Hot and cold cognitive disturbances in antidepressant-free patients with major depressive disorder: a NeuroPharm study

V. H. Dam1,2, D. S. Stenbæk1, K. Köhler-Forsberg1,2,3, C. Ip1,2,4, B. Ozenne1,5, B. J. Sahakian6,7, G. M. Knudsen1,2, M. B. Jørgensen3 and V. G. Frokjaer1,3

Abstract

Background. Cognitive disturbances are common and disabling features of major depressive disorder (MDD). Previous studies provide limited insight into the co-occurrence of hot (emotion-dependent) and cold (emotion-independent) cognitive disturbances in MDD. Therefore, we here map both hot and cold cognition in depressed patients compared to healthy individuals.

Methods. We collected neuropsychological data from 92 antidepressant-free MDD patients and 103 healthy controls. All participants completed a comprehensive neuropsychological test battery assessing hot cognition including emotion processing, affective verbal memory and social cognition as well as cold cognition including verbal and working memory and reaction time.

Results. The depressed patients showed small to moderate negative affective biases on emotion processing outcomes, moderate increases in ratings of guilt and shame and moderate deficits in verbal and working memory as well as moderately slowed reaction time compared to healthy controls. We observed no correlations between individual cognitive tasks and depression severity in the depressed patients. Lastly, an exploratory cluster analysis suggested the presence of three cognitive profiles in MDD: one characterised predominantly by disturbed hot cognitive functions, one characterised predominantly by disturbed cold cognitive functions and one characterised by global impairment across all cognitive domains. Notably, the three cognitive profiles differed in depression severity.

Conclusion. We identified a pattern of small to moderate disturbances in both hot and cold cognition in MDD. While none of the individual cognitive outcomes mapped onto depression severity, cognitive profile clusters did. Overall cognition-based stratification tools may be useful in precision medicine approaches to MDD.

Introduction

Disturbance of cognitive functioning is a common feature of major depressive disorder (MDD) and has been proposed as an important treatment target (Collins et al., 2011). Cognitive symptoms including inability to concentrate or difficulty making decisions are listed among the diagnostic criteria for MDD (APA, 2013; WHO, 2007). Investigations into cognitive disturbances in MDD have typically focused on either so-called ‘hot’ or ‘cold’ cognitive functions (Roiser & Sahakian, 2013). Hot cognition describes mental functions that involve the processing of emotionally salient information (e.g., identifying emotional facial expressions) or emotional responses (e.g., reward-driven behaviours). In particular negative affective biases, i.e., the subconscious allocation of more attention and mental resources to the processing of negative information over positive information, have been associated with MDD psychopathology (Elliott, Zahn, Deakin, & Anderson, 2011; Miskowiak & Carvalho, 2014) and may play a key role in the onset and maintenance of depressive symptoms (Roiser, Elliott, & Sahakian, 2012). Another hot cognitive domain which may be impaired in MDD is social cognition which includes functions such as seeing oneself in the ‘other’, i.e., Theory of Mind (Bora & Berk, 2015; Wolkenstein, Schonenberg, Schirm, & Hautzinger, 2011), interpretation of social situations and excessive experiences of negative social emotions such as shame and guilt (Kim et al., 2015).

Cold cognition describes mental processes that include emotionally neutral information and do not directly involve activation of emotional states (Roiser & Sahakian, 2013). Recent meta-analyses suggest that cold cognitive deficits in MDD are predominantly found in domains of attention, learning and memory and executive functions (Goodall et al., 2018; Lee, Hermens, Porter, & Redoblado-Hodge, 2012; Rock, Roiser, Riedel, & Blackwell, 2014). Slowed reaction time
has also been reported for depressed patients and is, along with agitated psychomotor function, considered a distinct symptom in MDD (Bennabi, Vandel, Papaxanthis, Pozzo, & Haffen, 2013).

The effects sizes reported for cognitive disturbances in MDD are typically small to moderate which is relatively modest compared to those reported for other serious neuropsychiatric disorders such as Alzheimer’s Disease and schizophrenia (Maruff & Jaeger, 2016). Nevertheless, their impact on daily life may be very disruptive. Both hot and cold cognitive disturbances in MDD have been found to be detrimental to the patient’s ability to engage successfully in work or educational activities as well as overall psychosocial functioning (Cambridge, Knight, Mills, & Baune, 2018; Weightman, Knight, & Baune, 2019). This is especially relevant as cognitive disturbances do not always resolve with the remission of core depressive symptoms (Hernaus, Gold, Waltz, & Frank, 2018).

Despite a growing body of data on specific cognitive deficits in MDD, we currently know little about the co-occurrence and magnitude of impairments across different types of cognitive domains. Few studies on MDD have included both hot and cold cognitive tasks and comparisons between studies are often hampered by differences in cohort characteristics such as medication/treatment status, comorbidity, age-range, chronicity and severity of current depressive episode. To address this, we therefore applied a broad range of both hot and cold cognitive tasks in a large cohort of well-characterised and antidepressant-free depressed patients.

**Methods**

**Participants and study design**

One hundred non-psychotic antidepressant-free patients suffering from a moderate to severe depressive episode lasting less than two years [Hamilton Depression Rating Scale-17 (HDRS17) ≥18] were included in a large multimodal neuroimaging clinical trial (NeuroPharm 1). Patients were eligible for inclusion if they had been antidepressant free for >2 months; had not previously exhibited non-response to SSRIs; and had not undergone more than one antidepressant treatment attempt in the current depressive episode. Patients were recruited through their primary care centre or a central referral site for ‘depression treatment packages’ at the Mental Health Services of the Capital Region of Copenhagen. MDD diagnosis was confirmed by a certified psychiatrist and corroborated by a Mini-International Neuropsychiatric Interview (MINI). Out of the 100 patients who entered the study, neuropsychological data was available from 92 patients (67 females); out of these, 41 patients had first-episode depression while 51 patients had recurrent depression. In addition, data from 100 healthy participants were collected as part of a validation study of the EMOTICOM test battery, a novel neuropsychological battery specifically designed to assess hot cognitive functions (Dam et al., 2019). A modified version of the Verbal Affective Memory Task 24 (Jensen et al., 2016), the Verbal Affective Memory Task 26 (VAMT-26), was used to assess affective memory biases. Biases were calculated by subtracting negative information scores from positive information scores (e.g. hit rate for recognition of happy faces minus hit rate for recognition of sad faces).

**Cold cognition**

We used three tasks to assess cold cognition: total word recall in the VAMT-26 was used to assess explicit non-affective verbal memory function; the Letter Number Sequence (LNS) task was used to assess working memory capacity; and the Simple Reaction Time (SRT) task was used to assess reaction time.

Note, a full description of all tasks including both primary and secondary outcomes can be found in online Supplementary Materials.

**Statistical analysis**

**Group differences**

Differences on cognitive performance between depressed patients and healthy controls were assessed with linear regression models with primary cognitive outcome as the dependent variable and age, sex and group coded as a categorical variable as independent variables. The reported p values were corrected for 11 tests using the Bonferroni–Holm method. The standardised effect size was estimated for each primary cognitive outcome by computing the Cohen’s d on the partial residuals relative to the group variable (i.e. after removing the age and sex effects from the cognitive outcome). Normality assumptions were assessed and found to be acceptable for all models except the SRT task outcome. Post hoc linear regression analyses were used to investigate secondary outcomes for tasks showing statistically significant group differences. The reported p values for the post hoc analyses were adjusted to control the family-wise error rate within each task: first the
Bonferroni–Holm method was applied (e.g. to eight tests for the ERT secondary outcomes). If the resulting \( p \) value was smaller than the adjusted \( p \) value for the primary task outcome, then the adjusted \( p \) value of the primary task outcome was used for this secondary outcome. This was done to ensure that the adjusted \( p \) values were coherent between the main and secondary outcomes when using significance threshold below 0.05.

Correlation with symptom severity
In MDD patients, Spearman’s rank correlation coefficient was used to assess the relationship between depressive symptom severity assessed with HDRS6 and HDRS17 and performance on primary cognitive outcomes. The reported \( p \) values were corrected for 11 tests using the Bonferroni–Holm method.

Clustering of cognitive profiles
In an exploratory post hoc analysis, we used a K-means cluster analysis to delineate potential groups of cognitive profiles within the depressed cohort (Clatworthy, Buick, Hankins, Weinman, & Horne, 2005). The input into the cluster analysis was restricted to the primary cognitive outcomes found to characterise the depressed state, i.e. those outcomes that differed with statistical significance between the depressed patients and healthy controls. The cognitive outcomes were standardised to \( z \)-scores and the number of clusters and the position used to initialise the K-means algorithm were obtained using a hierarchical clustering algorithm (see online Supplementary Material) (Milligan, 1980). The groups derived from the K-means analysis were compared on clinical and demographic factors using Analysis of Variance (ANOVA) for continuous outcomes and the \( \chi^2 \) test for binary outcomes (i.e. sex).

Outliers and missing data
Outliers were defined as observations 1.5 Interquartile Range (IQR) above or below the 1\textsuperscript{st} quartile or 3\textsuperscript{rd} quartile respectively. Each outlier was qualitatively evaluated based on notes from the testing session and congruency with scores on outcomes from the same task as well as outcomes from similar tasks. In total, 38 outlying data points were detected (representing 1.8% of all observations across the 11 primary cognitive outcomes); 37 were deemed to be ‘true outliers’ and kept in the analysis while one patient outlier from the LNS task was excluded as the patient had misunderstood the test instructions. Importantly, we found that none of the reported estimates changed critically when all outliers were removed. Note, in the SRT task two patient scores were so extreme (8.1 and 16.1 IQR above the 1\textsuperscript{st} quartile respectively) that independent of their potential neurobiological meaningfulness they were capped to one and two units above the third highest score respectively; this allowed the data to be included without losing their rank or impacting the group estimates unduly. Missing data included IM data from one patient, VAMT-26 data from six patients, SIP data from three controls, LNS data from nine patients and one control and SRT data from one patient and 37 controls.

Results

Descriptive
While there was no significant age difference between the two groups, the proportion of females was significantly higher in the patient group (73.9%) compared to the control group (50.4%) (Table 1). This reflects the well-documented prevalence of depression in females in the general population; although notably the proportion of females in the present depressed sample was higher than the ~60% reported for European countries in a recent report by the World Health Organization (WHO, 2017). In accordance with the inclusion criteria, all depressed patients had a HDRS17 score above 17, indicating moderate to severe depression. IQ was assessed with the Reynolds Intellectual Screening Test (Reynolds, 2011) and all study participants scored within the normal range (see Supplementary Materials). Cognitive performance did not differ between patients with first-episode (\( N = 41 \)) and recurrent depression (\( N = 51 \)) on any task outcome (all \( p_{\text{corrected}} > 0.75 \); see online Supplementary Materials for a full overview).

Group differences

Hot cognition: In the ERT task, the affective bias expressed by the depressed patients was 11.1 percentage points more negative for recognition rates (\( p_{\text{corrected}} = 0.03 \)) and 7.8 percentage points more negative for misattribution rates (\( p_{\text{corrected}} = 0.02 \)) compared to healthy controls. Likewise, the affective bias for emotion detection threshold in the IM task was 7.9 percentage points more negative for the depressed patients (\( p_{\text{corrected}} < 0.001 \)) while no substantial difference in bias was observed for affective verbal memory in the VAMT-26 (\( p_{\text{corrected}} = 1.00 \)). When asked to identify with cartoon characters in negative social situations in the ME task, the depressed patients also reported stronger experiences of negative moral emotions equivalent to 0.5 points on a seven-point Likert scale for both guilt (\( p_{\text{corrected}} < 0.001 \)) and shame (\( p_{\text{corrected}} < 0.001 \)) compared to healthy controls. We observed no substantial group differences in choice of social information (\( p_{\text{corrected}} = 1.00 \)) nor bias in interpretation of social situations (\( p_{\text{corrected}} = 1.00 \)) in the SIP task.

Cold cognition: The depressed patients recalled a total of 2.6 fewer words in the VAMT-26 (\( p_{\text{corrected}} < 0.001 \)) independent on affective valence; successfully sorted 1.7 fewer sequences on the working memory task (LNS, \( p_{\text{corrected}} = 0.002 \)); and exhibited 30.7 ms slower reaction time (SRT, \( p_{\text{corrected}} = 0.006 \)) compared to the healthy controls. Note, as model assumptions for normality were not met for the SRT outcome, we used bootstrapping to determine the reported \( p \) value. In addition, we also conducted a quantile regression analysis to assess whether the reported results were robust to outliers and found a similar effect (estimated group effect = 20.7 ms, \( p \) value = 0.003) (Fig. 1).

### Table 1. Descriptive data

<table>
<thead>
<tr>
<th></th>
<th>Depressed patients (( n = 92 ))</th>
<th>Healthy controls (( n = 103 ))</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years</td>
<td>27.3 ± 8.1 (18–57)</td>
<td>28.7 ± 7.3 (18–48)</td>
<td>0.19</td>
</tr>
<tr>
<td>Male/female</td>
<td>25/68</td>
<td>51/52</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MDI</td>
<td>34.5 ± 7.2 (16–50)(^a)</td>
<td>4.9 ± 3.9 (0–20)(^b)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HDRS(_6)</td>
<td>12.4 ± 1.6 (7–17)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>HDRS(_{17})</td>
<td>22.8 ± 3.4 (18–31)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\(^a\)\( N = 90 \) due to missing questionnaire data.
\(^b\)\( N = 102 \) due to missing questionnaire data.

---

https://doi.org/10.1017/S0033291720000938 Published online by Cambridge University Press
Fig. 1. Group differences on affective, social and cold cognitive outcomes between depressed patients and healthy controls. (I) Affective cognition: Recognition = affective bias for hit rate in the Emotional Recognition Task (patients \( n \) = 92, controls \( n \) = 103); Misattribution = affective bias for false alarm rate in the Emotional Recognition Task (patients \( n \) = 92, controls \( n \) = 103); Detection threshold = affective bias for the Intensity Morphing Task (patients \( n \) = 91, controls \( n \) = 103); Affective memory = affective bias for the Verbal Affective Memory Task 26 (patients \( n \) = 86, controls \( n \) = 103). (II) Social cognition: Guilt = average ratings of guilt in the Moral Emotions task (patients \( n \) = 91, controls \( n \) = 103); Shame = average ratings of shame in the Moral Emotions task (patients \( n \) = 91, controls \( n \) = 103); Information preference = choice of theory of mind-related information relative to facts in the Social Information Preference task (patients \( n \) = 89, controls \( n \) = 100); Interpretation bias = affective bias in choice of outcome in the Social Information Preference task (patients \( n \) = 89, controls \( n \) = 100). (III) Cold cognition: Verbal memory = Total recall score for the Verbal Affective Memory Task (patients \( n \) = 85, controls \( n \) = 103); Working memory = Letter-Number Sequence task (patients \( n \) = 83, controls \( n \) = 103); Reaction time = Simple Reaction Time (patients \( n \) = 91, controls \( n \) = 66). All models were corrected for age and sex. *\( p \) < 0.05, **\( p \) < 0.01, ***\( p \) < 0.001.

https://doi.org/10.1017/S0033291720000938 Published online by Cambridge University Press
Compared to healthy controls, MDD patients were more likely to incorrectly identify other emotions as sadness in the ERT task and continued to perceive sadness at lower intensity levels (decrease condition) in the IM task. In the ME task, patients also reported higher levels of guilt and shame in social scenarios where they identified with characters who accidentally harmed another person and increased guilt and scenarios where they identified with the victim of either accidental or intentional harm. Lastly, patients exhibited deficits for both immediate, short-term and long-term non-affective verbal memory compared to controls in the VAMT-26 (Fig. 2).

Correlation with depression severity
Correlations between cognitive performance and clinically rated depressive symptoms within the depressed group ranged from weak to negligible on all tasks and were statistically non-significant [HDRS6, $\rho (-0.2; 0.2)$, all $p_{corrected} > 0.42$; HDRS17, $\rho (-0.2; 0.2)$, all $p_{corrected} > 0.44$] (see online Supplementary Materials for a full overview).

Clustering of cognitive profiles
Based on the eight cognitive outcomes which showed a significant group difference, an initial hierarchical cluster analysis was run that indicated a three-cluster solution for cognitive profiles within the depressed group. The clustering centroids from the hierarchical cluster analysis were subsequently used to initialise a K-means analysis that converged within six iterations (Fig. 3).

There were no statistically significant differences between the clusters on age ($p = 0.58$) or sex ($p = 0.71$). The three clusters differed significantly on severity of core depressive symptoms indexed with HDRS6 [F(2, 90) = 4.1, $p = 0.02$] with patients from Cluster C (13.1 ± 1.9, mean ± s.d.) having higher scores than Cluster A (12.2 ± 1.6, mean ± s.d.) and B (11.9 ± 1.3, mean ± s.d.). The same pattern was present in severity of broad depressive symptoms indexed with HDRS17 (Cluster A, HDRS17 scores = 22.1 ± 3.2, mean ± s.d.; Cluster B, HDRS17 scores = 23.0 ± 3.1, mean ± s.d.; Cluster C, HDRS17 scores = 23.6 ± 3.8, mean ± s.d.), but did not reach statistical significance ($p = 0.19$).

Discussion
We here map the presence and magnitude of both hot and cold cognitive disturbances in a large cohort of antidepressant-free patients with a moderate to severe depressive episode. We found small to moderately sized negative biases in emotion processing but not in explicit verbal memory, large increases in experience of negative social emotions but no detectable differences in preference between social and non-social information or interpretation of ambiguous social situations. We also observed moderate impairment of cold cognitive functions including working and verbal memory and moderate slowing of reaction time. We found no direct link between depressive symptom severity and patient performance on any of the single task domains. Using an exploratory and data-driven approach, we identified three clusters with distinct cognitive profiles within the cohort of depressed patients: Cluster A was characterised by disturbances in hot but not cold cognition; Cluster B was characterised by positive biases and moderate deficits in cold cognitive domains; and Cluster C was characterised by large deficits across both hot and cold cognitive domains including extreme scores of guilt and shame.

Affective biases
As expected, the depressed patients exhibited negative affective biases across all emotion processing outcomes including...
recognition, misattribution and perceptual detection threshold. While abnormal processing of facial expressions is well-established in the MDD literature, the underlying cognitive mechanisms are still unclear (Elliott et al., 2011). Indeed, the findings from our study emphasise that the negative biases exhibited by the depressed patients must be understood relative to healthy controls. For example, while the depressed patients exhibited a clear negative affective bias in emotion misattribution, the negative bias observed for emotion detection threshold predominantly reflected the loss of positive bias exhibited by the healthy controls (see Fig. 1). Another line of research from attention paradigms suggests that affective bias in emotion processing is related to reduced orientation towards positive stimuli combined with an inability to disengage from negative stimuli (Armstrong & Olatunji, 2012). This aligns with our findings from the IM task as patients continued to perceive sadness at much lower intensity levels in the decrease condition compared to healthy controls.

Notably, we did not observe any negative affective bias in verbal memory performance. The concept of a mood-congruent memory bias was first proposed by Bower (1981) and posits that individuals will remember information that matches their current emotional state better than information that is not mood-congruent. Although this theory is relatively well supported by empirical studies investigating autobiographical (Köhler et al., 2015) and implicit (Gaddy & Ingram, 2014) types of memory, evidence for a bias in explicit, non-self-referential memory remains inconclusive. In fact, only a handful of studies have specifically investigated this type of memory (typically using valanced word-lists) in depression with some reporting a negative bias (Bradley, Mogg, & Williams, 1995; Neshat-Doost, Taghavi, Moradi, Yule, & Dalgleish, 1998; Watkins, Mathews, Williamson, & Fuller, 1992) while others suggest a positive bias (Calev, 1996; Danion, Kauffmann-Müller, Grange, Zimmermann, & Greth, 1995; Zupan, Žeželj, & Andjelković, 2017). In contrast to word memory, there have been attentional biases to negative words reported for unmedicated depressed subject (Beavers et al., 2013). The majority of memory studies had very small sample sizes and used different cognitive tasks and different depression criteria, which may contribute to the lack of consensus. We here present data from one of the largest study populations to date of well-characterised depressed patients which suggest that MDD symptomatology is not related to mood-congruent memory bias in explicit non-self-referential affective memory. Future studies should therefore consider using cognitive tasks assessing autobiographical and implicit memory as they may be more sensitive to affective memory disturbances in MDD.

Social cognition

Feelings of excessive shame and guilt are common in MDD and critically contribute to low self-esteem and social withdrawal (Mills et al., 2015). In the most severe cases they can even reach the threshold of psychosis (Lake, 2008). In particular contextual-maladaptive guilt (i.e. exaggerated guilt related to uncontrollable events) and generalised guilt (i.e. guilt divorced from concrete contexts) as well as external shame (i.e. shame based on beliefs about others’ opinions) are strongly associated with depressive symptoms (Kim, Thibodeau, & Jorgensen, 2011). Notably, in the ME task we observed most group differences between the depressed patients and the healthy controls in moral scenarios where the participants were asked to identify...
with the victim of harm (see Table 2). In fact, the level of shame and guilt reported by the patients appeared most pronounced when they were the victim of intentional rather than accidental harm, likely reflecting a disengaging of other-blaming schemata in favour of maladaptive self-blaming and internalising schemata. Clearly, this represents a critical target to address in psychotherapy, e.g. cognitive behavioural therapy (CBT).

Recent evidence suggests that ToM, i.e. the ability to attribute mental states to other people, is impaired in MDD and linked to depressive symptom severity (Bora & Berk, 2015). While we did not have a direct measure ToM in the present study, the SIP task indexes the preference for choosing social information (facts) when interpreting socially ambiguous situations. We did

Table 2. Group differences between depressed patients and healthy controls on secondary cognitive outcomes

<table>
<thead>
<tr>
<th>Emotion recognition task</th>
<th>Depressed patients</th>
<th>Healthy controls</th>
<th>B</th>
<th>p value</th>
<th>p_corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition – Happy</td>
<td>72.4 ± 17.8</td>
<td>78.0 ± 16.7</td>
<td>-5.0</td>
<td>5.25 × 10⁻²</td>
<td>0.21</td>
</tr>
<tr>
<td>Recognition – Sad</td>
<td>76.6 ± 16.6</td>
<td>71.3 ± 18.9</td>
<td>6.2</td>
<td>2.01 × 10⁻²</td>
<td>0.14</td>
</tr>
<tr>
<td>Recognition – Angry</td>
<td>62.5 ± 12.8</td>
<td>66.0 ± 11.4</td>
<td>-3.7</td>
<td>4.15 × 10⁻²</td>
<td>0.21</td>
</tr>
<tr>
<td>Recognition – Fearful</td>
<td>73.5 ± 11.7</td>
<td>75.4 ± 14.8</td>
<td>-2.4</td>
<td>2.18 × 10⁻¹</td>
<td>0.44</td>
</tr>
<tr>
<td>Misattribution – Happy</td>
<td>10.9 ± 8.4</td>
<td>15.0 ± 12.0</td>
<td>-3.5</td>
<td>2.37 × 10⁻²</td>
<td>0.14</td>
</tr>
<tr>
<td>Misattribution – Sad</td>
<td>19.8 ± 9.3</td>
<td>15.1 ± 8.6</td>
<td>4.3</td>
<td>1.52 × 10⁻³</td>
<td>0.03</td>
</tr>
<tr>
<td>Misattribution – Angry</td>
<td>3.0 ± 3.7</td>
<td>3.7 ± 5.8</td>
<td>-0.5</td>
<td>4.58 × 10⁻¹</td>
<td>0.46</td>
</tr>
<tr>
<td>Misattribution – Fearful</td>
<td>3.9 ± 6.0</td>
<td>2.6 ± 3.8</td>
<td>1.4</td>
<td>5.13 × 10⁻²</td>
<td>0.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intensity morphing taska</th>
<th>Depressed patients</th>
<th>Healthy controls</th>
<th>B</th>
<th>p value</th>
<th>p_corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase – Happy</td>
<td>51.1 ± 16.1</td>
<td>46.9 ± 14.9</td>
<td>4.5</td>
<td>5.09 × 10⁻²</td>
<td>0.31</td>
</tr>
<tr>
<td>Increase – Sad</td>
<td>53.9 ± 14.0</td>
<td>60.2 ± 15.2</td>
<td>-5.0</td>
<td>2.08 × 10⁻²</td>
<td>0.15</td>
</tr>
<tr>
<td>Increase – Angry</td>
<td>52.8 ± 13.2</td>
<td>55.6 ± 15.6</td>
<td>-1.9</td>
<td>3.68 × 10⁻¹</td>
<td>0.82</td>
</tr>
<tr>
<td>Increase – Fearful</td>
<td>56.1 ± 16.2</td>
<td>61.0 ± 16.7</td>
<td>-2.9</td>
<td>2.31 × 10⁻¹</td>
<td>0.82</td>
</tr>
<tr>
<td>Increase – Disgusted</td>
<td>55.5 ± 15.3</td>
<td>57.2 ± 14.7</td>
<td>0.06</td>
<td>9.80 × 10⁻¹</td>
<td>0.98</td>
</tr>
<tr>
<td>Decrease – Happy</td>
<td>29.0 ± 13.2</td>
<td>30.9 ± 16.2</td>
<td>-2.8</td>
<td>2.06 × 10⁻¹</td>
<td>0.82</td>
</tr>
<tr>
<td>Decrease – Sad</td>
<td>22.9 ± 12.2</td>
<td>31.9 ± 12.3</td>
<td>-9.2</td>
<td>1.00 × 10⁻⁶</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Decrease – Angry</td>
<td>21.9 ± 13.5</td>
<td>25.5 ± 12.2</td>
<td>-3.5</td>
<td>7.29 × 10⁻²</td>
<td>0.36</td>
</tr>
<tr>
<td>Decrease – Fearful</td>
<td>24.4 ± 13.6</td>
<td>29.7 ± 11.4</td>
<td>-5.0</td>
<td>7.37 × 10⁻³</td>
<td>0.06</td>
</tr>
<tr>
<td>Decrease – Disgusted</td>
<td>17.2 ± 9.9</td>
<td>22.0 ± 12.4</td>
<td>-4.7</td>
<td>6.29 × 10⁻³</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moral emotions task</th>
<th>Depressed patients</th>
<th>Healthy controls</th>
<th>B</th>
<th>p value</th>
<th>p_corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent/intentional – Guilt</td>
<td>6.3 ± 0.6</td>
<td>6.1 ± 0.8</td>
<td>0.1</td>
<td>2.00 × 10⁻¹</td>
<td>0.40</td>
</tr>
<tr>
<td>Agent/accident – Guilt</td>
<td>6.1 ± 0.8</td>
<td>5.6 ± 0.8</td>
<td>0.4</td>
<td>4.53 × 10⁻⁴</td>
<td>0.002</td>
</tr>
<tr>
<td>Victim/intentional – Guilt</td>
<td>2.5 ± 1.1</td>
<td>1.7 ± 0.7</td>
<td>0.8</td>
<td>1.67 × 10⁻⁸</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Victim/accident – Guilt</td>
<td>2.0 ± 0.9</td>
<td>1.5 ± 0.6</td>
<td>0.5</td>
<td>8.00 × 10⁻⁶</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Agent/intentional – Shame</td>
<td>6.1 ± 0.8</td>
<td>6.0 ± 0.9</td>
<td>0.03</td>
<td>8.00 × 10⁻¹</td>
<td>0.80</td>
</tr>
<tr>
<td>Agent/accident – Shame</td>
<td>5.9 ± 0.8</td>
<td>5.5 ± 0.9</td>
<td>0.3</td>
<td>7.68 × 10⁻³</td>
<td>0.02</td>
</tr>
<tr>
<td>Victim/intentional – Shame</td>
<td>3.5 ± 1.3</td>
<td>2.5 ± 1.0</td>
<td>0.9</td>
<td>2.48 × 10⁻⁷</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Victim/accident – Shame</td>
<td>2.1 ± 1.1</td>
<td>1.5 ± 0.6</td>
<td>0.6</td>
<td>7.72 × 10⁻¹</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verbal affective memory taskb</th>
<th>Depressed patients</th>
<th>Healthy controls</th>
<th>B</th>
<th>p value</th>
<th>p_corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate recall</td>
<td>14.6 ± 3.3</td>
<td>16.0 ± 3.0</td>
<td>-1.6</td>
<td>7.17 × 10⁻⁴</td>
<td>0.001</td>
</tr>
<tr>
<td>Short-term recall</td>
<td>14.7 ± 4.6</td>
<td>17.2 ± 4.5</td>
<td>-2.8</td>
<td>3.40 × 10⁻³</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>15.0 ± 4.9</td>
<td>17.9 ± 4.2</td>
<td>-3.2</td>
<td>3.00 × 10⁻⁶</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Raw p values as well as corrected p values are reported (see Method sections for description). β Values represent difference in scores between patients and healthy controls once age and sex has been accounted for.

dDepressed patients n = 91.

Depressed patients n = 86.

https://doi.org/10.1017/S0033291720000938 Published online by Cambridge University Press
not observe any differences between patients and healthy controls, suggesting that it is not a lack of attention towards or a preference away from social information that is causing the reported ToM deficits in MDD. This aligns with reports that dysphoria is associated with slightly increased sensitivity to social cues required for ToM (Harkness, Sabbagh, Jacobson, Chowdrey, & Chen, 2005). Our depressed patients did not exhibit negative bias in the interpretation of social situations. We speculate that this may partially be related to the task design; in several scenarios the negative interpretation had paranoid components (e.g., believing a colleague is poisoning a cup of tea) and might be too extreme to capture the more subtle negative biases in MDD.

Cold cognition

We were able to replicate previously reported impairments in verbal memory, working memory and reaction time (Bennabbi et al., 2013; Goodall et al., 2018; Lee et al., 2012; Marazziti, Consoli, Picchetti, Carlini, & Faravelli, 2010; Rock et al., 2014) showing moderate effect sizes. The difference in number of remembered words in the VAMT-26 between patients and healthy controls also appeared to become progressively larger across the three time points (immediate recall, 1.6 words; short-term recall 2.8 words, and long-term recall, 3.2 words). This could potentially indicate that initial learning is less affected than long-term memory or alternatively reflect effects of fatigue and/or apathy in the depressed patients (Marazziti et al., 2010).

Correlation with depression severity

We were unable to identify a clear association between any of the individual cognitive task domains and depression severity indexed with HDRS$_{6}$ or HDRS$_{17}$ scores. This suggests that cognitive disturbances may not simply be an extension of the ‘classic’ core mood and somatic symptoms in MDD but rather represent distinct characteristics of the depressive pathology. Furthermore, while cognitive disturbances may be largely independent from symptom severity during the depressive episode, they are still associated with long-term clinical and functional outcomes (Cicchetti, 1994; Collins et al., 2011) and thus hold promise as a relevant stratification tool for the identification of clinically meaningful subgroups in MDD. Indeed, empirical trials are currently investigating the benefits of cognition-based tools for optimising treatment in MDD (Kingslake et al., 2017). However, more work is still needed to evaluate the clinical value of such approaches.

Clustering of cognitive profiles

In an exploratory analysis, we identified three clusters of cognitive profiles in the depressed patients. Cluster A was the largest group ($n = 38$) and characterised patients with strong negative biases in emotion recognition and misattribution but no substantial deficits in cold cognitive domains apart from slowed reaction time. Cluster B ($n = 28$) conversely characterised patients with positive biases in emotion processing and moderate deficits across all cold cognition domains. Lastly, Cluster C ($n = 26$) characterised patients who had large deficits across both hot and cold cognitive domains and in particular extremely high ratings of shame and guilt. These findings not only suggest a dissociation between the presence of hot and cold cognitive deficits in MDD, as illustrated by the differences between Cluster A and B, but also the existence of a subgroup of patients with severe global cognitive deficits represented by Cluster C.

To our knowledge, only two other studies have used cluster analysis to identify cognitive profiles in MDD: the large iSPOT-D trial ($N = 1008$) (Etkin et al., 2015) and a smaller study ($N = 50$) in patients with first-episode depression (Vicent-Gil et al., 2018). Both studies identified two clusters based on performance on cold cognitive tasks: a large cluster of cognitive intact patients and a smaller cluster of cognitive impaired patients. Notably, the proportion of cognitive impaired patients reported in both studies was 25–26%, closely matching the size of the globally impaired Cluster C in the present cohort (~28%). This further aligns with previous reports that only a small proportion of patients experience pronounced impairments in cognitive performance with estimates ranging between 21 and 44% depending on the cognitive measures and cut-off criteria used (Gualtieri & Morgan, 2008; Iverson, Brooks, Langenecker, & Young, 2011; McIntyre et al., 2017). Importantly, none of the above studies included measures of affective biases or social cognition and may therefore have overlooked the presence of Cluster A type patients who exhibit strong negative biases in emotion processing but little to no deficits in cold cognitive domains. This highlights the importance of characterising both cold and hot cognitive disturbances in MDD concurrently.

Interestingly, the degree of cognitive disturbances across the three clusters partly mirrored the severity of depressive symptoms within the clusters, i.e. Cluster C had overall higher levels of depressive symptoms compared to Cluster A and B. This indicates that these cognitive profiles are able to capture MDD characteristics not captured by any individual task domain. Future studies should evaluate whether such cluster labelling, in addition to single cognitive domain information, may be useful for guiding antidepressants treatment choices and/or identify patients who will benefit from augmentation with e.g. cognitive remediation (Maruff & Jaeger, 2016) or cognitive enhancers (Bowie, Gupta, & Holshauser, 2013).

Methodological considerations

Some methodological limitations should be considered: (1) The depressed patients and healthy controls were unevenly matched on sex and because of the recruitment and inclusion procedures, the healthy controls likely represent very high-functioning individuals which may have inflated the observed differences on cognitive outcomes between the two groups. (2) We did not correct for the effect of IQ or education in the analyses as previous reports indicate that IQ measurement (Goss, Kaser, Costafreda, Sahakian, & Fu, 2013; Miskowiak et al., 2014) and education dropout rates (Marazziti et al., 2010) are affected by depressive symptoms. None of the reported estimates changed critically when IQ or education were included in the models (for corrected estimates see online Supplementary Materials). (3) Because the wordlist in the VAMT-26 contained both positively and negatively valanced words, the total word recall score does not represent a ‘pure’ cold measure of explicit memory. (4) Due to the limited stamina of the MDD patients, we had to restrict the number of cognitive domains tested; as a consequence, we did not collect data on e.g. attention or higher-executive functions despite their relevance in MDD pathology.

Conclusion

The current study represents one of the most comprehensive investigations into hot and cold cognitive impairment in a...
large, well-characterised and antidepressant-free cohort of depressed patients to date. This allowed us to assess and directly compare the magnitude and patterns of impairment across a broad range of cognitive domains as well as investigate the presence of clusters of distinct cognitive profiles in depression. It is also the first time tasks from the EMOTICOM test battery have been applied and shown to be sensitive to MDD pathology in a patient cohort. Overall, our findings highlight the importance of including both hot and cold cognitive domains in investigations into MDD and further suggest that cognitive measures capture features beyond those reflected by depression severity. While cognitive disturbances are not present in all patients, they do represent significant impairments in identifiable and large subgroups of patients that may benefit from augmentation with cognition targeted treatments. Thus, we argue that cognition-based tools hold promise as clinically useful stratification aids in the care of depressed patients.

**Supplementary material.** The supplementary material for this article can be found at [https://doi.org/10.1017/S0033291720000938](https://doi.org/10.1017/S0033291720000938).

**Acknowledgements.** We would like to thank the patients and healthy participants for their time. The data from the study is available upon request from the CIMBI database (http://www.cimbi.dk/db).

**Financial support.** This study was supported by the Augustinus Foundation for Sage Therapeutics. The remaining authors all declare no conflict of interest.


