

# Toxic Effects of Depression on Brain Function: Impairment of Delayed Recall and the Cumulative Length of Depressive Disorder in a Large Sample of Depressed Outpatients

Philip Gorwood, M.D., Ph.D.

Emmanuelle Corruble, M.D., Ph.D.

Bruno Falissard, M.D., Ph.D.

Guy M. Goodwin, D.Phil., F.Med.Sci.

**Objective:** An important current hypothesis suggests that the relationship between severe depression and the hippocampus is essentially toxic. The purpose of this study was to assess the generalizability of the impact of depression on hippocampal function.

**Method:** Participants were 8,229 outpatients who 1) fulfilled DSM-IV criteria for major depressive disorder based on clinical assessment and 2) were tested for delayed recall, a memory function that is particularly related to hippocampal integrity in humans, during two visits several weeks apart.

**Results:** As expected, at presentation with depression, the subjects' current illness severity was the major determinant of performance, as opposed to the intensity of their previous depressive history (the number and length of past episodes). However, following clinical response at

the second visit, the length of previous depressive history became more significant than current symptoms. The following factors had significant, independent impact: age, education level, and profession.

**Conclusions:** Previous studies of small samples assessed for memory function, more or less specific to the hippocampus, have shown great variability in age, gender, education level, and the length and intensity of depressive episode. Hence, a very large sample was required to disentangle the central effect of previous depressive history. As demonstrated in a general practice sample in this study, the hypothesis that the length of past depression impairs memory performance is supported, suggesting that there is a toxic link between the burden of depression and cognition. This finding has important implications for public health.

(*Am J Psychiatry* 2008; 165:731–739)

It is a widely accepted current hypothesis that the relationship between severe depression and the hippocampus is essentially toxic. Hence, the more intense the history of depression, the smaller the hippocampus. However, there is currently little evidence to consider this to be a general effect that is present across the spectrum of depression diagnoses, which are regarded as highly heterogeneous. This is particularly true in primary care, in which there is often skepticism that depression represents anything more than normal human distress. Accordingly, we have sought to provide evidence of the generalizability of depression's impact on hippocampal function. A large sample size would be required to account for many potentially major confounds, with an assessment that could be repeated before and after treatment in order to control for the impact of the acute depressive state and provide a simple assay of hippocampal function.

The majority of studies linking the hippocampus and major depressive illness have employed quantitative structural magnetic resonance imaging (MRI), functional MRI (fMRI), or positron emission tomography (PET),

methods that are not easy to access in ordinary clinical practice. However, hippocampal size has usually correlated inversely with illness duration. This was demonstrated in a chronically depressed sample (1) and in patients with variable illness durations seen in secondary care (2). Hippocampal size may also be related to other measures of illness intensity, such as the number of past hospitalizations (3) and recurrence of the disorder (4). Moreover, hippocampal abnormalities have been observed in the early years after illness onset (5). Meta-analyses have confirmed hippocampal volume reduction, and the total number of depressive episodes may be particularly correlated with right hippocampal volume (6, 7).

The samples in imaging studies have often been rather small and potentially unrepresentative of the majority of outpatients with depression. We have been interested in measures of cognitive function that assay the function of the hippocampus but are much easier to conduct on a very large scale in everyday medical care. The choice of measures can be informed by imaging evidence of fundamental hippocampal involvement. Thus, activation of the

This article is featured in this month's AJP [Audio](#).

hippocampus has been observed with tasks such as word-stem completion (8–10), success of word retrieval (11–14), emotional valence (15), and encoding (16–22). Even more consistent hippocampal activations have been shown in healthy subjects with a paragraph encoding task (23) that involves the encoding of complex and integrated information, which is hypothesized to be a core role for the hippocampus (24) and classically impaired in patients with known hippocampal lesions. Thus, there is considerable evidence to suggest that delayed paragraph recall is particularly related to hippocampal function in humans (e.g., 9, 25, 26).

While memory impairment may reasonably be taken to assay hippocampal function, its relationship to any current depressive episode has at least two (nonexclusive) aspects. 1) If the characteristics of the present depressive episode (such as symptom severity) predominate, memory function is likely to be simply a “state marker,” reflecting the direct cognitive impact of current mood. 2) Alternatively, if the dominant factor is a lifetime cumulative impact of mood disorder (such as the total length or number of past episodes), memory impairment will act as a “trait marker” of enduring toxic effects of depression on brain function. In depressed individuals, both would potentially contribute. In recovered individuals, deficits would most likely reflect the enduring brain changes seen with brain imaging.

The present study is of a large sample of outpatients who fulfilled DSM-IV criteria for major depressive disorder and were assessed for memory function during two visits several weeks apart. We anticipated that some clinical factors, such as mood state markers (the length and severity of current depressive episode), choice of antidepressant, and variation of testing methodology in a nonspecialist setting (as well as other factors such as older age, lower education level, and unemployment), would potentially confound the results at both visits. We hypothesized that the length and number of past episodes would also be involved in correct delayed recall but were much more likely to be discernible at a second clinic visit after symptom resolution, when the severity of current depression would be reduced.

## Method

### Participants

A list of 4,849 medical doctors were contacted via mail in France and asked to participate in a short-term follow-up protocol of depressed patients. Of these, 3,375 physicians (69.6%) agreed to participate. At least two contacts (usually via telephone) were made to each participating investigator: at the beginning of the protocol (in order to ensure that the protocol was clear and to explain how to assess memory recall) and at the close of trial entry (in order to verify the data received). By the end of the study, 1,844 clinicians (38.0%) had included at least one patient (a maximum number of five patients was requested to avoid center effects), who was followed up with a delay of at least 6 weeks between the two visits. The participating clinicians were experienced (mean age=49.9

years [SD=11.8]), with 7.8% practicing in a hospital, 46.6% in private practice, and 45.6% in a group practice.

Clinicians were asked to include consecutive patients for whom a new (or different) antidepressant had to be prescribed for a major depressive episode. In addition, the patients were required to be older than 18 years, speak fluent French, possess a social security number, and give informed consent. Patients had to be included by their clinicians during a 3-month interval. After complete description of the study was given to the subjects, written informed consent was obtained. A total of 9,515 patients were included in the study.

Exclusion criteria were a diagnosis of bipolar disorder and the use of a mood stabilizer during treatment. All antidepressants (in accordance with the French Food and Drug Administration) were accepted in order to reflect usual clinical practice. Any change of antidepressant, an increase in the dosage, or the addition of a benzodiazepine was recorded at the second visit.

### Instruments

The Hospital Anxiety and Depression Scale was chosen as a self-report instrument to measure symptom severity because of its rapidity and simplicity of rating. The scale was completed by all patients at the first and second visits. A score above 8 for the depression domain was required as an inclusion criterion. Hospital Anxiety and Depression Scale anxiety scores (score above 8 also required for inclusion) were analyzed separately because anxiety may also have an impact on hippocampal volumes, according to animal models (27, 28) as well as studies in humans (29).

The criteria for a major depressive episode were examined by the clinician, and the duration of each symptom was recorded during the two face-to-face visits. The presence of five or more symptoms (i.e., a DSM-IV diagnosis of major depressive disorder) was required for inclusion. The initial assessment also included the number of past depressive episodes, either treated or not treated with an antidepressant, and the cumulative length of past mood disorder.

The delayed paragraph recall index from the Wechsler Memory Scale—Revised (30) was employed as a valid (8), sensitive (31, 32) measure of verbal declarative memory and a surrogate marker of hippocampal function. This subtest was administered according to the standardized protocol outlined in the Wechsler Memory Scale—Revised Manual (1987). A different story at each visit was read aloud to the subject by the clinician. After hearing the story, the subject was asked to repeat it using as many of the same words as he or she could recall from memory. One point was given for each verbatim or acceptable alternative response phrase. After at least a 10-minute delay (when the subject was distracted to complete the Hospital Anxiety and Depression Scale and provide other information), the subject was asked to recall the story again. The clinician tabulated the scores for each story as well as the total score for the immediate and delayed trials.

The existence of a practice effect was indirectly assessed in the subgroup of patients with identical levels of depression at both visits (the same Hospital Anxiety and Depression Scale global scores and number of DSM-IV symptoms of major depressive episodes). In this subgroup of 33 patients, the difference between the number of correct delayed recall responses for the two visits (mean=0.029 [SD=3.6967]) was not significantly different from 0 ( $t=0.046$ ,  $p=0.963$ ).

The level of memory impairment can be stratified according to the original description by Russel (33), who suggested that delayed recall can be classified, out of the initial 25 elements, as follows: 24 to 25=“better than normal,” 20 to 23=“normal,” 15 to 19=“mild,” 9 to 14=“mild to moderate,” 4 to 8=“moderate to severe,” and 0 to 3=“severe.” An age-corrected standard score can be computed on the basis of the Wechsler delayed paragraph recall index, but we preferred to treat age as one of several potential con-

**TABLE 1. Clinical and Cognitive Characteristics of Patients With Major Depressive Disorder Before and After Antidepressant Treatment**

Characteristic	Assessment						
	Visit 1		Visit 2		Analysis		
	Mean	SD	Mean	SD	t	df	p
Number of DSM-IV criteria for depression	6.62	1.14	3.25	1.77	142.86	16, 456	<0.0001
Hospital Anxiety and Depression Scale score for depression	14.62	3.24	9.94	4.22	79.83	16, 456	<0.0001
Hospital Anxiety and Depression Scale score for anxiety	14.12	3.01	8.92	3.40	103.86	16, 456	<0.0001
Number of correct delayed recall responses	9.98	4.57	12.04	4.85	27.63	16, 456	<0.0001

**TABLE 2. Level of Memory Impairment for Delayed Recall in Patients With Major Depressive Disorder Before Visit 1 and 6 Weeks After Antidepressant Treatment (visit 2)<sup>a</sup>**

Classification	Assessment					
	Visit 1 (N=8,221)		Visit 2 (N=8,221)		Visit 2 Responders (N=1,895) <sup>b</sup>	
	N	%	N	%	N	%
Better than normal (24–25 elements)	19	0.24	74	0.93	26	1.37
Normal (20–23 elements)	180	2.25	520	6.52	162	8.55
Mild (15–19 elements)	1,239	15.47	1,961	24.61	554	29.23
Mild to moderate (9–14 elements)	3,267	40.80	3,279	41.13	754	39.79
Moderate to severe (4–8 elements)	2,793	34.88	1,979	24.82	340	17.94
Severe (0–3 elements)	510	6.37	160	2.01	14	0.74

<sup>a</sup> According to Russel classification (33).

<sup>b</sup> Subsample of patients after 6 weeks of treatment who had treatment response (i.e., at least a 50% decrease of their initial Hospital Anxiety and Depression Scale score for depression and less than five diagnostic criteria).

founding factors in our analysis. Nevertheless, age-corrected versus age-uncorrected scores of the Russel classification were very similar and identified the same group three standard deviations below the mean ( $\kappa=0.829$ ,  $\text{var}[\kappa]<0.001$ ).

### Statistics

Variables were examined for the normality of distribution before using parametric statistics. Given the sample size, rejection of a normal distribution (Kolmogorov-Smirnov test) for the factors analyzed was expected for all variables. However, even if significantly different from 0 ( $p<0.001$ ), the statistic was small for major covariables such as age (0.048), Hospital Anxiety and Depression Scale depression score at visit 1 (0.085) and visit 2 (0.057), Hospital Anxiety and Depression Scale anxiety score at visit 1 (0.075) and visit 2 (0.089), and the number of correct delayed recall responses at visit 1 (0.076) and visit 2 (0.073). A graphical appreciation was used to assess the normality of distribution of delayed recall (at the first and second visits). Thus, Q-Q (quantile) plots showing that dependent variables were close enough to the normal distribution were created.

Parametric correlation (Pearson test) was used to compare two continuous variables, and analysis of variance was used to analyze the role of a qualitative factor to explain continuous parameters. Since some parameters directly influenced correct recall in our sample (such as age and Hospital Anxiety and Depression Scale scores), we further analyzed the correlation between the number of potential recall responses and past episodes (their nonindependence being our main hypothesis) in different ranges.

Structural equation modeling was used on the basis of SPSS and SAS PROC CALIS to disentangle the respective role of clustered variables, including state-dependent variables (such as the number of symptoms and Hospital Anxiety and Depression Scale scores) and state-independent variables (such as the total length of depressive history and number of past major depressive episodes) (34).

Patients with "severe" impairment (according to the Russel [33] classification) are at higher risk of neurological defects, and thus structural equation modeling analyses at the first and second visits were conducted, omitting this subgroup of patients. A further

structural equation modeling analysis separated anxiety and depression scores on the Hospital Anxiety and Depression Scale and omitted the number of DSM-IV symptoms in order to avoid reinforcing the role of depression versus anxiety.

Last, we further investigated our main finding, the relationship between delayed recall and the number of past major depressive episodes in patients with treatment response, using a linear regression analysis.

## Results

### Sample

A total of 9,515 depressed patients entered the study. The final sample for analysis consisted of 8,229 patients (86.48%). Subjects were excluded if the Hospital Anxiety and Depression Scale score was below 8 for depression or anxiety (644 patients) or data characterizing the patient were not correctly or completely saved. The subsample of subjects excluded 1) had fewer DSM-IV symptoms of depression ( $t=4.25$ ,  $df=9512$ ,  $p<0.0001$ ); 2) had a shorter length of the present episode ( $t=2.15$ ,  $df=6798$ ,  $p=0.0159$ ); 3) were more frequently men ( $\chi^2=172.9$ ,  $df=8$ ,  $p<0.0001$ ); and 4) had better final Hospital Anxiety and Depression Scale scores for depression ( $t=3.974$ ,  $df=9512$ ,  $p<0.0001$ ) and anxiety ( $t=5.212$ ,  $df=9512$ ,  $p<0.0001$ ), and thus they were more frequently responders ( $\chi^2=3.71$ ,  $df=1$ ,  $p=0.0124$ ).

Women comprised 70.37% of the final sample, and the average age of the sample was 48.02 years ( $SD=14.09$ ). In this population, 1,115 patients were 65 years old or older, and 407 patients were over 75 years old. Civil status among these patients was as follows: married, 48.37%; single, 15.41%; divorced, 15.57%; and widowed, 7.31%. Education levels (middle=high school graduate) were as follows: low, 49.10%; middle, 30.06%; and high, 20.84%. Employment status was as follows: active employment, 57.86%; unem-

**TABLE 3.** Association of Correct Delayed Recall Responses With Different Characteristics in the Final Sample of Patients (N=8,229) With Major Depressive Disorder

Characteristic	Association With the Number of Correct Delayed Recall Responses	
	r (df=8, 227)	p
Number of DSM-IV criteria for symptoms of depression at inclusion	-0.076	<0.001
Length of present depressive episode	-0.010	0.442
Number of past depressive episodes	-0.056	<0.001
Total length of depressive illness	-0.032	0.011
Age	-0.016	<0.001
Education level	0.171	<0.001
Hospital Anxiety and Depression Scale score for depression	-0.146	<0.001
Hospital Anxiety and Depression Scale score for anxiety	-0.059	<0.001
Delay between visit 1 and visit 2	-0.013	0.229
Number of minutes before delayed recall	-0.004	0.712
	F (df=1, 8227)	p
Gender	0.587	0.556
Marital status	15.384	<0.001
Professional activity	46.729	<0.001

ployed, 12.43%; retired, 19.07%; student, 0.11%; and another type of professional activity, 10.53%.

The Hospital Anxiety and Depression Scale score for depression was below 11 in only 10.48% of patients. The duration of current major depressive episode was on average 8.37 weeks (SD=10.84). Among the patients, 18.38% fulfilled five DSM-IV criteria for depression, 31.82% fulfilled six criteria, 27.76% fulfilled seven criteria, 15.44% fulfilled eight criteria, and 6.60% fulfilled nine criteria (Table 1). This was the first episode of depression for 49.64% of patients, the second episode for 25.30%, the third episode for 6.05%, and between the fourth and 13th episode for the remaining patients (4.67%). The length of unipolar depression, except the present episode, was on average 12.55 weeks (SD=34.91), and the lifetime duration (or number of weeks depressed) was 21.23 weeks (SD=33.91).

The second visit was on average 42 days (SD=8.93) after the first visit (between 3 and 20 weeks). At the second visit, the number of depressive symptoms from the list of DSM-IV criteria decreased (Table 1). According to these criteria, 76.39% of patients no longer fulfilled the diagnosis of major depressive episode (i.e., fulfilled less than five criteria). According to Hospital Anxiety and Depression Scale self-rating, 76.44% of patients were responders (i.e., patients who had at least a 50% decrease of the depression score between the two visits).

The number of correct answers for immediate recall of the Weschler paragraph recall index tested at the first visit was between one and 24, with an average of 11.75 appropriate answers (SD=4.72) (see the table in the data supplement accompanying the online version of this article). The delay between immediate and delayed recall was on average 14.15 minutes (SD=7.73; range=5–56). Accordingly, the delay between immediate and delayed recall was introduced into multivariate analyses, as was the delay between the first and second visits. The patients recalled approximately 10 correct details of the paragraph (i.e., 85.89% of the initial immediate recall) (Table 1).

During the second visit, a global improvement of immediate recall was observed (13.24 correct answers; [SD=

4.79; range=1–24]) (see the table in the data supplement accompanying the online version of this article). On average, 13.58 minutes later (SD=7.32; range=5–65), delayed recall was correct for approximately 12 items (Table 1).

The level of memory impairment at both visits as well as in the sample of patients with treatment response at the second visit is detailed in Table 2.

Benzodiazepines were prescribed for 52.51% of patients, but treatment was unchanged during the entire observation period for 90.73% of patients (reflecting the short delay between the first and second visits), and the dosage was fixed for the vast majority of patients (93.30%). Nevertheless, coprescription of benzodiazepines was included in the analyses as a potential confound.

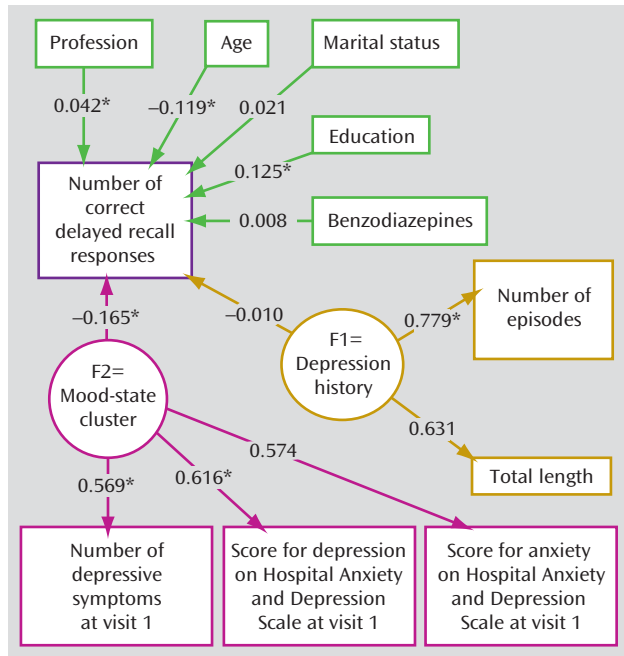
At the initial visit, the 8,229 patients who had a Hospital Anxiety and Depression Scale score above 8 for both depression and anxiety and who fulfilled DSM-IV criteria for major depressive disorder according to the clinician were tested for paragraph delayed recall.

The relationship between the main variables at the first visit is presented as uncorrected correlations to aid in understanding of the data structure. The number of correct responses was significantly correlated with the following (Table 3): younger age, higher education level, presence of professional activity, marital status (positive correlation), number of DSM-IV symptoms of depression, Hospital Anxiety and Depression Scale scores for depression and anxiety, and number of past episodes (negative correlation) ( $p<0.001$ ). A significant but weaker negative correlation was shown for the total length of depressive episodes ( $p<0.012$ ). Gender, delay between the first and second visits, and the time between immediate and delayed recall were not predictive of the number of correct delayed recall responses. When the role of depression and anxiety scores of the Hospital Anxiety and Depression Scale were analyzed separately, depression was significantly correlated with the number of correct delayed recall responses ( $r=-0.125$ ,  $p<0.001$ ), whereas anxiety was not ( $r=-0.016$ ,  $p=0.579$ ).

To account for the existence of different clusters of variables influencing the number of correct responses for



**FIGURE 1. Structural Equation Model Accounting for the Number of Correct Delayed Recall Responses in 8,229 Patients With Major Depressive Disorder<sup>a</sup>**



<sup>a</sup> The numbers are correlation coefficients (statistically different from 0). The following two clusters were proposed according to the initial hypothesis: 1) a cluster of trait variables (F1=depression history) related to the cumulative length of the time of depressive episodes and the number of past major depressive episodes and 2) a cluster of state-dependent variables (F2=mood-state cluster) related to the severity of the depressive state at the initial visit, relating the number of depressive symptoms and Hospital Anxiety and Depression Scale scores for both depression and anxiety. All explanatory variables (depression history, mood state, age, and education) are freely correlated and therefore related by arrows. At the first visit, the mood-state cluster was much more important than the depression history cluster.

\*p<0.05.

paragraph delayed recall, structural equation modeling was performed. At the first visit, mood-state dependent variables had the largest effect on the number of correct delayed recall responses, together with smaller effects of age and level of education (Figure 1). In contrast, the length and number of past episodes of major depression were noncontributory (Figure 1).

The model fit was good, since the upper portion of the 95% confidence interval of the root mean square error of approximation was smaller than 0.05 (0.0436) and the Bentler and Bonnet normed fit index was higher than 0.95 (0.974).

Given the changes in mood and cognitive performance at the second visit, we predicted that the role of past episodes in the second analysis would be strengthened in the subsample of patients who most fully responded. A correlation between the number of past episodes and the number of correct delayed recall responses for the final sample of analysis (N=8,229) was observed ( $r=-0.063$ ,  $p<0.001$ ) and was more notable in younger patients (Table 4). More precisely, among those 1,895 patients whose Hospital Anx-

**TABLE 4. Correlation Between the Number of Past Major Depressive Episodes and the Number of Correct Delayed Recall Responses According to Age in the Final Sample of Patients (N=8,229) After 6 Weeks of Antidepressant Treatment (visit 2)**

Age	Number of Patients	Analysis	
		r	p
18–31	933	-0.105	0.002
32–37	1,032	-0.076	0.016
38–42	1,175	-0.053	0.071
43–47	1,067	-0.064	0.040
48–51	923	-0.023	0.502
52–56	1,045	-0.051	0.103
57–65	1,035	-0.030	0.347
66–97	1,020	-0.018	0.568

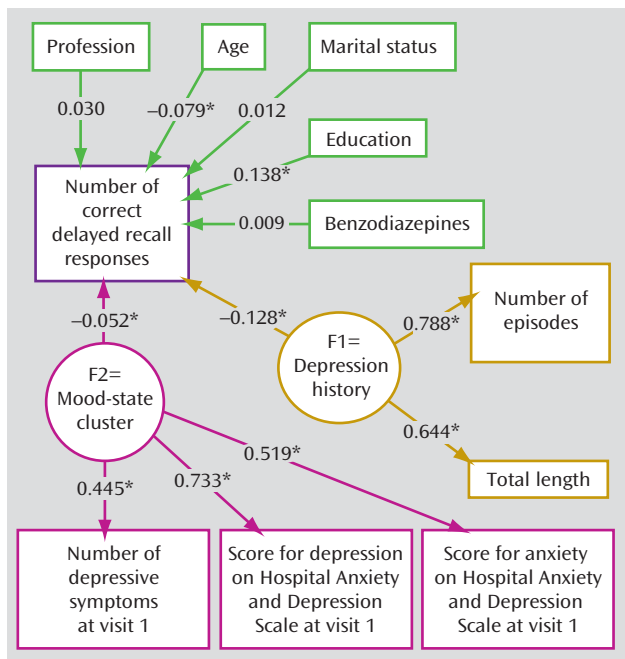
iety and Depression Scale depression score decreased by 50% at the second visit and who had less than five diagnostic criteria, the structural equation modeling showed that the number of past episodes and total length of major depressive disorder constituted a cluster that was significantly correlated with the number of correct delayed recall responses (Figure 2). Thus, there was a significant relationship between past depression and memory performance as proposed in our hypothesis. Moreover, the path coefficient for the past depression cluster was substantially higher than the path coefficient observed between mood-dependent parameters (constituting the other cluster) (Figure 2). None of the Hospital Anxiety and Depression Scale scores for depression and anxiety (omitting the mood-state cluster) were significantly correlated with the number of correct delayed recall responses ( $r=-0.052$ ,  $p=0.067$ ;  $r=-0.016$ ,  $p=0.579$ , respectively).

Again, the model fit was good, with the upper portion of the 95% confidence interval of the root mean square error of approximation equal to 0.0301 and the Bentler and Bonnet normed fit index equal to 0.977. As expected, age and education remained contributory.

We then used a linear regression analysis to examine whether the correlation between the number of correct delayed recall responses and the number of past major depressive episodes was greater at the second visit relative to the first visit. The three most important variables that were significantly correlated with the number of correct delayed recall responses were 1) the visit number (visit 1 versus visit 2;  $F=10,400$ ,  $df=15, 286$ ,  $p<0.001$ ), 2) the number of past depressive episodes ( $F=34.71$ ,  $df=15, 286$ ,  $p<0.001$ ), and 3) the interaction between the former and latter variables ( $F=32.28$ ,  $df=15, 286$ ,  $p<0.001$ ). Age, gender, education, and Hospital Anxiety and Depression Scale score for depression were also significantly correlated ( $F>6.91$ ,  $df=15, 286$ ,  $p<0.009$ ).

Finally, at the first visit we examined the sample of patients who were in remission at the time of the second visit. This allowed us to analyze whether the significant role of the total length of mood disorder and number of past episodes detected in the 1,895 patients who were in remission at the second visit was obscured by current de-

FIGURE 2. Structural Equation Model Accounting for the Number of Correct Delayed Recall Responses in Patients After 6 Weeks of Successful Antidepressant Treatment for Major Depressive Disorder<sup>a</sup>

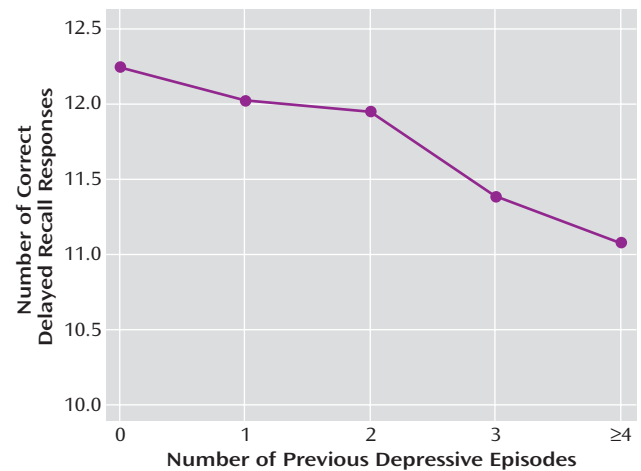


<sup>a</sup> Patients had a Hospital Anxiety and Depression Scale score for depression below 11. The numbers are correlation coefficients (statistically different from 0). The following two clusters were proposed according to the initial hypothesis: 1) a cluster of trait variables (F1=depression history) related to the cumulative length of time of depressive episodes and the number of past major depressive episodes and 2) a cluster of state-dependent variables (F2=mood-state cluster) related to the severity of the depressive state at the initial visit, relating the number of depressive symptoms and Hospital Anxiety and Depression Scale scores for both depression and anxiety. All explanatory variables (depression history, mood state, age, and education) are freely correlated and therefore related by arrows. At the second visit, the mood-state cluster was much less important than past depression history. Thus, past depression history was not a significant predictor of memory function.  
\* $p < 0.05$ .

pression at the first visit. The structural equation modeling showed that the number of past episodes and total length of major depressive disorder constituted a cluster that was no longer significantly correlated with the number of correct delayed recall responses ( $r = -0.019$ ,  $p = 0.59$ ), whereas the mood-state dependent variable was significantly predictive ( $r = -0.249$ ,  $p < 0.001$ ). Age ( $r = -0.121$ ,  $p < 0.001$ ) and education level ( $r = 0.118$ ,  $p < 0.001$ ) were also contributory, and the model fit was satisfying, with the upper portion of the 95% confidence interval of the root mean square error of approximation equal to 0.023 and the Bentler and Bonnet normed fit index equal to 0.972.

Omitting 510 patients at the first visit and 160 patients at the second visit (Table 2) who were classified as having “severe” impairment did not significantly change the results (the state cluster was correlated [ $r = -0.128$ ,  $p < 0.001$ ] with the number of correct delayed recall responses at the first visit,

FIGURE 3. Number of Correct Delayed Recall Responses at the Second Visit According to the Number of Previous Depressive Episodes in Patients With Major Depressive Disorder



but the trait cluster was not [ $r = -0.016$ ,  $p = 0.185$ ], whereas the state cluster was not correlated [ $r = -0.035$ ,  $p = 0.163$ ] with the number of correct delayed recall responses at the second visit, but the trait cluster was correlated [ $r = -0.123$ ,  $p < 0.001$ ]). Therefore, the initial results shown in Figure 1 and Figure 2 reflect a general effect and were not unduly influenced by the “severe” impairment group.

## Discussion

Memory function was critically influenced by a range of clinical variables in a very large sample of outpatients with major depressive disorder. At presentation with depression, current illness severity was an important determinant of performance, while the intensity of previous depressive history (the number and length of past episodes) was not. At the follow-up and after significant clinical response, the intensity of previous depressive history was more significant than current symptoms. Inevitably, age, education level, and profession had significant impact, regardless of the subsample and stage of testing.

The present study is uniquely large. It therefore permits generalization of conclusions pertaining to memory function previously proposed from much smaller, more intensively studied samples of often more severely ill patients. The findings support the hypothesis that the intensity of past depression contributes to the impairment of memory performance when patients are recovered. This effect was most notable in younger patients, and it is difficult to determine whether this is because the reporting of previous depression is more accurate among patients of younger age groups or because other factors are more significant in older patients. It is also noteworthy that the effects of illness intensity were differentiated in patients who showed the most complete clinical response. The contribution in partial responders or chronically depressed individuals may be

greater, but this is difficult to estimate because they are always confounded by ongoing symptoms of depression.

There are several limitations to this study that should be acknowledged. First, the delayed recall was assessed by informed but untrained clinicians. The Wechsler paragraph recall index belongs to a package that usually requires specific training, and the quality of assessment is an important factor, especially in depressed patients in whom motivation and concentration are decreased. (On the other hand, the clinicians were unbiased and indifferent to our primary hypothesis, which are advantages. Moreover, the impact of trait markers of depression upon delayed memory was large enough to be detected, even in an informal clinical setting.) Second, associated neurological defects, such as dementia, may affect memory functions independent of the impact of depression. While not specifically excluded at screening, omitting patients with severe memory impairments did not alter our findings.

The majority of previous studies of paragraph delayed recall focused on geriatric depression (35, 36), since the cognitive impact of mood disorder is believed to be more important in this population (37). Nevertheless, associations have been found between depression and cognitive function related to the hippocampus in young (38) and middle-aged (39) patients as well. Age, gender, education level, length and intensity of depressive episode, and illness may be relatively interrelated, and thus a large sample is needed to disentangle the central effects. There are other potential concerns. Two out of three studies have relied on elderly patients, in whom remission frequently takes longer to attain (40). Low quality of remission has a direct impact on motivation and many cognitive functions, which interferes with the assessment of delayed recall. The adverse impact of antidepressant treatments on memory is also an inevitable confounder in most studies but is usually found to be modest in effect size (28, 37, 38).

The present results are comparable with one of only a few follow-up studies focusing on hippocampal volume changes during and after a major depressive episode (39). In the study conducted by Bremner et al., no significant difference was observed between depressed patients and comparison subjects at baseline, but at the 1-year follow-up, a subgroup of nonremitting patients had smaller hippocampal volumes relative to comparison subjects and remitting patients. Atrophy of the hippocampus is one of the most consistent imaging findings in major depressive disorder. This has highlighted the potential role of physiological stress as a core mediating factor (often assuming a neurotoxic effect of cortisol on the hippocampus) and has emphasized the need for antidepressant treatments that might prevent or even reverse hippocampal atrophy (41). The toxicity of stress on the hippocampus may depend on the period of time depressed, with an effect size that remains to be assessed. Large effect sizes may be observed when analyses are based on the direct measurement of the

## Patient Perspective

“Mr. X,” a 57-year-old divorced man who worked as a driver, was depressed for 4½ months and decided to visit his general practitioner primarily because of fatigue, insomnia, and sadness. These symptoms reminded him of a 6-month depressive episode that he experienced 13 years prior, which was successfully treated by a selective serotonin reuptake inhibitor. The general practitioner made the diagnosis of major depressive episode, with a Hospital Anxiety and Depression Scale score of 28. The patient agreed to participate in the present protocol. He was asked to concentrate on the following story and to remember as many details as possible: “Miss Anne Boiron, from the 20th arrondissement of Paris, a cook in a school canteen, went to the police close to a City Hall to report that she had been attacked in the street the preceding night and that she was robbed of 30 euros. As the rent was unpaid and she had four young children going hungry, the policeman was moved by the story of this lady and organized a collection for her.” Of the 24 items of detail in this story, Mr. X was able to directly quote eight. Ten minutes after completing the Hospital Anxiety and Depression Scale, he could recollect only five items of detail from the story.

Six weeks later, the patient’s Hospital Anxiety and Depression Scale score decreased to 17. When asked to remember a story containing exactly the same number of items but on a different topic, he could immediately remember 12 items. After a delay of 10 minutes, he could recall 10 items of detail.

hippocampus with MRI in severely ill patients with a variable length of mood disorder (2).

Our hypothesis proposed that depression is directly involved in hippocampal atrophy. However, it is possible that anxiety plays a role in influencing limbic anatomy. Such an effect could be independent from the impact of depression (28) or reversed relative to (27, 29) the impact of depression. Anxiety and depression are highly interconnected dimensions in our sample, and therefore it is of interest that depressive symptoms seemed to have the dominant effect on delayed recall at the initial visit. However, we studied major depressive episodes, and thus the possibility that anxiety disorders have specific implications for memory function cannot be excluded.

We expressed the detected association between depression and hippocampal atrophy in a statistical model that takes into account the contributions of many factors to the observed effects. The strong statistical significance of these factors reflects sample size rather than the magnitude of the effect of any individual factor. However, if the impact of repeated episodes of major depression is expressed as an effect size, we can argue that memory performance is impaired by 2%–3% for each previous depressive episode up to four episodes (Figure 3). This ignores

other potentially correlated variables, but it captures an important clinical reality.

Our findings may have important public health implications. We conclude that frequent, long, or chronic states of depression have impairing effects on brain function. Therapeutically, there may be advantages in early treatment, for a sufficient period, and in not tolerating chronicity or even partial remission. Failure to treat adequately may impair the global outcome of major depressive disorder because brain recovery may be incomplete. The existence of even small cognitive effects has important implications pertaining to how we feel about depression in the general population. There is a current tendency to demean the significance of depressive symptoms as evidence of distress rather than illness. We would not seek to dispute the distress, but our data support the hypothesis that recurrent or prolonged depression has effects on the brain that make it a significant and disabling illness.

---

Received April 3, 2007; revisions received Oct. 10 and Dec. 11, 2007; accepted Jan. 17, 2008 (doi: 10.1176/appi.ajp.2008.07040574). From INSERM U675, Institut Fédératif de Recherche (IFR 02), Paris; Assistance Publique-Hôpitaux de Paris, Service de Psychiatrie Adulte, Louis Mourier Hospital (Paris 7), Colombes, France; Assistance Publique-Hôpitaux de Paris, Department of Psychiatry, Bicêtre University Hospital, Le Kremlin Bicêtre, France; INSERM U 669, Paris; University Paris-Sud and Paris Descartes, UMR-S0669, Paris; Assistance Publique-Hôpitaux de Paris, Paul Brousse Hospital, Public Health Department, Villejuif, France; and the Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK. Address correspondence and reprint requests to Dr. Gorwood, INSERM U675, Faculty of Medicine Bichat (IFR02), 16 rue Henri Huchard, 75018 Paris, France; philip.gorwood@lmr.aphp.fr (e-mail).

Dr. Gorwood has received research grants from Eli Lilly; he has received speakers honoraria from Lundbeck, Servier, and UCB-Pharma; and he has served on the advisory boards for Janssen, Servier, and Wyeth. Dr. Corruble has received speakers honoraria from Servier, Eli Lilly, Wyeth, Bristol-Myers Squibb, and UCB-Pharma; and he has served on the advisory boards of Servier, Janssen, and Wyeth. Dr. Falissard has received speakers honoraria from Bristol-Myers Squibb, Servier, Pierre Fabre, and Eli Lilly; and he has served on the advisory boards of Eli Lilly, Janssen Cilag, and Sanofi-Aventis. Dr. Goodwin has received research grants from Sanofi-Aventis and Servier; he has received speakers honoraria from AstraZeneca, Bristol-Myers Squibb, Eisai, Lundbeck, Sanofi-Aventis, and Servier; and he has served on the advisory boards of AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Lundbeck, P1Vital, Sanofi-Aventis, Servier, and Wyeth.

Supported by a research grant from Servier. Servier had no involvement in the design, organization, analysis, or preparation for publication of the study.

## References

- Shah PJ, Ebmeier KP, Glabus MF, Goodwin GM: Cortical grey matter reductions associated with treatment-resistant chronic unipolar depression: controlled magnetic resonance imaging study. *Br J Psychiatry* 1998; 172:527–532
- Sheline YI, Sanghavi M, Mintun MA, Gado MH: Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *J Neurosci* 1999; 19:5034–5043
- Bearden CE, Glahn DC, Monkul ES, Barrett J, Najt P, Kaur S, Sanches M, Villarreal V, Bowden C, Soares JC: Sources of declarative memory impairment in bipolar disorder: mnemonic processes and clinical features. *J Psychiatr Res* 2006; 40:47–58
- Rapp MA, Dahlman K, Sano M, Grossman HT, Haroutunian V, Gorman JM: Neuropsychological differences between late-onset and recurrent geriatric major depression. *Am J Psychiatry* 2005; 162:691–698
- MacQueen GM, Campbell S, McEwen BS, Macdonald K, Amano S, Joffe RT, Nahmias C, Young LT: Course of illness, hippocampal function, and hippocampal volume in major depression. *Proc Natl Acad Sci U S A* 2003; 100:1387–1392
- Campbell S, Marriott M, Nahmias C, MacQueen GM: Lower hippocampal volume in patients suffering from depression: a meta-analysis. *Am J Psychiatry* 2004; 161:598–607
- Videbech P, Ravnkilde B: Hippocampal volume and depression: a meta-analysis of MRI studies. *Am J Psychiatry* 2004; 161:1957–1966
- Squire LR, Ojemann JG, Miezin FM, Petersen SE, Videen TO, Raichle ME: Activation of the hippocampus in normal humans: a functional anatomical study of memory. *Proc Natl Acad Sci U S A* 1992; 89:1837–1841
- Schacter DL, Alpert NM, Savage CR, Rauch SL, Albert MS: Conscious recollection and the human hippocampal formation: evidence from positron emission tomography. *Proc Natl Acad Sci U S A* 1996; 93:321–325
- Schacter DL, Savage CR, Alpert NM, Rauch SL, Alpert MS: The role of hippocampus and frontal cortex in age-related memory changes: a PET study. *Neuroreport* 1996; 7:1165–1169
- Grasby PM, Frith CD, Friston KJ, Frackowiak RSJ, Dolan RJ: Activation of the human hippocampal formation during auditory verbal long-term memory function. *Neurosci Lett* 1993; 163:185–188
- Nyberg L, McIntosh AR, Houle S, Nilsson L-G, Tulving E: Activation of medial temporal structures during episodic memory retrieval. *Nature* 1996; 380:715–717
- Grasby PM, Frith CD, Friston KJ, Bench C, Frackowiak RSJ, Dolan RJ: Functional mapping of brain areas implicated in auditory-verbal memory function. *Brain* 1993; 116:1–20
- Dolan RJ, Fletcher PC: Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature* 1997; 388:582–585
- Bremner JD, Soufer R, McCarthy G, Delaney RC, Staib LH, Duncan JS, Charney DS: Gender differences in cognitive and neural correlates of remembrance of emotional words. *Psychopharmacol Bull* 2001; 35:55–87
- Grady CL, McIntosh AR, Horwitz B, Maisog JM, Ungerleider LG, Mentis MJ, Pietrini P, Schapiro MB, Haxby JV: Age-related reductions in human recognition memory due to impaired encoding. *Science* 1995; 269:218–221
- Haxby JV, Horwitz B, Ungerleider LG, Maisog JM, Pietrini P, Grady CL: The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J Neurosci* 1994; 14:6336–6353
- Haxby JV, Ungerleider LG, Horwitz B, Maisog JM, Rapoport SI, Grady CL: Face encoding and recognition in the human brain. *Proc Natl Acad Sci U S A* 1996; 93:922–927
- Kapur N, Friston KJ, Young A, Frith CD, Frackowiak RSJ: Activation of human hippocampal formation during memory for faces: a PET study. *Cortex* 1995; 31:99–108
- Schacter DL, Reiman E, Uecker A, Polster MR, Yun LS, Cooper LA: Brain regions associated with retrieval of structurally coherent visual information. *Nature* 1995; 376:587–590
- Ghaem O, Mellet E, Crivello F, Tzourio N, Mazoyer B, Berthoz A, Denis M: Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport* 1997; 8: 739–744
- Maguire EA, Frackowiak RSJ, Frith CD: Learning to find your way: a role for the human hippocampal retrieval of structurally coherent visual information. *Nature* 1995; 376:587–590



23. Bremner JD, Vythilingam M, Vermetten E, Southwick SM, McGlashan T, Nazeer A, Khan S, Vaccarino LV, Soufer R, Garg PK, Ng CK, Staib LH, Duncan JS, Charney DS: MRI and PET study of deficits in hippocampal structure and function in women with childhood sexual abuse and posttraumatic stress disorder. *Am J Psychiatry* 2003; 160:924–932
24. Zola-Morgan SM, Squire LR: The primate hippocampal formation: evidence for a time-limited role in memory storage. *Science* 1990; 250:288–290
25. Buckner RL, Petersen SE, Ojemann JG, Miezin FM, Squire LR, Raichle ME: Functional anatomical studies of explicit and implicit memory retrieval tasks. *J Neurosci* 1995; 15:12–29
26. Fletcher PC, Henson RN: Frontal lobes and human memory: insights from functional neuroimaging. *Brain* 2001; 124:849–881
27. Vyas A, Mitra R, Shankaranarayana Rao BS, Chattarji S: Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *J Neurosci* 2002; 22: 6810–6818
28. Kalisch R, Schubert M, Jacob W, Kessler MS, Hemauer R, Wigger A, Landgraf R, Auer DP: Anxiety and hippocampus volume in the rat. *Neuropsychopharmacology* 2006; 31:925–932
29. Rusch BD, Abercrombie HC, Oakes TR, Schaefer SM, Davidson RJ: Hippocampal morphometry in depressed patients and control subjects: relations to anxiety symptoms. *Biol Psychiatry* 2001; 50:960–964
30. Wechsler D: Wechsler Memory Scale–Revised. New York, Psychological Corp, 1987
31. Butters N, Sax D, Montgomery K, Tarlow S: Comparison of the neuropsychological deficits associated with early and advanced Huntington's disease. *Arch Neurol* 1978; 35:585–589
32. Storandt M, Botwinick J, Danziger WL, Berg L, Hughes CP: Psychometric differentiation of mild senile dementia of the Alzheimer type. *Arch Neurol* 1984; 41:497–499
33. Russel EW: A multiple scoring method for the assessment of complex memory functions. *J Consult Clin Psychol* 1973; 43: 800–809
34. Aitkin M, Francis B: *Statistical Modelling in GLIM4*. Oxford, UK, Oxford University Press, 2005
35. de Asis JM, Stern E, Alexopoulos GS, Pan H, Van Gorp W, Blumberg H, Kalayam B, Eidelberg D, Kiosses D, Silbersweig DA: Hippocampal and anterior cingulate activation deficits in patients with geriatric depression. *Am J Psychiatry* 2001; 158:1321–1323
36. Golomb J, Kluger A, de Leon MJ, Ferris SH, Convit A, Mittelman MS, Cohen J, Rusinek H, De Santi S, George AE: Hippocampal formation size in normal human aging: a correlate of delayed secondary memory performance. *Learn Mem* 1994; 1:45–54
37. Steffens DC, Otey E, Alexopoulos GS, Butters MA, Cuthbert B, Ganguli M, Geda YE, Hendrie HC, Krishnan RR, Kumar A, Lopez OL, Lyketsos CG, Mast BT, Morris JC, Norton MC, Peavy GM, Petersen RC, Reynolds CF, Salloway S, Welsh-Bohmer KA, Yesavage J: Perspectives on depression, mild cognitive impairment, and cognitive decline. *Arch Gen Psychiatry* 2006; 63:130–138
38. Blumberg HP, Kaufman J, Martin A, Whiteman R, Zhang JH, Gore JC, Charney DS, Krystal JH, Peterson BS: Amygdala and hippocampal volumes in adolescents and adults with bipolar disorder. *Arch Gen Psychiatry* 2003; 60:1201–1208
39. Bremner JD, Vythilingam M, Vermetten E, Vaccarino V, Charney DS: Deficits in hippocampal and anterior cingulate functioning during verbal declarative memory encoding in midlife major depression. *Am J Psychiatry* 2004; 161:637–645
40. Gorwood P, Weiller E, Lemming O, Katona C: Escitalopram prevents relapse in older patients with major depressive disorder. *Am J Geriatric Psychiatry* 2007; 15:581–593
41. Vythilingam M, Vermetten E, Anderson GM, Luckenbaugh D, Anderson ER, Snow J, Staib LH, Charney DS, Bremner JD: Hippocampal volume, memory, and cortisol status in major depressive disorder: effects of treatment. *Biol Psychiatry* 2004; 56:101–112