment on brain structure, function and connectivity. *Nature reviews neuroscience*, *17*(10), 652.

Treadway, M. T., Waskom, M. L., Dillon, D. G., Holmes, A. J., Park, M. T. M., Chakravarty, M. M., ... & Pizzagalli, D. A. (2015). Illness progression, recent stress, and morphometry of hippocampal subfields and medial prefrontal cortex in major depression. *Biological psychiatry*, *77*(3), 285-294.

Van Dam, N. T., Rando, K., Potenza, M. N., Tuit, K., & Sinha, R. (2014). Childhood maltreatment, altered limbic neurobiology, and substance use relapse severity via trauma-specific reductions in limbic gray matter volume. *JAMA psychiatry*, *71*(8), 917-925.

Vanderkolk BA, Perry JC and Herman JL (1991) Childhood origins of self- destructive behavior. American Journal of Psychiatry 148, 1665–1671.

Videbech, P., & Ravnkilde, B. (2004). Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry*, *161*(11), 1957-1966.

Vythilingam, M., Heim, C., Newport, J., Miller, A. H., Anderson, E., Bronen, R., ... & Bremner, J. D. (2002). Childhood trauma associated with smaller hippocampal volume in women with major depression. *American Journal of Psychiatry*, *159*(12), 2072-2080.

Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., & Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: data from the randomized international study to predict optimized treatment for depression. *Translational psychiatry*, *6*(5), e799-e799.

World Health Organization. (1994). Composite International Diagnostic Interview (CIDI) researcher's manual (Version 1.1, 1994). World Health Organization; American Psychiatric Association.

Xie, P., Wu, K., Zheng, Y., Guo, Y., Yang, Y., He, J., ... & Peng, H. (2018). Prevalence of childhood trauma and correlations between childhood trauma, suicidal ideation, and social support in patients with depression, bipolar disorder, and schizophrenia in southern China. *Journal of affective disorders*, *228*, 41-48.

Yang, S., Cheng, Y., Mo, Y., Bai, Y., Shen, Z., Liu, F., ... & Xu, X. (2017). Childhood maltreatment is associated with gray matter volume abnormalities in patients with first-episode depression. *Psychiatry Research: Neuroimaging*, *268*, 27-34.

Yuan, M., Rubin-Falcone, H., Lin, X., Rizk, M. M., Miller, J. M., Sublette, M. E., ... & Mann, J. J. (2020). Smaller left hippocampal subfield CA1 volume is associated with reported childhood physical and/or sexual abuse in major depression: a pilot study. *Journal of affective disorders*, *272*, 348-354.

Zhao, X., Zhang, Y., Li, L., Zhou, Y., Li, H., Yang, S., 2005. Reliability and validity of the Chinese version of childhood trauma questionnaire. Chin. J. Clin. Rehabil. 9, 105–107.

Study 3: Decreased subfield volume in multiple regions of the hippocampal head, including cornu ammonis, dentate gyrus, and presubiculum, in unmedicated female participants with early life stress, independent of depression

Franziska Goer[1], Emily L. Belleau[2,3], Diego A. Pizzagalli[2,3], Rebecca Elliott[1]
[1]University of Manchester, Oxford Road, Manchester M13 9PL, UK
[2]McLean Hospital, 115 Mill St, Belmont, MA 02478, USA
[3]Harvard Medical School, 25 Shattuck St, Boston, MA 02115, USA

Author Contributions:

Franziska Goer contributed to study design, assisted in recruitment of participants and data collection, conducted all analyses, wrote the manuscript and completed all revisions following review by the co-authors.

Professor Rebecca Elliott provided methodological guidance and supervisory input, gave guidance and feedback on analyses, and critically reviewed and provided feedback on drafts of the manuscript prepared by FG.

Dr. Diego Pizzagalli designed the overall study from which data was used. Dr. Pizzagalli and Dr. Emily Belleau gave guidance and supervisory input for the project, and helped design and gave feedback on all analyses.

Funding Information:

This study was study was supported by a grant from NIMH, awarded to Dr. Diego Pizzagalli, grant number 5R01 MH095809. Franziska Goer was supported by funding from the Medical Research Council and University of Manchester as part of Franziska Goer's Medical Research Council (MRC) Doctoral Training Scholarship and MRC DTP flexible training grant.

ABSTRACT

Background: Studies have consistently found reductions in hippocampal volume in major depressive disorder (MDD) and in early life stress (ELS), however, few studies have included measures of both MDD and ELS in their samples, making it difficult to distinguish individual contribution of each or possible interactions/additive effects. Studies of hippocampal subfields have implicated regions of the cornu ammonis (CA1 -CA3) and the dentate gyrus (DG) in both MDD and ELS. Findings in amygdala volume have been highly heterogenous to date in both MDD and ELS and few studies exist investigating amygdala nuclei subregions.
Methods: 75 unmedicated female participants (26.51 ± 6.32 years of age) were recruited into four groups, depending on presence/absence of MDD and presence/absence of childhood sexual abuse (CSA), and completed 3T MRI anatomical scanning. T1-weighted MPRAGE scans were processed and segmented into hippocampal and amygdala subfields using the FreeSurfer 7.0 processing stream.
Results: Results indicated significant volume reductions in several hippocampal head subfields in CSA compared to no CSA, independent of MDD, specifically in the bilateral GC-ML-DG head (4.5% reduction in left & 4.6% in right), the left presubiculum head (5.2% reduction), left CA3 head (5.1% reduction), right CA4 head (4% reduction), and the left whole hippocampal head (3.9% reduction). No significant volume reductions were seen in the whole amygdala or its nuclei.
Conclusion: Findings suggesting that volume reductions in subfields of the hippocampal head may be driven primarily by CSA rather than MDD diagnosis may shed a light on the large literature that has found hippocampal volume reduction in MDD without controlling for ELS as a potential confounding factor. This may highlight the need for new avenues for treatment and timely interventions to prevent/lessen the detrimental effects of CSA (and perhaps more generally ELS) on hippocampal volume and subsequent negative consequences for illness progression and treatment response. Further research in larger sample sizes is required to determine whether this effect replicates and whether timing of abuse and type of ELS may have an effect on volume of the hippocampus (and possibly also the amygdala) and their respective subregions.

1. Introduction

Early life stress (ELS) is known to significantly increase the risk of various psychiatric disorders including depression (Lindert et al., 2014; McCrory et al., 2017); in fact, a large global epidemiological study calculated population-attributable risk proportions and concluded that eradication of childhood adversity/trauma would lead to a 23% reduction in mood disorders (Kessler et al., 2010). While childhood sexual abuse (CSA) usually occurs alongside other types of abuse, studies suggest that CSA is associated with an increased risk of psychiatric disorders during adulthood, even when controlling for other types of abuse (Ferguson et al., 1996; Molnar et al., 2001). The precise mechanism underlying the increased risk for depression and other psychiatric disorders following ELS exposure remains unknown; one postulated theory is that stress-induced structural changes to the hippocampus (and possibly also the amygdala) contribute to the pathogenesis of major depressive disorder (MDD).

1.1. Hippocampal volume in ELS and MDD

Studies in MDD have overwhelmingly reported hippocampal volume reductions compared to healthy controls (HC) (Bremner et al., 2000; Campbell et al., 2004; Videbech & Ravnkilde, 2004). Notable exceptions exist, with some studies failing to find any differences in hippocampal volume between MDD and HC (Ashtari et al., 1999; Posener et al., 2003; Vakili et al., 2000). Similarly to findings in MDD, reduced hippocampal volumes have been consistently linked to ELS (for systematic reviews and meta-analyses, see Calem et al., 2017; Paquola et al., 2016). It is important to note, however, that a caveat in these meta-analyses, and more generally ELS research, is the different definitions and measures of ELS

implemented in studies. While some have used global measures encompassing various subtypes of ELS including physical/emotional/sexual abuse and physical/emotional neglect (such as total score on the childhood trauma questionnaire (CTQ; Bernstein et al., 2003), others have focussed more exclusively on one or a small subset of these types of childhood trauma. Since different types of ELS have been found to differentially affect brain structure (Cassiers et al., 2018; Heim et al., 2013; Teicher et al., 2016), this complicates conclusions and comparisons that can be drawn between studies of ELS and grey matter volume (GMV). Furthermore, timing of ELS may be another key factor, in addition to type of ELS, that may affect GMV of certain structures. For instance, initial evidence suggests that there may be sensitive periods during brain development at which CSA has the greatest detrimental effect on the hippocampus, with one study reporting attenuated hippocampal volume following CSA at ages 3-5 and 11-13 compared to healthy controls (HC) (Andersen et al., 2008).

1.2. Amygdala volume in MDD and ELS

Findings in amygdala volume in depression have been heterogenous and may in part reflect differences depending on the duration of the disorder, with individuals with recurrent MDD demonstrating a relative reduction in amygdala volume (Sheline et al., 1998) while individuals during a first depressive episode demonstrating increased amygdala volume (as compared to recurrent MDD or HC; Frodl et al., 2003). A meta-analysis of amygdala volume in MDD found that inconsistent findings may also be partially explained by medication status, with studies of medicated patients finding an aggregate increased amygdala volume, whilst studies of unmedicated MDD showing the opposite (Hamilton et al., 2008). This, however, does not appear to fully account for the differing findings of studies and more research is needed to determine the directionality and specificity of structural changes to

the amygdala in depression. There have been fewer studies investigating amygdala volume in individuals with a history of ELS. One study found increased right, but not left, amygdala volume in adults with ELS exposure aged 10-11 (predominantly in the form of emotional neglect, parental verbal abuse, and physical neglect; Pechtel et al., 2014). This is in line with animal studies that have found amygdala hypertrophy in macaques (Coplan et al., 2014) and mice (Cohen et al., 2013) exposed to ELS, as well as rats exposed to 10-day chronic stress paradigms (Vyas et al., 2003). However, a meta-analysis of 13 studies found contrary evidence of reductions in amygdala volume in adults with a history of ELS (Paquola et al., 2016), though the effects were less marked than for reductions in hippocampal volume. Similarly, a recent study found ELS, in the form of exposure to violence, to be associated with decreased hippocampal and amygdala volumes (Weissman et al., 2020). Interestingly, studies of posttraumatic stress disorder (PTSD) have yielded similarly incongruent results of amygdala volume, with several studies finding no difference to HC (Woon & Hedges, 2009), attenuated volume compared to HC (Pavlisa et al., 2006; Rogers et al., 2009), and increased volume compared to HC (Kuo et al., 2012). Part of the heterogeneity in amygdala volume findings may stem from differences in type of ELS included in studies as well as inconsistencies in controlling for potential confounding variables such as sex or psychiatric disorders (Paquola et al., 2016).

1.3. Disentangling the effect of ELS and MDD on hippocampal and amygdala volume

While studies investigating changes to amygdala volume in ELS and MDD have produced highly heterogenous results, likely due to differing methods and MDD patient characteristics/type of ELS, reductions in hippocampal volume have been consistently reported in both MDD and ELS samples. However, the vast majority of studies have merely

investigated either MDD or ELS in isolation, without controlling for the other as a potential mediating variable, and thereby making it impossible to disentangle the effects of ELS and MDD on GMV changes. Initial findings from a small number of studies analysing both ELS and MDD in the same sample appear to suggest that ELS may be driving hippocampal reductions observed in MDD. In a study of 63 depressed patients, ELS was associated with smaller hippocampal volume compared to absence of ELS in men, but not women (Colle et al., 2017), though no control group was studied so conclusions were limited to patients with current MDD. In a study comparing patients with MDD and HCs, Chaney and colleagues (2014) found that ELS exposure predicted decreased hippocampal volume, independent of depression, age, or sex. This finding was replicated in a study which found that hippocampal volume reductions in MDD compared to HC were no longer significant when controlling for ELS (measured as total CTQ score), which in turn independently predicted lower hippocampal volumes in both MDD and HC (Opel et al., 2014). Similarly, another study found significant reduction of left hippocampal volume in MDD participants with childhood physical and/or sexual abuse compared to non-abused MDD participants and healthy controls, while non-abused depressed subjects showed similar bilateral hippocampal volumes to healthy participants (Vythilingam et al., 2002). Despite this evidence for ELS possibly accounting for hippocampal GMV reductions observed in MDD, two studies failed to replicate these findings, including a study reporting hippocampal volume reductions in exclusively MDD but not ELS participants (Gerritsen et al., 2015), and a whole brain volumetric analysis of 3036 participants in the ENIGMA study network, which found ELS to predict decreased caudate volume (but no other subcortical structures), independent of MDD diagnosis (Frodl et al., 2017). It should be noted, however, that the former consisted of a significantly older and predominantly male (81%) sample, potentially complicating

comparisons to other studies, and the latter included study samples of both medicated and unmedicated MDD patients and heterogeneous exclusion criteria pertaining to present comorbidities of psychiatric disorders. Furthermore, it should be noted that the reviewed studies of hippocampal GMV in ELS and MDD included different age ranges of ELS exposure and types of ELS, including death of a caregiver (Colle et al., 2017), any combination of physical/emotional/sexual abuse and/or emotional/physical neglect (Chaney et al., 2014; Frodl et al., 2017; Gerritsen et al., 2015; Opel et al., 2014), and restricted to physical and/or sexual abuse (Vythilingham et al., 2002), and as such complicate study comparisons and conclusions about whether effects are general to ELS or specific to a particular type of maltreatment.

1.4. Hippocampal subfields and amygdala nuclei

Some of the reported inconsistencies of studies investigating amygdala and hippocampal GMV in ELS and MDD may also be due to a lack of specificity in analysing the whole amygdala/hippocampus structure rather than its anatomical subfields, which may be differentially affected. Preclinical animal studies of chronic stress (Samuels et al., 2015; Vyas et al., 2002; Vyas et al., 2003) and postmortem studies of depression (Boldrini et al., 2013) have found dendritic atrophy and decreased arborisation in the cornu ammonis (CA1-CA3) and dentate gyrus (DG) subfields of the hippocampus, suggesting that the neurotoxic effect of stress may be specific to certain hippocampal subfields. These findings have been replicated in a handful of studies that have investigated hippocampal subfields in MDD patients which have found volumetric reductions of the DG (Huang et al., 2013; Treadway et al., 2015; Lindqvist et al., 2014; Han et al., 2016) and the lower cornu ammonis (CA1-3) (Huang et al., 2013; Lim et al., 2012; Lindqvist et al., 2014). Similar subfields may be affected

in ELS: a study of adolescent girls found significantly reduced volume of left CA3 in girls with emotional trauma exposure (Malhi et al., 2019). Additional hippocampal subfields have also been implicated in MDD, including decreased volumes of the presubiculum and subiculum (Han et al., 2016) compared to HC.

Studies investigating amygdala nuclei GMV in ELS or MDD are sparse to date. One study found a significant negative correlation between depressive symptom severity and GMV in various amygdala nuclei (right lateral nucleus, left cortical nucleus, left accessory basal nucleus, and bilateral corticoamygdaloid transition area) but failed to identify any group differences between MDD and HC (Brown et al., 2019). Young adults at high risk for psychiatric disorders were found to have significant negative correlation between severity of childhood trauma (in particular CSA and physical abuse) and various amygdala nuclei, particularly the basal regions (Nogovitsyn et al., 2020). However, since these studies investigated either MDD or ELS exclusively and did not measure or control for the other, it is unclear whether they may represent independent, cumulative, or mediating effects on GMV in these subfields.

To our knowledge, only five studies have analysed hippocampal and/or amygdala subfields in studies reporting at least some type of measure for both ELS and MDD. A study in healthy young adults found a significant negative relationship between ELS exposure and volumes of several hippocampal subfields, most notably the left CA2/CA3 and CA4/DG, independent of lifetime MDD (Teicher et al., 2012). It must be noted, however, that the study consisted of a community sample with low levels of ELS and lifetime MDD and it did not differentiate between past or current MDD (Teicher et al., 2012). Similar findings were reported in a

study using manual delineation which found a significant negative correlation between total CTQ score in MDD participants and the bilateral cornu ammonis, right DG, and left subiculum (Aghamohammadi-Sereshki et al., 2021). However, no comparison was made to the HC group (as no CTQ scores were collected for HC) limiting these findings to within MDD. Interestingly, while no group differences emerged between MDD and HC for total and subfield amygdala volumes, CTQ scores were negatively correlated with right total amygdala volume (but no subfields) in MDD (Aghamohammadi-Sereshki et al., 2021). Further evidence for ELS driving volumetric reductions in hippocampal subfields, independent of MDD, was found by Yuan and colleagues (2020) who reported no group differences in hippocampal volume (whole and subfields) between MDD (~40% of whom met for ELS) and HC but found significantly decreased left CA1 volume in abused versus non-abused MDD. ELS was defined as having experienced physical and/or sexual abuse before age 15, however, no additional measures of ELS were collected in either group, making it impossible to ascertain whether other types of abuse/neglect may have been present in MDD or HC and may have contributed to the findings.

It is important to note that two studies did not find evidence for ELS driving volumetric reductions in hippocampal subfields in MDD. Mikolas and colleagues (2019) found overall reductions in hippocampal volumes in MDD versus HC in the cornu ammonis (CA1, CA3, CA4) and DG subfields (GCDG, ML) while, contrary to predictions, MDD patients with ELS showed increased volumes in CA1, CA3 and ML. Importantly though, the majority of participants were on antidepressant treatment and the study did not actively recruit for patients with a history of significant ELS and as such had comparatively lower levels of ELS. A relatively large (*N*=117) recent 7-Tesla imaging study found no differences in hippocampal

subfields or global hippocampal volume in individuals with MDD or ELS compared to HC (Tannous et al., 2020). These findings were unexpected given previous studies and meta-analyses demonstrating hippocampal volume reduction in both MDD and ELS. Again, a significant proportion (23%) of MDD participants were on antidepressant medication which may partially explain null findings as antidepressant medication has been found to reverse hippocampal GMV reductions (Frodl et al., 2008). Furthermore, the study did not actively recruit for a HC/ELS group (i.e. high ELS exposure but no lifetime psychiatric disorder) and hence, unsurprisingly, the MDD group had far higher levels of ELS than the HC ($p<0.001$), making it more difficult to accurately tease out effects of MDD versus ELS on GMV. Finally, it should be noted that the studies reviewed had relatively low levels of ELS as they did not specifically recruit for high ELS exposure in either HC or MDD groups and included a large variety of types of ELS, including predominantly physical and parental verbal abuse (Teicher et al., 2012) total CTQ score (Aghamohammadi-Sereshki et al., 2021; Tannous et al., 2020) and physical and/or sexual abuse (Yuan et al., 2020).

1.5. Aim and hypotheses

Whilst several studies have investigated GMV changes in the amygdala and hippocampus in ELS and MDD, few studies have included measures of both in the same sample, making it difficult to ascertain what is driving the observed effect. Furthermore, initial studies including subfield analyses of these regions in ELS and MDD suggest effects may be localised to particular subregions that are perhaps more susceptible to the deleterious effects of stress induced glucocorticoids. Again, very few studies have included measures of both MDD and ELS in their subfield analyses and those have yielded incongruent findings, likely due to crucial differences in study design and variations in type of ELS studied. The aim of

this original research study was to disentangle the effects of MDD and ELS on GMV in the amygdala and hippocampus (both whole and their respective nuclei/subfields). Since different types of ELS, age of exposure, and gender may affect brain structure including GMV (Cassiers et al., 2018; Heim et al., 2013; Teicher et al., 2016), we aimed to more specifically investigate the effect of CSA exposure at ages 5-14 (a potential sensitive period for CSA's effect on brain development; Andersen et al., 2008) on GMV in the amygdala and hippocampus and their respective subfields. Furthermore, to more accurately attempt to disentangle the relative individual contribution or interaction between CSA and MDD, we recruited 4 groups with each combination of MDD/no MDD and CSA/no CSA. Based on the literature, we hypothesised that reductions in GMV in the hippocampus, particularly in the cornu ammonis and DG subfields, in MDD would be at least partially, if not fully, mediated by CSA. Given the highly inconsistent previous findings pertaining to the amygdala and its subfields, we did not have specific a priori directional hypotheses for this structure or its nuclei.

2. Methods

2.1. Participants

Participants comprised 75 unmedicated women, aged 19-44, recruited from the Boston area. Three additional women were enrolled in the study but were excluded from analyses due to comorbid bipolar disorder (*n*=1), unreliable data ("professional" research subject, *n*=1), and brain abnormalities discovered on the structural scan (confirmed by a radiologist, *n*=1). Participants were assigned to four groups: (1) healthy controls with no history of lifetime psychiatric disorder or ELS (HC, *n*=23), (2) individuals with current MDD and CSA (MDD/CSA, *n*= 18), (3) individuals with current MDD but no ELS (MDD, *n*=18), and (4)

healthy controls with CSA but no history of lifetime psychiatric disorder (HC/CSA, *n*=16). Presence of current and past psychiatric disorders were evaluated via Structured Clinical Interview for the DSM-IV-TR Non-Patient Edition (SCID-IV-N/IP; First et al., 2002). Participant groups did not significantly differ on demographic factors including age, race, years of education or socioeconomic status (see **Table 2** in the results section).

2.1.1. Inclusion/exclusion criteria

All participants were right-handed due to potential differences in brain morphology according to handedness (Jang et al., 2017). Furthermore, a relatively limited age range (19-44) was recruited due to the well documented effects of normal ageing on brain volume and morphology (Fostenos et al., 2005; Galluzzi et al., 2008; Giorgio et al., 2010; Peters, 2006). Exclusively female participants were recruited due to higher incidence rates of CSA in females than males (Molnar et al., 2001; Pereda et al., 2009) and to control for possible sex differences in brain morphometry (Ruigrok et al., 2014). In the MDD groups, participants had to be at least 2 weeks off any psychotropic medication (6+ weeks for fluoxetine and 6+ months for neuroleptics). General exclusion criteria for all groups included pregnancy/risk of pregnancy (determined by urine pregnancy test), MRI counter-indications, serious or unstable medical illnesses (e.g. cardiovascular, hepatic, renal, respiratory, endocrine, neurologic, or hematologic disease), seizure disorder, head injury or loss of consciousness, abnormal results on metabolic panel or blood count (conducted at initial screening session), and abnormal EKG. The following psychiatric history led to exclusion from any group: current or past diagnoses of obsessive–compulsive disorder (OCD), attention deficit hyperactivity disorder (ADHD), bipolar disorder, any psychotic disorder, substance dependence, cocaine or stimulant abuse or use; any substance abuse within past 12

months; current primary diagnosis of phobia, social anxiety, or generalised anxiety disorder (GAD); lifetime diagnosis of anorexia; bulimia diagnosis within last 2 years. Diagnoses of PTSD were permitted in the MDD and MDD/CSA groups if the traumatic incident occurred after age 18. Please see **appendix Table 2** for clinical comorbidities present in the MDD and MDD/CSA group. A urine drug screen was conducted at each study visit and strict current and past drug use criteria were applied (see appendix for details).

2.2. Procedures

The study took place at McLean Hospital, Belmont, MA, USA, and was approved by both the Partners Healthcare and McLean Hospital Institutional Review Boards. Individuals completed a comprehensive initial screening session during which eligibility and childhood trauma were evaluated. Eligible participants were invited back for a second session during which participants completed an MRI scan including anatomical scanning. The study was part of a larger multimodal fMRI, EEG, and PET study with separate aims of investigating reward-learning and stress response in women with childhood trauma.

2.3. Measures

2.3.1. Childhood sexual abuse and ELS

CSA was assessed in the interview version of the Traumatic Antecedents Questionnaire (TAQ; Herman et al., 1989; Vanderkolk et al., 1991). Individuals were classed as meeting criteria for CSA if they experienced at least one incident of contact sexual abuse between the ages of 5-14 years. Contact sexual abuse was operationalised as reporting being forced, coerced, or induced into contact (e.g., touching) activities that are reportedly sexual in nature. Examples from the study include direct fondling, being coerced to engage in oral sex

acts, and being coerced to engage in vaginal sex. Events varied in severity/level of physical intrusion and may have varied in subjective stressfulness. Severity was rated on a scale of 1 to 5, with 1 being 'experienced as not upsetting' to 5 'experienced as extremely severe'. Additional instances of other types of abuse (e.g. physical and emotional) were not exclusionary in the HC/CSA or CSA/MDD groups. The age range of 5-14 years for experience of CSA was applied due to previous research highlighting this timeframe as a sensitive window during which CSA appears to have significant effects on brain development, associated with reductions in GMV in the hippocampus, corpus callosum, and frontal cortex (Andersen et al., 2008). The TAQ was used to evaluate CSA age of onset, severity, and duration for participants in the HC/CSA and MDD/CSA groups. The TAQ was also used to confirm absence of any incidents of childhood sexual, verbal or physical abuse in the MDD only and HC groups. The short form of the CTQ (Bernstein et al., 2003), a retrospective self-report questionnaire that assesses 5 subtypes of childhood trauma (physical/emotional/sexual abuse and physical/emotional neglect), was only collected partway through the study and hence was missing for a third of participants, including the vast majority of HC and MDD subjects as well as for 4 MDD/CSA and 1 HC/CSA. It hence was used for descriptive purposes only to indicate levels of other forms of trauma in the CSA groups. Previously validated and widely applied cut-off scores were calculated to indicate significant levels of abuse/neglect in each subscale: ≥8 for physical abuse, ≥8 for sexual abuse, ≥10 for emotional abuse, ≥8 for physical neglect, and ≥15 for emotional neglect (Bernstein & Fink, 1998; Walker et al., 1999).

2.3.2. Neuroimaging data acquisition and processing

After passing both a urine drug and pregnancy screen, participants completed an anatomical MRI scan at the McLean Imaging Centre. A Siemens TrioTim 3T scanner with 32-channel head coil was used to collect high-resolution T1-weighted MPRAGE scans (TR=2200ms, TE=3.36ms, voxel size = 1.2 x 1.2 x 1.2 mm, flip angle = 7.0 deg, base resolution = 192, slices per slab = 144, field of view = 230mm). FreeSurfer software (version 7.0; Fischl, 2012) was used to process structural T1 data and calculate subcortical brain volume. Volumetric neuroimaging data were pre-processed using the standard semi-automated processing pipeline in FreeSurfer, separating grey matter from white matter and subcortical structures (Maksimovskiy, 2019b). Following the cortical reconstruction process, quality checks and manual inspection of each of the 75 subjects were conducted by the first author (FG) in FreeView (FreeSurfer's graphical visualisation tool). Any uncertainties were discussed and resolved with ELB. The five possible errors include skull strip errors, segmentation errors, intensity normalisation errors, pial surface misplacement and topological defect (white surface errors). All individual slices in each plane (axial, saggital, coronal) were carefully inspected for such errors. Minor manual corrections were needed for 38 subjects (50.6%) of which the majority included edits to the brainmask only for skull strip and pial corrections (*n*=34, 45.3%) and only 4 subjects (5.3%) required edits to the white matter volume due to issues with white matter segmentation. All manual edits were made according to the FreeSurfer Tutorial and clear outliers (as seen when overlaying segmentations over the original T1 data for participants) were re-run (*n*=1) through the pre-processing stage to remove errors. Subcortical volumes were calculated using the Desikan-Killiany atlas (Desikan et al., 2006). Subcortical structures included in analyses were corrected for total intracranial volume (eTIV, as computed by FreeSurfer) to adjust for any

differences in head size and total GMV between subjects (Maksimovskiy, 2019a; Maksimovskiy, 2019b).

2.3.2.1 Hippocampal subfield and amydala nuclei segmentation

The recently improved semi-automated algorithm from FreeSurfer 7.0 was applied to conduct hippocampal subfields segmentation (Iglesias et al., 2015) and amygdala nuclei segmentation (Saygin et al., 2017), which have been found to have excellent numerical and spatial reliability (Buser et al., 2020). Joint, simultaneous segmentation of both the hippocampus and amygdala prevents overlap and gaps between structures, a clear improvement of the newly developed segmentation algorithm compared to previous iterations of the FreeSurfer software. A list of the 19 hippocampal subfields generated is listed in **Table 1** (see appendix for a visualisation of these subfields)**.** Amygdala segmentation generated volumes for 9 nuclei: lateral nucleus, basal nucleus, accessory basal nucleus, anterior amygdaloid area AAA, central nucleus, medial nucleus, cortical nucleus, cortico amygdaloid transition, and the paraliminar nucleus.

Parasubiculum	HEAD
Presubiculum head	
Subiculum head	
CA1 head	
CA3 head	
CA4 head	
GC-ML-DG head	
Molecular layer head	
HATA	
Presubiculum body	BODY
Subiculum body	
CA1 body	
CA3 body	
CA4 body	
GC-ML-DG body	

Molecular layer HP body	
Fimbria	
Hippocampal tail	TAIL
Hippocampal fissure	FISSURE

Table 1: Hippocampal Subfields
Note: The CA2 is included in the CA3. Abbreviations: CA, cornu ammonis; GC-ML-DG, granule cell (GC) and molecular layer (ML) of the dentate gyrus (DG); HATA, hippocampus-amygdala-transition-area.

2.4. Statistical Analyses

All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) software, version 25. Repeated measures ANOVAs, with ROI (subfields) as within-subjects factors and CSA and MDD as between-subject factors, were performed for the left and right hemisphere for both amygdala and hippocampal subfields. Post-hoc independent t-tests were performed to follow up any significant interactions obtained from the repeated measures ANOVA. Separate univariate ANOVAs were performed for each hemisphere for both the whole amygdala and the whole hippocampus. Significance threshold was set at $p<0.05$. All means of hippocampal and amygdala volumes are given in mm^3.

3. Results

3.1. Demographics and clinical characteristics

Groups did not significantly differ on demographic factors including age, race, education, or socioeconomic status (see **Table 2**). While there was no significant difference in CSA severity between HC/CSA and MDD/CSA groups, the duration of CSA was longer in MDD/CSA (M = 5.18, SD = 3.046) than HC/CSA (M = 2.88, SD = 3.053), $F(1,31)= 4.696$, $p= 0.038$, $\eta p^2 = 0.132$, and CSA onset was earlier in MDD/CSA (M = 6.59, SD = 1.734) than HC/CSA (M = 8.94, SD = 3.549), $F(1,31) = 5.949$, $p=0.021$, $\eta p^2 = 0.161$. There was a significant difference in the percentage of participants meeting cut-offs for significant abuse/neglect in each CTQ

subscale, with a significantly higher percentage of MDD/CSA participants meeting cut-offs than HC/CSA (p<0.05 for all, see **Table 2**), apart from CSA which was met by 100% of participants who completed the CTQ in both CSA groups. As expected, there was a significant overall difference in total score of the Beck Depression Inventory-II (BDI-II; Beck et al, 1996) between all four groups, importantly however, there was no significant difference between HC (M = 1.461, SD = 1.825) and HC/CSA (M = 2.260, SD = 2.939), $F(1,36) = 1.067$, $p = 0.309$ or between MDD (M = 29.577, SD = 8.651) and MDD/CSA (M = 28.093, SD = 10.119), $F(1,34) = 0.224$, $p= 0.639$. Similarly, total score on the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960) was not significant different between HC (M = 1.00, SD = 1.508) and HC/CSA (M = 1.56, SD = 2.502), $F(1, 37) = 0.767$, $p= 0.387$ or between MDD (M = 14.11, SD = 4.886) and MDD/CSA (M = 15.38, SD = 3.284), $F(1.34) = 0.763$, $p= 0.389$. There was no significant difference in number of current major depressive episode (MDE) symptoms between the MDD only (M = 6.00, SD = 1.414) and MDD/CSA groups (M = 5.118, SD = 1.35), $F(1,33) = 3.551$, $p = 0.068$, $\eta p^2 = 0.097$). There was also no significant difference in number of lifetime MDEs between MDD (M = 3.35, SD = 3.427) and MDD/CSA groups (M = 3.58, SD = 3.029), $F(1,27) = 0.035$, $p = 0.853$, $\eta p^2 = 0.001$).

	Participants, Mean (SD)						
Characteristic	HC (n=23)	MDD (n=18)	HC/CSA (n=16)	MDD/CSA (n=18)	Statistic	*P* value	Partial Eta Squared
Age (years)	26.52 (7.006)	24.94 (5.418)	24.88 (5.365)	29.50 (6.410)	F =2.165	p= 0.100	0.084
Caucasian (%)	73.9	61.1	56.3	41.2	X^2 = 4.439	p = 0.218	
Education (years)	14.870 (4.7223)	16.500 (2.9556)	15.750 (1.6125)	15.361 (2.3998)	F =0.859	p = 0.467	0.035
SES (yearly income, $, in thousands)	47.636 (28.380)	51.083 (35.699)	47.188 (22.983)	45.528 (33.278)	F 0.105	p = 0.957	0.004

TAQ CSA severity	NA	NA	3.38 (1.025)	4.00 (1.118)	F = 2.792	p = 0.105	0.083
TAQ CSA duration (years)	NA	NA	2.88 (3.052)	5.18 (3.046)	F = 4.696	p = 0.038	0.132
TAQ CSA onset (age in years)	NA	NA	8.94 (3.549)	6.59 (1.734)	F =5.949	p = 0.021	0.161
# (%) meeting CTQ cut-off for EA	NA	NA	4 (25%)	9 (50%)	X^2 = 4.144	p = 0.042	
# (%) meeting cut-off for PA	NA	NA	2 (12.5%)	7 (38.9%)	X^2 = 4.549	p = 0.033	
# (%) meeting CTQ cut-off for SA	NA	NA	15 (100%)	14 (100%)			
# (%) meeting CTQ cut-off for EN	NA	NA	2 (12.5%)	7 (38.9%)	X^2 = 4.549	p = 0.033	
# (%) meeting CTQ cut-off for PN	NA	NA	2 (12.5%)	8 (44.5%)	X^2 = 6.152	p = 0.013	
BDI-II*	1.461 (1.825)	29.577 (8.651)	2.260 (2.939)	28.09 (10.119)	F = 98.34	p =0.00	0.808
HRSD**	1.00 (1.508)	14.11 (4.886)	1.56 (2.502)	15.38 (3.284)	F = 107.159	p =0.00	0.823

Table 2. Demographics and clinical characteristics
Abbreviations: SES, socioeconomic status; BDI-II, Beck Depression Inventory-II (Beck et al, 1996); HRSD, Hamilton Rating Scale for Depression (Hamilton et al., 1960); MDD, major depressive disorder; CSA, childhood sexual abuse; HC, healthy control; NA, not applicable; TAQ, Traumatic Antecedents Questionnaire (Herman et al., 1989; Vanderkolk et al., 1991); CTQ, childhood trauma questionnaire; EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect. *Note:* CTQ scores were missing for 1 HC/CSA and 4 MDD/CSA.

3.2. Hippocampal subfields GMV

3.2.1. Left hemisphere hippocampal subfields

While there was no significant ROI*MDD or ROI*MDD*CSA interaction in left hippocampal subfields (all p>0.05), there was a significant ROI*CSA interaction, $F(20,1420)=2.402$, $p<0.001$, $\eta p^2 = 0.033$. Pairwise follow up comparisons of each subfield revealed smaller volume in CSA (M=129.909, SD=11.369) than no CSA (M=137.096, SD=14.443) in the left

presubiculum head: $t(73)=2.358$, $p=0.021$; smaller volume in CSA (M=150.022, SD=13.527) than no CSA (M=157.052, SD=15.529) in the left GC-ML-DG head, $t(73)=2.068$, $p=0.042$; and smaller volume in CSA (M=117.620, SD=12.767) than no CSA (M=123.948, SD=13.335) in the left CA3 head, $t(73)=2.085$, $p=0.041$. The left whole hippocampal head (consisting of all head subfields) was significantly smaller in CSA (M=1629.042, SD=129.241) than no CSA (M=1695.182, SD=137.698), $t(73)=2.129$, $p=0.037$. While not statistically significant, there were also trends in the left CA4 head, with smaller volume in CSA (M=125.615, SD=11.471) than no CSA (M=130.998, SD=12.189), $t(73)=1.955$, $p=0.053$; and the left molecular layer HP head, once again with smaller volume in CSA (M=319.005, SD=26.968) than no CSA (M=331.233, SD=26.466), $t(73)=1.975$, $p=0.052$. In percentages, CSA was associated with volumetric reductions of 5.24% in the left presubiculum head, 4.476% in the left GC-ML-DG head, 5.105% in the left CA3 head, and 3.902% in the left whole hippocampal head. The trends observed in the left CA4 head and the left molecular layer HP head represented 4.109% and 3.691% reductions in volume in CSA compared to no CSA, respectively.

3.2.2. Right hemisphere hippocampal subfields

As in the left hemisphere, there were no significant ROI*MDD or ROI*MDD*CSA interactions in right hippocampal subfields (all $p>0.05$), but again there was a significant ROI*CSA interaction, $F(20,1420)=1.683$, $p=0.030$, $\eta p^2 = 0.023$. Post hoc follow up tests for each subfield found significant volumetric differences in the right GC-ML-DG head, which was smaller in CSA (M=153.409, SD=12.666) than no CSA (M=160.783, SD=15.683), $t(73)=2.252$, $p=0.027$, and in the right CA4 head, which was also smaller in CSA (M=127.689, SD=10.416) than no CSA (M=132.968, SD=12.382), $t(73)=2.005$, $p=0.049$. While not statistically significant, there were also trends in the right whole hippocampal head, with smaller

volume in CSA (M=1677.429, SD=227.649) than no CSA (M=1734.383, SD=125.076), *t*(73)=1.945, *p*=0.056; and the right CA3 head, again with smaller volume in CSA (M=123.402, SD=12.191) than no CSA (M=129.813, SD=16.254), *t*(73)=1.898, *p*=0.055. In percentages, CSA was associated with volumetric reductions of 4.586% in the right GC-ML-DG head and 3.970% in the right CA4 head. The trends observed in the right whole hippocampal head and the right CA3 head represented 3.284% and 4.939% reductions in CSA compared to no CSA, respectively.

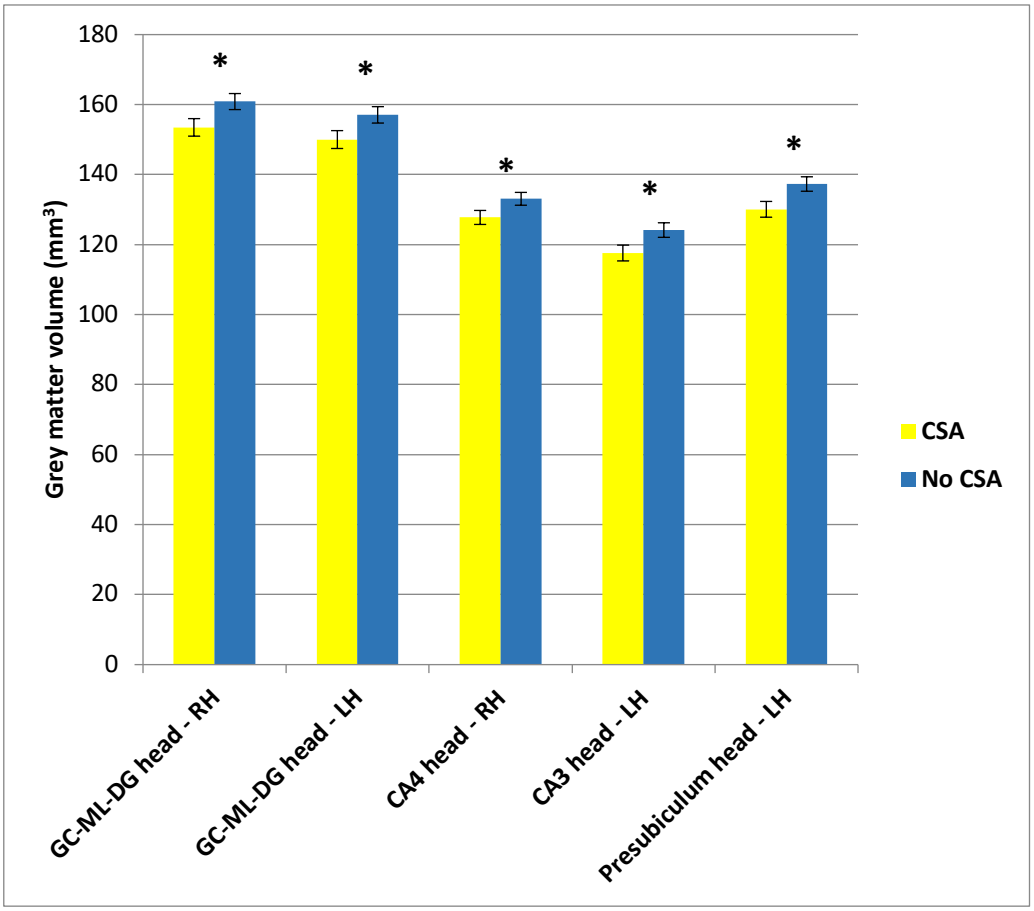


Figure 1. Hippocampal subfield volumes in CSA and no CSA. * indicates significance, *p*<0.05, LH = left hemisphere, RH = right hemisphere.

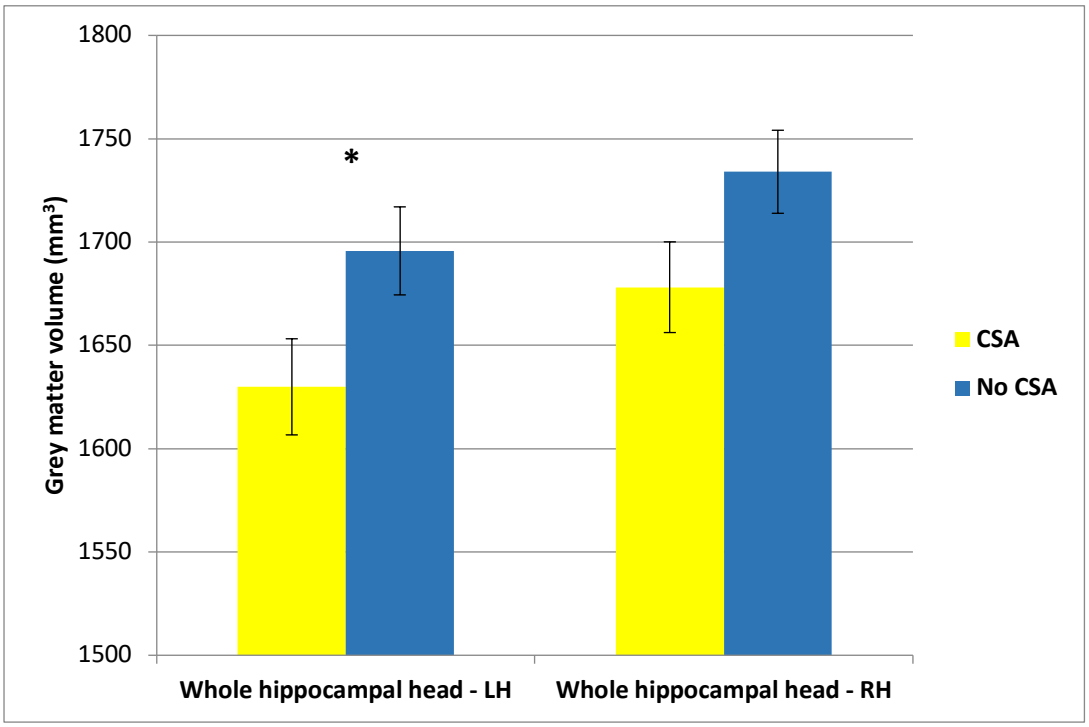


Figure 2. Left & right whole hippocampal head volume in CSA and no CSA. * indicates significance, $p<0.05$, LH = left hemisphere, RH = right hemisphere.

3.3. Amygdala nuclei GMV

There were no significant interactions for left hemisphere nuclei, however, there was a significant ROI*CSA interaction in the right hemisphere nuclei, $F(8,568)=2.336$, $p=0.018$, $\eta p^2 = 0.032$. Post hoc pairwise comparisons for each right hemisphere nuclei yielded no significant differences, however there was a trend for smaller basal nucleus volume in CSA (M=415.798, SD=33.898) versus no CSA (M=429.599, SD=36.892), $t(73)=1.673$, $p=0.096$ and similarly a trend for smaller paralaminar nucleus volume in CSA (M=46.756, SD= 3.811) versus no CSA (M=48.403, SD=4.269), $t(73)=1.745$, $p=0.082$. Inspection of means (see **Table 3**) demonstrates that the majority of nuclei (with the exception of central, medial and cortical nuclei) were smaller in CSA than absence of CSA, albeit not reaching statistical significance in the aforementioned post hoc tests for each nuclei.

	ROI (right hemisphere)	Mean	Std. Error
No CSA	Lateral nucleus	646.006	8.238
	Basal nucleus	429.924	5.663
	Accessory Basal nucleus	257.044	3.589
	Anterior amygdaloid area AAA	55.619	1.135
	Central nucleus	41.028	0.933
	Medial nucleus	18.628	0.806
	Cortical nucleus	24.373	0.5
	Cortico amygdaloid transition	175.108	2.777
	Paralaminar nucleus	48.427	0.647
CSA	Lateralnucleus	627.237	8.995
	Basalnucleus	415.906	6.183
	Accessory Basal nucleus	251.155	3.919
	Anterior amygdaloid area AAA	52.723	1.24
	Central nucleus	41.546	1.019
	Medial nucleus	19.072	0.88
	Cortical nucleus	24.974	0.546
	Cortico amygdaloid transition	172.492	3.032
	Paralaminar nucleus	46.78	0.707

Table 3. GMV of right amygdala nuclei by CSA presence/absence

3.4. Whole amygdala and hippocampus GMV

There was no significant main effect or interaction of MDD or CSA on whole amygdala volume or on whole hippocampus volume for either the right or left hemisphere ($p>0.05$ for all).

4. Discussion

Analyses of hippocampal subfields revealed a pattern of decreased volume in CSA versus no CSA in various locations of the hippocampal head, independent of MDD. Specifically, GMV was significantly reduced in the bilateral GC-ML-DG head, the left presubiculum head, left CA3 head, and right CA4 head, in CSA versus no CSA. In line with findings that individual subfields reaching significance were all located in the hippocampal head (and none in the body or tail), the left whole hippocampal head (and trending in the right), showed

decreased volume in CSA versus no CSA. There was no main effect of MDD or interaction between MDD and CSA in any subfields, suggesting that, at least in this sample, CSA, and not MDD, was the driving factor behind volumetric reductions in several hippocampal subfields.

As expected, the subfields demonstrating volumetric reductions in CSA included subfields of the dentate gryus (GC-ML-DG head and CA4 head) and cornu ammonis (CA3). These subregions appear to be particularly sensitive to the deleterious effects of stress as demonstrated by dendritic atrophy in translational animal studies of chronic stress (Samuels et al., 2015; Vyas et al., 2002; Vyas et al., 2003). This has been replicated in several studies of MDD which found volumetric reductions of the DG (Huang et al., 2013; Treadway et al., 2015; Lindqvist et al., 2014; Han et al., 2016) and the lower cornu ammonis (CA1-3) (Huang et al., 2013; Lim et al., 2012; Lindqvist et al., 2014), and ELS (Malhi et al., 2019; Travis et al., 2016). As discussed previously, these studies did not include measures of both depression and ELS, making it impossible to discern the relative contribution of each on hippocampal subfields GMV reductions. The few studies that did include measures of both ELS and MDD have yielded inconsistent findings, though this may be due to differences in sample characteristics, such as including patients on anti-depressant medication and with low levels of ELS (Tannous et al., 2020; Mikolas et al., 2019). Our findings are in line with a study by Teicher and colleagues (2012) in which ELS (predominantly in the form of physical abuse and parental verbal abuse) was associated with reductions in the left CA2/3 and CA4/DG, independent of lifetime MDD diagnosis. The size of these volumetric reductions were comparable to those in the present investigation in which we found a 5.24% reduction in volume of the left CA3 head, a 4.476% and 4.586% reduction in the left and right GC-ML-DG

head, respectively, and a 3.870% reduction in right CA4 head. Interestingly, Teicher and colleagues (2012) also identified a significant reduction in presubiculum volume in ELS which was comparable to our finding of a 5.24% reduction in the left presubiculum head in CSA versus no CSA. A recent study in MDD also found reductions in presubiculum volume (in addition to cornu ammonis, GC-ML-DG, and CA4) though no measure of ELS was included (Han et al., 2019). There is some evidence that the presubiculum and subiculum, especially ventral regions (another term for the head, Zeidman & Maguire, 2016), play a key role in regulation of the HPA-axis (Herman et al., 1998) and contain a high concentration of glucocorticoid binding sites (Sarrieau et al., 1986), however, future studies are needed to investigate this potential link underlying volumetric reduction in the presubiculum/subiculum following ELS. The present investigation extends the findings obtained by Teicher et al. (2016) by specifically recruiting MDD and CSA groups (rather than a community sample with low levels of both MDD and ELS) and applying a newer iteration of FreeSurfer software which further divides the cornu ammonis and dentate gyrus into more specific subfields (such as the GC-ML-DG) and includes distinctions between the tail and head.

Our results indicated that the subfields in the head of the hippocampus, including the bilateral GC-ML-DG head, the left presubiculum head, left CA3 head, right CA4 head, as well as the left whole hippocampal head (and trending for right), were exclusively implicated in volumetric reductions in CSA versus no CSA (as oppose to the hippocampal body or tail). Few studies in MDD and ELS have parcelled out hippocampal subfields into head/body/tail divisions. Initial evidence suggests a particular emphasis on the hippocampal head in childhood trauma. Reduced volume of the bilateral hippocampal head following childhood

emotional abuse (compared to those with no ELS) in first-degree relatives of MDD has been reported, suggesting that these structural changes may be present in high-risk individuals prior to symptom onset (Carballedo et al., 2012). Within MDD participants, particularly marked effects of ELS (in the form of physical and/or sexual abuse) on left hippocampal head volume reduction have been observed (Vythilingam et al., 2002) and total CTQ score has been found to significantly correlate with reductions in hippocampal head (but not body or tail) volume (Aghamohammadi-Sereshki et al., 2021). These findings are in line with results from the present investigation and suggest that the hippocampal head may be particularly vulnerable to deleterious effects of ELS during brain development. However, more studies delineating the hippocampal head, tail and body are necessary to verify if this effect replicates in larger sample sizes and whether it is a product of general ELS or driven by specific subtypes.

Particularly of note is that our findings of hippocampal subfield volume reductions appeared to be driven by CSA, independent of MDD. While studies of hippocampal volume in samples including measures of both depression and ELS (such as CSA) are still limited in number, initial evidence points to ELS as a potential underlying factor explaining observed hippocampal volume reductions in MDD, with findings of volumetric reductions of both hippocampal subfields (Teicher et al., 2012; Yuan et al., 2020) and the whole hippocampus (Chaney et al., 2014; Opel et al., 2014; Vythilingam et al, 2002) in ELS independent of MDD. Similarly, whilst not finding significant differences for hippocampal volume, a large (N=3036) whole brain analysis found a strong correlation between ELS and reduced caudate volume, independent of MDD (Frodl et al., 2017). Importantly, these studies all used differing criteria for ELS in their samples, with some, such as the present investigation, focusing on one form

of ELS in particular, whilst others used a global measure of ELS such as total CTQ score - further studies are needed to investigate whether certain types of ELS may have greater effect on hippocampal subfield volumes or whether there may be an additive effect of multiple types of ELS. Combined, these studies, including the present investigation, may offer initial support for the theory that some structural changes, including volumetric reductions of subfields of the hippocampus, may be driven by exposure to ELS rather than primarily caused by MDD. This may shed light on the vast literature that has found similar findings of hippocampal volume reduction in both ELS and MDD but rarely has controlled for the other factor, thereby largely making it impossible to determine to which factor to attribute the effect. Given the high prevalence of ELS in MDD (studies have found a greater than 2-fold increase of significant ELS exposure in MDD participants compared to HC with a prevalence between 50.5% to 62.5% of significant childhood trauma reported in MDD participants; Williams et al., 2016; Xie et al., 2018) it is possible that high levels of unmeasured ELS in MDD samples may account for reductions in hippocampus volume reported in studies of MDD patients. ELS as a driving factor underlying GMV changes may also serve as an explanation of similar findings observed in other psychiatric disorders which are also associated with high levels of ELS. Reductions of cornu ammonis, dentate gyrus, and (pre)subiculum subfields have been reported in alcohol use disorder (Zahr et al., 2019), schizophrenia (Ota et al., 2017), and bipolar disorder (Han et al., 2019). Certainly further studies, particularly those in large sample sizes, are needed to establish whether this effect replicates, but if so, it could have notable ramifications for a field that has consistently documented hippocampal subfield reductions in MDD but predominantly did not control for ELS exposure, thereby potentially misattributing the effect to MDD alone.

This would also potentially open up new avenues for treatment and intervention, highlighting the importance of prevention of childhood trauma and timely interventions that may help reduce the consequences of ELS on brain development. Importantly, a meta-analysis of 16 epidemiological studies (N= 23,544) found that depressed individuals with a history of ELS have significantly worse illness and treatment outcomes than those with no history of trauma (Nanni et al., 2012). Though studies are needed to investigate underlying mechanisms, reduction in hippocampal volume following ELS may represent one such factor. This is supported by a study by Frodl et al. (2010) that found that ELS and hippocampal volume reduction predicted more severe illness course in depressed individuals. Similar findings have been identified in substance use disorder, with ELS exposure correlating with decreased GMV in various hippocampal subfields, which in turn was found to predict risk of relapse (Dam et al., 2014). This suggests that GMV reductions in hippocampal regions due to ELS exposure may underlie poorer outcomes in a variety of psychiatric disorders and as such make it an even more important target for early identification and intervention. Reduced hippocampal volume has further been associated with poorer treatment response in MDD (Frodl et al., 2008) and may even function as an accurate state biomarker for acute suicide attempts in MDD patients (Colle et al., 2014). On a hopeful note, antidepressant medication has been found to partially reverse hippocampal GMV reductions (Frodl et al., 2008) which may be the result of antidepressants increasing neurogenesis in hippocampal cells in individuals with depression (Boldrini et al., 2009). A similar effect of increased hippocampal volume has been observed following clinical improvement after receiving cognitive behavioural therapy in individuals with PTSD (Levy-Gigi et al., 2013).

One of the key mechanisms underlying hippocampal volume reduction following ELS appears to involve dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis, which has been consistently implicated following ELS (Heim et al., 2000). Hyperactivity of the HPA axis (Krishnan & Nestler, 2011) and neural atrophy in the hippocampus and amygdala (Lupien et al., 2008; Magarin & McEwen, 1995) have also been observed in animal models of stress exposure. High levels of chronic stress are thought to play a key contributing factor to HPA-axis dysregulation, as stressors lead to increased secretion of neurotoxic glucocorticoids (Burke et al., 2005; Pittinger et al., 2008) to which the hippocampus appears particularly sensitive (Lupien et el., 2008) as can be seen in atrophy of dendritic neurons and reductions in GMV (McEwen, 1999; Sapolsky, 2000). The deleterious effect of stress induced glucorcorticoids may particularly affect the cornu ammonis (CA1-CA3) and DG hippocampal subfields due to their high density of glucocorticoid receptors, as has been observed in animal studies that have identified dendritic atrophy and decreased arborisation in these subfields following chronic stress paradigms (Samuels et al., 2015; Vyas et al., 2002; Vyas et al., 2003). Similar effects have been observed in studies of humans following ELS, with specific reductions in CA3 and DG areas following various types of ELS (Malhi et al., 2019; Teicher et al., 2012). A longitudinal study has provided further evidence of a potential causal link between ELS and hippocampal volume: increased cortisol levels were found to predict decreased hippocampal volume in a longitudinal study of youth with ELS and post-traumatic stress symptoms compared to healthy controls (Carrion & Wong, 2012).

Sensitisation of the HPA-axis as a result of exposure to severe childhood stressors is also a key theory of the link between childhood trauma and development of depression (Heim et al., 2008). Stress sensitisation of the HPA axis may explain why individuals with a history of

significant ELS are more likely to develop a major depressive episode in response to stressful life events than those without ELS exposure (Hammen et al., 2000). It is important to note, however, that multiple biological processes, in addition to HPA-axis dysregulation, may interact and contribute to the observed structural changes in ELS and MDD, such as inflammation, oxidative stress, and neurotransmitter abnormalities (Belleau et al., 2018). Additionally, it remains unclear whether hippocampal volume changes precede and contribute to development of MDD or rather represent a scarring effect of the disorder, as some studies have reported decreased hippocampal volume in healthy individuals at high risk of depression (Amico et al., 2010; Chen et al., 2010; Rao et al., 2010) whilst others have found no indication of volume reductions prior to symptom onset (Chan et al., 2016), though definitions of high-risk groups differed between these studies. Furthermore, precisely how ELS may modify this relationship remains unclear – initial evidence from longitudinal studies suggest that smaller hippocampal volumes following ELS exposure may increase susceptibility to the development of MDD during adolescence (as assessed at 5 year follow up; Rao et al., 2010).

Contrary to expectations, we did not observe any effect of MDD and/or ELS on whole hippocampal volume. While a vast majority of studies have reported decreased hippocampal volume in MDD (Bremner et al., 2000; Campbell et al., 2004; Videbech & Ravnkilde, 2004) and ELS (Calem et al., 2017; Paquola et al., 2016), several studies have failed to find any difference to HC (Ashtari et al., 1999; Posener et al., 2003; Vakili et al., 2000; Frodl et al., 2017). Given our findings, one possible theory may be that reductions in only certain hippocampal subfields contribute to the overall reduction in hippocampal volume observed and may go undetected when only evaluating the hippocampus whole.

Volumetric analyses of the whole amygdala revealed no significant findings for a main effect of CSA or MDD or an interaction. While there appeared to be a significant ROI*CSA interaction for amygdala nuclei, post hoc follow up tests for individual nuclei did not reach significance, though visual inspection of means suggest this may be driven by lower mean volumes in CSA than no CSA present in 6 out of 9 nuclei. While preclinical animal studies have shown a pattern of enhanced amygdala arborisation and hypertrophy following various stress paradigms (Cohen et al., 2013; Coplan et al., 2014; Vyas et al., 2002; Vyas et al., 2003), human studies in ELS and MDD have been highly heterogenous, and may depend on a complex set of factors such as timing and type of ELS exposure (Pechtel et al., 2014), MDD illness duration (Frodl et al., 2003; Sheline et al., 1998) and antidepressant use (Hamilton et al., 2008). A meta-analysis found significant reduction of amygdala volume in ELS, though the effect was less marked and less robust than volume reductions in the hippocampus (Paquola et al., 2016). Furthermore, larger effect sizes for decreased amygdala volume were found amongst studies with greater mean age and higher proportion of male participants, which may also partially explain the lack of clear results in our sample which was relatively young (M=26.51 , SD=6.32) and exclusively female.

The present investigation had several limitations that should be considered and addressed by future studies. While our sample specifically recruited and matched four groups with varying levels of CSA and MDD to allow for close control of these variables in statistical analyses, our two CSA groups (HC/CSA and MDD/CSA) differed significantly in CSA duration (longer in MDD/CSA) and age of onset (older in HC/CSA) though they did not differ in CSA severity. Furthermore, whilst our trauma groups focused on CSA, other types of ELS were

not excluded and were present at significant levels in a subset of CSA participants, hence we cannot isolate effects from our study to CSA alone. While a significantly higher proportion of the MDD/CSA versus MDD only group met the cut-off for significant levels of other types of abuse and neglect, this should not affect our results as we did not observe any MDD* CSA interaction and instead found our results to be driven by CSA alone. Nevertheless, future studies should ideally balance co-occurring abuse/neglect in both groups to avoid higher levels of overall ELS in one group over the other which could potentially affect results should there be an MDD*CSA interaction on GMV. A further limitation was that the frequently applied CTQ measure was only collected for a small subset of our participants and hence we were not able to include this in our analyses (apart from descriptive statistics for prevalence of other types of trauma for a subset of participants in the CSA groups). Given the frequent use of this measure in ELS studies, future studies should aim to collect this data to aid in comparisons across studies with potentially varying definitions of ELS. Due to a relatively small sample size, the high number of participants with missing CTQ scores and limitations provided by the TAQ measure, our analyses were not able to tease apart the effects of different types of ELS and age of ELS occurrence. This may be particularly relevant for future studies to include, as there is preliminary evidence for sensitive periods during brain development during which CSA may have greater effect on particular subcortical regions, including the hippocampus (Andersen et al., 2008). Another limitation in our study design is that we collected cross-sectional data and as such cannot draw causational conclusions – this would ideally be addressed by large-scale longitudinal studies with both ELS and MDD measures and regular intervals of GMV assessments throughout childhood and at adulthood. In general, a larger sample size would certainly be beneficial, as our study consisted of four relatively small subgroups, and hence may be underpowered to detect

smaller effect sizes. Given that meta-analyses have found smaller effect sizes for amygdala than hippocampal volume reductions following ELS (Paquola et al., 2016), this may perhaps explain our lack of significant findings in the amygdala and its nuclei, and highlight the need for larger well-powered studies.

5. Conclusion

Our study applied the recent FreeSurfer 7.0 processing stream to analyse amygdala nuclei and hippocampal subfield volumes in depression and CSA. Results indicated significant volume reductions in several hippocampal head subfields in CSA, independent of MDD, specifically in the bilateral GC-ML-DG head, the left presubiculum head, left CA3 head, right CA4 head, and the left whole hippocampal head. No clear significant results emerged in the whole amygdala or its nuclei and there were no effects or interactions of MDD and CSA on whole hippocampal volume. Further research in larger sample sizes is required to determine whether these findings replicate and whether timing and type of ELS may have an effect on GMV of the hippocampus and amygdala and their respective subregions.

6. APPENDIX

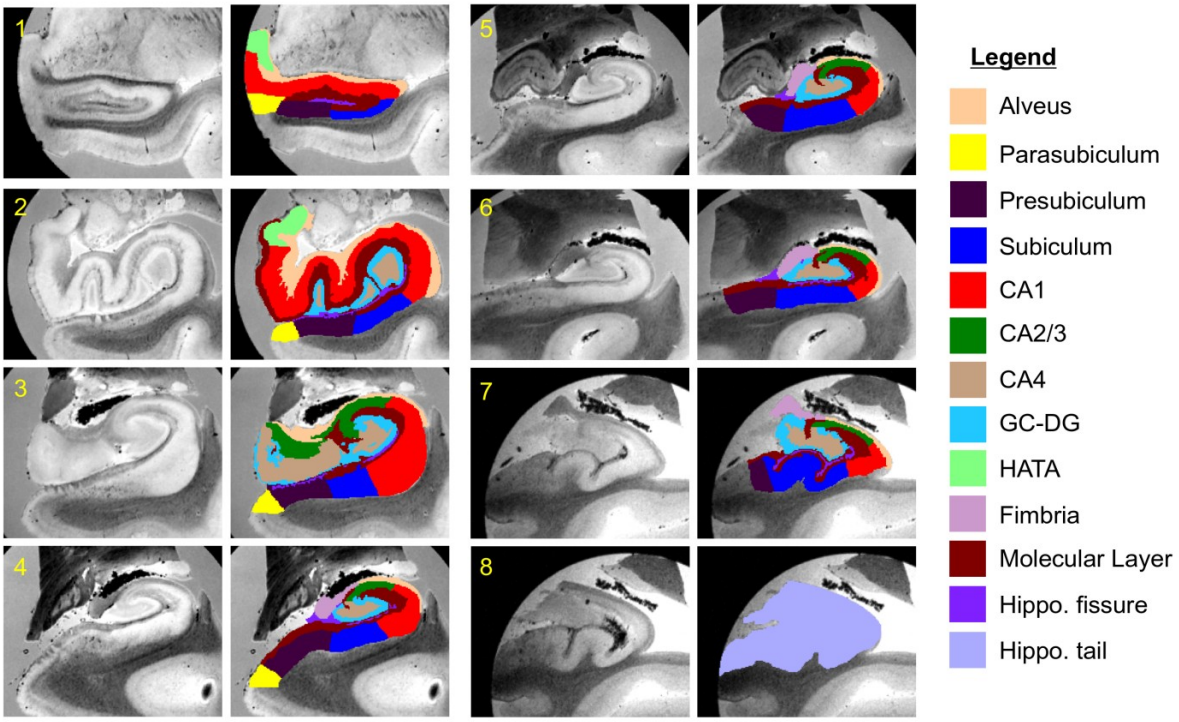
2.1.1. Inclusion/exclusion criteria

Type of drug	Number of occasions used leading to exclusion
Crack, crystal meth	>1 occasion
IV drug use	>1 occasion
Inhalants	>1 occasion
Anxiolytics, sedatives & hypnotics	>5 instances over lifetime
Hallucinogens: mushrooms	> 10 instances over lifetime
Hallucinogens: other (e.g., LSD, Ecstasy)	>5 instances over lifetime
Cocaine	>5 instances over lifetime
Opioids (e.g., oxycodone)	>5 instances over lifetime
Stimulants (e.g., amphetamine)	>5 instances over lifetime
Prescription psychostimulants (e.g. Adderall)	> 5 instances over lifetime

Appendix Table 1. Exclusion criteria for substance use

For marijuana and other cannabis products, the abuse/dependence guidelines apply, and any use within past 2 weeks or regular use commenced before age 15 lead to exclusion. This is due to the long-term lasting consequences of regular marijuana use prior to age 15 (Fontes et al., 2011). All participants had to pass a negative urine drug screen on the date of both the screening session and MRI scan.

2.3.2.1 Hippocampal subfield and amydala nuclei segmentation



Appendix Figure 1. **Visualisation of hippocampal subfields in 8 coronal slices** (Iglesias et al., 2015)[4]

[4] Check for permission completed with Elsevier June 2021, no permission necessary for non-commercial use.

3.1. Demographics and clinical characteristics

	MDD (*n*=18)	MDD/CSA (*n*=18)
Current Axis-I Comorbidities		
GAD	3	1
Social phobia	2	4
Dysthymia	1	3
Panic disorder WOA	0	2
Specific Phobia	0	3
PTSD	0	1
Lifetime Axis-I Comorbidities		
GAD	2	2
Social phobia	3	4
Dysthymia	2	6
Alcohol abuse	2	4
Panic disorder WOA	2	4
Specific Phobia	1	1
PTSD	0	5
Substance abuse	0	1

Appendix Table 2. Current (at time of study) and lifetime (past) Axis-I comorbidities of MDD and MDD/CSA groups. As assessed using the Structured Clinical Interview for the DSM-IV-TR Non-Patient Edition (SCID-IV-N/IP; First et al., 2002). Note that frequently the same participant met for current and lifetime (past) diagnosis of the same disorder so the above does not represent the number of participants with comorbidities. All Axis-I diagnoses in the SCID-IV-N/IP were assessed, only those participants met criteria for are listed above.

REFERENCES

Aghamohammadi-Sereshki, A., Coupland, N. J., Silverstone, P. H., Huang, Y., Hegadoren, K. M., Carter, R., ... & Malykhin, N. V. (2021). Effects of childhood adversity on the volumes of the amygdala subnuclei and hippocampal subfields in individuals with major depressive disorder. *Journal of psychiatry and neuroscience*, *46*(1), E186-E195.

Amico, F., Meisenzahl, E., Koutsouleris, N., Reiser, M., Möller, H. J., & Frodl, T. (2011). Structural MRI correlates for vulnerability and resilience to major depressive disorder. *Journal of psychiatry & neuroscience: JPN*, *36*(1), 15.

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of neuropsychiatry and clinical neurosciences*, *20*(3), 292-301.

Ashtari, M., Greenwald, B. S., Kramer-Ginsberg, E., Hu, J., Wu, H., Patel, M., ... & Pollack, S. (1999). Hippocampal/amygdala volumes in geriatric depression. *Psychological medicine*, *29*(3), 629-638.

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

Belleau, E. L., Treadway, M. T., & Pizzagalli, D. A. (2019). The impact of stress and major depressive disorder on hippocampal and medial prefrontal cortex morphology. *Biological psychiatry*, *85*(6), 443-453.

Bernstein, D. P., & Fink, L. (1998). CTQ: Childhood trauma questionnaire manual.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child abuse & neglect*, *27*(2), 169-190.

Boldrini, M., Santiago, A. N., Hen, R., Dwork, A. J., Rosoklija, G. B., Tamir, H., ... & Mann, J. J. (2013). Hippocampal granule neuron number and dentate gyrus volume in antidepressant-treated and untreated major depression. *Neuropsychopharmacology*, *38*(6), 1068-1077.

Boldrini, M., Underwood, M. D., Hen, R., Rosoklija, G. B., Dwork, A. J., Mann, J. J., & Arango, V. (2009). Antidepressants increase neural progenitor cells in the human hippocampus. *Neuropsychopharmacology*, *34*(11), 2376-2389.

Bremner, J. D. (2003). Functional neuroanatomical correlates of traumatic stress revisited 7 years later, this time with data. *Psychopharmacology bulletin*, *37*(2), 6-25.

Bremner, J. D., Narayan, M., Anderson, E. R., Staib, L. H., Miller, H. L., & Charney, D. S. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry*, *157*(1), 115-118.

Brown, S. S. G., Rutland, J. W., Verma, G., Feldman, R. E., Alper, J., Schneider, M., ... & Balchandani, P. (2019). Structural MRI at 7T reveals amygdala nuclei and hippocampal subfield volumetric association with major depressive disorder symptom severity. *Scientific reports*, *9*(1), 1-10.

Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology*, *30*(9), 846-856.

Calem, M., Bromis, K., McGuire, P., Morgan, C., & Kempton, M. J. (2017). Meta-analysis of associations between childhood adversity and hippocampus and amygdala volume in non-clinical and general population samples. *NeuroImage: Clinical*, *14*, 471-479.

Campbell, S., Marriott, M., Nahmias, C., & MacQueen, G. M. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry*, *161*(4), 598-607.

Carballedo, A., Lisiecka, D., Fagan, A., Saleh, K., Ferguson, Y., Connolly, G., ... & Frodl, T. (2012). Early life adversity is associated with brain changes in subjects at family risk for depression. *The World Journal of Biological Psychiatry*, *13*(8), 569-578.

Carrion, V. G., & Wong, S. S. (2012). Can traumatic stress alter the brain? Understanding the implications of early trauma on brain development and learning. *Journal of adolescent health*, *51*(2), S23-S28.

Cassiers, L. L., Sabbe, B. G., Schmaal, L., Veltman, D. J., Penninx, B. W., & Van Den Eede, F. (2018). Structural and functional brain abnormalities associated with exposure to different childhood trauma subtypes: A systematic review of neuroimaging findings. *Frontiers in psychiatry*, *9*, 329.

Chan, S. W., Harmer, C. J., Norbury, R., O'Sullivan, U., Goodwin, G. M., & Portella, M. J. (2016). Hippocampal volume in vulnerability and resilience to depression. *Journal of affective disorders*, *189*, 199-202.

Chaney, A., Carballedo, A., Amico, F., Fagan, A., Skokauskas, N., Meaney, J., & Frodl, T. (2014). Effect of childhood maltreatment on brain structure in adult patients with major depressive disorder and healthy participants. *Journal of psychiatry & neuroscience: JPN*, *39*(1), 50.

Chen, M. C., Hamilton, J. P., & Gotlib, I. H. (2010). Decreased hippocampal volume in healthy girls at risk of depression. *Archives of general psychiatry*, *67*(3), 270-276.

Cohen, M. M., Jing, D., Yang, R. R., Tottenham, N., Lee, F. S., & Casey, B. J. (2013). Early-life stress has persistent effects on amygdala function and development in mice and humans. *Proceedings of the National Academy of Sciences*, *110*(45), 18274-18278.

Colle, R., Chupin, M., Cury, C., Vandendrie, C., Gressier, F., Hardy, P., ... & Corruble, E. (2015). Depressed suicide attempters have smaller hippocampus than depressed patients without suicide attempts. *Journal of psychiatric research*, *61*, 13-18.

Colle, R., Segawa, T., Chupin, M., Dong, M. N. T. K. T., Hardy, P., Falissard, B., ... & Corruble, E. (2017). Early life adversity is associated with a smaller hippocampus in male but not female depressed in-patients: a case–control study. *BMC psychiatry*, *17*(1), 1-8.

Colodro-Conde, L., Couvy-Duchesne, B., Zhu, G., Coventry, W. L., Byrne, E. M., Gordon, S., ... & Martin, N. G. (2018). A direct test of the diathesis–stress model for depression. *Molecular psychiatry*, *23*(7), 1590-1596.

Coplan, J. D., Fathy, H. M., Jackowski, A. P., Tang, C. Y., Perera, T. D., Mathew, S. J., ... & Kaufman, J. (2014). Early life stress and macaque amygdala hypertrophy: preliminary evidence for a role for the serotonin transporter gene. *Frontiers in behavioral neuroscience*, *8*, 342.

Corbo, V., Salat, D. H., Amick, M. M., Leritz, E. C., Milberg, W. P., & McGlinchey, R. E. (2014). Reduced cortical thickness in veterans exposed to early life trauma. *Psychiatry Research: Neuroimaging*, *223*(2), 53-60.

De Bellis, M. D., & Zisk, A. (2014). The biological effects of childhood trauma. *Child and Adolescent Psychiatric Clinics*, *23*(2), 185-222.

Edwards, V. J., Holden, G. W., Felitti, V. J., & Anda, R. F. (2003). Relationship between multiple forms of childhood maltreatment and adult mental health in community respondents: results from the adverse childhood experiences study. *American Journal of Psychiatry*, *160*(8), 1453-1460.

Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *American journal of Psychiatry*, *164*(10), 1476-1488.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1996). Childhood sexual abuse and psychiatric disorder in young adulthood: II. Psychiatric outcomes of childhood sexual abuse. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(10), 1365-1374.

Finkelhor, D. (1994). The international epidemiology of child sexual abuse. *Child abuse & neglect*, *18*(5), 409-417.

Fontes, M. A., Bolla, K. I., Cunha, P. J., Almeida, P. P., Jungerman, F., Laranjeira, R. R., ... & Lacerda, A. L. (2011). Cannabis use before age 15 and subsequent executive functioning. *The British Journal of Psychiatry*, *198*(6), 442-447.

Frodl, T., Jäger, M., Smajstrlova, I., Born, C., Bottlender, R., Palladino, T., ... & Meisenzahl, E. M. (2008). Effect of hippocampal and amygdala volumes on clinical outcomes in major depression: a 3-year prospective magnetic resonance imaging study. *Journal of psychiatry & neuroscience: JPN*, *33*(5), 423.

Frodl, T., Janowitz, D., Schmaal, L., Tozzi, L., Dobrowolny, H., Stein, D. J., ... & Grabe, H. J. (2017). Childhood adversity impacts on brain subcortical structures relevant to depression. *Journal of psychiatric research*, *86*, 58-65.

Frodl, T., Meisenzahl, E. M., Zetzsche, T., Born, C., Jäger, M., Groll, C., ... & Möller, H. J. (2003). Larger amygdala volumes in first depressive episode as compared to recurrent major depression and healthy control subjects. *Biological psychiatry*, *53*(4), 338-344.

Frodl, T., Reinhold, E., Koutsouleris, N., Reiser, M., & Meisenzahl, E. M. (2010). Interaction of childhood stress with hippocampus and prefrontal cortex volume reduction in major depression. *Journal of psychiatric research*, *44*(13), 799-807.

Gerritsen, L., Van Velzen, L., Schmaal, L., van der Graaf, Y., van der Wee, N., Van Tol, M. J., ... & Geerlings, M. (2015). Childhood maltreatment modifies the relationship of depression with hippocampal volume. *Psychological medicine*, *45*(16), 3517.

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, *23*(1), 56.

Hamilton, J. P., Siemer, M., & Gotlib, I. H. (2008). Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Molecular psychiatry*, *13*(11), 993-1000.

Hammen, C., Davila, J., Brown, G., Ellicott, A., & Gitlin, M. (1992). Psychiatric history and stress: predictors of severity of unipolar depression. *Journal of abnormal psychology*, *101*(1), 45.

Hammen, C., Henry, R., & Daley, S. E. (2000). Depression and sensitization to stressors among young women as a function of childhood adversity. *Journal of consulting and clinical psychology*, *68*(5), 782.

Han, K. M., Won, E., Sim, Y., & Tae, W. S. (2016). Hippocampal subfield analysis in medication-naive female patients with major depressive disorder. *Journal of affective disorders*, *194*, 21-29.

Han, K., Kim, A., Kang, W., Kang, Y., Kang, J., Won, E., . . . Ham, B. (2019). Hippocampal subfield volumes in major depressive disorder and bipolar disorder. European Psychiatry, 57, 70-77. doi:10.1016/j.eurpsy.2019.01.016

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. (2008). The link between childhood trauma and depression: insights from HPA axis studies in humans. *Psychoneuroendocrinology*, *33*(6), 693-710.

Heim, C. M., Mayberg, H. S., Mletzko, T., Nemeroff, C. B., & Pruessner, J. C. (2013). Decreased cortical representation of genital somatosensory field after childhood sexual abuse. *American Journal of Psychiatry*, *170*(6), 616-623.

Herman, J. P., Dolgas, C. M., & Carlson, S. L. (1998). Ventral subiculum regulates hypothalamo–pituitary–adrenocortical and behavioural responses to cognitive stressors. *Neuroscience*, *86*(2), 449-459.

Hoy, K., Barrett, S., Shannon, C., Campbell, C., Watson, D., Rushe, T., ... & Mulholland, C. (2012). Childhood trauma and hippocampal and amygdalar volumes in first-episode psychosis. *Schizophrenia bulletin*, *38*(6), 1162-1169.

Huang, Y., Coupland, N. J., Lebel, R. M., Carter, R., Seres, P., Wilman, A. H., & Malykhin, N. V. (2013). Structural changes in hippocampal subfields in major depressive disorder: a high-field magnetic resonance imaging study. *Biological psychiatry*, *74*(1), 62-68.

Iglesias, J. E., Augustinack, J. C., Nguyen, K., Player, C. M., Player, A., Wright, M., ... & Alzheimer's Disease Neuroimaging Initiative. (2015). A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. *Neuroimage*, *115*, 117-137.

Jaworska-Andryszewska, P., & Rybakowski, J. K. (2019). Childhood trauma in mood disorders: neurobiological mechanisms and implications for treatment. *Pharmacological Reports*, *71*(1), 112-120.

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., ... & Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British journal of psychiatry*, *197*(5), 378-385.

Krishnan, V., & Nestler, E. J. (2011). Animal models of depression: molecular perspectives. *Molecular and functional models in neuropsychiatry*, 121-147.

Kuo, J. R., Kaloupek, D. G., & Woodward, S. H. (2012). Amygdala volume in combat-exposed veterans with and without posttraumatic stress disorder: a cross-sectional study. *Archives of General Psychiatry*, *69*(10), 1080-1086.

Levy-Gigi, E., Szabó, C., Kelemen, O., & Kéri, S. (2013). Association among clinical response, hippocampal volume, and FKBP5 gene expression in individuals with

posttraumatic stress disorder receiving cognitive behavioral therapy. *Biological psychiatry*, *74*(11), 793-800.

Lim, H. K., Hong, S. C., Jung, W. S., Ahn, K. J., Won, W. Y., Hahn, C., ... & Lee, C. U. (2012). Automated hippocampal subfields segmentation in late life depression. *Journal of affective disorders*, *143*(1-3), 253-256.

Lindert, J., von Ehrenstein, O. S., Grashow, R., Gal, G., Braehler, E., & Weisskopf, M. G. (2014). Sexual and physical abuse in childhood is associated with depression and anxiety over the life course: systematic review and meta-analysis. *International journal of public health*, *59*(2), 359-372.

Lindqvist, D., Mueller, S., Mellon, S. H., Su, Y., Epel, E. S., Reus, V. I., ... & Wolkowitz, O. M. (2014). Peripheral antioxidant markers are associated with total hippocampal and CA3/dentate gyrus volume in MDD and healthy controls–preliminary findings. *Psychiatry Research: Neuroimaging*, *224*(3), 168-174.

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.

Magarin, A. M., & McEwen, B. S. (1995). Stress-induced atrophy of apical dendrites of hippocampal CA3c neurons: involvement of glucocorticoid secretion and excitatory amino acid receptors. *Neuroscience*, *69*(1), 89-98.

Maksimovskiy, A. L., Fortier, C. B., Milberg, W. P., & McGlinchey, R. E. (2019a). A structural MRI study of differential neuromorphometric characteristics of binge and heavy drinking. *Addictive behaviors reports*, *9*, 100168.

Maksimovskiy, A. L., Oot, E. N., Seraikas, A. M., Rieselbach, M., Caine, C., Sneider, J. T., ... & Silveri, M. M. (2019b). Morphometric biomarkers of adolescents with familial risk for alcohol use disorder. *Alcoholism: Clinical and Experimental Research*, *43*(11), 2354-2366.

Malhi, G. S., Das, P., Outhred, T., Irwin, L., Gessler, D., Bwabi, Z., ... & Mannie, Z. (2019). The effects of childhood trauma on adolescent hippocampal subfields. *Australian & New Zealand Journal of Psychiatry*, *53*(5), 447-457.

McCrory, E. J., Gerin, M. I., & Viding, E. (2017). Annual research review: childhood maltreatment, latent vulnerability and the shift to preventative psychiatry–the contribution of functional brain imaging. *Journal of child psychology and psychiatry*, *58*(4), 338-357.

McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual review of neuroscience*, *22*(1), 105-122.

Mikolas, P., Tozzi, L., Doolin, K., Farrell, C., O'Keane, V., & Frodl, T. (2019). Effects of early life adversity and FKBP5 genotype on hippocampal subfields volume in major depression. *Journal of affective disorders*, *252*, 152-159.

Molnar, B. E., Buka, S. L., & Kessler, R. C. (2001). Child sexual abuse and subsequent psychopathology: results from the National Comorbidity Survey. *American journal of public health*, *91*(5), 753.

Nanni, V., Uher, R., & Danese, A. (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *American Journal of Psychiatry*, *169*(2), 141-151.

Nogovitsyn, N., Addington, J., Souza, R., Placsko, T. J., Stowkowy, J., Wang, J., ... & MacQueen, G. (2020). Childhood trauma and amygdala nuclei volumes in youth at risk for mental illness. *Psychological Medicine*, 1-8.

Opel, N., Redlich, R., Zwanzger, P., Grotegerd, D., Arolt, V., Heindel, W., ... & Dannlowski, U. (2014). Hippocampal atrophy in major depression: a function of childhood maltreatment rather than diagnosis?. *Neuropsychopharmacology*, *39*(12), 2723-2731

Ota, M., Sato, N., Hidese, S., Teraishi, T., Maikusa, N., Matsuda, H., ... & Kunugi, H. (2017). Structural differences in hippocampal subfields among schizophrenia patients, major depressive disorder patients, and healthy subjects. *Psychiatry Research: Neuroimaging*, *259*, 54-59.

Paquola, C., Bennett, M. R., & Lagopoulos, J. (2016). Understanding heterogeneity in grey matter research of adults with childhood maltreatment—A meta-analysis and review. *Neuroscience & Biobehavioral Reviews*, *69*, 299-312.

Pariante, C. M., & Lightman, S. L. (2008). The HPA axis in major depression: classical theories and new developments. *Trends in neurosciences*, *31*(9), 464-468.

Pavliša, G., Papa, J., Pavić, L., & Pavliša, G. (2006). Bilateral MR volumetry of the amygdala in chronic PTSD patients. *Collegium antropologicum*, *30*(3), 565-568.

Pechtel, P., Lyons-Ruth, K., Anderson, C. M., & Teicher, M. H. (2014). Sensitive periods of amygdala development: the role of maltreatment in preadolescence. *Neuroimage*, *97*, 236-244.

Pittenger, C., & Duman, R. S. (2008). Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology*, *33*(1), 88-109.

Posener, J. A., Wang, L., Price, J. L., Gado, M. H., Province, M. A., Miller, M. I., ... & Csernansky, J. G. (2003). High-dimensional mapping of the hippocampus in depression. *American Journal of Psychiatry*, *160*(1), 83-89.

Rao, U., Chen, L. A., Bidesi, A. S., Shad, M. U., Thomas, M. A., & Hammen, C. L. (2010). Hippocampal changes associated with early-life adversity and vulnerability to depression. *Biological psychiatry*, *67*(4), 357-364.

Rogers, M. A., Yamasue, H., Abe, O., Yamada, H., Ohtani, T., Iwanami, A., ... & Kasai, K. (2009). Smaller amygdala volume and reduced anterior cingulate gray matter density associated with history of post-traumatic stress disorder. *Psychiatry Research: Neuroimaging*, *174*(3), 210-216.

Samuels, B. A., Leonardo, E. D., & Hen, R. (2015). Hippocampal subfields and major depressive disorder. *Biological psychiatry*, *77*(3), 210.

Sapolsky, R. M. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of general psychiatry*, *57*(10), 925-935.

Sarrieau, A., Dussaillant, M., Agid, F., Philibert, D., Agid, Y., & Rostene, W. (1986). Autoradiographic localization of glucocorticosteroid and progesterone binding sites in the human post-mortem brain. *Journal of steroid biochemistry*, *25*(5), 717-721.

Saygin, Z. M., Kliemann, D., Iglesias, J. E., van der Kouwe, A. J., Boyd, E., Reuter, M., ... & Alzheimer's Disease Neuroimaging Initiative. (2017). High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: manual segmentation to automatic atlas. *Neuroimage*, *155*, 370-382.

Sheffield, J. M., Williams, L. E., Woodward, N. D., & Heckers, S. (2013). Reduced gray matter volume in psychotic disorder patients with a history of childhood sexual abuse. *Schizophrenia research*, *143*(1), 185-191.

Sheline, Y. I., Gado, M. H., & Price, J. L. (1998). Amygdala core nuclei volumes are decreased in recurrent major depression. *Neuroreport*, *9*(9), 2023-2028.

Tannous, J., Godlewska, B. R., Tirumalaraju, V., Soares, J. C., Cowen, P. J., & Selvaraj, S. (2020). Stress, inflammation and hippocampal subfields in depression: A 7 Tesla MRI Study. *Translational psychiatry*, *10*(1), 1-7.

Teicher, M. H., Anderson, C. M., & Polcari, A. (2012). Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proceedings of the National Academy of Sciences*, *109*(9), E563-E572.

Teicher, M. H., Samson, J. A., Anderson, C. M., & Ohashi, K. (2016). The effects of childhood maltreatment on brain structure, function and connectivity. *Nature reviews neuroscience*, *17*(10), 652.

Travis, S. G., Coupland, N. J., Hegadoren, K., Silverstone, P. H., Huang, Y., Carter, R., ... & Malykhin, N. V. (2016). Effects of cortisol on hippocampal subfields volumes and memory performance in healthy control subjects and patients with major depressive disorder. *Journal of affective disorders*, *201*, 34-41.

Treadway, M. T., Waskom, M. L., Dillon, D. G., Holmes, A. J., Park, M. T. M., Chakravarty, M. M., ... & Pizzagalli, D. A. (2015). Illness progression, recent stress, and morphometry of hippocampal subfields and medial prefrontal cortex in major depression. *Biological psychiatry*, *77*(3), 285-294.

Vakili, K., Pillay, S. S., Lafer, B., Fava, M., Renshaw, P. F., Bonello-Cintron, C. M., & Yurgelun-Todd, D. A. (2000). Hippocampal volume in primary unipolar major depression: a magnetic resonance imaging study. *Biological psychiatry*, *47*(12), 1087-1090.

Van Dam, N. T., Rando, K., Potenza, M. N., Tuit, K., & Sinha, R. (2014). Childhood maltreatment, altered limbic neurobiology, and substance use relapse severity via trauma-specific reductions in limbic gray matter volume. *JAMA psychiatry*, *71*(8), 917-925.

Videbech, P., & Ravnkilde, B. (2004). Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry*, *161*(11), 1957-1966.

Vyas, A., Bernal, S., & Chattarji, S. (2003). Effects of chronic stress on dendritic arborization in the central and extended amygdala. *Brain research*, *965*(1-2), 290-294.

Vyas, A., Mitra, R., Rao, B. S., & Chattarji, S. (2002). Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *Journal of Neuroscience*, *22*(15), 6810-6818.

Vythilingam, M., Heim, C., Newport, J., Miller, A. H., Anderson, E., Bronen, R., ... & Bremner, J. D. (2002). Childhood trauma associated with smaller hippocampal volume in women with major depression. *American Journal of Psychiatry*, *159*(12), 2072-2080.

Walker, E. A., Unutzer, J., Rutter, C., Gelfand, A., Saunders, K., VonKorff, M., ... & Katon, W. (1999). Costs of health care use by women HMO members with a history of childhood abuse and neglect. *Archives of general psychiatry*, *56*(7), 609-613.

Weissman, D. G., Lambert, H. K., Rodman, A. M., Peverill, M., Sheridan, M. A., & McLaughlin, K. A. (2020). Reduced hippocampal and amygdala volume as a mechanism underlying stress sensitization to depression following childhood trauma. *Depression and anxiety*, *37*(9), 916-925.

Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., & Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: data from the randomized international study to predict optimized treatment for depression. *Translational psychiatry*, *6*(5), e799-e799.

Woon, F. L., & Hedges, D. W. (2009). Amygdala volume in adults with posttraumatic stress disorder: a meta-analysis. *The Journal of neuropsychiatry and clinical neurosciences*, *21*(1), 5-12.

Xie, P., Wu, K., Zheng, Y., Guo, Y., Yang, Y., He, J., ... & Peng, H. (2018). Prevalence of childhood trauma and correlations between childhood trauma, suicidal ideation, and social support in patients with depression, bipolar disorder, and schizophrenia in southern China. *Journal of affective disorders*, *228*, 41-48.

Yuan, M., Rubin-Falcone, H., Lin, X., Rizk, M. M., Miller, J. M., Sublette, M. E., ... & Mann, J. J. (2020). Smaller left hippocampal subfield CA1 volume is associated with reported childhood physical and/or sexual abuse in major depression: A pilot study. *Journal of affective disorders*, *272*, 348-354.

Zeidman, P., & Maguire, E. A. (2016). Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nature Reviews Neuroscience*, *17*(3), 173-182.

Study 4: Increased medial orbitofrontal cortex and insula cortical thickness in depression is independent of early life stress

Franziska Goer[1], Emily L. Belleau[2,3], Diego A. Pizzagalli[2,3], Rebecca Elliott[1]
[1]University of Manchester, Oxford Road, Manchester M13 9PL, UK
[2]McLean Hospital, 115 Mill St, Belmont, MA 02478, USA
[3]Harvard Medical School, 25 Shattuck St, Boston, MA 02115, USA

Author Contributions:

Franziska Goer contributed to study design, assisted in recruitment of participants and data collection, conducted all analyses, wrote the manuscript and completed all revisions following review by the co-authors.

Professor Rebecca Elliott provided methodological guidance and supervisory input, gave guidance and feedback on analyses, and critically reviewed and provided feedback on drafts of the manuscript prepared by FG.

Dr. Diego Pizzagalli designed the overall study from which data was used. Dr. Pizzagalli and Dr. Emily Belleau gave guidance and supervisory input for the project, and helped design and gave feedback on all analyses.

Funding Information:

This study was study was supported by a grant from NIMH, awarded to Dr. Diego Pizzagalli, grant number 5R01 MH095809. Franziska Goer was supported by funding from the Medical Research Council and University of Manchester as part of Franziska Goer's Medical Research Council (MRC) Doctoral Training Scholarship and MRC DTP flexible training grant.

ABSTRACT

Background: Studies of cortical thickness in major depressive disorder (MDD) and early life stress (ELS) have produced heterogenous findings, and very few studies to date have accounted for both variables in the same sample, thereby making it difficult to differentiate potential effects of MDD and/or ELS contributing to changes in brain morphology.
Aim/Methods: This study aimed to tease out potential effects of MDD and childhood sexual abuse (CSA) on cortical thickness in five a priori selected regions of interest (ROIs). Anatomical scans of 75 women (aged 19-44), divided into four groups according to presence/absence of MDD and CSA, were collected using 3T MRI. Semi-automated FreeSurfer processing pipelines were applied to compute cortical thickness of the parahippocampal gyrus, medial orbitofrontal cortex (mOFC), posterior cingulate cortex, anterior cingulate cortex, and insula.
Results: The right mOFC (p = 0.003) and left insula (p = 0.002) showed significantly greater cortical thickness in MDD than healthy controls, irrespective of CSA. Trends were observed for other regions but did not survive conservative Bonferroni corrections. We also found tentative evidence (p = 0.006) for reductions in cortical thickness of the parahippocampal gyrus in CSA, though this also did not survive correction for multiple comparisons.
Conclusion: While limited by a relatively small sample size, our findings suggest that MDD, independent of CSA, may be associated with increased cortical thickness in certain regions. Future studies of greater sample size and including whole brain analyses are needed to investigate whether these observations replicate, are generalizable to other types of ELS, and may extend to other brain regions not included in this study's analysis.

1. Introduction

Studies of cortical thickness in major depressive disorder (MDD) have produced highly heterogenous results. Two recent meta-analyses have aimed to address these discrepancies but nevertheless yielded differing findings: Suh and colleagues (2019) found cortical thinning in frontal and occipital regions and thickening in one parietal area in MDD while Li et al. (2020) found increased cortical thickness in posterior and anterior cingulate cortices (PCC and ACC, respectively) and the ventromedial prefrontal cortex (vmPFC) and decreased cortical thickness in the gyrus rectus, superior frontal gyrus and middle temporal gyrus in unmedicated MDD compared to HC. While the former included studies of both medicated and unmedicated MDD in their meta-analysis, this is unlikely to fully explain the difference in findings as the authors were able to partially replicate their findings in a subgroup of unmedicated samples only (Suh et al., 2019). Studies have also identified thinning of the parahippocampal gyrus in MDD (which was also present, to a lesser degree, in unaffected young adults at high familial risk of depression; Papmeyer et al., 2015).

Studies of cortical thickness in healthy adults exposed to early life stress (ELS) report cortical thinning in many cortical and subcortical structures (Bounoua et al., 2020; Dannlowski et al., 2012; Gold et al., 2016; Heim et al., 2013; Kelly et al., 2013; Mclaughlin et al., 2014; Monninger et al., 2020). However, as is frequently an issue in ELS research, these studies differed in definition and measures of ELS and few controlled for psychiatric symptoms/history. While some studies measured total ELS exposure encompassing emotional/physical/sexual abuse and physical/emotional neglect (Dannlowski et al., 2012; Heim et al., 2013; Kelly et al., 2013) others focused their ELS analyses on specific trauma subtypes, such as physical and/or sexual abuse (Gold et al., 2016), assaultive trauma

(including physical violence such as assault/robbery, exposure to an active war zone, or physical/sexual abuse; Bounoua et al., 2020), or neglect (in the form of institutional rearing; McLaughlin et al., 2014). Heim and colleagues (2013) conducted additional analyses to investigate whether findings of cortical thinning in ELS (using a composite score of 5 main subtypes of abuse and neglect) may be driven specifically by childhood sexual abuse (CSA) and found that CSA, independent of other forms of abuse or depressive symptoms, explained cortical thinning of the somatosensory cortex and bilateral parahippocampal gyrus.

To our knowledge, only one study has analysed cortical thickness in a sample measuring both MDD and ELS, which found a complex relationship between ELS, age of onset of MDD, and MDD diagnosis on the right frontal pole, however, despite a very large number of statistical tests performed bilaterally on each region provided in the FreeSurfer output, no correction for multiple comparisons was performed (Jaworska et al., 2014). Furthermore, cortical thickness in ELS was only analysed within the depressed group (with no comparison to the HC group) making it more difficult to differentiate effects of MDD versus ELS. Analyses compared abused MDD (exposed to physical and/or sexual abuse) to non-abused MDD, however, both groups had similar levels of other forms of ELS, including emotional neglect/abuse and physical neglect, suggesting that perhaps the findings were exclusive the physical and/or sexual abuse rather than general ELS.

1.1. Aim and hypotheses

The aim of the cortical thickness analyses in this study was to investigate the effects and potential interaction of MDD and ELS on a priori selected regions of interest (ROIs). Since

different types of trauma, age of exposure, and gender may have differential effects on brain morphometry and structure (Cassiers et al., 2018; Teicher et al., 2016), we chose to investigate the effect of CSA exposure in women between ages 5-14 (previously identified as a potential sensitive period during which CSA may have particularly deleterious effects on the developing brain; Andersen et al., 2008) on cortical thickness in five a priori selected ROIs that have been implicated in cortical thickness changes in MDD and ELS.

Given previous literature, we hypothesised that MDD would be characterised by increased cortical thickness in the ACC, PCC, medial orbitofrontal cortex (mOFC), and insula, and cortical thinning in the parahippocampal gyrus. We hypothesised that CSA would be associated with widespread cortical thinning of structures, particularly the parahippocampal gyrus which has been specifically implicated in CSA. Given the dearth of previous literature investigating cortical thickness in samples measuring and analysing both ELS and MDD, we did not have specific predictions on possible interactions between CSA and MDD on cortical thickness.

2. Methods

2.1. Participants

Detailed description of participants, demographics, inclusion/exclusion criteria, and measures can be found in **(**Goer et al., unpublished findings, **Study 3**). The study took place at McLean Hospital, Belmont, MA, USA, and was approved by both the Partners Healthcare and McLean Hospital Institutional Review Boards. In brief, 75 unmedicated women, aged 19-44 (M=26.51 , SD=6.32) were recruited, including 23 healthy controls (HC), with no CSA or lifetime psychiatric disorder; 18 MDD, with current MDD but no CSA; 16 healthy controls,

with CSA but no lifetime psychiatric disorder (HC/CSA); and 18 MDD/CSA, with both current MDD and CSA. The groups did not significantly differ on demographic characteristics including age, race, education or income (see Goer et al., unpublished findings, **Study 3**). CSA was assessed in the interview version of the Traumatic Antecedents Questionnaire (TAQ; Herman et al., 1989; Vanderkolk et al., 1991). Individuals were classed as meeting criteria for CSA if they experienced at least one incident of contact sexual abuse between the ages of 5-14 years. As CSA rarely occurs in a vacuum and frequently co-occurs with other types of ELS, we permitted additional types of abuse and neglect in the CSA groups while any type of ELS was exclusionary for our MDD and HC groups. The short form of the CTQ (Bernstein et al., 2003), a 28-item questionnaire that retrospectively assesses 5 subtypes of childhood trauma (physical abuse, emotional abuse, sexual abuse, physical neglect, emotional neglect) was collected only partway through the study and hence a third of participants (25/75), including 4 MDD/CSA and 1 HC/CSA did not have CTQ scores. Therefore, CTQ scores were analysed for descriptive purposes in the CSA groups only to assess levels of other subtypes of ELS in addition to CSA. Previously validated and widely applied cut-off scores were calculated to indicate significant levels of abuse/neglect in each subscale: ≥8 for physical abuse, ≥8 for sexual abuse, ≥10 for emotional abuse, ≥8 for physical neglect, and ≥15 for emotional neglect (Bernstein & Fink, 1998; Walker et al., 1999). Presence of current and past psychiatric disorders was evaluated via Structured Clinical Interview for the DSM-IV-TR Non-Patient Edition (SCID-IV-N/IP; First et al., 2002).

2.2. Neuroimaging data acquisition and processing

A Siemens TrioTim 3T scanner with 32-channel head coil was used to collect high-resolution T1-weighted MPRAGE scans (TR=2200ms, TE=3.36ms, voxel size = 1.2 x 1.2 x 1.2 mm, flip

angle = 7.0 deg, base resolution = 192, slices per slab = 144, field of view = 230mm). The FreeSurfer (version 7.0; Fischl, 2012) semi-automated workflow was used to process structural T1 data. Following the cortical reconstruction process, quality checks and manual inspection of each of the 75 subjects were conducted and manual edits of the brainmask and white matter volume were applied where necessary.

2.2.1. Cortical thickness

Cortical thickness was determined using the validated and highly accurate automated processing pipelines in FreeSurfer 7.0 which calculate the shortest possible distance between the pial surface and grey/white matter at each point (for details, see Fischl & Dale, 2000). All cortical thickness measures are given in mm.

2.2.2. ROIs and analyses

We conducted analyses of cortical thickness in 5 bilateral ROIs selected a priori based on MDD and ELS literature (and availability of these regions in the automatically computed cortical thickness in FreeSurfer 7.0): (1) parahippocampal gyrus (associated with reduced cortical thickness in MDD (Peng et al., 2015) and ELS (Gold et al., 2016)), (2) mOFC (associated with increased cortical thickness in MDD (Grieve et al., 2013; Qiu et al., 2014)), (3) PCC and (4) ACC (both associated with increased cortical thickness in MDD (Li et al., 2020)) and (5) insula (increased cortical thickness implicated in current MDD (Zorlu et al., 2017) and possibly predictive of MDD risk (Jones et al., 2019; Foland-Ross et al., 2015)). It is important to note that, as discussed in the introduction, several additional ROIs have been implicated in cortical thickness in ELS and MDD, however, we limited ourselves to 5 key

areas identified in the literature to avoid excessive multiple comparisons and achieve satisfactory power to identify group differences.

Each hemisphere of each ROI was analysed with a 2 (MDD: present versus absent) x 2 (CSA: present versus absent) between-subjects ANOVA. The conservative Bonferroni correction was implemented as a stringent correction for multiple comparisons, resulting in a new *p* value threshold of $p<0.005$ for statistical significance (original alpha of 0.05 / 10 comparisons = 0.005 corrected alpha).

3. Results

3.1 ELS and clinical characteristics

There was no difference in CSA severity between CSA groups ($p>0.05$), however, the duration of CSA was significantly longer for MDD/CSA (M = 5.18, SD = 3.046) than HC/CSA (M = 2.88, SD = 3.053), $F(1,31)= 4.696$, $p= 0.038$, $\eta p^2 = 0.132$, and onset of CSA was earlier in MDD/CSA (M = 6.59, SD = 1.734) than HC/CSA (M = 8.94, SD = 3.549), $F(1,31) = 5.949$, $p=0.021$, $\eta p^2 = 0.161$. As expected, 100% of CSA participants, regardless of diagnosis, met the CTQ cut-off for significant sexual abuse. Of note, a significantly higher number of CSA/MDD than HC/CSA participants met for additional other types of ELS, including emotional abuse, physical abuse, emotional neglect, and physical neglect (see **Table 1** for detailed statistics). While, as expected, there was a significant difference between all four groups in depressive symptoms as measured using the Beck Depression Inventory-II (BDI-II; Beck et al, 1996) and Hamilton Rating Scale for Depression (HRSD; Hamilton et al., 1960), these measures did not significantly differ between HC versus HC/CSA and MDD versus MDD/CSA. Specifically, for BDI-II scores, there was no significant difference between HC (M

= 1.461, SD = 1.825) and HC/CSA (M = 2.260, SD = 2.939), $F(1,36) = 1.067$, $p = 0.309$ or between MDD (M = 29.577, SD = 8.651) and MDD/CSA (M = 28.093, SD = 10.119), $F(1,34) = 0.224$, $p = 0.639$. Likewise, total scores on the HRSD did not significantly differ between HC (M = 1.00, SD = 1.508) and HC/CSA (M = 1.56, SD = 2.502), $F(1, 37) = .767$, $p = 0.387$ or between MDD (M = 14.11, SD = 4.886) and MDD/CSA (M = 15.38, SD = 3.284), $F(1.34) = .763$, $p = 0.389$. Furthermore, MDD groups were well matched in number of symptoms within the current major depressive episode (MDE) and number of lifetime MDEs (see **Table 1** for statistics).

	Participants, Mean (SD)						
Characteristic	HC (*n*=23)	MDD (*n*=18)	HC/CSA (*n*=16)	MDD/CSA (*n*=18)	Statistic	*P* value	Partial Eta Squared
TAQ CSA severity	NA	NA	3.38 (1.025)	4.00 (1.118)	$F = 2.792$	$p = 0.105$	0.083
TAQ CSA duration (years)	NA	NA	2.88 (3.052)	5.18 (3.046)	$F = 4.696$	**$p = 0.038$***	0.132
TAQ CSA onset (age in years)	NA	NA	8.94 (3.549)	6.59 (1.734)	$F = 5.949$	**$p = 0.021$***	0.161
# (%) meeting CTQ cut-off for EA	NA	NA	4 (25%)	9 (50%)	$X^2 = 4.144$	**$p = 0.042$***	
# (%) meeting cut-off for PA	NA	NA	2 (12.5%)	7 (38.9%)	$X^2 = 4.549$	**$p = 0.033$***	
# (%) meeting CTQ cut-off for SA	NA	NA	15 (100%)	14 (100%)			
# (%) meeting CTQ cut-off for EN	NA	NA	2 (12.5%)	7 (38.9%)	$X^2 = 4.549$	**$p = 0.033$***	
# (%) meeting CTQ cut-off for PN	NA	NA	2 (12.5%)	8 (44.5%)	$X^2 = 6.152$	**$p = 0.013$***	
BDI-II	1.461 (1.825)	29.577 (8.651)	2.260 (2.939)	28.09 (10.119)	$F = 98.34$	**$p = 0.00$***	0.808
HRSD	1.00 (1.508)	14.11 (4.886)	1.56 (2.502)	15.38 (3.284)	$F = 107.159$	**$p = 0.00$***	0.823

Number of symptoms in current MDE	NA	6.00 (1.414)	NA	5.118 (1.35)	*F* = 3.551	*p* = 0.068	0.097
Lifetime number of MDEs	NA	M = 3.35, SD = 3.427	NA	3.58, SD = 3.029	*F* = 0.035	*p* = 0.853	0.001

Table 1. ELS and clinical characteristics.
Mean values are displayed with standard deviations in parentheses where applicable.
***** indicates statistical significance at p<0.05
Abbreviations: BDI-II, Beck Depression Inventory-II (Beck et al, 1996); HRSD, Hamilton Rating Scale for Depression (Hamilton et al., 1960); MDD, major depressive disorder; MDE, major depressive episode; CSA, childhood sexual abuse; HC, healthy control; NA, not applicable; TAQ, Traumatic Antecedents Questionnaire (Herman et al., 1989; Vanderkolk et al., 1991); CTQ, childhood trauma questionnaire; EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect. *Note:* CTQ scores were missing for 1 HC/CSA and 4 MDD/CSA.

3.2. Medial orbitofrontal cortex

There was a significant main effect of MDD in the right (but not left) mOFC, $F(1,71)=9.816$, $p=0.003$, $\eta p^2 = 0.121$, with significantly greater cortical thickness in MDD (M=2.524, SE= 0.023) versus no MDD (M=2.424, SE = 0.022).

3.3. Insula

There was a significant main effect of MDD in the left insula, $F(1,71)=10.746$, $p=0.002$, $\eta p^2 = 0.131$, with greater thickness in MDD (M=3.102, SE=0.025) than no MDD (M=2.987, SE=0.024). Cortical thickness was similarly increased in MDD versus no MDD subjects in the right insula, though not significant at the corrected alpha level, $F(1,71)=7.558$, $p=0.008$, $\eta p^2 = 0.096$.

3.4. Parahippocampal gyrus

There was a main effect of MDD (greater cortical thickness in MDD) in the left ($p=0.010$) and right ($p=0.037$) hemisphere of the parahippocampal gyrus, but neither of these survived

Bonferroni corrections. There was also a main effect of CSA in the left parahippocampal gyrus (reduced cortical thickness in CSA) but again this failed to survive Bonferroni correction $F(1,71)=7.992$, $p=0.006$, $\eta p^2=0.090$.

3.5. Posterior cingulate cortex

There was a trend of greater cortical thickness in MDD in the right posterior cingulate at uncorrected alpha level ($p=0.055$), but no effects emerged that met statistical significance in either hemisphere.

3.6. Anterior cingulate cortex

There was a main effect of MDD (greater cortical thickness in MDD) in the left (but not right) ACC, $p=0.022$, which did not survive Bonferroni correction.

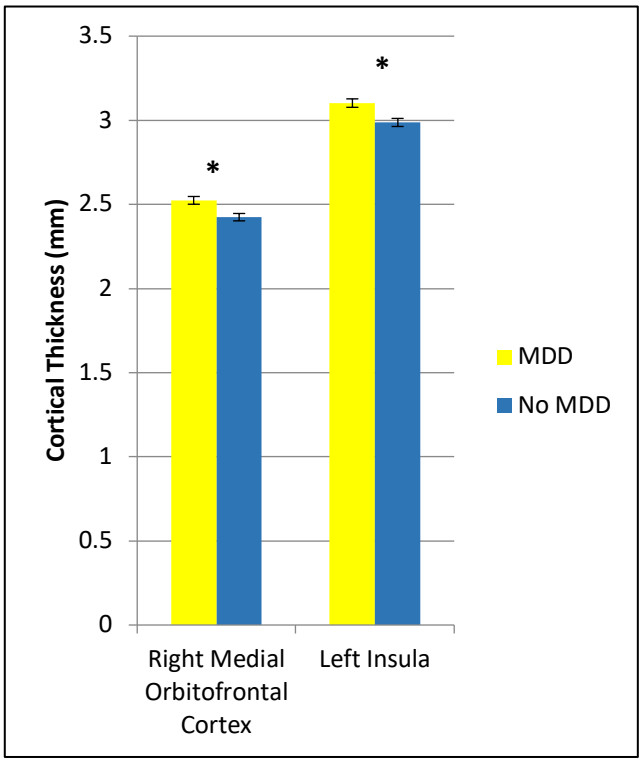


Figure 1. Depressed individuals had a significantly thicker right medial orbitofrontal cortex (p=0.003) and left insula (p=0.002) than individuals with no MDD. *indicates significance at corrected alpha, p<0.005.

4. Discussion

The right mOFC cortex and left insula showed significantly greater cortical thickness in MDD than HC, independent of CSA (see **Figure 1**). A subregion of the vmPFC (Gourley et al., 2016), the mOFC serves several key functions in goal-directed decision making and reward-learning. Interestingly, activity in the anteromedial OFC (located within the vmPFC) has been found to be heightened in current MDD and is positively correlated with depression severity, while successful antidepressant treatment has been found to reverse this effect (Drevets, 2007). In terms of cortical thickness, increased thickness in the mOFC has previously been identified in MDD versus HC (Grieve et al., 2013; Qiu et al., 2014) and a recent meta-analysis identified the overarching vmPFC as an area of increased cortical thickness in unmedicated MDD (Li et al., 2020). Increased cortical thickness of the insula has also been previously observed in MDD (Qiu et al., 2014; Zorlu et al., 2017) and has been found to predict development of MDD risk in middle aged adults (Jones et al., 2019) and adolescent girls (Foland-Ross et al., 2015). These findings suggest that increased insula thickness may constitute a risk factor for the development of depression, rather than a result of long-term depression. However, a study examining cortical thickness prior to antidepressant treatment found thinner insula in MDD versus HC (Järnum et al., 2011), indicating the need for additional studies, ideally with larger sample sizes, to aim to resolve inconsistent findings.

While not statistically significant after a conservative Bonferroni correction for multiple comparisons, we identified several trends that further suggest cortical thickness in these 5 ROIs may be increased in MDD, independently of CSA. The left and right parahippocampal gyrus, right posterior cingulate, left ACC and right insula all showed trends of increased cortical thickness in MDD versus no MDD, irrespective of CSA. While increased thickness in the PCC and ACC in MDD is in line with findings from a recent large meta-analysis (Li et al., 2020), the opposite finding (decreased thickness) has previously been observed in the parahippocampal gyrus (Peng et al., 2015; Qiu et al., 2014).

The mechanism underlying increased cortical thickness in certain brain areas in MDD remains unclear. Li and colleagues (2020) posited that this may be due to increased inflammation in MDD which in turn may result in astrocyte hypertrophy as a potential compensatory attempt. However further studies are needed to investigate this potential link. Further insight into cortical thickening in certain areas in MDD has come from a longitudinal study which found that while reduced cortical thickness in frontal and temporal brain regions appeared to constitute a risk factor to the development of MDD, the onset of MDD was associated with an increased thickening of these areas over time compared to high-risk patients who did not develop MDD (Papmeyer et al., 2015). The authors posit that while high risk patients who do not go on to develop MDD show normal cortical thinning of these structures over time (likely representing successful synaptic pruning), this mechanism may be partially disrupted or delayed by the onset of MDD and hence lead to relative cortical thickening in MDD patients.

While several studies have reported cortical thinning of various brain structures following ELS, we only found a trend of a main effect of CSA in the parahippocampal gyrus (reduced cortical thickness in CSA, p=0.006). As previously discussed, studies of cortical thickness in ELS have used a large variety of measures and definitions to define ELS and this may explain differences in brain areas identified. Interestingly, other studies that have also focused on CSA have similarly identified cortical thinning of the parahippocampal gyrus, both in a sample with physical and/or sexual abuse (Gold et al., 2016) and more specifically in a study that found cortical thinning of the bilateral parahippocampus to be one of two regions exclusively affected by CSA, independent of other types of abuse (Heim et al., 2013). Two studies in women with PTSD have found no effect of sexual abuse on cortical thickness in either adults (Landré et al., 2010) or adolescents (Rinne-Albers et al., 2020), however, the former was not restricted to CSA (inclusion criteria for sexual abuse did not specify an age range of occurrence) and the former selected a limited number of ROIs (including the vmPFC, ACC, and insula, but not the parahippocampal gyrus). Null findings regarding the insula, vmPFC and ACC following CSA in PTSD (Rinne-Albers et al., 2020) mirror our findings. Taken together, these initial findings highlight the importance of differentiating between types of childhood trauma in ELS studies as they may have differential effects on brain morphometry, with CSA perhaps specifically implicated in cortical thinning of the parahippocampal gyrus. Further studies of larger sample size are necessary to permit whole brain analyses of cortical thickness to identify whether potentially additional brain regions may be particularly affected by CSA.

It must be noted that the present investigation was limited by several factors. For one, our analyses were restricted by a relatively small sample size – it has been suggested that well-

powered studies of cortical thickness require at least 50 participants per group to detect a 0.25mm cortical thickness difference (Pardoe et al., 2012), and hence our analyses were likely underpowered. Future studies of cortical thickness in ELS and MDD should aim for significantly larger sample sizes to achieve sufficient statistical power to detect cortical thickness differences between groups at a whole brain level. Larger sample sizes would also allow for a more detailed analysis of different types of childhood trauma and age at occurrence to determine whether different types of ELS at particular ages during brain development may affect cortical thickness differentially. Our study was also limited by its cross-sectional design - the finding that cortical thinning in certain temporal and frontal areas (which have also been documented following ELS) may constitute a risk factor for the development of MDD is intriguing (Papmeyer et al., 2015), and further longitudinal studies measuring cortical thickness, ELS and MDD at regular intervals are needed to understand whether there may be a possible link between cortical thickness and risk for MDD following ELS. Further studies of cortical thickness before and after MDD onset are also needed to test the theory that increased cortical thickness may develop as MDD progresses (Papmeyer et al., 2015).

Finally, while all participants in both CSA groups met cut-offs for significant CSA using both the CTQ and TAQ measures and had comparable levels of CSA severity, there were significant differences in duration of CSA (longer in MDD/CSA) and age of onset (younger in MDD/CSA). Ideally, these factors would be matched between groups in future studies to rule out a possible contribution of age of onset or duration of CSA on findings. Furthermore, the MDD/CSA group had higher levels of other types of abuse/neglect (as measured by the CTQ) than the HC/CSA group. Since no MDD by CSA interactions emerged in the analyses,

this difference in ELS between groups should not affect the results of the present investigation, however, future studies should ideally recruit participants with comparable levels of ELS subtypes across diagnostic groups to avoid any possible confounding factors. Moreover, since different types of abuse frequently co-occur in the same individual and hence make it extremely difficult to recruit groups with exclusively one type of ELS exposure, future studies should ideally collect data on multiple types of abuse (e.g. via the CTQ) for all participants in order to run specific analyses aimed at investigating the unique contributions and/or additive effects of various types of ELS on cortical thickness. A small sample size and missing CTQ data points (including for 5 CSA participants) unfortunately made this unfeasible in the current study.

5. Conclusion

Depressed individuals had a significantly thicker right mOFC (p=0.003) and left insula (p=0.002) than individuals with no MDD, independent of CSA, suggesting that, at least in this sample and ROIs examined, changes in cortical thickness may be primarily driven by MDD diagnosis, rather than CSA exposure. However, this is likely an oversimplification, as a trend was found for decreased cortical thickness in the parahippocampal gyrus in CSA (irrespective of MDD diagnosis), replicating findings of previous studies of CSA. Taken together, these findings suggest that MDD and CSA likely have differential, independent effects on cortical thickness in the ROIs examined. Future studies, ideally with substantially larger sample sizes and longitudinal designs, are needed to clarify the effect of MDD and CSA on cortical thickness on a whole brain level, and ascertain whether perhaps other variables, such as type and timing of ELS and MDD age of onset/duration, may affect cortical thickness and identify whether these changes precede or follow development of MDD

symptoms.

REFERENCES

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of neuropsychiatry and clinical neurosciences*, *20*(3), 292-301.

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

Bernstein, D. P., & Fink, L. (1998). CTQ: Childhood trauma questionnaire manual.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child abuse & neglect*, *27*(2), 169-190.

Bounoua, N., Miglin, R., Spielberg, J. M., & Sadeh, N. (2020). Childhood assaultive trauma and physical aggression: Links with cortical thickness in prefrontal and occipital cortices. *Neuroimage: clinical*, *27*, 102321.

Cassiers, L. L., Sabbe, B. G., Schmaal, L., Veltman, D. J., Penninx, B. W., & Van Den Eede, F. (2018). Structural and functional brain abnormalities associated with exposure to different childhood trauma subtypes: A systematic review of neuroimaging findings. *Frontiers in psychiatry*, *9*, 329.

Dannlowski, U., Stuhrmann, A., Beutelmann, V., Zwanzger, P., Lenzen, T., Grotegerd, D., ... & Kugel, H. (2012). Limbic scars: long-term consequences of childhood maltreatment revealed by functional and structural magnetic resonance imaging. *Biological psychiatry*, *71*(4), 286-293.

Drevets, W. C. (2007). Orbitofrontal cortex function and structure in depression. *Annals of the New York Academy of Sciences*, *1121*(1), 499-527.

First M, Spitzer R, Gibbon M and Williams J (2002) Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-Patient Edition. (SCID-I/NP). New York: Biometrics Research, New York State Psychiatric Institute.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050-11055.

Fischl B. (2012). FreeSurfer. NeuroImage, 62(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Foland-Ross, L. C., Sacchet, M. D., Prasad, G., Gilbert, B., Thompson, P. M., & Gotlib, I. H. (2015). Cortical thickness predicts the first onset of major depression in adolescence. *International Journal of Developmental Neuroscience*, *46*, 125-131.

Gold, A. L., Sheridan, M. A., Peverill, M., Busso, D. S., Lambert, H. K., Alves, S., ... & McLaughlin, K. A. (2016). Childhood abuse and reduced cortical thickness in brain regions involved in emotional processing. *Journal of Child Psychology and Psychiatry*, *57*(10), 1154-1164.

Gourley, S. L., Zimmermann, K. S., Allen, A. G., & Taylor, J. R. (2016). The medial orbitofrontal cortex regulates sensitivity to outcome value. *Journal of Neuroscience*, *36*(16), 4600-4613.

Grieve, S. M., Korgaonkar, M. S., Koslow, S. H., Gordon, E., & Williams, L. M. (2013). Widespread reductions in gray matter volume in depression. *NeuroImage: Clinical*, *3*, 332-339.

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, *23*(1), 56.

Heim, C. M., Mayberg, H. S., Mletzko, T., Nemeroff, C. B., & Pruessner, J. C. (2013). Decreased cortical representation of genital somatosensory field after childhood sexual abuse. *American Journal of Psychiatry*, *170*(6), 616-623.

Herman JL, Perry JC and Vanderkolk BA (1989) Childhood trauma in bor- derline personality-disorder. American Journal of Psychiatry 146, 490–495.

Järnum, H., Eskildsen, S. F., Steffensen, E. G., Lundbye-Christensen, S., Simonsen, C. W., Thomsen, I. S., ... & Larsson, E. M. (2011). Longitudinal MRI study of cortical thickness, perfusion, and metabolite levels in major depressive disorder. *Acta Psychiatrica Scandinavica*, *124*(6), 435-446.

Jaworska, N., MacMaster, F. P., Gaxiola, I., Cortese, F., Goodyear, B., & Ramasubbu, R. (2014). A preliminary study of the influence of age of onset and childhood trauma on cortical thickness in major depressive disorder. *BioMed research international*, *2014*.

Jones, E. C., Liebel, S. W., Hallowell, E. S., & Sweet, L. H. (2019). Insula thickness asymmetry relates to risk of major depressive disorder in middle-aged to older adults. *Psychiatry Research: Neuroimaging*, *283*, 113-117.

Kelly, P. A., Viding, E., Wallace, G. L., Schaer, M., De Brito, S. A., Robustelli, B., & McCrory, E. J. (2013). Cortical thickness, surface area, and gyrification abnormalities in children exposed to maltreatment: neural markers of vulnerability?. *Biological psychiatry*, *74*(11), 845-852.

Landré, L., Destrieux, C., Baudry, M., Barantin, L., Cottier, J. P., Martineau, J., ... & El Hage, W. (2010). Preserved subcortical volumes and cortical thickness in women with sexual abuse-related PTSD. *Psychiatry Research: Neuroimaging*, *183*(3), 181-186.

Li, Q., Zhao, Y., Chen, Z., Long, J., Dai, J., Huang, X., ... & Sweeney, J. A. (2020). Meta-analysis of cortical thickness abnormalities in medication-free patients with major depressive disorder. *Neuropsychopharmacology*, *45*(4), 703-712.

McLaughlin, K. A., Sheridan, M. A., Winter, W., Fox, N. A., Zeanah, C. H., & Nelson, C. A. (2014). Widespread reductions in cortical thickness following severe early-life deprivation: a neurodevelopmental pathway to attention-deficit/hyperactivity disorder. *Biological psychiatry*, *76*(8), 629-638.

Monninger, M., Kraaijenvanger, E. J., Pollok, T. M., Boecker-Schlier, R., Jennen-Steinmetz, C., Baumeister, S., ... & Holz, N. E. (2020). The long-term impact of early life stress on orbitofrontal cortical thickness. *Cerebral Cortex*, *30*(3), 1307-1317.

Papmeyer, M., Giles, S., Sussmann, J. E., Kielty, S., Stewart, T., Lawrie, S. M., ... & McIntosh, A. M. (2015). Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder. *Biological psychiatry*, *78*(1), 58-66.

Pardoe, H. R., Abbott, D. F., Jackson, G. D., & Alzheimer's Disease Neuroimaging Initiative. (2013). Sample size estimates for well-powered cross-sectional cortical thickness studies. *Human brain mapping*, *34*(11), 3000-3009.

Peng, D., Shi, F., Li, G., Fralick, D., Shen, T., Qiu, M., ... & Fang, Y. (2015). Surface vulnerability of cerebral cortex to major depressive disorder. *PLoS One*, *10*(3), e0120704.

Qiu, L., Lui, S., Kuang, W., Huang, X., Li, J., Zhang, J., ... & Gong, Q. (2014). Regional increases of cortical thickness in untreated, first-episode major depressive disorder. *Translational psychiatry*, *4*(4), e378-e378.

Rinne-Albers, M. A., Boateng, C. P., van der Werff, S. J., Lamers-Winkelman, F., Rombouts, S. A., Vermeiren, R. R., & van der Wee, N. J. (2020). Preserved cortical thickness, surface area and volume in adolescents with PTSD after childhood sexual abuse. *Scientific reports*, *10*(1), 1-9.

Suh, J. S., Schneider, M. A., Minuzzi, L., MacQueen, G. M., Strother, S. C., Kennedy, S. H., & Frey, B. N. (2019). Cortical thickness in major depressive disorder: a systematic review and meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *88*, 287-302.

Teicher, M. H., Samson, J. A., Anderson, C. M., & Ohashi, K. (2016). The effects of childhood maltreatment on brain structure, function and connectivity. *Nature reviews neuroscience*, *17*(10), 652.

Vanderkolk BA, Perry JC and Herman JL (1991) Childhood origins of self- destructive behavior. American Journal of Psychiatry 148, 1665–1671.

Walker, E. A., Unutzer, J., Rutter, C., Gelfand, A., Saunders, K., VonKorff, M., ... & Katon, W. (1999). Costs of health care use by women HMO members with a history of childhood abuse and neglect. *Archives of general psychiatry*, *56*(7), 609-613.

Zorlu N, Cropley VL, Zorlu PK, Delibas DH, Adibelli ZH, Baskin EP, et al. Effects of cigarette smoking on cortical thickness in major depressive disorder. J Psychiatry Res. 2017;84:1–8.

Discussion

1. Summary of main findings

The overarching aim of my PhD was to investigate the relative contributions (and possible interactions) of MDD and ELS on cognitive and neurobiological measures, specifically affective cognition and morphological measures (GMV and cortical thickness). A summary of findings is given below, with reference to hypotheses presented in the introduction.

Study 1 employed a novel standardised affective cognition test battery (EMOTICOM, Bland et al., 2016) to assess various constructs of affective cognition in individuals with MDD and HC with varying degrees of ELS (diagnostic groups were matched for ELS and included both participants with low/absent ELS exposure and with significant ELS of multiple subtypes). Our results were in line with our general a priori hypothesis that ELS and MDD independently affect different aspects of affective cognition. Constructs particularly sensitive to MDD diagnosis (as indicated by significant group differences: MDD versus HC) included emotion recognition and categorisation, incentive motivation/effort, and emotional memory. On the other hand, several tasks yielded no group effects (MDD versus HC) but instead revealed significant correlations with CTQ. Tasks more sensitive to ELS over diagnosis included measures of attentional bias, value-based choice, and moral emotions. Overall null results were recorded for the MIR task (a measure of reward/punishment sensitivity), and the remaining three social cognition tasks (including measures of theory of mind and social economic exchange games).

Our initial specific hypothesis that MDD would be associated with negative affective biases (negative versus positive) while ELS would be characterised by a threat (anger/fear) processing bias was not quite supported. Rather, both ELS and MDD were characterised by both threat biases and negative affective biases in various tasks. The exact nature of the relationship observed depended on the task/construct measured – both MDD diagnosis and PA were associated with a negative affective bias in the ERT, while the Emotional Intensity Morphing task revealed both a threat bias (for fearful faces) and negative affective bias (for sad faces) in MDD, independent of ELS. A negative affective bias in the FAGN task (measuring attentional bias) was, on the other hand, exclusively associated with various types of ELS (subtypes & total CTQ), independent of diagnosis. Hence our a priori hypothesis that ELS and MDD would be distinguished by different types of affective bias was not supported. Instead, both ELS and MDD were found to be associated with both threat and negative affective biases in different tasks of the EMOTICOM test battery. It should also be noted that contradictory results (not supporting a negative affective bias) were observed for the emotional memory task, which found that MDD and EA were associated with poorer memory for negative words, divergent from evidence from previous literature.

Our second specific hypothesis pertaining to the EMOTICOM test battery was that depression would be characterised by reduced motivation/reward function, which we predicted may be, at least partially, moderated by ELS. Our results largely confirmed this hypothesis, and again emphasised that effects of MDD and ELS depend on the affective cognition construct measured. Specifically, while MDD appeared to be associated with reduced motivation/effort as measured during the PRT, independent of ELS (though analyses were severely limited by task ceiling effects), ELS independently predicted reduced

optimisation of value-based choice in the adapted Cambridge Gambling Task. The MIR task, on the other hand, yielded null results, suggesting that perhaps reward/punishment sensitivity as measured using this task is less sensitive to ELS/MDD (or more difficult to detect).

The third specific hypothesis had predicted heightened levels of negative moral emotions in MDD, possibly mediated by ELS. No group differences emerged (MDD versus HC), but multiple subtypes and cumulative ELS were associated with higher levels of shame, guilt, and negative ratings, particularly of victims of harm. These effects of ELS on moral emotions appeared particularly prevalent within MDD participants.

Finally, null results emerged for theory of mind and social economic exchange games for which we had not made previous hypotheses based on limited and conflicting previous literature.

Study 2 consisted of a systematic review of GMV changes in MDD and ELS (including only studies that assessed both MDD and ELS in the same sample). Of 5129 records identified through database searches (including grey/unpublished literature), 20 studies met full eligibility criteria for inclusion in the systematic review. Analyses of studies revealed that the hippocampus was the most frequently studied and identified brain structure. Moreover, hippocampal volume reductions (both for the whole hippocampus and several subfields, particularly the cornu ammonis, dentate gyrus, and subiculum) appeared to be driven primarily by ELS, independent of MDD diagnosis. This was similarly observed for several other brain regions, including the caudate, orbitofrontal cortex, and posterior cingulate

cortex, which were all found to have lower GMV in ELS, independent of MDD diagnosis. These findings confirmed our initial hypothesis, as set out in the general introduction, that ELS partially or fully drives the reduction of GMV often attributed to MDD, particularly in the hippocampus and its subfields. However, it should be noted, that many studies included in the systematic review either altogether lacked a HC group with ELS or had very low levels of ELS in their HC comparison group, thereby limiting findings and making it more difficult to draw firm conclusions about relative contributions of ELS versus MDD on GMV changes.

The results from **Study 2** informed our decision to investigate GMV changes in the hippocampus (both whole and subfields) in a highly controlled experimental setup in **Study 3**, aiming to meticulously disentangle MDD and ELS effects on GMV. In particular, we were interested in hippocampal subfields as initial evidence from the systematic review pointed to an effect of ELS (independent of MDD) on specific hippocampal subfields, however this had only been investigated in 5 studies (and only 3 included measures of ELS in the HC group). Due to a novel FreeSurfer development allowing for segmentation of the amygdala and its nuclei, this was also included as an exploratory analysis that may help shed more light on the to date mixed amygdala GMV findings in MDD and ELS. Given evidence that both type and timing of ELS differentially affects GMV (Cassiers et al., 2018; Anderson et al., 2008), ELS was specified as childhood sexual abuse (CSA; though other types of abuse were not excluded for ecological validity and practical limitations in recruitment), and confined to ages 5-14 (which may represent a sensitive time period for CSA on hippocampal development, Anderson et al., 2008).

Results largely confirmed our hypothesis that reductions in hippocampal GMV and its subfields (particularly the cornu ammonis [CA] and dentate gyrus [DG]) would be partially or fully mediated by CSA. While we did not find any evidence for either a main effect of MDD or CSA (or interaction) on the whole hippocampus, several regions of the hippocampal head showed reduced GMV related to CSA exposure, independent of diagnosis. Subfields identified included the cornu ammonis (specifically CA3 and CA4 head), dentate gyrus (specifically the GC-ML-DG, granule cell and molecular layer of the DG), presubiculum head, and whole hippocampal head.

No a priori hypothesis had been formulated for the amygdala or its nuclei given inconsistent previous findings and results only revealed a ROI*CSA interaction for the right hemisphere amygdala nuclei (though no significant follow up tests emerged, making the finding difficult to interpret). Based on visual inspection of means, this appeared to be driven by decreased GMV for the majority of amygdala nuclei in CSA compared to no CSA (irrespective of MDD), though none of the post hoc tests of individual nuclei was significant. Further research is needed to investigate this potential link between CSA and amygdala nuclei GMV, ideally with larger sample sizes with the necessary power to detect and fully analyse any potential effects. Overall, **Study 3** demonstrated that ELS appears to drive GMV reduction in several hippocampal subfields, independent of MDD diagnosis and may have a similar effect on amygdala nuclei.

Finally, **Study 4** investigated another morphological index, cortical thickness. Neuroimaging data from the same 75 unmedicated female participants was analysed using FreeSurfer software. Five a priori regions of interest were selected based on previous findings of

cortical thickness in MDD and ELS in an effort to reduce multiple comparisons. While limited by a fairly small sample size for identifying differences in cortical thickness (Pardoe et al., 2012), results indicated that cortical thickness may be more sensitive to MDD diagnosis, rather than CSA, with the exception of cortical thinning of the parahippocampal gyrus which showed the opposite relationship (though this was not statistically significant after applying conservative Bonferroni correction for multiple comparisons). In particular, the right medial orbitofrontal cortex and left insula showed significantly greater cortical thickness in MDD than HC, irrespective of CSA. Additional trends in the same direction (increased thickness in MDD versus HC, irrespective of CSA) were found for the posterior and anterior cingulate cortex, though findings were not significant after Bonferroni correction. These findings were predominantly in line with a priori hypothesis, in which we predicted increased cortical thickness in all ROIs apart from the parahippocampal gyrus. Furthermore, our prediction that cortical thinning of the parahippocampal gyrus may be driven primarily by CSA, rather than diagnosis, appeared to be supported, though results were just short of statistical significance after Bonferroni correction.

The main results from each study are summarised in **Table 1** below.

Experimental Measure (Task/Brain region)	**Main construct measured**	**MDD** (versus HC)	**ELS**
Study 1. EMOTICOM			
Emotional Recognition Task (ERT)	Emotion recognition/ categorisation	✓	✓ Correlations with EA*, PA
Emotional Intensity Morphing Task	Emotion recognition	✓	✓ Correlations with SA (but opposite effect to MDD)
Face Affective Go No-Go (FAGN) Task	Attentional bias	✗	✓ PA, SA, EA, CTQ total

Emotional Memory Recognition Task	Emotional memory	✓	✓ Correlation with EA*
Monetary Incentive Reward (MIR) Task	Reward/ punishment sensitivity	✗	✗
Progressive Ratio Task (PRT)	Incentive motivation / effort	✓	✗
Adapted Cambridge Gambling Task (CGT)	Value-based choice	✗	✓ All subscales & CTQ total
Moral Emotions Task	Moral emotion	✗	✓ EA, PA, EN, PN, CTQ Total
Social Information Preference Test	Theory of mind	✗	✗
Prisoners' Dilemma (PD)	Social economic exchange game	✗	✗
Ultimatum Game (UG)	Social economic exchange game	✗	✗
Study 2. GMV systematic review			
Hippocampus	GMV (Whole hippocampal volume and subfields)		✓ ↓
Caudate	GMV		✓ ↓
Orbitofrontal cortex	GMV		✓ ↓
Posterior Cingulate Cortex	GMV		✓ ↓
Study 3. GMV of hippocampus and amygdala (including subfields/nuclei)			
Hippocampus whole	GMV	✗	✗
Hippocampal subfields	GMV		✓ ↓ (several hippocampal head subfields)
Amygdala whole	GMV	✗	✗
Amygdala subfields	GMV	✗	ROI*CSA interaction for right hemisphere nuclei but no significant post hoc follow up tests for individual nuclei
Study 4. Cortical thickness of 5 ROIs			
Medial Orbitofrontal Cortex	Cortical thickness	✓ ↑	✗

Insula	Cortical thickness	✓ ↑	✗
Parahippocampal gyrus	Cortical thickness	✗	✓ ↓ (trend, *p*=0.006, did not survive Bonferroni correction)
Anterior Cingulate Cortex	Cortical thickness	✓ ↑ (trend, *p*=0.022, did not survive Bonferroni correction)	✗
Posterior Cingulate Cortex	Cortical thickness	✓ ↑ (trend, *p*=0.055, did not survive Bonferroni correction)	✗

Table 1. Overview of Main Results from Studies 1-4. Green checkmarks indicate statistically significant findings (unless otherwise indicated), red crosses indicate null results. Blue arrows are used to indicate direction (increase or decrease) for neuroimaging findings.
*For study 1, correlations with EA should be cautiously interpreted for tasks also showing an effect of diagnosis (i.e. ERT and Emotional Memory Recognition Task) since EA was significantly higher in MDD versus HC group.
Note that ELS for studies 3 and 4 was defined as at least 1 instance of CSA aged 5-14 (other types of abuse were not exclusionary), while study 1 included all subtypes of ELS as measured by the CTQ (EA, PA, SA, EN, PN). Studies included in the systematic review (study 2) used a variety of ELS definitions (predominantly total CTQ scores).
Abbreviations: EA, emotional abuse; PA, physical abuse; SA, sexual abuse; EN, emotional neglect; PN, physical neglect; MDD, major depressive disorder; HC, healthy control; CTQ, childhood trauma questionnaire CTQ; (Bernstein et al., 2003); ROI, region of interest; GMV, grey matter volume.

1.1. Novel contributions to knowledge base

Studies 1-4 each generated novel evidence and extended the knowledge base from existing literature. Specifically, **Study 1** demonstrated that affective cognition is a multifaceted construct that is differentially affected by MDD and ELS and hence requires extensive, thorough, and systematic assessment. It was the first study to apply a standardised comprehensive testing battery of affective cognition in a sample measuring both MDD and ELS. Similarly, **Study 2** was the first systematic review, to the authors knowledge, to assess GMV changes in samples measuring both ELS and MDD. Analyses confirmed emerging theoretical approaches that ELS may be a driving factor of neurobiological changes observed

in MDD: ELS appeared to independently predict GMV reductions of the hippocampus (including its subfields) and various other brain structures often implicated in MDD. Critical review of studies included in the systematic review informed us in the study design and analyses of a cross-sectional neuroimaging study of GMV in MDD and ELS (**Study 3**). This study added substantially to the existing knowledge base by applying new analysis methods for amygdala nuclei segmentation and hippocampal subfields and investigating GMV changes in these structures in MDD and ELS. While findings for the amygdala were unclear, several hippocampal subfields, previously implicated in MDD (and also in animal models of depression and ELS), showed attenuated GMV in ELS, independently of diagnosis. This builds on similar findings from the systematic review (**Study 2**) and highlights the role of ELS in structural brain changes in MDD. Finally, **Study 4** investigated cortical thickness, a morphological measure previously only assessed in a single study measuring both MDD and ELS (Jaworska et al., 2014). Our study was the first to include a HC group with matching levels of ELS (in the form of CSA) to the MDD group. This enabled analyses to disentangle the effect of MDD and CSA and demonstrate that changes in cortical thickness appear to be primarily sensitive to depression status, independent of CSA. Overall, the studies included in the PhD offer a novel evidence base for differential unique effects of MDD and ELS on various cognitive and neurobiological measures, using highly controlled experimental designs.

2. Overarching themes, mechanisms, and implications

2.1. As demonstrated by **Studies 1-4**, ELS likely accounts for several of the cognitive and neurobiological changes previously often attributed to MDD alone. This has large implications for both past and future studies, as the vast majority of previous studies of

depression have not controlled for ELS in their samples. This may have resulted in a misattribution of certain characteristics to MDD, when in fact high levels of ELS could potentially have been driving (or at least mediating) this effect. This is particularly concerning given the high levels of ELS in depression (Williams et al., 2016; Xie et al., 2018), which may be contributing to experimental variables without being measured or controlled for. Importantly, however, the influence of ELS and MDD on various morphological measures and constructs of affective cognition appeared to be specific, i.e. while some measures were more sensitive to ELS (independent of diagnosis), others showed the opposite pattern (an effect of MDD but not ELS), an effect of both (for some affective cognition constructs), or null results for either. As hypothesised, the relationship between ELS and MDD and various neural and cognitive markers appears to be complex and dependent on specific measures/constructs.

Constructs particularly sensitive to ELS (independently of MDD diagnosis) included attentional bias, value-based choice, and moral emotions, as well as GMV changes of various brain areas (in particular the hippocampus and its subfields). Interestingly, negative affective biases as observed for ELS in the FAGN task (measuring attentional bias) were also observed for MDD in other tasks that focussed more on emotion identification and categorisation. This suggests that negative affective biases in emotion processing may be affected by both ELS and MDD independently, depending on the specific task and construct measured. This is in line with previous research that has repeatedly highlighted the role of affective biases in MDD (Elliott et al., 2011), and ELS (Pechtel & Pizagalli, 2011). The fact that ELS appeared to be driving impaired optimisation of bets (for both win and loss conditions in the adapted CGT) is in line with research that has found similar patterns of impairment in

medicated, unmedicated and remitted MDD (Murphy et al., 2001; Rawal et al., 2013; Roiser & Sahakian, 2013) suggesting that current depressive symptoms are not driving this effect. **Study 1** hence provided novel evidence that ELS may in fact be the key factor leading to abnormal value-based choice previously attributed to MDD. While only investigated by a handful of previous studies, findings that heightened moral emotions in MDD may are in fact driven by ELS, are in line with emerging evidence that self-blame may play a key role in mediating the link between ELS exposure and later life adverse psychiatric outcomes, such as PTSD and depression (Dorresteijn et al., 2019; Sharma-Patel & Brown, 2016).

In regard to morphological indices examined, **Study 2** and **3** (consisting of a systematic review and cross-sectional study), both highlighted the key role of ELS (independent of MDD diagnosis) on decreased GMV in several brain regions, in particular of the hippocampus and several subfields (comprising the cornu ammonis, dentate gyrus, and presubiculum). Hippocampal GMV reduction has been consistently cited in MDD, and is seen as one of the largely undisputed structural brain changes of depression (Bremner et al., 2000; Campbell et al., 2004; Videbech & Ravnkilde, 2004). Findings from both the systematic review (**Study 2**) and the original research study (**Study 3**) included in this PhD add to the emerging, but to date sparse, literature emphasising the role of ELS in GMV abnormalities in MDD. Future studies hence should take care to include measures of ELS in all structural studies of depression, and control for these in analyses, to avoid potential misattribution of findings.

Interestingly, other measures of affective cognition and brain structure showed the reverse pattern, and seemed predominantly sensitive to MDD diagnosis over ELS. In regard to affective cognition, this included tasks measuring emotion recognition/categorisation,

incentive motivation/effort and emotional memory. However, it should be noted that some of these tasks were less reliable or findings were not in line with previous literature, and hence should be cautiously interpreted. Specifically, the progressive ratio task (PRT) designed to assess motivation/effort demonstrated extremely high ceiling effects, with over 88% of participants completing the full task and hence not having a breakpoint. While the small subset of participants with breakpoints was analysed and led to the conclusion that a significantly greater number of MDD than HC participants quit the task, this should be cautiously interpreted as evidence for an independent or significant contribution of MDD on reduced motivation/effort. While the finding is in line with a robust previous literature, these studies have primarily employed other paradigms, such as the Effort- Expenditure for Rewards Task (EEfRT; Treadway et al., 2009) which may represent a better measure for motivation in MDD. Similarly, cautious interpretation of the effect of MDD on memory observed in **Study 1** should be employed, as the result was unexpected given previous literature. Rather than confirm either a reduced positive affective bias or increased negative affective bias in memory recall (both of which have been consistently implicated in MDD, e.g. Hamilton & Gotlib, 2008; Ridout et al., 2003), we observed the opposite effect of poorer memory recall for negative stimuli specifically, in MDD relative to HC. Hence, while the general finding of MDD being sensitive to affective memory biases in MDD was confirmed in our study, the direction was opposite to our hypothesis based on an extensive previous literature. More conclusive, and in line with previous studies, was the negative affective bias observed in MDD (versus HC) in emotion recognition and identification, which was found to be independent of ELS. The fact that MDD appears to be driving this effect is consistent with studies of antidepressants that have shown reversal of such negative affective biases in MDD, even prior to any mood or symptom changes (Harmer et al., 2009).This further

indicates that MDD is likely central to these observed negative affective biases of emotion processing, rather than a pre-existing environmental risk factor such as ELS.

Finally, while evidence from this PhD indicates that GMV changes in MDD may be attributable to ELS, investigations of cortical thickness (in **Study 4)** indicated the opposite pattern. With the exception of the parahippocampal gyrus (which demonstrated cortical thinning likely driven independently by ELS), all other regions of interest (ROIs) were associated with increased cortical thickness in MDD, irrespective of ELS. Previous studies of cortical thickness in MDD have been inconsistent, with two recent meta-analyses reporting divergent findings (Li et al., 2020; Suh et al., 2019). One possibility raised for inconsistent findings by the authors had been differences in medication status in included studies, however controlling for antidepressant use did not seem to explain these differences. Given the findings from **Study 4**, it is possible that differences in previous studies of cortical thickness in MDD may be in part due to not controlling for ELS in experimental samples. Given that ELS has been associated with the opposite findings, namely widespread cortical thinning in various brain regions (Bounoua et al., 2020; Dannlowski et al., 2012; Gold et al., 2016; Heim et al., 2013; Kelly et al., 2013; Mclaughlin et al., 2014; Monninger et al., 2020), this may potentially interfere with findings in MDD samples with high levels of unmeasured ELS.

In summary, the main overarching conclusion of the studies included in this PhD is that different constructs of affective cognition and different measures of structural brain changes appear to differentially and specifically be affected by ELS or MDD. Given that our analyses all controlled for ELS and ensured matched HC groups with comparable ELS (or

specifically CSA) levels were included, findings can be more firmly attributed to either MDD or ELS than has often been the case in previous studies. This distinction is important as it may inform mechanisms underlying ELS as a risk factor for depression and elucidate more targeted treatments/interventions.

2.1. Mechanisms

Mechanisms of how affective cognition and structural brain changes may confer risk from ELS exposure to later development of MDD have been discussed in more detail in the relevant included studies, but a brief overview (to avoid unnecessary repetition) is given here.

From a neurobiological perspective, reductions in hippocampal volume following ELS has been largely attributed to dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis (Heim et al., 2000). HPA axis hyperactivity and atrophy of the hippocampus have also been observed in animals following stress exposure (Krishnan & Nestler, 2011; Lupien et al., 2008; Magarin & McEwen, 1995). Stress exposure leads to increased secretion of neurotoxic glucocorticoids (Burke et al., 2005; Pittinger et al., 2008) which appear to have particularly deleterious effects on the hippocampus (Lupien et el., 2008; McEwen, 1999; Sapolsky, 2000). This also may explain the particular sensitivity of hippocampal subfields implicated in **Study 3**, as the cornu ammonis and dentate gyrus have particularly high densities of glucocorticoid receptors, which has been theorised to underlie dendritic atrophy in these hippocampal subfields following chronic stress exposure in animal studies (Samuels et al., 2015; Vyas et al., 2002; Vyas et al., 2003). Sensitisation of the HPA-axis following ELS may not only explain hippocampal GMV loss in ELS, but also provide a link to development of

MDD. This mechanism may explain why those with significant ELS histories are more likely to develop a major depressive episode in response to stressful life events than those without ELS exposure (Hammen et al., 2000). However, the exact mechanism underlying the connection between ELS, GMV hippocampal abnormalities, and MDD is still unknown, and likely depends on a complex interaction of various biological processes, including inflammation, oxidative stress, and neurotransmitter abnormalities, in addition to HPA-axis sensitisation (Belleau et al., 2018).

The mechanisms underlying abnormalities of cortical thickness in MDD and ELS are less well studied and understood. In general, it has been posited that exposure to significant ELS during sensitive periods of brain development interferes with various essential processes of brain development, including neurogenesis and pruning of synapses (Pechtel & Pizzagalli, 2011; Teicher et al., 2006a; Teicher et al., 2006b). Independent of ELS, cortical thickness changes in MDD have been hypothesised to be related to increased inflammation in current depression, which may lead to astrocyte hypertrophy as a compensatory mechanism, thus leading to increases in cortical thickness (Li et al., 2020). However, further research is needed to test this hypothesis.

Disruptions in affective cognition may represent another key mechanism linking ELS and MDD. As shown in **Study 1**, several affective cognition constructs appeared primarily modulated by ELS, irrespective of MDD. The changes observed, including a negative affective bias of attention, heightened negative moral emotions, and abnormalities in value-based choice, have previously been implicated in MDD and cognitive models of depression, that highlight the role of such changes in the aetiology and maintenance of the disorder (for

a review see Elliott et al., 2011; Roiser & Sahakian, 2013). These models emphasise the interplay between affective cognitive changes and several key brain areas, including the amygdala and ventromedial frontal cortex (Elliott et al., 2011). Further evidence of the key role of affective cognition changes in MDD have come from studies of antidepressant drugs and psychological interventions (reviewed below), which have suggested that reversal of affective biases may represent a key aspect of the efficacy of such treatments.

2.2. Clinical Implications

The results from this PhD have several clinical implications. Previous research has proposed that affective cognition changes in MDD may be a prime target for antidepressant treatment (Harmer et al., 2011; Clark et al., 2009). Extensive research has demonstrated that affective biases observed in MDD can be reversed using both pharmacological and psychological treatments (see Browning et al., 2010 for a review). Of note, studies of antidepressant drugs (including selective serotonin reuptake inhibitors [SSRIs] such as citalopram, and selective norepinephrine reuptake inhibitors [SNRIs] such as reboxetine) have demonstrated improved affective memory bias (increased recall for positive versus negative stimuli) following administration in HC (Harmer et al., 2004), a finding that was replicated using autobiographical stimuli (rather than experimental stimuli such as emotional words; Papadatou-Pastou et al., 2012). Interestingly, different types of antidepressant treatment may have unique effects on affective cognition. For instance, one study found that selective serotonin reuptake inhibitor (SSRI) led to decreased neural activation in response to both pleasant and aversive stimuli (McCabe et al., 2010) which authors propose may relate to SSRIs potentially limited benefit for patients with anhedonia. SSRIs have also been found to affect emotion processing, specifically in an emotional face

morphing task in which healthy female participants had a higher rate of detection for fearful and happy faces (and quicker reaction times) than participants given placebo (saline solution), while no changes in anger, disgust, or sadness was observed (Harmer et al., 2003).

Similar effects to antidepressant treatments may be observed in psychological interventions designed to reverse affective cognitive biases. A study using positive mental imagery as part of a cognitive bias modification (CBM) in depression found significant reduction in cognitive bias and depressive symptoms following treatment, relative to depressed individuals who were assigned to the control condition and did not undergo CBM (Lang et al., 2012).

Studies of the effect of antidepressant treatment on measures of affective cognition in MDD have, to the authors knowledge, not included measures of ELS. It would be interesting to investigate whether potentially the beneficial effects of antidepressant treatment in reversing affective biases may be specific to those constructs that showed heightened sensitivity to MDD, irrespective of ELS. Some evidence for this may be inferred, as many affective processes studied in the aforementioned studies focussed on emotional memory, and emotion processing (including emotion categorisation/identification), which we found to be influenced by MDD diagnosis, independently of ELS, in **Study 1**.

Few studies have specifically investigated treatment responses for depressed patients with high levels of ELS. A large multicentre study of almost 700 participants found that psychotherapy alone (consisting of a blend of cognitive behavioural therapy and interpersonal therapy) was superior to antidepressant treatment alone (using the atypical antidepressant Nefazadone) in individuals with current MDD and a history of ELS (Nemeroff

et al., 2005). A combination of both treatments was the most effective, though not substantially more so than psychotherapy alone. While the authors conclude that this highlights the need for psychotherapy in MDD cohorts with significant ELS, it should be noted that this study limited its investigation to a single type of antidepressant medication (Nefazadone) and hence cannot generalise to other types of antidepressant drugs. Furthermore, the study used a unique definition of ELS and the second most common type of ELS reported, after PA, was parental loss before age 15. Hence findings might not be quite comparable to other studies focusing on CTQ subscales of ELS. Further research is clearly needed to investigate whether specific types of antidepressant treatment may be particularly beneficial for patients with high levels of ELS.

In addition to being a target of antidepressant treatment, changes in affective cognition may also represent a potential future biomarker used for early detection of later depression (Roiser & Sahakian, 2013). In particular for tasks sensitive to risk factors of MDD, such as ELS, this may provide a unique opportunity to identify individuals who may benefit from treatment to prevent the sequalae that may lead to later onset depression. However, this is merely a hypothetical proposal, and significant research into the viability of tasks, such as from the EMOTICOM test battery, to serve as individual predictors of later adverse outcomes is needed.

Finally, in regard to structural brain changes associated with ELS and MDD, some research has indicated that antidepressant treatment may partially reverse hippocampal GMV reductions observed in studies of MDD (Frodl et al., 2008). While the precise mechanism underlying this effect is unclear, some have proposed that this may be the result of

antidepressants increasing neurogenesis in hippocampal cells, thereby leading to an increase in GMV in this region (Boldrini et al., 2009). These studies, have however, not included measures of ELS, so it is unclear how these factors may interact. Future studies are needed to investigate the use of antidepressants in reversing hippocampal atrophy in individuals with and without MDD and with and without significant levels of ELS, to better understand potential benefits and mechanisms.

2.3. Implications for future research

Several implications for future research have already been noted. First and foremost, this PhD has highlighted the need for future studies of depression and affective cognition and/or brain structure to include measures of ELS and control for this in analyses. Furthermore, several findings from past studies of depression, in particular relating to hippocampal volume reductions and certain constructs of affective cognition, may need to be re-evaluated, as emerging evidence (including from this PhD) suggests that several neural and cognitive constructs may have been misattributed to MDD (when in fact ELS may be driving these observed changes). However, future studies are needed to replicate these findings and elucidate the mechanisms underlying observed changes.

It must be noted that ELS is by no means an exclusive risk factor for MDD. In fact, ELS confers heightened risk for the development of multiple psychiatric disorders, including psychosis (Varese et al., 2012), bipolar disorder (Palmier-Claus et al., 2016), anxiety disorders, substance use, and eating disorders, amongst others (Green et al., 2010). Specific disorders may be primarily associated with certain subtypes of ELS, as indicated by a meta-analysis that identified physical/sexual abuse and neglect as most prevalent in depression

and emotional abuse as most frequent in schizophrenia (Carr et al., 2013). ELS may represent a transdiagnostic risk factor, possibly via shared mechanisms giving rise to different symptoms. Altered affective cognition may underlie this link as a construct known to be affected by ELS (as seen in **Study 1**) and implicated in various disorders in addition to MDD. There is therefore a need to replicate the approach of this PhD in other mental health problems, which raises issues of feasibility and tolerability.

2.3.1. Schizophrenia and affective cognition – Pilot Study

For instance, schizophrenia, similarly to MDD, has been consistently associated with various impairments in affective cognition. Specifically, schizophrenia has been associated with difficulty recognizing and labelling facial expressions accurately regardless of valence (Marsh and Williams, 2006; Pomaroi-Clotet et al., 2010), reduced memory recall for positive (but not neutral or negative) stimuli (Herbener, 2008; Herbener et al., 2007), dysfunctional reinforcement learning (Murray et al., 2008; Roiser et al., 2009; Waltz et al., 2007), reduced incentive motivation (Gold et al., 2013), and marked deficits in social cognition, including theory of mind (Bora et al., 2009; Brüne, 2005). Despite a comparably small lifetime risk of schizophrenia of approximately 1% versus approximately 25% for MDD (Beddington et al., 2008; Green et al., 2005), the resulting disability is so severe that it accounts for an estimated annual £11.8 billion societal and £7.2 billion public sector cost in the UK alone (Andrew et al., 2012). In addition to severely reduced quality of life, schizophrenia is also associated with a more than 2.5 times greater mortality rate than the general UK population (Andrew et al., 2012). Moreover, while antipsychotic drugs commonly prescribed for schizophrenia and other psychotic disorders may successfully reduce positive symptoms (such as hallucinations and delusions) they appear to have little effect on negative

symptoms, cognition, or functional outcomes (Green et al., 2012; Wunderink et al., 2013). Similarly, while randomised controlled trials of cognitive behavioural therapy in schizophrenia have demonstrated some success in symptom reduction, generalizability to social functioning outcomes and quality of life is poor (Cather et al., 2005; Garety et al., 1997; Gumley et al., 2003). Affective cognition may represent a unique avenue to explore a potential mechanism linking ELS exposure to development of schizophrenia, and possibly offer insight into novel treatment targets, similarly as for MDD.

A small pilot study was conducted as part of the PhD, to assess feasibility of assessment of affective cognition using the novel EMOTICOM test battery (Bland et al., 2016) in a cohort of participants with schizophrenia. 6 participants, with a primary diagnosis of schizophrenia, were recruited and completed both clinical evaluation (including the The Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998), Positive and Negative Symptoms Scale (PANSS; Kay et al., 1987), the Calgary Depression Scale for Schizophrenia (Addington et al., 1993), and the same EMOTICOM test battery employed in **Study 1**. Participants also completed the childhood trauma questionnaire (CTQ, Bernstein et al., 2003) and provided demographic information. All 6 participants completed the full study and task performance was checked to confirm non-random responding. The only significant difference to the MDD and HC participants taking part in the same experimental setup in **Study 1** was longer time to complete the cognitive testing. The need for frequent breaks and reduced attention after longer stretches of cognitive testing meant that participants completed the clinical interview and tasks in 3-4 separate study sessions (rather than 2 as for the HC and MDD participants). This may also be partially attributable to the testing environment, which was in a communal space of a mental health rehabilitation unit, rather than a dedicated

separate testing room. Due to small numbers, unfortunately no meaningful statistical analyses could be conducted, however, the smooth process of data collection and full datasets obtained provide promising feedback for the use of the EMOTICOM test battery in schizophrenia. Future studies should seek to obtain data from a larger sample of individuals with schizophrenia, to assess the hypothesis that ELS may represent a transdiagnostic risk factor affecting various constructs of affective cognition and possibly relate to treatment response and/or functional outcomes.

3. Strengths and limitations

3.1. Strengths

The studies included in this PhD had several strengths. Firstly, the multi-method approach adopted enabled us to analyse possible effects of ELS and MDD on various different experiment variables. Furthermore, highly controlled study designs with matching levels of ELS across both HC and MDD groups were a unique advantage compared to many previous studies in this area, as it allowed us to disentangle effects of ELS and MDD on neural and cognitive variables of interest. In regard to affective cognition, the use of the validated, standardised EMOTICOM test battery allowed us to comprehensively and systematically assess multiple types of affective cognition, rather than limiting analyses to a single or few tasks, as has often been previously done. In regard to neuroimaging analyses of structural brain changes, use of FreeSurfer software, version 7.0, enabled segmentation of hippocampal and amygdala subfields, and as such offered a more nuanced and detailed analysis of structures of interest. Another strength of the PhD was the thorough clinical assessment and inclusion of exclusively medication free MDD participants, ensuring no possible confounding factors through antidepressant medication use, which have been

shown to affect both affective cognition and morphological indices. Finally, since ELS cannot be viewed as a single monolithic construct, we took care to report and analyse subtypes of ELS (EA, PA, SA, EN, PN) or limit inclusion criteria to specific type and timing of abuse (such as CSA in **Studies 3** and **4**). Many studies in ELS focus solely on cumulative ELS exposure before age 18 (in the form of total CTQ score), and as such may miss more intricate findings that may be specific to certain types of abuse.

3.2. Limitations

While the studies included in this paper generated multiple lines of novel evidence, several limitations should be noted. Specific limitations pertaining to measures/tasks used in the included studies have been described in detail in the individual papers and throughout the discussion (and hence shall not be restated to avoid unnecessary repetition), however, there are a few overarching limitations shared by all studies that are mentioned here.

A key limitation of this PhD, and common in ELS research more generally, pertains to retrospective assessment of ELS. Prospective longitudinal studies remain the gold standard, however, these are often not feasible due to time and financial constraints. Some prospective studies have confirmed the reliability of the CTQ, such as a study using a longitudinal design which measured actual exposure to physical abuse (assessed in both face to face and computer based self-interviews) at regular intervals during childhood (aged 9-16), and compared these with CTQ scores collected during a later follow-up (aged 18-23), and found the latter to provide a reliable indication (consistent with previous reporting during childhood) of early life physical abuse (Liebschutz et al., 2018). Other studies have suggested that ELS may be underreported in retrospective recall (such as self-report

surveys), particularly in individuals with a history of childhood sexual abuse (Williams, 1994; Widom & Morris, 1997) and physical abuse (Widom & Shephard, 1996). In a large study of children who experienced ELS and were followed up approximately 20 years later, authors found that 38% of women with documented childhood sexual abuse did not recall the abuse at the time of follow up (Williams, 1994). Finally, there is some evidence to suggest that current depression may slightly increase reporting of childhood trauma (though it is not clear if this is due to previous suppression and then remembering, or incorrect remembering during the depressive state) – a longitudinal study of 7466 adults completing Canada's National Population Health Survey found that individuals reporting no mental health problems at baseline (1994/95) who went on to develop depression at the second timepoint (2006/07) were significantly more likely to report additional instances of ELS that had not been reported at baseline (Colman et al., 2016). The measure used did not specify subscales of abuse comparable to the CTQ (other than physical abuse and loose correlates to emotional neglect and physical neglect), and additional studies would be necessary to determine whether certain subtypes of abuse may be more sensitive to this effect. Due to feasibility, we were limited to retrospective recall of ELS in the studies included in this PhD. We chose the validated and widely used CTQ in all studies, in addition to an interview measure of ELS (the Traumatic Antecedents Questionnaire (TAQ; Herman et al., 1989; Vanderkolk et al., 1991)) in **Studies 3** and **4.** The vast majority of studies included in the systematic review (**Study 2**) similarly used the CTQ, indicating the widespread use of the measure which may improve comparability between studies.

Another limitation of the PhD was that several studies, particularly **Study 1** and **Study 4** were underpowered for their designed analyses. This is discussed in more detail in the

relevant discussion sections of these studies; however, it should be noted again here to emphasise that conclusions drawn from these findings should hence be cautiously interpreted, and future studies designed to investigate whether observed effects replicate, are needed. Due to limited sample size, we were not able to further break down our analyses by timing of abuse. We did limit timing of abuse in our eligibility criteria of both neuroimaging studies (**Study 3** and **Study 4**) in which we specified a timeframe for occurrence of CSA (5-14 years of age), based on previous studies implicated this time frame in hippocampal GMV (Andersen et al., 2008). Future studies, in particular those with larger sample sizes, should further investigate the effect of timing of abuse on cognitive and neural markers. The recently validated Maltreatment and Abuse Chronology of Exposure (MACE; Teicher & Parigger, 2015) may provide a good tool to assess timing and type of abuse in far greater detail.

Furthermore, another limitation that should be noted is that due to practical and funding constraints, studies focused on two aspects of cognition and neural markers known to be implicated in both ELS and MDD – affective cognition and brain structure (specifically GMV and cortical thickness). However, as previously mentioned, other measures, including cold cognition and functional brain measures, also play a key role in both MDD and ELS (Roiser & Sahakian, 2013; Pechtel & Pizzagalli, 2011). Functional neuroimaging studies may be especially well suited to study the neural underpinnings of affective cognition, and future studies may consider investigating this using a similar study design (in particular with matching ELS exposure in MDD and HC groups) as employed in this PhD. Similarly, while tasks in the EMOTICOM test battery have been designed to be largely unaffected by potential differences in cold (non-affective) cognition (or analysed in such a way that a

potential confounding factor of cold cognition is accounted for, such as by controlling for reaction time differences), future studies would ideally separately assess cold cognition (e.g. using the CANTAB test battery) to control for this in analyses more directly.

Future research would also benefit from exploring other potential factors that have been associated with both early life stress and MDD. Given the complexity of this research area and the many variables involved, it may be difficult for any one study to incorporate all of these factors, nevertheless, these should be considered and investigated in regard to cognitive and neural changes associated with both ELS and MDD. While not an exhaustive list of additional concepts of interest within this research domain, a brief mention of some of the additional key variables not explored in this PhD, are included here to highlight the complexity of the field and identify key areas of future research.

One such concept is resilience, defined as the capacity to maintain normal psychological functioning and not develop psychopathology in the face of extreme stress and trauma (Russo et al., 2012). In this PhD project, the HC/CSA group included in Studies 3 and 4 could be considered resilient, given that participants in this group were classified as experiencing significant CSA without any current or past psychopathology. While more research is needed to understand the mechanisms underlying resilience, studies have found that resilience in the face of ELS exposure may be an important mitigating factor in the development of later depressive symptoms (Wingo et al., 2010). The neurobiological underpinnings of resilience remain poorly understood though increasing evidence has implicated complex mechanisms involving epigenetics, immune pathways, and various neural signatures (Murrough & Russo, 2019). A related topic concerns resilience to

recurrence in MDD, defined as continued remission in remitted MDD, which may itself be associated with a unique neural signature (Workman et al,, 2017). Further research on resilience is needed to better understand its role as a mitigating factor of ELS in the development of MDD and other psychiatric disorders. Future research may hence benefit including a validated measure of resilience, such as the frequently used Connor-Davidson Resilience Scale (Campbell-Sills & Stein, 2007; Connor & Davidson, 2003) in studies of ELS and MDD. An improved understanding of both the neural and behavioural/cognitive mechanisms underlying resilience may provide important new avenues for treatment and intervention, designed to prevent the development of psychopathology (such as MDD) following ELS.

Another concept which was outside the scope of the current PhD but warrants further study in this field pertains personality disorders. Some research has highlighted that in particular emotional abuse may be linked to development of personality disorders, often in conjunction with MDD (Carr et al., 2013; Kim et al., 2018). Interestingly, similar findings as those published frequently in MDD, namely volumetric reductions of the hippocampus and amygdala, have also been identified in borderline personality disorder (BPD) and appear to be correlated with the severity of ELS experienced (Driessen et al., 2000). Given that ELS has been identified as a risk factor for several personality disorders, including BPD (Ball & Links, 2009) and schizotypal personality disorder (Velikonja et al., 2019) and the high comorbidity between personality disorders and MDD, estimated to be as a high as 50% of patients with depression (Van & Kool, 2018), future studies should take care to include validated assessments of personality disorders. This would allow additional analyses controlling for potential personality disorders amongst participants with MDD and enable investigations of

how ELS, MDD, and personality disorders may interact. It is well-established that individuals with both MDD and personality disorders are associated with poorer clinical prognosis (Van & Kool, 2018), further highlighting the importance of better understanding possible cognitive and neural mechanisms as well as potential interactions with ELS.

Many more factors have been identified as potential links between ELS and later life MDD, including inflammatory responses (Pace et al., 2006), gene-environment interactions and epigenetics (Heim & Binder, 2012), and lifestyle factors such as exercise (Masrour et al., 2018), to name but a few. It has become increasingly evident that the relationship between ELS and MDD is extremely complex and likely involves the interplay of multiple factors. Further research is needed to better understand these mechanisms in the hope of ultimately identifying novel targets for successful interventions and treatments.

4. Conclusion

Overall, results from **Studies 1-4,** indicated that several findings previously attributed to MDD, including specific constructs of affective cognition and GMV of the hippocampus and its subfields (in addition to several other brain regions) are in fact better explained by ELS. On the other hand, other factors, including cortical thickness and other constructs of affective cognition, were driven primarily by MDD diagnosis, independent of ELS. These findings highlight the specificity of effects of ELS and MDD on measures of affective cognition and brain structure, and emphasise the need for inclusion of ELS measures in studies of depression. Furthermore, results may shed light on potential mechanisms underlying the link between ELS as a risk factor for later life depression with implications for both pharmacological and psychological treatment mechanisms.

REFERENCES

Addington, D., Addington, J., & Maticka-Tyndale, E. (1993). Assessing depression in schizophrenia: the Calgary Depression Scale. *The British journal of psychiatry*, *163*(S22), 39-44.

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of neuropsychiatry and clinical neurosciences*, *20*(3), 292-301.

Andrew, A., Knapp, M., McCrone, P. R., Parsonage, M., & Trachtenberg, M. (2012). Effective interventions in schizophrenia: the economic case.

Ball, J. S., & Links, P. S. (2009). Borderline personality disorder and childhood trauma: evidence for a causal relationship. *Current psychiatry reports*, *11*(1), 63-68.

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., ... & Thomas, S. M. (2008). The mental wealth of nations. *Nature*, *455*(7216), 1057-1060.

Belleau, E. L., Treadway, M. T., & Pizzagalli, D. A. (2019). The impact of stress and major depressive disorder on hippocampal and medial prefrontal cortex morphology. *Biological psychiatry*, *85*(6), 443-453.

Bentall, R. P., Wickham, S., Shevlin, M., & Varese, F. (2012). Do specific early-life adversities lead to specific symptoms of psychosis? A study from the 2007 the Adult Psychiatric Morbidity Survey. *Schizophrenia bulletin*, *38*(4), 734-740.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child abuse & neglect*, *27*(2), 169-190.

Boldrini, M., Underwood, M. D., Hen, R., Rosoklija, G. B., Dwork, A. J., Mann, J. J., & Arango, V. (2009). Antidepressants increase neural progenitor cells in the human hippocampus. *Neuropsychopharmacology*, *34*(11), 2376-2389.

Bora, E., Yucel, M., & Pantelis, C. (2009). Theory of mind impairment in schizophrenia: meta-analysis. *Schizophrenia research*, *109*(1), 1-9.

Bounoua, N., Miglin, R., Spielberg, J. M., & Sadeh, N. (2020). Childhood assaultive trauma and physical aggression: Links with cortical thickness in prefrontal and occipital cortices. *Neuroimage: clinical*, *27*, 102321.

Bremner, J. D., Narayan, M., Anderson, E. R., Staib, L. H., Miller, H. L., & Charney, D. S. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry*, *157*(1), 115-118.

Browning, M., Holmes, E. A., & Harmer, C. J. (2010). The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(1), 8-20.

Brüne, M. (2005). "Theory of mind" in schizophrenia: a review of the literature. *Schizophrenia Bulletin*, *31*(1), 21-42.

Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology*, *30*(9), 846-856.

Campbell, S., Marriott, M., Nahmias, C., & MacQueen, G. M. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry*, *161*(4), 598-607.

Campbell-Sills, L., & Stein, M. B. (2007). Psychometric analysis and refinement of the connor–davidson resilience scale (CD-RISC): Validation of a 10-item measure of resilience. *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, *20*(6), 1019-1028.

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.

Cassiers, L. L., Sabbe, B. G., Schmaal, L., Veltman, D. J., Penninx, B. W., & Van Den Eede, F. (2018). Structural and functional brain abnormalities associated with exposure to different childhood trauma subtypes: A systematic review of neuroimaging findings. *Frontiers in psychiatry*, *9*, 329.

Cather, C., Penn, D., Otto, M. W., Yovel, I., Mueser, K. T., & Goff, D. C. (2005). A pilot study of functional Cognitive Behavioral Therapy (fCBT) for schizophrenia. *Schizophrenia research*, *74*(2), 201-209.

Clark L, Chamberlain SR, Sahakian BJ (2009). Neurocognitive mechanisms in depression: implications for treatment. *Annu Rev Neurosci*32: 57–74.

Colman, I., Kingsbury, M., Garad, Y., Zeng, Y., Naicker, K., Patten, S., ... & Thompson, A. H. (2016). Consistency in adult reporting of adverse childhood experiences. *Psychological medicine*, *46*(3), 543-549.

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and anxiety*, *18*(2), 76-82.

Dannlowski, U., Stuhrmann, A., Beutelmann, V., Zwanzger, P., Lenzen, T., Grotegerd, D., ... & Kugel, H. (2012). Limbic scars: long-term consequences of childhood maltreatment revealed

by functional and structural magnetic resonance imaging. *Biological psychiatry*, *71*(4), 286-293.

Dorresteijn, S., Gladwin, T. E., Eekhout, I., Vermetten, E., & Geuze, E. (2019). Childhood trauma and the role of self-blame on psychological well-being after deployment in male veterans. *European journal of psychotraumatology*, *10*(1), 1558705.

Driessen, M., Herrmann, J., Stahl, K., Zwaan, M., Meier, S., Hill, A., ... & Petersen, D. (2000). Magnetic resonance imaging volumes of the hippocampus and the amygdala in women with borderline personality disorder and early traumatization. *Archives of general psychiatry*, *57*(12), 1115-1122.

Elliott, R., Zahn, R., Deakin, J. W., & Anderson, I. M. (2011). Affective cognition and its disruption in mood disorders. *Neuropsychopharmacology*, *36*(1), 153-182.

Frodl, T., Jäger, M., Smajstrlova, I., Born, C., Bottlender, R., Palladino, T., ... & Meisenzahl, E. M. (2008). Effect of hippocampal and amygdala volumes on clinical outcomes in major depression: a 3-year prospective magnetic resonance imaging study. *Journal of psychiatry & neuroscience: JPN*, *33*(5), 423.

Garety, P., Fowler, D., Kuipers, E., Freeman, D., Dunn, G., Bebbington, P., ... & Jones, S. (1997). London-East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis. II: Predictors of outcome. *The British Journal of Psychiatry*, *171*(5), 420-426.

Gold, A. L., Sheridan, M. A., Peverill, M., Busso, D. S., Lambert, H. K., Alves, S., ... &

Green, H., McGinnity, Á., Meltzer, H., Ford, T., & Goodman, R. (2005). Mental health of children and young people in Great Britain, 2004.

Green, J. G., McLaughlin, K. A., Berglund, P. A., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2010). Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: associations with first onset of DSM-IV disorders. *Archives of general psychiatry*, *67*(2), 113-123.

Green, M. F., Hellemann, G., Horan, W. P., Lee, J., & Wynn, J. K. (2012). From perception to functional outcome in schizophrenia: modeling the role of ability and motivation. *Archives of General Psychiatry*, *69*(12), 1216-1224.

Gumley, A., O'Grady, M., McNay, L., Reilly, J., Power, K., & Norrie, J. (2003). Early intervention for relapse in schizophrenia: results of a 12-month randomized controlled trial of cognitive behavioural therapy. *Psychological medicine*, *33*(03), 419-431.

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.

Hammen, C., Henry, R., & Daley, S. E. (2000). Depression and sensitization to stressors among young women as a function of childhood adversity. *Journal of consulting and clinical psychology*, *68*(5), 782.

Harmer, C. J., Bhagwagar, Z., Perrett, D. I., Völlm, B. A., Cowen, P. J., & Goodwin, G. M. (2003). Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology*, *28*(1), 148-152.

Harmer, C. J., Cowen, P. J., & Goodwin, G. M. (2011). Efficacy markers in depression. *Journal of Psychopharmacology*, *25*(9), 1148-1158.

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., ... & Cowen, P. J. (2009). Effect of acute antidepressant administration on negative affective bias in depressed patients. *American Journal of Psychiatry*, *166*(10), 1178-1184.

Harmer, C. J., Shelley, N. C., Cowen, P. J., & Goodwin, G. M. (2004). Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. American Journal of Psychiatry, 161(7), 1256-1263.

Heim, C., & Binder, E. B. (2012). Current research trends in early life stress and depression: Review of human studies on sensitive periods, gene–environment interactions, and epigenetics. *Experimental neurology*, *233*(1), 102-111.

Heim, C. M., Mayberg, H. S., Mletzko, T., Nemeroff, C. B., & Pruessner, J. C. (2013). Decreased cortical representation of genital somatosensory field after childhood sexual abuse. *American Journal of Psychiatry*, *170*(6), 616-623.

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.

Herbener, E. S. (2008). Emotional memory in schizophrenia. *Schizophrenia Bulletin*, *34*(5), 875-887.

Herbener, E. S., Rosen, C., Khine, T., & Sweeney, J. A. (2007). Failure of positive but not negative emotional valence to enhance memory in schizophrenia. *Journal of abnormal psychology*, *116*(1), 43.

Herman, J. P., Dolgas, C. M., & Carlson, S. L. (1998). Ventral subiculum regulates hypothalamo–pituitary–adrenocortical and behavioural responses to cognitive stressors. *Neuroscience*, *86*(2), 449-459.

Jaworska, N., MacMaster, F. P., Gaxiola, I., Cortese, F., Goodyear, B., & Ramasubbu, R. (2014). A preliminary study of the influence of age of onset and childhood trauma on cortical thickness in major depressive disorder. *BioMed research international*, *2014*.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin*, *13*(2), 261-276.

Kelly, P. A., Viding, E., Wallace, G. L., Schaer, M., De Brito, S. A., Robustelli, B., & McCrory, E. J. (2013). Cortical thickness, surface area, and gyrification abnormalities in children exposed to maltreatment: neural markers of vulnerability?. *Biological psychiatry*, *74*(11), 845-852.

Kim, M. K., Kim, J. S., Park, H. I., Choi, S. W., Oh, W. J., & Seok, J. H. (2018). Early life stress, resilience and emotional dysregulation in major depressive disorder with comorbid borderline personality disorder. *Journal of affective disorders*, *236*, 113-119.

Lang, T. J., Blackwell, S. E., Harmer, C. J., Davison, P., & Holmes, E. A. (2012). Cognitive bias modification using mental imagery for depression: Developing a novel computerized intervention to change negative thinking styles. *European Journal of Personality*, *26*(2), 145-157.

Li, Q., Zhao, Y., Chen, Z., Long, J., Dai, J., Huang, X., ... & Sweeney, J. A. (2020). Meta-analysis of cortical thickness abnormalities in medication-free patients with major depressive disorder. *Neuropsychopharmacology*, *45*(4), 703-712.

Liebschutz, J. M., Buchanan-Howland, K., Chen, C. A., Frank, D. A., Richardson, M. A., Heeren, T. C., ... & Rose-Jacobs, R. (2018). Childhood Trauma Questionnaire (CTQ) correlations with prospective violence assessment in a longitudinal cohort. *Psychological assessment*, *30*(6), 841.

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.

Marsh, P. J., & Williams, L. M. (2006). ADHD and schizophrenia phenomenology: visual scanpaths to emotional faces as a potential psychophysiological marker?. *Neuroscience & Biobehavioral Reviews*, *30*(5), 651-665.

Masrour, F. F., Peeri, M., Hosseini, M. J., & Azarbayjani, M. A. (2018). Exercise during adolescence attenuated depressive-like behaviors and hippocampal mitochondrial dysfunction following early life stress in adult male rats. *Iranian journal of pharmaceutical research: IJPR*, *17*(Suppl2), 124.

McCabe, C., Mishor, Z., Cowen, P. J., & Harmer, C. J. (2010). Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. *Biological psychiatry*, *67*(5), 439-445.

McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual review of neuroscience*, *22*(1), 105-122.

McLaughlin, K. A. (2016). Childhood abuse and reduced cortical thickness in brain regions involved in emotional processing. *Journal of Child Psychology and Psychiatry*, *57*(10), 1154-1164.

McLaughlin, K. A., Sheridan, M. A., Winter, W., Fox, N. A., Zeanah, C. H., & Nelson, C. A. (2014). Widespread reductions in cortical thickness following severe early-life deprivation: a neurodevelopmental pathway to attention-deficit/hyperactivity disorder. *Biological psychiatry*, *76*(8), 629-638.

Monninger, M., Kraaijenvanger, E. J., Pollok, T. M., Boecker-Schlier, R., Jennen-Steinmetz, C., Baumeister, S., ... & Holz, N. E. (2020). The long-term impact of early life stress on orbitofrontal cortical thickness. *Cerebral Cortex*, *30*(3), 1307-1317.

Murphy, F. C., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., Paykel, E. S., & Sahakian, B. J. (2001). Decision-making cognition in mania and depression. *Psychological medicine*, *31*(4), 679-693.

Murray, G. K., Cheng, F., Clark, L., Barnett, J. H., Blackwell, A. D., Fletcher, P. C., ... & Jones, P. B. (2008). Reinforcement and reversal learning in first-episode psychosis. *Schizophrenia bulletin*, *34*(5), 848-855.

Murrough, J. W., & Russo, S. J. (2019). The neurobiology of resilience: complexity and hope. *Biological psychiatry*, *86*(6), 406-409.

Nemeroff, C. B., Heim, C. M., Thase, M. E., Klein, D. N., Rush, A. J., Schatzberg, A. F., ... & Keller, M. B. (2005). Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Focus*, *100*(1), 14293-135.

Pace, T. W., Mletzko, T. C., Alagbe, O., Musselman, D. L., Nemeroff, C. B., Miller, A. H., & Heim, C. M. (2006). Increased stress-induced inflammatory responses in male patients with major depression and increased early life stress. *American Journal of Psychiatry*, *163*(9), 1630-1633.

Palmier-Claus, J. E., Berry, K., Bucci, S., Mansell, W., & Varese, F. (2016). Relationship between childhood adversity and bipolar affective disorder: systematic review and meta-analysis. *The British Journal of Psychiatry*, *209*(6), 454-459.

Papadatou-Pastou, M., Miskowiak, K. W., Williams, J. M. G., Harmer, C. J., & Reinecke, A. (2012). Acute antidepressant drug administration and autobiographical memory recall: a functional magnetic resonance imaging study. *Experimental and clinical psychopharmacology*, *20*(5), 364.

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: an integrated review of human literature. *Psychopharmacology*, *214*(1), 55-70.

Pittenger, C., & Duman, R. S. (2008). Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology*, *33*(1), 88-109.

Pomarol-Clotet, E., Hynes, F., Ashwin, C., Bullmore, E. T., McKenna, P. J., & Laws, K. R. (2010). Facial emotion processing in schizophrenia: a non-specific neuropsychological deficit?. *Psychological medicine*, *40*(06), 911-919.

Rawal, A., Collishaw, S., Thapar, A., & Rice, F. (2013). 'The risks of playing it safe': a prospective longitudinal study of response to reward in the adolescent offspring of depressed parents. *Psychological medicine*, *43*(1), 27-38.

Ridout, N., Astell, A., Reid, I., Glen, T., & O'Carroll, R. (2003). Memory bias for emotional facial expressions in major depression. *Cognition and Emotion*, *17*(1), 101-122.

Roiser, J. P., & Sahakian, B. J. (2013). Hot and cold cognition in depression. *CNS spectrums*, *18*(3), 139-149.

Roiser, J. P., Stephan, K. E., Den Ouden, H. E. M., Barnes, T. R. E., Friston, K. J., & Joyce, E. M. (2009). Do patients with schizophrenia exhibit aberrant salience?. *Psychological medicine*, *39*(02), 199-209.

Russo, S. J., Murrough, J. W., Han, M. H., Charney, D. S., & Nestler, E. J. (2012). Neurobiology of resilience. *Nature neuroscience*, *15*(11), 1475-1484.

Samuels, B. A., Leonardo, E. D., & Hen, R. (2015). Hippocampal subfields and major depressive disorder. *Biological psychiatry*, *77*(3), 210.

Sapolsky, R. M. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of general psychiatry*, *57*(10), 925-935.

Sharma-Patel, K., & Brown, E. J. (2016). Emotion regulation and self blame as mediators and moderators of trauma-specific treatment. *Psychology of Violence*, *6*(3), 400.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. In *Journal of Clinical Psychiatry* (Vol. 59, pp. 22–33).

Simpson, J., Helliwell, B., Varese, F., & Powell, P. (2020). Self-disgust mediates the relationship between childhood adversities and psychosis. *British Journal of Clinical Psychology*, *59*(2), 260-275.

Suh, J. S., Schneider, M. A., Minuzzi, L., MacQueen, G. M., Strother, S. C., Kennedy, S. H., & Frey, B. N. (2019). Cortical thickness in major depressive disorder: a systematic review and meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *88*, 287-302.

Teicher M, Samson J, Polcari A, McGreenery C (2006a) Sticks, stones, and hurtful words: relative effects of various forms of childhood maltreatment. Am J Psychiatry 163:993–1000

Teicher M, Tomoda A, Andersen S (2006b) Neurobiological con- sequences of early stress and childhood maltreatment: are results from human and animal studies comparable? Ann NY Acad Sci 1071:313–323

Teicher, M. H., & Parigger, A. (2015). The 'Maltreatment and Abuse Chronology of Exposure'(MACE) scale for the retrospective assessment of abuse and neglect during development. *PLoS one*, *10*(2), e0117423.

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PloS one*, *4*(8), e6598.

Vanderkolk BA, Perry JC and Herman JL (1991) Childhood origins of self- destructive behavior. American Journal of Psychiatry 148, 1665–1671.

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., ... & Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective-and cross-sectional cohort studies. *Schizophrenia bulletin*, *38*(4), 661-671.

Velikonja, T., Velthorst, E., McClure, M. M., Rutter, S., Calabrese, W. R., Rosell, D., ... & Perez-Rodriguez, M. M. (2019). Severe childhood trauma and clinical and neurocognitive features in schizotypal personality disorder. *Acta Psychiatrica Scandinavica*, *140*(1), 50-64.

Videbech, P., & Ravnkilde, B. (2004). Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry*, *161*(11), 1957-1966.

Vyas, A., Bernal, S., & Chattarji, S. (2003). Effects of chronic stress on dendritic arborization in the central and extended amygdala. *Brain research*, *965*(1-2), 290-294.

Vyas, A., Mitra, R., Rao, B. S., & Chattarji, S. (2002). Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *Journal of Neuroscience*, *22*(15), 6810-6818.

Waltz, J. A., Frank, M. J., Robinson, B. M., & Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biological psychiatry*, *62*(7), 756-764.

Widom, C. S., & Morris, S. (1997). Accuracy of adult recollections of childhood victimization, part 2: childhood sexual abuse. *Psychological assessment*, *9*(1), 34.

Widom, C. S., & Shepard, R. L. (1996). Accuracy of adult recollections of childhood victimization: Part 1. Childhood physical abuse. *Psychological assessment*, *8*(4), 412.

Williams, L. M. (1994). Recall of childhood trauma: a prospective study of women's memories of child sexual abuse. *Journal of consulting and clinical psychology*, *62*(6), 1167.

Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., & Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: data from the randomized international study to predict optimized treatment for depression. *Translational psychiatry*, *6*(5), e799-e799.

Wingo, A. P., Wrenn, G., Pelletier, T., Gutman, A. R., Bradley, B., & Ressler, K. J. (2010). Moderating effects of resilience on depression in individuals with a history of childhood abuse or trauma exposure. *Journal of affective disorders*, *126*(3), 411-414.

Workman, C. I., Lythe, K. E., McKie, S., Moll, J., Gethin, J. A., Deakin, J. F., ... & Zahn, R. (2017). A novel resting-state functional magnetic resonance imaging signature of resilience to recurrent depression. *Psychological medicine*, *47*(4), 597-607.

Wunderink, L., Nieboer, R. M., Wiersma, D., Sytema, S., & Nienhuis, F. J. (2013). Recovery in remitted first-episode psychosis at 7 years of follow-up of an early dose reduction/discontinuation or maintenance treatment strategy: long-term follow-up of a 2-year randomized clinical trial. *JAMA psychiatry*, *70*(9), 913-920.

Xie, P., Wu, K., Zheng, Y., Guo, Y., Yang, Y., He, J., ... & Peng, H. (2018). Prevalence of childhood trauma and correlations between childhood trauma, suicidal ideation, and social support in patients with depression, bipolar disorder, and schizophrenia in southern China. *Journal of affective disorders*, *228*, 41-48.