Affective Enhancement of Working Memory Is Maintained in Depression

Susanne Schweizer  
Medical Research Council, Cambridge, United Kingdom

Lauren Navrady  
University of Edinburgh

Lauren Breakwell and Rachel M. Howard  
Medical Research Council, Cambridge, United Kingdom

Aliza Werner-Seidler  
Prince of Wales Hospital, Sydney, Australia, and University of New South Wales

Tim Dalgleish  
Medical Research Council, Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, United Kingdom

We currently know little about how performance on assessments of working memory capacity (WMC) that are designed to mirror the concurrent task demands of daily life are impacted by the presence of affective information, nor how those effects may be modulated by depression—a syndrome where sufferers report global difficulties with executive processing. Across 3 experiments, we investigated WMC for sets of neutral words in the context of processing either neutral or affective (depressogenic) sentences, which had to be judged on semantic accuracy (Experiments 1 and 2) or self-reference (Experiment 3). Overall, WMC was significantly better in the context of depressogenic compared with neutral sentences. However, there was no support for this effect being modulated by symptoms of depression (Experiment 1) or the presence of recurrent major depressive disorder (MDD; Experiments 2 and 3). Implications of these findings for cognitive theories of the role of WM in depression are discussed in the context of a growing body of research showing no support for a differential impact of depressogenic compared with neutral information on WM accuracy.

Keywords: working memory capacity, emotion, depression, complex span

Supplemental materials: http://dx.doi.org/10.1037/emo0000306.supp

Individuals suffering from major depressive disorder (MDD) reliably show difficulties with executive functioning (Snyder, 2013). Yet the origin and impact of these difficulties on everyday mental life remains underexplored. Executive functioning problems, especially in affective disorders, have been proposed to arise as a result of cognitive resources being preferentially deployed toward the processing of affectively salient material (Gotlib & Joormann, 2010). That is, resources are recruited by affectively laden thoughts, feelings, and behavioral impulses, which although relevant to individuals’ immediate goal states (e.g., attempting to reduce negative affect), are often less relevant to the executive task at hand and therefore have the capacity to hinder performance on that task (Mason et al., 2007; Pessoa, 2009). Interestingly however, until recently these executive functions have typically been investigated in the laboratory using affect-neutral tasks (Snyder, 2013). Explicitly investigating executive functioning within the context of affective information in the laboratory arguably provides a more naturalistic evaluation of the problems that individuals suffering from emotional disorders have with everyday mental operations in the context of their negative mood states and associated cognitions and behavioral tendencies (Schweizer & Dalgleish, 2011, 2016).

Affective Working Memory in MDD

Preliminary advances in the investigation of executive functioning in affective contexts have been made in the area of working memory (WM), a capacity-limited system that consists of an executive control component interacting with one or more storage systems to transiently maintain and store information in the service of other forms of cognition (Baddeley, 2003). Specifically, WM
performance in MDD has been assessed in contexts where either affective material has to be remembered as part of the task (e.g., Joormann & Gotlib, 2008) or where the affective material is present in the form of task-irrelevant distractor stimuli (e.g., Bertocci et al., 2012). These studies have produced equivocal findings. No studies found support for a differential impact of affective (positive or negative) material on WM accuracy in individuals currently suffering from MDD (Berman et al., 2011; Bertocci et al., 2012; Foland-Ross et al., 2013; Joormann & Gotlib, 2008; Joormann, Levens, & Gotlib, 2011; Joormann, Nee, Berman, Jonides, & Gotlib, 2010; Ladouceur et al., 2005; Levens & Gotlib, 2009, 2010; Tavitian et al., 2014; Yoon, LeMoult, & Joormann, 2014)\(^1\) nor in those in remission from their latest episode of MDD (De Lissnyder et al., 2012; Demeyer, De Lissnyder, Koster, & De Raedt, 2012; Kerestes et al., 2012; Levens & Gotlib, 2015) compared with healthy controls. For some of these tasks, WM reaction time (RT) appeared to be a more sensitive measure of WM performance with individuals suffering from MDD showing slowed RTs on WM tasks in the context of negative material compared with positive and/or neutral information (De Lissnyder et al., 2012; Demeyer et al., 2012; Joormann et al., 2011, 2010; Ladouceur et al., 2005; Levens & Gotlib, 2010, 2015; Yoon et al., 2014). However, even in the case of RT, the data was mixed with several studies finding no support for affective influences (Foland-Ross et al., 2013; Joormann & Gotlib, 2008; Levens & Gotlib, 2009; Tavitian et al., 2014).

**Is Affective Working Memory Capacity a More Sensitive Measure?**

A potential account of these equivocal findings is that the WM measures employed (such as the n-back task and the modified Sternberg task) were measures of WM storage rather than assessments of what has been called WM capacity (WMC; Conway et al., 2005). WMC tasks seek to index the executive control component of WM described above, as opposed to simply storage. That is, individuals’ capacity to attend and respond to task-relevant information while inhibiting attention and prepotent responses to task-irrelevant distractions (Engle, 2002; Engle & Kane, 2003). WMC has been shown to be predictive of individual differences on a large number of higher-cognitive functions (Barrett, Tugade, & Engle, 2004). Individual differences in performance on a task assessing WMC in the context of affective distraction are therefore arguably predictive of functioning on a variety of cognitive operations (e.g., remembering the grocery list or keeping track of proceedings at a work meeting) requiring WMC in everyday life in affective contexts, such as depressed feeling states and thoughts.

Preliminary research in analogue samples examining the effects of depressed mood on WM in the context of affective information appear to provide clearer support for the view that the presence of depressive symptoms would modulate the effects of affective material on WMC. In particular, a recent pair of studies showed that affective WMC in dysphoric undergraduates was impaired compared with nondysphoric students on an affective complex span task (Hubbard, Hutchison, Hambrick, & Rypma, 2016; Hubbard, Hutchison, Turner, et al., 2016). Complex span tasks are considered a gold-standard for operationalizing WMC and are defined as tasks that require simultaneous performance of a target storage task and some form of operation task (Conway et al., 2005). In both studies, Hubbard and colleagues showed that dysphoric participants remembered fewer numbers (storage task) than nondysphoric controls when they were required to simultaneously make judgments of self-relevance about depressogenic statements derived from symptom measures of depression (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Lako et al., 2014; Radloff, 1977). There were no group differences in numbers recalled when the operation task required participants to make semantic decisions about neutral sentences. While these findings suggest that measures of WMC in affective contexts may be most sensitive to WM deficits reported in MDD, there are two methodological considerations that together merit a note of caution. First, it is unclear differential performance across the conditions was due to differences in the affective significance of the task stimuli (cf., Pessoa, 2009) or differences in the type of decision made (i.e., a semantic judgment vs. a self-relevance judgment). Second, these studies were conducted in seemingly high functioning healthy individuals (i.e., undergraduates) who reported dysphoric mood states. The underlying processes leading to WM deficits in these dysphoric students on this task may therefore be markedly different to the processes operating in MDD.

The Present Study

In the present series of experiments, we therefore aimed to address the question of whether there is impact of affective information on WM in MDD when performing an affective WMC task (aWMC-task). We used the affective reading span task (Schweizer & Dalgleish, 2011), modified to include depressogenic material. The task required participants to recall words presented at the end of either neutral or depressogenic statements, which they had to judge in both cases in terms of semantic accuracy (Experiments 1 and 2) or self-relevance (Experiment 3). The task has previously been shown to be sensitive to differences in WM in affective context between healthy individuals and those with emotional disorders (i.e., posttraumatic stress disorder (PTSD); Schweizer & Dalgleish, 2011). Specifically, in line with the dual competition framework (Pessoa, 2009) healthy individuals showed improved WMC for words following affective (relative to neutral) task-relevant information. The enhancement effect, however, was shown to be attenuated in individuals with a lifetime history of PTSD (Schweizer & Dalgleish, 2011).

Pessoa’s (2009) dual competition framework of the impact of affective significance on executive control can also be used to guide the predictions for the current experiments. According to the theory, low-threat (i.e., affective significance) affective information is prioritized at both the perceptual and executive levels of processing, thereby improving executive performance on tasks that include task-relevant affective information (Pessoa, 2009; Vuilleumier, 2005). The framework further suggests that stimuli high in affective significance hijack executive control resources thereby impairing performance on concurrent executive tasks. A recent meta-analysis has generated broad support for these predictions (Schweizer et al., 2016).

Based on the dual competition framework, we predicted that WMC in the presence of depressogenic compared with neutral

---

\(^1\) A study in a sample of youth with MDD showed impaired (relative to healthy controls) WM for neutral trials (Tavitian et al., 2014).
sentences should be improved overall in a nonclinical community sample. However, we further hypothesized that individuals with high levels of depressive symptoms in that sample (Experiment 1) and those with MDD (Experiments 2 and 3) would be more impaired by negative, depressogenic sentences compared with those with lower levels of symptoms (Experiment 1) or never-depressed controls (Experiments 2 and 3), because the negative sentences arguably hold greater affective significance for depressed individuals.

**Experiment 1**

Experiment 1 tested the hypothesis that performance on the aWMC-task would vary as a function of the valence (depressogenic vs. neutral) of the sentences that proceeded the memoranda. Based on the dual competition framework (Pessoa, 2009) reviewed above, we predicted that overall participants would show better WMC capacity in the presence of depressogenic material relative to neutral material (Hypothesis 1). However, aWMC (operation-alized as performance in the depressogenic condition after subtracting performance in the neutral condition) was also predicted (Hypothesis 2) to show a negative association with depressive symptoms, as depressive symptomatology arguably increases the distractor sentences’ affective significance. That is, we expected those with higher levels of depressive symptoms to be more impaired in their WMC in the context of depressogenic sentences relative to neutral. We investigated these hypotheses in a community sample selected to provide a range of self-reported depression symptom severity across the non- to mild to severely depressed range. To further approximate a clinical case-control design, we also investigated whether individuals in this community sample scoring in the “depressed” range on the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996) would perform differently on the aWMC-task compared with those scoring in the non-depressed range.

**Method**

**Participants.** Exactly 123 participants (age range: 18–64 years, M = 41 years, sd = 14.21, 83 women) were recruited from the Medical Research Council Cognition and Brain Sciences Unit volunteer panel. To be included, participants needed to be between 18 and 65 years old, speak English with native fluency, have normal (or corrected to normal) vision, and be neurologically healthy (no history of neurological disorders or head injuries). Because of the focus of the study on depressive symptomatology, we used a stratified approach to recruit participants from across the whole range of scores on the BDI using scores recorded during previous assessments as an initial screen. The scale was then readministered in the current study. Specifically, we sourced participants classified as “nondepressed” (BDI scores of 0–13), alongside participants in the “depressed” range (>14), based on these previous BDI scores (Dozois, Dobson, & Ahnberg, 1998; Segal, Coolidge, Cahill, & O’Riley, 2008). The included sample presented with a good range of BDI scores (0–36; the maximum range is 0–63), with n = 83 (67.5%) in the nondepressed range, and n = 40 (32.5%) in the depressed range (of whom some, n = 3 (2.4%), would be classified as severe in clinical terms).

**Measure: aWMC-task.** To measure aWMC, we modified the valenced reading span task from Schweizer and Dalgleish (2011). In keeping with the original reading span task (Daneman & Carpenter, 1980), the valenced version is a complex span task (Conway et al., 2005), which requires the storage of single words (for a detailed description of the words included see: Schweizer & Dalgleish, 2011), in tandem with an operation component (evaluating sentences on their semantic accuracy) that potentially disrupts participants’ ability to memorize the to-be-remembered material. Valence is manipulated by including blocks of neutral and negative sentences. In this version of the aWMC-task, the negative sentences were relevant to depression and were derived from the 100-item Dysfunctional Attitude Scale (DAS; Weissman & Beck, 1978). The DAS has been shown to be sensitive to cognitive, genetic, and physiological vulnerabilities to depression (Beeser, Strong, Meyer, Pillkonis, & Miller, 2007; Meyer et al., 2003; Oliver, Murphy, Ferland, & Ross, 2007; Zuroff, Igreja, & Mon-
Procedure. The following procedure was broadly consistent across all three experiments; any minor deviations are noted in the appropriate sections. After providing informed consent, participants completed the BDI-II and NART, before performing the aWMC-task. Participants were compensated for their time (£6 per hour). All 3 experiments were approved by the Cambridge Local Research Ethics Committee.

Results and Discussion

Given research showing that WMC deteriorates with age (e.g., Campbell & Charness, 1990) and the possible influence of verbal IQ on this verbal version of the WMC task, we first explored potential correlations between the measures. aWMC correlated significantly with neither age, r(123) = −.14 nor verbal IQ r(112) = .03, with trivial effect sizes, and so we did not include either variable in our regression modeling.

Overall, in line with our first hypothesis, participants showed significantly better WMC for words following depressogenic ($M = .54, SD = .18$) compared with neutral ($M = .50, SD = .16$) sentences, $F(1, 122) = 22.17, p < .001, \eta^2 = 0.15$. The moderate to large effect was confirmed with Bayesian analysis. The posterior probability of an effect of valence was $p(H1, D) = 1$. The posterior probability of the experimental hypothesis indicates the likelihood of the effect being significant if this analysis were repeated in a different hypothetical sample. That is, the likelihood for an effect of valence on WMC in a different hypothetical sample is 100%. As noted in the Introduction, this is in line with Pessoa’s (2009) dual competition framework for the effects of affective significance on executive control performance. These findings further concur with evidence from the long-term memory literature, which shows a significant memory enhancement for affective stimuli and events (Buchanan & Adolphs, 2002; Phelps, 2004) as well as with our previous work in PTSD showing an affective enhancement effect on this task in trauma-exposed participants (Schweizer & Dalgleish, 2011). Accuracy on the semantic judgment task was high (98%), and did not differ across valence (neutral: $M = .99, SD = .09$; depressogenic: $M = .98, SD = .04$), $F(1, 122) = 1.70, p = .20, \eta^2 = .01$. This was confirmed by Bayesian analyses with the Bayes Factor providing moderate support for the null hypothesis, BF$_{01}$ = 3.22. That is, the effects of valence on WMC accuracy are unlikely to be accounted for by the semantic decision being more cognitively demanding across valence conditions.

In contrast with our second hypothesis, symptoms of depression did not significantly predict aWMC, $b = .001, 95\%$ CI [−0.001, 0.002], $t = 0.71, p = .478, R^2 = .005$. In order to examine the extent to which the null hypothesis of no effect of depressive symptoms on aWMC could be supported, this absence of an effect was again investigated using Bayesian analysis. The Bayes Factor for the null hypothesis was BF$_{01}$ = 3.89 whereas the posterior probability an effect of depression was $p(H1, D) = 0.20$. The BF$_{01}$ here provides moderate (values between 3 and 10 are considered to provide moderate support) evidence in support of the null hypothesis (Jeffreys, 1998). And the likelihood of finding a significant effect of depressive symptoms on aWMC in a different hypothetical sample was only 20%.

Finally, a comparison of aWMC in individuals scoring in the depressed range (above the cutoff of ≥14 on the BDI-II; Carney, Ulmer, Edinger, Krystal, & Knauss, 2009; Leentjens, Verhey, Luijckx, & Troost, 2000; n = 40, $M_{BDI} = 20.63, SD = 5.49$) with those who scored in the nondepressed range (n = 83; $M_{BDI} = 5.28, SD = 4.51$) revealed no significant effect of BDI-cutoff group on aWMC $F ≤ 1, \eta^2 = 0.01$ (see Table 1). These results were again supported by Bayesian analyses (see Table 1).

These findings were surprising because, based on the dual competition framework, we expected the affective significance of the negative statements to increase as a function of depressive symptomatology and aWMC to concurrently deteriorate. Moreover, we had expected the negative statements to engender depressogenic (e.g., ruminative) processing in those with depressive tendencies and thereby recruit cognitive resources away from the storage component of the aWMC-task. The results are also at odds with the complex span data in dysphoric samples reported previously (Hubbard, Hutchison, Hambrick, et al., 2016; Hubbard, Hutchison, Turner, et al., 2016), although as noted there are also methodological issues surrounding the design of those studies. However, it should be reiterated that the current pattern of data is consistent with other studies looking at WM in affective contexts in depression using simple span WM tasks (Berman et al., 2011; Bertocci et al., 2012; De Lissnyder et al., 2012; Demeyer et al., 2012; Foldand-Ross et al., 2013; Joormann & Gotlib, 2008; Joormann et al., 2011, 2010; Ladouceur et al., 2005; Levens & Gotlib, 2009, 2010, 2015; Tavitian et al., 2014; Yoon et al., 2014).

One possible account of the present finding is that these ruminative processes and other mechanisms, which may impact on WMC, only impair concurrent cognitive processing in those with recurrent clinical depression, which is associated with more ruminative responses compared with subclinical dysphoria (Garnefski et al., 2002). That is, the increased (relative to subclinical depression) tendency to ruminate on the depressive sentences reduces the cognitive resources available to perform the WM task. Specifically, rumination and WMC are both dependent on executive control processes (Cooney, Joormann, Eugène, Dennis, & Gotlib, 2010; Demeyer et al., 2012). We therefore explored aWMC-task performance in individuals with a current diagnosis of recurrent MDD compared with never-depressed individuals.

Experiment 2

Experiment 2 employed a case-control design using the same procedure as Experiment 1 to compare aWMC in never-depressed participants against individuals with a diagnosis of recurrent MDD. We predicted greater aWMC in the never depressed individuals (ND group) relative to individuals in the MDD group.

---

1 The correlations were also nonsignificant between BDI-II scores and the individual condition WMC scores (i.e., depressogenic WMC and neutral WMC), r’s (112) = −.05–.18, p values > .05.
2 NART scores were missing for 9 participants. However, given the nonsignificant association between verbal IQ and aWMC, which meant there was no case for covarying for verbal IQ, we included these participants in the overall analyses. We ran all analyses excluding these participants as well, and the pattern of results remained unchanged.
3 Differential accuracy (depressogenic—neutral) on the semantic operation task was not predicted by symptoms of depression, $b = .000, 95\%$ CI [−0.002, 0.002], $t = 0.14, p = .89, R^2 = .000$. 

---

*Schweizer, Edinger, Krystal, & Knauss, 2009; Leentjens, Verhey, Luijckx, & Troost, 2000; $n = 40, M_{BDI} = 20.63, SD = 5.49$; those who scored in the nondepressed range ($n = 83; M_{BDI} = 5.28, SD = 4.51$) revealed no significant effect of BDI-cutoff group on aWMC $F ≤ 1, \eta^2 = 0.01$ (see Table 1). These results were again supported by Bayesian analyses (see Table 1).
Table 1
WM Performance across Depressive State and Condition

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;BDI-II cutoff</td>
<td>&gt;BDI-II cutoff</td>
<td>Never depressed</td>
</tr>
<tr>
<td></td>
<td>n = 83</td>
<td>n = 40</td>
<td>n = 14</td>
</tr>
<tr>
<td>Neutral (SD)</td>
<td>.53 (.17)</td>
<td>.48 (.13)</td>
<td>.50 (.15)</td>
</tr>
<tr>
<td>Depressogenic (SD)</td>
<td>.56 (.19)</td>
<td>.52 (.16)</td>
<td>.53 (.22)</td>
</tr>
<tr>
<td>aWMC</td>
<td>.03 (.09)</td>
<td>.04 (.08)</td>
<td>.03 (.12)</td>
</tr>
<tr>
<td>Effect sizes (η^2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valence</td>
<td>η^2 = .15</td>
<td>η^2 = .13</td>
<td>η^2 = .04</td>
</tr>
<tr>
<td>Group</td>
<td>η^2 = .02</td>
<td>η^2 = .04</td>
<td>η^2 = .01</td>
</tr>
<tr>
<td>Valence × Group</td>
<td>η^2 = .01</td>
<td>η^2 = .02</td>
<td>η^2 = .01</td>
</tr>
</tbody>
</table>

Bayes factor in support of null hypothesis (BF_01) and posterior probability for the experimental hypothesis p(H1, D)

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>BF_{01} = 4.53; p(H1, D) = 1</td>
<td>BF_{01} = .58; p(H1, D) = .63</td>
<td>BF_{01} = .88; p(H1, D) = .52</td>
</tr>
<tr>
<td>WMC: Group</td>
<td>BF_{01} = 1.87; p(H1, D) = .35</td>
<td>BF_{01} = 1.87; p(H1, D) = .35</td>
<td>BF_{01} = 1.13; p(H1, D) = .47</td>
</tr>
<tr>
<td>Valence × Group</td>
<td>BF_{01} = 3.25; p(H1, D) = .23</td>
<td>BF_{01} = 2.61; p(H1, D) = .28</td>
<td>BF_{01} = 3.26; p(H1, D) = .23</td>
</tr>
</tbody>
</table>

Note. Neutral = proportion of words recalled correctly in the context of neutral sentences; depressogenic = proportion of words recalled correctly in the context of depressogenic sentences; < BDI-II cutoff = individuals scoring 14 or lower on the BDI-II; > BDI-II cutoff = individuals scoring 15 or higher on the BDI-II; never depressed = individuals with no history of MDD; current MDD = individuals currently suffering from MDD; remitted MDD = individuals in remission from MDD; valence = effect comparing negative versus neutral trials; group = effect of group on WMC; Valence × Group = interaction between valence and group.

Method

Participants. Participants with recurrent MDD were recruited through newspaper and other local advertisements. Never-depressed participants were recruited through the departmental volunteer panel. Recurrent MDD was assessed upon arrival for testing using the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders (SCID; First, Spitzer, Gibbon, & Williams, 1995). The total sample in Experiment 2 included 36 participants (age range: 19–65, M = 51 years, sd = 10.77; 17 women), with 14 in the ND group and 22 in the MDD group. For one participant in the MDD group, the computer malfunctioned. The final sample therefore included 21 MDD participants (of whom 13 were experiencing a current episode of depression and 8 were in remission).

However, compared with the ND group (93%, n = 13), there were significantly fewer women in the MDD group (50%, n = 11), X^2 (1) = 5.19, p = .023 (with Yates continuity correction). Subsequent group comparisons therefore included gender as a covariate.

Results and Discussion

In line with Experiment 1, there was a significant effect of valence (neutral: M = .46, SD = .16; depressogenic: M = .49, SD = .20) on WMC with better performance in the presence of depressogenic sentences, F(1, 34) = 4.88, p = .034, η^2 = 0.13. p(H1, D) = .63. As in Experiment 1, the effect of valence on WMC accuracy does not seem to be accounted for by differential difficulty of the semantic decision across conditions with high overall accuracy (97%), and no significant difference across conditions (neutral: M = .96, SD = .04; depressogenic: M = .97, SD = .04), F(1, 34) = 3.32, p = .08, η^2 = .09. Though it should be noted, Bayesian analyses supported neither experimental BF_{10} = 1.09 nor null hypothesis, BF_{01} = 0.99. There was no significant effect of group, F(1, 32) = 1.434, p = .240, η^2 = 0.04; p(H1, D) = .35 on WMC. Finally, failing to support our prediction, there was no significant interaction between group and valence, BF_{10} = 0.02 with scores on the index of aWMC (WMC in depressogenic contexts minus WMC in neutral contexts) indicating comparable enhancement in depressogenic contexts across the MDD and ND groups (see Table 1 for means and Bayesian statistics). Because the MDD group also included individuals who were not currently experiencing an episode (n = 8), we repeated these analyses comparing the ND group only to individuals who were currently in episode. Again, there was no significant difference between groups on aWMC, F(1, 25) < 1; BF_{01} = 4.37; p(H1, D) = .19 (for means and effects of valence and group on WMC, see online supplemental materials). In other words, there was strong support for a null effect and the likelihood of an effect of diagnosis of MDD (current vs. never) on aWMC in a different sample was only 19%.

As noted, the absence of an effect of recurrent MDD on aWMC in Experiment 2 and the lack of association between aWMC and depressive symptoms in Experiment 1 is in contrast with Hubbard and colleagues (Hubbard, Hutchison, Hambrick, et al., 2016; Hubbard, Hutchison, Turner, et al., 2016), who showed a dissociation of WMC performance for affective compared with neutral contexts in their dysphoric student samples. A possible account for this discrepancy may be that the material in the present experiment was processed in a semantic rather than self-referential manner (Wisco, 2009). This hypothesis was investigated in Experiment 3.

Moreover, there was no significant group-by-valence interaction for accuracy in semantic decision making, F(1, 33) = 0.30, p = .59; η^2 = 0.01.
Experiment 3

Self-referential processing in depression has been argued to activate self-schemas and bias information processing toward self-referent material (Alloy, Abramson, Murray, Whitehouse, & Hogan, 1997). Previous research in individuals with depression has shown a recall advantage for words that were previously processed in a self-referential manner (Alloy et al., 1997; Kuiper & Derry, 1982; Kuiper, Olinger, MacDonald, & Shaw, 1985). In line with this, the long-term memory literature shows a mood-congruent recall bias for negative, depressogenic information in depression (for reviews see: Dalgleish & Watkins, 1990; Gotlib & Krasnoperova, 1998), particularly for the most self-referent of memories: autobiographical memories (see Dalgleish & Werner-Seidler, 2014).

Here we investigated the effect of self-referential processing of depressogenic material on aWMC. We used the same case-control design as for Experiment 2 but modified the operation component of the aWMC-task. In this third experiment, participants were required to make judgments of self-reference instead of semantics regarding the statements that they read as part of the operation task. As the memorandum in the current task (the neutral words presented alongside the sentences) were unrelated to the sentences that required a self-reference judgment, the expectation was that drawing attention toward the sentences would therefore impair memory for the storage component of the aWMC.

In addition to these competing hypotheses regarding the effect of self-referential processing on aWMC in those with MDD currently in episode, in this study we also chose to explore aWMC in individuals with MDD in remission. It is well established that the performance of the never-depressed comparison conditions suggesting that nonendorsement of such statements are consider-ably slower to generate in remitted individuals relative to compar-

Method

Participants. Recruitment procedures were identical to Ex-

Results and Discussion

Self-reference ratings. As would be expected, analyses of the rating scores showed a significant interaction between group and endorsement, F(2, 66) = 30.92, p < .001, η²p = 0.48, with the ND and MDDC-groups showing similar response patterns, and only the MDDDc-group showing greater relative endorsement of depressogenic sentences (see Figure 1). The finding suggests that participants were processing the sentence content appropriately and that the sentences included were sensitive to active depressive concerns. Moreover, the finding is in line with research showing that endorsement of the DAS is mood-dependent (Miranda et al., 1990).

aWMC performance. In line with Experiments 1 and 2, when performance on endorsed and nonendorsed sentences were included together, there was a significant effect of valence (neu-

gender distribution. There was, however, a significant difference in age, F(1, 67) = 5.21, p = .008, d = 0.79 with the remitted group (M = 51.17, SD = 16.77), being the oldest and individuals in the MDDC-group (M = 44.15, SD = 14.30) being older than the ND group (M = 39.95, SD = 13.21). All group comparisons therefore included age as a covariate.

aWMC-task. As before, the aWMC-task required the storage of single words (presented at the end of each sentence), which had to be recalled at the end of each trial. In tandem with the storage task, participants had to perform an operation task that potentially disrupts their ability to memorize the to-be-remembered material. In the self-reference version of the aWMC, the operation task probed participants to rate how much they agreed with the state-

memory for the storage component of the aWMC.

In this third experiment, participants were required to make judgments of self-reference instead of semantic judgments that they had just read out-loud (instead of making a semantic judgment about these sentences as in Experiments 1 and 2). Agreement ratings were provided on a Likert scale ranging from 1 (not at all) to 7 (very much). The depressogenic trials used the same sentences that were used in the aWMC for Experiments 1 and 2. The neutral sentences were modified to include neutral statements about which participants could make a judgment of self-reference (e.g., “I plan out what I am going to do during the day” or “I try to get at least seven hours sleep every night”). As before, the neutral sentences were matched for length and reading ease with the depressogenic sentences.

Recruitment procedures were identical to Ex-

While there was no significant effect of group on WMC, F(1, 67) = 4.09, p = .047, η²p = 0.06, with better performance again in the presence of depressogenic sentences. As in Experiment 2, there was no significant effect of group on WMC, F(2, 64) = 1.31, p = .244, η²p = 0.01. The finding was confirmed by Bayesian analyses, BF01 = 2.19, p(H1, D) = .31. There was also no significant group by valence interaction, F(2, 64) = 1.21, p = .305, η²p = 0.03, BF01 = 4.08, p(H1, D) = .20. Examining the aWMC scores again revealed comparable performance across the ND and the MDD groups with the absence of a group by valence interaction again confirmed by Bayesian analyses with moderate support for the null hypothesis BF01 = 4.08 and an interaction only likely to be found in 20% of hypo-

6 However see: Haefel et al. (2005) for a failure to replicate the predictive validity of the DAS in a remitted sample.

7 We also ran all analyses excluding the 7 participants that were older than 65 years, and this did not change the pattern of results.
A potential account for the absence of a three-way interaction is that the affective salience of endorsed items is insufficient to adversely impact aWMC differentially in those with depression. Although the depressed group endorsed more of the depressogenic sentences as self-relevant, the mean strengths of endorsement were only $M = 5.00^{\text{SD}} (SD = 0.87)$ on a 7-point scale. Although lack of affective impact is therefore an important factor to consider, it should also be noted that, in the context of the differential activation hypothesis, discussed above, the impact of depressogenic statements on cognitive processing has been observed in other cognitive tasks in the absence of endorsement of these statements. For example, Sheppard and Teasdale (2004) showed that response latencies on a semantic decision making task in currently depressed and remitted individuals were comparable despite remitted individuals endorsing the statements less strongly compared with currently depressed individuals.

Finally, given the findings by Hubbard and colleagues (Hubbard, Hutchison, Hambrick, et al., 2016, Hubbard, Hutchison, Turner, et al., 2016), which showed a difference between groups only for the depressogenic condition, we investigated the interaction of the effects of group and endorsement in depressogenic trials only. Again there was no significant interaction between group and endorsement when including all three groups, $F(2, 64) < 1; \eta^2_p = 0.02; \text{BF}_{01} = 5.06; p(H(1, D) = .17, \text{nor when comparing the MDD}_c$ and ND groups only, $F(1, 41) < 1; \eta^2_p = 0.02; \text{BF}_{01} = 2.18; p(H(1, D) = .31$.

**Meta-Analytic Investigation of the Effect Size Across Experiments 1–3**

**Valence.** There was good support of an overall effect of valence across the three experiments with a pooled effect size of $\eta^2_p = 0.11$, which translates to a Cohen’s $d$ of 0.70. Bayesian analyses confirmed the conclusion from the pooled analyses with conclusive support for a rejection of the null hypothesis $\text{BF}_{01} = 2.12^{\text{a}}$ and a posterior probability of $p(H(1, D) = 1$ for an effect of valence. That is, using this task an effect of valence on WMC should be found in 100% of hypothetical future samples.

**Effect of clinical levels of depression on aWMC.** aWMC was unaffected by clinical levels of depression as was evidenced by the pooled effect size across all three experiments of $\eta^2_p = 0.00$ (Cohen’s $d = 0.06$). Again these meta-analytic results were confirmed by Bayesian analyses on the pooled sample with moderate support in favor of accepting the null hypothesis $\text{BF}_{01} = 6.38$ and a posterior probability of $p(H(1, D) = .14$ for an effect of clinical depression on aWMC.

**General Discussion**

The first aim of the present series of experiments was to investigate the hypothesis that WMC would be enhanced in the context...
Table 2
Proportion of Words Recalled across Valence, Group and Endorsement

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Never depressed</th>
<th>Remitted MDD</th>
<th>Current MDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral endorsed M (SD)</td>
<td>.61 (.16)</td>
<td>.55 (.20)</td>
<td>.51 (.17)</td>
</tr>
<tr>
<td>Neutral not-endorsed M (SD)</td>
<td>.51 (.15)</td>
<td>.50 (.19)</td>
<td>.45 (.17)</td>
</tr>
<tr>
<td>Depressogenic endorsed M (SD)</td>
<td>.60 (.19)</td>
<td>.52 (.22)</td>
<td>.51 (.18)</td>
</tr>
<tr>
<td>Depressogenic not-endorsed M (SD)</td>
<td>.60 (.16)</td>
<td>.55 (.16)</td>
<td>.52 (.21)</td>
</tr>
</tbody>
</table>

Note. Neutral endorsed = proportion of correctly recalled words that followed neutral sentences, which participants endorsed (i.e., rated 5 or higher); neutral not-endorsed = proportion of correctly recalled words that followed neutral sentences, which participants did not endorse (i.e., rated 4 or less); depressogenic endorsed = proportion of correctly recalled words that followed depressogenic sentences, which participants endorsed (i.e., rated 5 or higher); depressogenic not-endorsed = proportion of correctly recalled words that followed depressogenic sentences, which participants did not endorse (i.e., rated 4 or less).

of affective compared with neutral information. Second, we aimed to resolve the discrepant findings in the literature on the effects of depression on aWM using a complex span task. In line with the first hypothesis, the results showed a moderate to large (i.e., Cohen’s $d = 0.50 – 0.84$) WM-recall advantage for words following depressogenic sentences (relative to neutral) across three different studies in analyses combining individuals with varying levels of depression (a meta-analytic review across the three experiments yielded a moderate effect size: Cohen’s $d = .70$). This was the case irrespective of whether the statements were judged on affect-neutral semantic characteristics (Experiments 1 and 2) or on their affectively laden level of self-reference (Experiment 3). The finding is in line with the dual competition framework arguing that across both perceptual and executive competition affective material will be preferentially processed thereby enhancing executive performance on tasks with task-relevant affective stimuli (Pessoa, 2009).

In contrast with the framework’s predictions, however, affective significance of the stimulus material did not appear to impact on WMC performance. That is, individuals with MDD for whom the DAS statements used in the depressogenic condition should have significantly greater affective significance did not differ in aWM performance compared with never-depressed groups. The finding was reliable (i.e., pooled trivial effect size of Cohen’s $d = 0.06$ and $BF_{10} = 6.38$ in support of the null hypothesis) across all three studies again irrespective of whether the operation task required a semantic or self-reference judgment. The finding contributes to a growing body of literature providing little support for the differential effects of affective material on WM accuracy in depression (Berman et al., 2011; Bertocci et al., 2012; Foland-Ross et al., 2013; Joormann & Gotlib, 2008; Joormann et al., 2011, 2010; Ladouceur et al., 2005; Levens & Gotlib, 2009, 2010; Tavitian et al., 2014; Yoon et al., 2014).

Findings from Experiment 3 are, however, in contrast with Hubbard and colleagues (Hubbard, Hutchison, Hambrick, et al., 2016; Hubbard, Hutchison, Turner, et al., 2016) who showed an impairing effect of depressogenic statements (relative to neutral) on WM in dysphoric undergraduates. While the present study offers support against an impairing effect, or attenuation of the enhancement effect in currently depressed individuals and those in remission from depression, it does not preclude such an impairing effect in subclinical dysphoria.

Interestingly, however, as noted in the Introduction, theories of WM in depression have been advanced on the assumption of such a WM deficit that is likely to be particularly pronounced for depressogenic material (Gotlib & Joormann, 2010; Koster, De Lissnyder, Derakshan, & De Raedt, 2011). One possibility is that, if such an effect exists, it may be more discernible when looking at performance on a complex span task that requires the inhibition of attention and responses toward truly task-irrelevant distractors as opposed to sentences that have to be processed as part of the task, as in the current design. Preliminary evidence to support this hypothesis stems from Schweizer and Dalgleish (2016) who showed that simultaneously performing a visuospatial search task (operation task) and a verbal storage task (remembering words) in the context of either neutral or negative task-irrelevant background images impaired WMC in both healthy and clinical (i.e., lifetime history of PTSD) samples reliably. In line with Schweizer and Dalgleish (2016), a recent meta-analysis investigating the effects of affective material on WM showed that negative distractors, in particular, had an impairing effect on WM performance (Schweizer et al., 2016).

Another possibility is that any effects of affective material on WM in depression may take the form of impoverished processing in the context of positive material, rather than differential processing of negative material (Rottenberg, 2007). However, to date there is little support for this view, with studies failing to show a differential effect of positive information on WM accuracy in those with depression compared with healthy controls (Bertocci et al., 2012; Joormann & Gotlib, 2008; Joormann et al., 2011; Kerestes et al., 2012; Ladouceur et al., 2005; Levens & Gotlib, 2009, 2010; Tavitian et al., 2014).

A further possibility is that, while WM accuracy may not be a sensitive measure of WMC deficits experienced in MDD, RT could be. Again as discussed at the outset, findings are equivocal with some (e.g., De Lissnyder et al., 2012) but not all (e.g., Foland-Ross et al., 2013) studies showing slowed WM RT for negative relative to neutral and/or positive material. One question that remains unanswered with this line of research into WM RT is how this relates to the executive functioning deficits experienced in everyday life. That is, a measure of WM accuracy is intuitively related to a quotidian task such as remembering a grocery list. In contrast, it is more difficult to hypothesize how differential RTs for negative versus neutral items on a WM task are related to remembering to buy dishwasher tablets and milk.

The present study showed an enhancing effect of affective information on WM in a pooled sample of nearly 300 participants, including over 100 participants endorsing clinically significant levels of depression, across three different experiments. Together the findings supported the hypothesis based on the dual competition framework (Pessoa, 2009) and insights from the long-term memory literature that task-relevant affective content improves WMC. However, in contrast with predictions from both the dual competition framework and theoretical models of cognitive vulnerabilities to depression, there was no support for reduced aWM in those currently suffering from MDD, in remission from recurrent MDD, or who self-reported symptoms of depression above the clinical cutoff for depression.
References


Received November 2, 2016
Revision received February 10, 2017
Accepted February 13, 2017