Autobiographical Episodic Memory in Major Depressive Disorder

Hedvig Söderlund  
Uppsala University

Morris Moscovitch  
Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada, and University of Toronto

Namita Kumar  
Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada

Zafiris J. Daskalakis  
Centre for Addiction and Mental Health, Clarke Division, Toronto, Ontario, Canada

Alastair Flint  
University Health Network, Toronto, Ontario, Canada, and University of Toronto

Nathan Herrmann  
Sunnybrook Hospital, Toronto, Ontario, Canada

Brian Levine  
Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada, and University of Toronto

Autobiographical memory in major depression has been characterized as overgeneralized, with patients recalling few episodic details, prioritizing general schematic events. However, whether this effect reflects impaired episodic or semantic memory, or domain-general cognitive processes, is unknown. We used the Autobiographical Interview (Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002) to derive episodic and semantic contributions to autobiographical memory in patients with severe major depression. We also assessed memory for public events and famous people. Depressed patients were impaired on episodic, but not semantic, autobiographical memory from 2 weeks to 10 years before testing. They were also impaired on memory for public events, possibly because they followed the news less than controls. Patients’ memory for famous names was not impaired, although this was strongly associated with non-episodic memories to a greater degree than in controls. The findings suggest a specific impairment of episodic autobiographical memory in depression that is not fully accounted for by domain-general processes involved in strategic retrieval.

Keywords: autobiographical memory, episodic memory, semantic memory, public events, depression

Supplemental materials: http://dx.doi.org/10.1037/a0035610.supp

Depression is the most common psychiatric disorder in the Western world, affecting between 4 and 17% of the population at some point in their lives, depending on country and location (World Health Organization, 2001). In addition to mood disturbances, reduced psychomotor speed, executive function, and attention is sometimes observed (Zakzanis, Leach, & Kaplan, 1998; Castaneda, Tuulio-Henriksson, Marttunen, Suvisaari, & Lönnqvist, 2008; McDermott & Ebmeier, 2009). One of the most common cognitive disturbances, however, is impaired memory (Zakzanis, Leach & Kaplan, 1998). Effects are small on semantic memory, short-term memory, and working memory, whereas large effects are found on episodic memory, the memory for personally experienced events (Zakzanis, Leach, & Kaplan, 1998; McDermott & Ebmeier, 2009). A classic and influential finding in the literature on depression and cognition is that depressed patients’ retrieval is overgeneral (Williams & Broadbent, 1986), such that when provided with an emotional cue-word (e.g., embarrassed) and asked to generate a specific memory in response to that cue-word, they...
come up with a general and repeated event (e.g., “It’s embarrassing to fall in front of other people”) rather than a specific event that occurred only once (e.g., “I vividly recall falling on my bike on an icy road last winter, feeling embarrassed in front of the people at the bus stop”). Overgeneral retrieval has also been related to the persistence of the depressive state (Brittlebank, Scott, Williams, & Ferrier, 1993; Scott, Williams, Brittlebank, & Ferrier, 1995), to a difficulty in generating specific simulations of the future (Williams et al., 1996), and persists beyond the acute depressive state (Brittlebank, Scott, Williams, & Ferrier, 1993; Mackinger, Pachinger, Leibetseder, & Fastacek, 2000; Spinhoen et al., 2006). It appears, however, that retrieval can be manipulated and made less categorical by means of cognitive intervention (Watkins, Teasdale, & Williams, 2000).

Given the importance attached to depressed patients having overgeneral memory, it is remarkable that nearly all of the research in this area has used a single test: the Autobiographical Memory Test (AMT; Williams & Broadbent, 1986; for exceptions, see Warren & Haslam, 2007 and Palombo, Williams, Abdi, & Levine, 2013), in which participants are asked to generate memories given an emotional cue-word. Although this test has been very useful, it has certain limitations. As the AMT uses a cue-word method of retrieval, it is affected by domain-general executive functions governing strategic generative retrieval, which may be affected in depression independently from memory per se (Dalgleish et al., 2007). The memories retrieved are also dependent on the nature of the cue-words, which are emotional and often related to the symptoms of depression, and thus may elicit memories that are schematic and pathological, giving rise to overgeneral retrieval (Dalgleish et al., 2007).

The distinction between episodic and semantic memory, influential in the behavioral (Tulving, 1985), neuropsychological (Kapur, 1999), and neuroimaging (Svoboda, McKinnon, & Levine, 2006) autobiographical memory (AM) literature, is not formally addressed by the AMT, although there is a loose correspondence between specific and general memories with episodic and semantic memory. Episodic AM entails recalling richly detailed perceptual information, thoughts, and feelings about temporally and spatially specific events, whereas semantic AM relates to retrieval of information that is not contextually bound, such as factual information about the world and oneself. Episodic AM is closely related to one’s identity and sense of self (Levine et al., 1998; Tulving, 1985), and is thought to entail the ability to see oneself as a continuous entity across time, a function termed autonoetic awareness (Tulving, 1985, 1989; Wheeler, Stuss, & Tulving, 1997).

In previous studies on AM and depression, no constraints were placed on the age of the memory that was elicited, thus providing little information on the temporal gradient of AM recall in depression. Some studies did report that more recent than remote memories were retrieved in older adults (Fromholt, Larsen, & Larsen, 1995) and adolescents (Kuyken & Howell, 2006), but they assessed memory age effects as an incidental factor in memory retrieval, rather than by probing lifetime periods in a controlled fashion. Such probing is standard in autobiographical memory research (e.g., Kopelman, Wilson, & Baddeley, 1989; Levine et al., 2002) as mechanisms supporting retrieval of recent and remote memories differ at both the psychological (Galton, 1879; Rubin & Schulkind, 1997) and neural (Svoboda, McKinnon, & Levine, 2006; Sheldon & Levine, 2013; Söderlund, Moscovitch, Kumar, Mandic, & Levine, 2012) levels.

In summary, it is unknown whether depressed patients’ autobiographical memory is overgeneral because of the effects of depression on domain-general resources, test characteristics, or a specific effect of depression on episodic or semantic mnemonic processes. Dominant theories of depression involve altered cognition concerning the self, including interpretation of past events (Beck, Rush, Shaw, & Emery, 1979). Improved specification of the mechanisms of overgeneral autobiographical memory, therefore, is important to understanding the genesis of depression, its relapse, and its treatment, which often involves how patients remember and interpret past events.

We sought to assess recent and remote memory comprehensively in depressed patients. To assess AM, we used the Autobiographical Interview (AI; Levine et al., 2002), an established tool for the assessment of episodic and semantic AM in various healthy and clinical populations (Addis, Moscovitch, & McAndrews, 2007; Berryhill, Pahuong, Picasso, Cabeza, & Olson, 2007; Davidson et al., 2008; Irish et al., 2011; Levine et al., 2002; McKinnon, Black, Miller, Moscovitch, & Levine, 2006; Milton et al., 2010; Rosenbaum et al., 2008; Steinworth, Levine, & Corkin, 2005; Willoughby, Desrocher, Levine, & Rovet, 2012). We controlled for domain-general strategic retrieval involved in event selection, as engaged by the cue-word method, by assisting participants in generating highly specific retrieval cues from autobiographical events from four different time periods. Whereas in the AMT, entire memories are classified as overgeneral or specific, in the AI the details of participants’ narratives collected after the event selection are scored individually by a reliable procedure that distinguishes between the episodic and semantic aspects of the memory. Reduced performance on the AI measure of episodic autobiographical memory would support the hypothesis that depressed patients have a specific deficit in retrieval of details from personal autobiographical episodes. Alternatively, no difference between patients and controls on episodic memory would suggest that their autobiographical deficit on the AMT can be accounted for by executive deficits (see Dalgleish et al., 2007). This conclusion would be further supported by an increase in non-episodic details, as we have found in other patients with executive deficits because of prefrontal dysfunction (e.g., McKinnon et al., 2008; Levine, 2004).

Overgeneral memory in depression has been reported for retrieval of personal memories, but it is unknown if this phenomenon extends to other types of remote memory, such as the recall of public events. Although there is evidence for overgeneralized recall of public events in depression (Warren & Haslam, 2007), a potential confound is the extent to which depression affects awareness and engagement with news events at the time of encoding. To assess public events recollection and how it may be related to news awareness, we scored participants’ recollection by adapting the AI method to public event recall, and had them rate their awareness of public events. We also assessed recognition memory for famous names, an alternative measure of semantic or public knowledge. Recall of public events and famous names can be affected by personal relevance, providing an alternative measure of autobiographical memory through its influence on a semantic task (Westmacott & Moscovitch, 2003; Westmacott, Black, Freedman, & Moscovitch, 2004; Petrican et al., 2010; Renoult, Davidson, Pal-
ombo, Moscovitch, & Levine, 2012). Therefore, we assessed both personal memory for these events and names and its relation to recognition. Such measures provide a useful adjunct to direct measures of autobiographical memory such as the AI as the retrieval cues are derived from shared knowledge rather than self-generated search and retrieval operations.

**Method**

**Participants**

We assessed 21 patients with major depressive disorder recruited from academic psychiatrics clinics in Toronto, Canada (see Table 1 for participant characteristics). All participants were severely depressed and, with two exceptions, referred for electroconvulsive therapy (ECT), making this a relatively homogenous sample of severely depressed patients. Diagnosis was made by the treating psychiatrist who prescribed the ECT, and confirmed with the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (SCID; First, Spitzer, Williams, & Gibbon, 1995). Participants with a history of severe unstable coexisting medical illness, neurologic disease, significant traumatic brain injury, schizophrenia, schizoaffective disorder, schizophreniform disorder, psychotic disorders not otherwise specified, organic affective disorder or dementia, substance abuse or dependence during the last 12 months, previous treatment with convulsive therapy (ECT), making this a relatively homogenous sample of severely depressed participants. A structured interview (Specific Probe) was administered to assess the effects of retrieval support on recall for depressed participants. A structured interview (Specific Probe) was employed for two control participants and no depressed participants. A structured interview (Specific Probe) was administered to assess the effects of retrieval support on recall via questions concerning event, time, place, perceptual, thought, and emotional details. The AI was audio-recorded for later transcription and scoring.

**Materials**

**Neuropsychological test battery.** A 1-hr battery of standardized neuropsychological tests was included to assess mnemonic and speeded processes potentially affected by depression. Vocabulary knowledge (Zachary, 1986) was used to estimate Verbal IQ. The Hopkins Verbal Learning Test – Revised (Benedict, Schretlen, Groninger, & Brandt, 1998) and the Brief Visual Memory Test – Revised (Benedict, 1997) assessed anterograde verbal and visual memory, respectively. Speeded information processing was assessed with the Trail Making Test, Parts A and B (Army Individual Test Battery), and the Symbol Digit Modalities Test (Smith, 1982). Phonemic and semantic word list generation assessment verbal fluency.

**The Autobiographical Interview (AI; Levine et al., 2002).** The AI, which has been described fully elsewhere (Levine et al., 2002), was used to derive measures of episodic and semantic AM. In brief, before testing, participants selected events specific in time and place in which they were personally involved and that they felt able to recall in detail. At testing, they were asked to describe each event (Free Recall), followed by nonspecific cueing to reinforce the test instructions (General Probe: “Is that everything you can say about it? I want to know all the details that come to mind.”). There is a 5-min cutoff at both the Free Recall and General Probe levels that is employed for those who produce extremely lengthy protocols. In practice, this cutoff is rarely employed. In the current sample, it was employed for two control participants and no depressed participants. A structured interview (Specific Probe) was administered to assess the effects of retrieval support on recall via questions concerning event, time, place, perceptual, thought, and emotional details. The AI was audio-recorded for later transcription and scoring.

Six events were collected for each of four time periods and randomized to three testing sessions, pre-ECT, post-ECT, and 6-month follow-up, each with two events per time period. The present paper is concerned only with the pre-ECT assessment, thus two events per time period. The four time periods were derived from the principle of exponential decay of memory over time (Rubin & Schulkind, 1997) to tap memories that show the most potential of differing in their memorability, including the past week (exclusive of the day of testing), the past month (range: 2–6 weeks), the past year (range: 6–18 months), and the past 10 years (range: 5–15 years). At the time of booking the appointment (i.e., within 1–2 days of testing for most patients, and within a week of testing for most controls), the research assistant described the criteria for event retrieval cues and encouraged the participant to generate them. To facilitate retrieval of AMs, participants were encouraged to consult their calendars and that of significant others. They were also provided with a list of typical life events (Levine et al., 2002), including events likely to have occurred in the previous week. Use of traumatic events was discouraged. Event

**Table 1**

<table>
<thead>
<tr>
<th>Table 1 Demographic and Clinical Characteristics of Patients and Healthy Controls (Means and Standard Deviations)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depressed patients</strong></td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
</tr>
<tr>
<td><strong>Sex (men/women)</strong></td>
</tr>
<tr>
<td><strong>Education (yrs)</strong></td>
</tr>
<tr>
<td><strong>BDI score</strong></td>
</tr>
<tr>
<td><strong>Time since diagnosis (yrs)</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

**Note.** BDI = Beck Depression Inventory, 0–9: not depressed, 10–18: mild-moderate depression, 19–29: moderate-severe depression, and 30–63: severe depression.

<sup>a</sup> Values are the median and range. **<sup>**</sup> Patients differed from controls at p < .0001.
cued cue selection was finalized at the beginning of the first test session. This process involved further event cue generation, depending on how many cues were generated by subjects before testing. Each event was rated for frequency of rehearsal (from 0: never; to 11: constantly), visualization (from 1: vague memory/no recollection to 6: extremely clear as if it is happening now), personal importance then and now (from 1: no importance at all; to 6: of great importance), valence (from 1: extremely sad; to 6: extremely happy), and emotional change after the event (from 1: no change in how I felt; to 6: underwent tremendous emotional change). The pretest generation of events was limited to generation of only the event titles and dates for the purposes of generating retrieval cues. Participants were instructed not to discuss the events with others before testing.

AI scoring consists of segmenting the transcribed protocol into informational bits or details. A detail is defined as a unique occurrence, observation, or thought, generally expressed as a grammatical clause. Details that relate directly to the event, that are specific to time and place, and that convey a sense of episodic reexperiencing are classified as “internal” or episodic and assigned to one of five categories (event, place, time, perceptual, and emotion/thought). If the recollected details do not fit into any of these categories, they are considered “external.” These consist of semantic facts (factual information or extended events that did not require recollection of a specific time and place), autobiographical events occurring at a different time than the main event, repetitions, or other metacognitive statements or editorializing. Details were tallied for each category and summed to form internal and external composites, which are the main variables of interest (for scoring example, see Supplemental Material). Scores were analyzed cumulatively across levels of recall. As the pattern of findings was not affected by probing level, they were collapsed into one total score.

To avoid bias in scoring, participants’ memories are stripped of information that would identify group membership (for the purposes of scoring, censored details were replaced with equivalent fictional details that did not reveal group membership), placed in a common pool (including memories from studies of other patient and healthy adult samples), and scored at random by eight scorers blind to group membership who scored an average of 41 memories each (SD = 17). Scorers were trained by the main developer of the instrument (B.L.) and had previously attained high interrater reliability (intraclass correlations >.90) using a standard set of previously scored memories (for methods, see Levine et al., 2002). Scoring fidelity was maintained by monitoring of randomly selected scored memories and ongoing scoring meetings.

Public Events Interview. For the Public Events Interview, event selection paralleled that described above for autobiographical events, except that participants selected cues from a list provided by the experimenters at the time of testing, and that they were encouraged to relate historical and current facts rather than autobiographical details, such as how they heard about the event. They were probed in a similar fashion to the AI, starting with Free Recall, followed by General Probe where they were asked to recall anything else they may remember. During Specific Probe they were asked about specific details of the event, such as when and where it took place, who was involved, and what it looked and sounded like (e.g., if they saw it on TV). Because of the relative paucity of recent public events, the recent time period was extended to the past 6 months. The past year and the past 10 years time periods were the same as described above for the AI, resulting in events from three time periods. A list of public events garnered from news, entertainment, and sports sources was provided for participants during event selection, and this list was continually updated with passing time. Scoring also paralleled the AI, but as there were floor-effects in the external details and they did not add any information to the internal details, all details were collapsed into a single category (for scoring example, see Supplemental Material).

As patients may differ from controls in how much they follow the news, participants were asked to rate to what extent they followed the news on the radio, on TV, in newspapers, and on the Internet, going from 0 (never) to 8 (several times a day) for each of the three time periods. Scores were averaged across news media for one news awareness score per time period.

Famous Names Test. To assess potential effects of episodic impairment on semantic memory, we developed a Famous Names Test (e.g., Westmacott et al., 2004; Renoult, Davidson, Palombo, Moscovitch, & Levine, 2012). This consisted of 60 names recognized as famous political, athletic, or entertainment figures from three time periods: the 1980s, the 1990s, and the 2000s, based on their initial entry into the public eye, and 20 lures. Participants indicated whether or not each name represented someone famous, then, if the name was classified as famous, the category of fame in multiple choice format (e.g., politician, actress, musician, athlete). Autobiographical memory contribution to recognition was assessed with a modified Remember/Know procedure (Tulving, 1985; Gardiner, 1988), by asking whether they had a personal memory associated with the name. Recognized names associated with specific AMs (as indicated by recollection of a specific event, including perceptual or mental contextual details) were classified as “remembered.” Those associated with knowledge about the famous person, but no AMs, were classified as “known,” providing indices of autobiographically influenced semantic memory across time periods. Know responses were statistically adjusted (adj. K = K/(1-R)) to estimate familiarity independently from recollection (Yonelinas & Jacoby, 1995).

Analyses

For the AI, Public Events Interview, and Famous Names test, a mixed (split-plot) design ANOVA was used with Group (depressed, controls) as a between-subjects factor and Time as a within-subjects factor. Time was assessed with a polynomial contrast with a hypothesized linear effect (Rubin & Schulkind, 1997). For the AI, detail type (i.e., internal, external) was included as a within-subjects effect. These AI analyses were followed up with mixed-design ANOVAs to probe potential group differences in specific categories within the internal and external composites. Group differences in participants’ ratings of their own autobiographical memories were also assessed with mixed design ANOVAs (separately for each rating). For the Famous Names Test, correlational analyses between Adjusted Recognition (hits minus false alarms) and Remember and Adjusted Know responses were conducted to assess the relationship of episodic memory to famous name recognition, with between-groups differences in these correlations assessed as described by Cohen, Cohen, West, and Aiken.
(2003). Group differences on neuropsychological tests were assessed with independent samples t tests.

Results

Neuropsychological Test Performance

Patients were slower than controls, taking significantly longer to complete the Trail Making Test, both A and B (see Table 2). Patients did not perform disproportionately worse on version B, however, suggesting that their slowing was not enhanced by the executive switching element of this test. Patients performed no differently from controls on immediate recall of verbal and visuospatial material (HVLT-R and BVMT-R), but they were impaired on BVMT-R delayed recall. There were no significant group differences for verbal fluency and verbal intelligence (Shipley Institute of Living Scale).

Autobiographical Memory

Overall, depressed patients recalled fewer details than controls (F[1, 37] = 9.24; p = .004, \( \eta_p^2 = 0.20 \)), but there was a Group by Detail interaction (F[1, 37] = 6.08, p = .01, \( \eta_p^2 = 0.15 \)) as patients recalled disproportionately fewer internal details (F[1, 37] = 12.7; p = .001, \( \eta_p^2 = 0.26 \)), but not fewer external details than controls (F[1, 37] = 3.36, p > .05 (p = .08, \( \eta_p^2 = 0.08 \); see Fig 1 for the group difference in internal but not external details). The polynomial contrast involving time revealed a significant interaction between the linear component of Time and Detail (F[1, 37] = 10.67, p = .002, \( \eta_p^2 = 0.22 \)) corresponding to the expected recency effect for internal, but not external details (Levine et al., 2002) because of greater vividness of recently experienced events (Rubin & Schulkind, 1997). The interaction between the linear component of Time and Group was not significant, F[1, 37] = 2.83, p = .10, \( \eta_p^2 = 0.07 \), however. The Time X Detail X Group interaction was also not significant.

We next assessed the effects of depression on detail categories within the AI. As the Time X Detail X Group interaction was not significant for composite AI internal and external detail scores, and because there were few details for certain combinations of detail categories and time periods, these analyses were collapsed across time periods. A significant interaction between Group and Detail Category (F[1.68, 61.96] = 8.62; p = .001, \( \eta_p^2 = 0.19 \)) for internal details was accounted for by significantly more event (F[1, 37] = 11.21, p = .002, \( \eta_p^2 = 0.23 \)), time (F[1, 37] = 4.4, p = .04, \( \eta_p^2 = 0.11 \)), perceptual (F[1, 37] = 8.57, p = .006, \( \eta_p^2 = 0.19 \)), and thought details (F[1, 37] = 7.34, p = .01, \( \eta_p^2 = 0.17 \)) in control participants relative to depressed patients; place was not significant.

As expected, there was no significant Group X Detail Category interaction for external details, (F[1,78, 65.72] = 1.19; ns). We nonetheless assessed group effects for these detail categories because the specific detail categories in the external composite are heterogeneous (i.e., including semantic along with other non-episodic categories; see Method) and, therefore, could obscure group differences when summed. The only external category to differ significantly between groups was repetitions (F[1, 37] = 8.05, p = .007, \( \eta_p^2 = 0.21 \)), which were higher in control participants relative to depressed patients. This effect can be interpreted as a byproduct of the longer protocols in controls, where repetitions of internal and external details occur naturally in discourse (means and SDs of total amount of details for controls and patients were for 2 weeks: 85.3 ± 29.9, 55.8 ± 24.6; 1 month: 82.9 ± 26.8, 54.9 ± 33.6; 1 year: 80.8 ± 26.1, 60.0 ± 26.4; 10 years: 73.6 ± 28.3, 55.2 ± 22.6, respectively).

Analysis of the ratings of memories selected and retrieved by patients and controls indicated that these were comparable in terms of such episodic features as emotionality and vividness, although there was an overall effect of Group in terms of how negative ratings were (F[1, 36] = 7.01, p = .012, \( \eta_p^2 = 0.16 \)) with patients overall having numerically higher ratings. Nevertheless, only the 2-week-old memories were statistically more negative in patients than in controls (F[1, 36] = 8.2, p = .007, \( \eta_p^2 = 0.19 \); see Supplementary Table 1).

Public Events

Patients recalled fewer public event details than did controls (F[1, 28] = 9.27, p = .005, \( \eta_p^2 = 0.25 \)) at all time periods (ps ≤ .01; see Table 3). Patients reported following the news to a lesser extent than controls, (F[1, 26] = 9.88, p = .004, \( \eta_p^2 = 0.28 \)), especially for events occurring in the last 6 months (F[1, 26] = 13.87, p = .001, \( \eta_p^2 = 0.35 \)), and to some degree for those occurring during the previous year (F[1, 26] = 7.16, p = .01, \( \eta_p^2 = 0.22 \)) and the last 10 years (F[1, 26] = 5.71, p < .05, \( \eta_p^2 = 0.18 \)). Correlating the number of details recalled within each time period with news awareness within each time period, for the two groups separately, showed that the extent to which participants followed the news was not associated with control participants’ recollection (ps > .10), but was associated with patients’ recollection of events from 6 months back (r = 0.69; p < .05) and 1 year back (r = 0.83; p < .005). Thus patients’ reduction in recall of details concerning public events appeared to be associated with reduced attention to these events, at least for the two recent time periods.

Famous Names Test

There were no significant group differences on the recognition, identification of profession, Remember/Know responses, or adjusted know responses (see Supplementary Table 2). To assess the contribution of autobiographical memory to famous name recognition across groups, we calculated the correlations between name
recognition and Remember/Know responses within each group, collapsing across time period (see Table 4). Patients’ recognition performance correlated with “Know” responses to a significantly higher degree than in controls, suggesting that patients’ performance was supported by familiarity to a greater degree than controls. By contrast, controls’ performance was significantly correlated with adjusted “Remember” but not “Know” responses.

Discussion

We assessed AM memory in severely depressed patients across remote and very recent time periods, with a view to differentiating between episodic and semantic memory for events with specific retrieval cues. We also examined memory for public events and for famous people. We found that depression was associated with a reduction in episodic, but not semantic, autobiographical details across all time periods. Depression was also associated with reduction in public event recall, although patients also reported following the news to a lesser extent than controls for two of the three periods. There were no group differences for recognition and recall of information regarding famous people, although patients’ famous name recognition was more strongly associated with familiarity than controls1.

Our findings are consistent with the classic observation of overgeneral memory as assessed by the AMT (Williams & Broadbent, 1986) in the sense that there is a depression-related reduction in specificity as reflected by reduced internal details relative to external details. Although direct comparison of the AMT and AI would entail administration of both measures in the same study, we note that the effect sizes for internal details were comparable to those observed for overgeneral memory (for the AI, Cohen’s d ranged from 0.69 to 1.09 for 10-year-old and 2-week-old memories, respectively; for the AMT, d ranged from 0.48 to 0.91 for negative and positive cues, respectively, and was 1.12 overall when 11 studies on MDD were considered; Williams et al., 2007). Our findings, however, extend those from the AMT in several ways. Overgeneral memory on the AMT could result from reduced episodic recall, increased semantic recall, or both, in combination with reduced domain-general resources required for strategic event generation (Dalgleish et al., 2007; Williams et al., 2007). We found that depression was not associated with a global reduction in AM. Rather, the effect of depression was specific to internal (episodic) AM details; there was no group difference for external (non-episodic) details. In particular, patients produced fewer details relating events or happenings, perceptual information, and thoughts and feelings relative to controls, whereas there were no differences for semantic details. This effect could not be accounted for by dysfunctional processes required for event selection because of the use of nonspecific retrieval cues, as in the AMT. The pattern observed here is different from the pattern observed in association with aging, mild cognitive impairment, and frontotemporal dementia (Levine et al., 2002; McKinnon et al., 2008; Murphy, Troyer, 2010).

Table 3

<table>
<thead>
<tr>
<th>Recall</th>
<th>News awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
</tr>
<tr>
<td>6 months</td>
<td>26.0 (10.9)</td>
</tr>
<tr>
<td>1 year</td>
<td>25.5 (10.0)</td>
</tr>
</tbody>
</table>

Note. Awareness scores ranges: patients: 0.6–3.9 (6 mo), 0.6–4.1 (1 yr), 0.4–4.6 (10 yrs); controls: 2.7–4.9 (6 mo), 2.6–4.7 (1 yr), 1.9–5.0 (10 yrs).

Table 4

| Correlations Between Adjusted Recognition (Hits Minus False Alarms) and Remember/Know Responses in Patients and Controls in the Famous Names Test Across All Time Periods Collapsed |
|---------------------------------|----------------|
| Patients (n = 14)               | Controls (n = 18) |
| Remembera | .50† | .71** |
| Knowb | .84*** | .32 |

a No significant difference between the groups’ correlation coefficients. b Patients’ correlation was significantly higher than that of controls at p < .05.

† p = .07. ‡ p = .001. ** p < .0001.
Depression was associated with reduced episodic autobiographical specificity across all time periods assessed (from 2 weeks to 10 years), suggesting a general retrieval deficit that is not specific to life period. Patients rated the most recent (2 weeks old) event as more negative and emotionally significant than did controls, likely owing to the selection of events surrounding their referral for ECT. These effects, however, do not change the overall interpretation of reduced episodic specificity across time periods. There were no group differences in memory characteristics for other periods, including the relatively recent 1-month period.

Whereas depressed patients show a negatively biased memory selection (Blaney, 1986; Burt, Zembar, & Niederehe, 1995), this was not the case for the 1-month, 1-year, and 10-year periods, possibly because of the constraints that the patients select a spatiotemporally specific event. The AMT emphasizes the generative aspects of AM retrieval by using common, yet evocative, words as AM retrieval cues that are presented for the first time at test. In the present study, participants were encouraged to use external aids to derive specific retrieval cues before testing, and it is these cues that were subsequently used to generate detailed AM protocols. Assistance in event retrieval was provided when necessary, reducing variance attributable to generative capacities. In this sense, the AM is less reliant on domain-general generative processes for memory specification than the AMT (see Dalgleish et al., 2007) and, therefore, is more specifically considered as a test of direct, rather than generative, AM retrieval (Conway, 1992; Moscovitch, 1992). Indeed, if our methods were more sensitive to domain-general processes, similar effects would have been observed for both internal and external details. Consistent with our findings, Watson, Berntsen, Kuyken, and Watkins (2013) found that specificity of voluntarily, but not involuntarily, retrieved autobiographical memories was affected in those with unremitting depression to a greater degree than in those with depression in remission or no history of depression.

The process of retrieval from autobiographical memory is multifaceted, involving cue specification, navigation through different levels of self-knowledge, reciprocal interaction with personal semantic memory, and manipulation of and monitoring of retrieval output within working memory (Conway & Pleydell-Pearce, 2000). Dalgleish et al. (2007) emphasize the effects of depression on the executive and generative components of autobiographical retrieval. Although the present findings do not contradict those of Dalgleish et al. (2007), our results suggest that depression is associated with deficient episodic retrieval even when domain-general executive demands are reduced by providing more specific retrieval cues. Thus depressed patients have inefficient access to specific episodic autobiographical details.

Depression is associated with structural and functional changes in both prefrontal (Beyer & Krishnan, 2002; Drevets, Savitz, & Trimble, 2008; Mayberg, 1994; Videbech, 2000) and medial temporal areas (Campbell, Marriott, Nahmias, & MacQueen, 2004; Drevets, 2003; Drevets, Savitz, & Trimble, 2008). Although discussion of the neuroanatomical mechanisms of the AM retrieval deficits observed here is beyond the scope of this paper, we note that the pattern of deficits resembles that of patients with medial temporal damage, and that reduced internal details on the AM are correlated with medial temporal damage in patients with epilepsy (Addis et al., 2007) and in patients with frontotemporal dementia (McKinnon et al., 2008) and focal frontal lesions (Levine, 2004). By comparison, frontal volume loss was not correlated with internal details (McKinnon et al., 2008). Considering our results together with Dalgleish et al.’s (2007), autobiographical retrieval deficits in depression may indeed reflect executive difficulties because of frontal lobe dysfunction, but these deficits are superimposed on impoverished representation of episodic autobiographical details consistent with medial temporal lobe dysfunction. Although our findings can be accounted for by inefficient representational and retrieval mechanisms, encoding was not assessed in this study and, therefore, cannot be ruled out as contributing to the pattern of results. Similarly, the treatment-refractory depression inherent in our ECT-referred sample may have conferred a greater burden of depression-related changes on our patients’ memory retrieval systems than would be the case in a sample with less severe depression (as in Dalgleish et al.’s 2007 study).

With the exception of delayed recall on the BVMT, we did not observe memory deficits on neuropsychological tests. In contrast to these measures, the AM assesses the multimodal richness of episodic memory associated with the self and mental time travel, which are at the core of episodic memory (Tulving, 2002), and fundamental to personal history, decision making, problem solving, and future thought (Conway & Pleydell-Pearce, 2000; Dudai & Carruthers, 2005; Klein, Loftus, & Kihlstrom, 2002; Klein, Robertson, & Delton, 2010; Rosenbaum, Gilboa, Levine, Winocur, & Moscovitch, 2009; Schacter & Addis, 2007; Sheldon, McAndrews, & Moscovitch, 2011; Wilson & Ross, 2003; but see Hasabis & Maguire, 2009). Episodic AM impairment, therefore, is likely to affect quality of life and well-being.

Depression was associated with impairment of episodic autobiographical memory and not semantic autobiographical memory. Similarly, patients were not impaired on a non-autobiographical test of semantic memory assessing knowledge of famous people and their professions. Although patients had fewer Remember responses than controls, indicating a paucity of personal memory associated with the famous person in question, this difference did not reach significance. Nevertheless, patients’ performance was more strongly associated with Know responses than in controls, suggesting that their famous name recognition may have been supported more by familiarity, whereas controls’ recognition was strongly associated with episodic recollection. It has been suggested that performance on tests of semantic memory can be enhanced by autobiographical significance, and that a reduction in episodic memory can correspondingly impair performance on tests of semantic memory (e.g., Westmacott et al., 2004; Renoult et al., 2012). Our finding suggests that although semantic memory may not be strongly impaired in depression, it may still be less robust than in controls and have fewer autobiographical associations, given that episodic memory is impaired.

Patients’ memory impairment was not restricted to memory for their own lives; they also recalled fewer details from public events (see Warren & Haslam, 2007, for a similar finding). Although this effect was apparent at all time periods, it was associated with...
reduced attention to news stories relative to controls for the 6-month and 1-year periods. Thus, an encoding deficit could not be ruled out as contributing to patients’ reduction in public event recall for recently occurring events.

The patients assessed in our study suffered from severe depression, and were prescribed ECT as they had not responded to psychopharmacological treatment. It is possible that the episodic autobiographical memory impairment observed here would not be as extensive in a less depressed group that still fulfills the criteria for clinical depression. Nevertheless, although our patients were indeed severely depressed, they had normal immediate memory, word fluency, and intelligence as measured by Shipley vocabulary scale, suggesting they were not impaired in all cognitive domains, and not even in all memory domains, considering their relatively preserved memory for external details and for famous people.

Using patients prescribed ECT for depression increased the homogeneity of our severely depressed sample relative to community samples defined by screening questionnaires identifying depressed mood only. In the present study, the relationship of the findings to depression is supported by the inclusion of only those patients with severe depression, excluding other conditions that are frequently comorbid with depression. Given our findings in severe depression, there is now a need to extend these to more typical community samples. The costs associated with this sample selection included a relatively small sample size owing to the challenges of recruiting patients both willing and available to undergo a rigorous test battery during the short time window between referral for and treatment with ECT. Although the smaller sample size affected power, comparisons involving the AI were powered similarly to other studies, with similar effect sizes (see above). For the public events and famous names tests, sample sizes were lower because of lack of knowledge of public information and fatigue related to the lengthy test battery, meaning the samples for these comparisons were both smaller and less representative. Considering the lower public knowledge of those who did not complete these tests, it is likely that there would have been larger or more significant effects on these measures had these patients been testable. In general, replication in samples with a wider range of depression symptoms is required to establish the generalizability of our findings.

In summary, using a tool that reduces domain-general event-selection demands, we have extended previous findings on overgeneral autobiographical memory in depression by showing for the first time a dissociation between episodic and semantic AM in severe, clinical depression. Depression is associated with a specific impairment of episodic autobiographical memory that extends back in time for at least 10 years, whereas semantic autobiographical memory remains relatively preserved. This loss of episodic specificity in depressed people extends, to different degrees, to memory for public events and for famous people, also encompassing at least the last 10 years. By assessing memory across time periods ranging from very recent to remote, we have demonstrated also, for the first time, a temporally extensive loss of specificity in retrieval details concerning personal and public events in severe depression.

It is unknown whether our findings would generalize to remote memory impairment in patients with remitted or milder depression as opposed to the severely depressed patients used here. Nonetheless, the negative self-appraisal, rumination, self-blame, and hopelessness that are core to depression involve dysfunctional cognitive processing of information retrieved about past events, and such information processing is targeted by effective psychotherapeutic interventions for depression and relapse prevention, such as cognitive–behavioral therapy (Beck et al., 1979) and mindfulness-based cognitive therapy (Teasdale et al., 2002), and ECT when other treatments fail. The fact that severely depressed patients are unable to retrieve episodic details of events, even with highly specific retrieval cues, may contribute to the vicious cycle of dysfunctional cognition that perpetuates depression (Teasdale, 1983).

References


Sheldon, S., & Levine, B. (2013). Same as it ever was: Vividness modulates the similarities and differences between the neural networks that