# Quantitative and qualitative differences in the verbal learning performance of elderly depressives and healthy controls

DEBORAH A. KING, <sup>1</sup> CHRISTOPHER COX, <sup>2</sup> JEFFREY M. LYNESS, <sup>1</sup> YEATES CONWELL, <sup>1</sup> AND ERIC D. CAINE <sup>1</sup>

<sup>1</sup>Department of Psychiatry, University of Rochester School of Medicine and Dentistry,

UR-NIMH Center for the Study of Psychopathology of the Elderly

<sup>2</sup>Department of Biostatistics, University of Rochester Medical Center

(Received May 11, 1995; Revised May 27, 1997; Accepted May 30, 1997)

#### Abstract

We compared the verbal learning and memory performance of 57 inpatients with unipolar major depression and 30 nondepressed control participants using the California Verbal Learning Test. The effect of age within this elderly sample was also examined, controlling for sex, educational attainment, and estimated level of intelligence. Except for verbal retention, the depressives had deficits in most aspects of performance, including cued and uncued recall and delayed recognition memory. As well, there were interactions between depression effects and age effects on some measures such that depressives' performance declined more rapidly with age than did the performance of controls. The results are discussed in the context of recent contradictory reports about the integrity of learning and memory functions in late-life depression. We conclude that there is consistent evidence, from this and other studies, that elderly depressed inpatients have significant deficits in a range of explicit verbal learning functions. (*JINS*, 1998, *4*, 115–126.)

Keywords: Depression in the elderly, Verbal learning and memory

## **INTRODUCTION**

Clinical investigators have explored the neuropsychological functioning of depressed patients for at least two decades, yet there is marked disagreement as to whether depression-related deficits exist and whether purported deficits are indicative of disturbed brain function (cf. Bieliauskas, 1993; King & Caine, 1996; LaRue, 1992; Poon, 1992). Moreover, studies are just beginning to clarify the nature of the relationships between depression, aging, and cognitive function (Boone et al., 1994; King et al., 1993; Lyness et al., 1994; Raskin, 1986). These controversies and unresolved questions are especially apparent when one reviews the literature on learning and memory. Is there an identifiable pattern of memory deficit associated with depression in late life? Are individuals with major depression

Alzheimer's disease (AD), on the basis of quantitative and/or qualitative aspects of episodic memory performance. These

studies have employed a variety of types of learning tasks and stimulus materials, including word-list learning (Hart et al., 1987b; Hill et al., 1993; King et al., 1991; Speedie et al., 1990), memory for line drawings of shapes or common objects (Hart et al., 1987a; King et al., 1991), and learning of common objects presented tactually and visually (LaRue, 1989). Deficits in acquisition and retrieval of information have been commonly associated with both depression and dementia, whereas retention deficits appeared to be more characteristic of dementia. In terms of qualitative aspects of performance, patients with dementia have been reported to make more errors of intrusion when attempting to recall words from a list (i.e., they include more

more vulnerable to the effects of aging than nondepressed

ical studies, investigators have consistently distinguished el-

derly depressives from patients with dementia, typically

Providing one area of uniform agreement among empir-

individuals?

Reprint requests to: Deborah A. King, Associate Professor of Psychiatry (Psychology), University of Rochester, Strong Memorial Hospital, 300 Crittenden Boulevard, Rochester, NY 14642-8408, USA.

extra-list words) and to recall a lower percentage of words from the beginning of the word list (Hill et al., 1993).

In striking contrast to the uniformity of findings regarding depression and dementia, there is considerable disagreement about the specific nature of elderly depressives' episodic memory deficits relative to control participants. A number of investigations reported elderly depressives to be deficient on tests of immediate and/or delayed recall of information, including word-list learning (Bäckman & Forsell, 1994; Hart et al., 1987a; Hill et al., 1993; King et al., 1991, 1993); story recall (King et al., 1995), recall of line drawings (Hart et al., 1987b; King et al., 1991, 1993) and learning of common objects (LaRue, 1989). As well, some have reported deficits in recognition memory on these tasks (Bäckman & Forsell, 1994; King et al., 1991; LaRue, 1989; Speedie et al., 1990). Others, using a similar variety of verbal memory and object learning tasks, found few if any depression-associated deficits (Bieliauskas, 1993; Miller & Lewis, 1977; Niederehe, 1986; Niederehe & Camp, 1985; Poon, 1992).

As pointed out previously by a number of authors (e.g., Burt et al., 1995; Niederehe, 1986; Raskin, 1986; Weingartner & Silberman, 1982), interpretation of these disparate findings has been hampered by inconsistent control of variables such as education and intelligence, and the use of samples that differed with respect to the severity and type of depressive disorder (e.g., inpatients vs. outpatients, patients with syndromically defined major depressive disorder vs. individuals with elevated scores on dimensional depression scales or heterogeneous samples of patients with major or minor depression). As well, studies have used a broad variety of tasks to assess memory performance that included different modes of stimulus presentation, different numbers of learning trials, varied amounts of time between acquisition and recall, and varied levels of processing or encoding specificity. All of these factors would be expected to result in varying degrees of difficulty and/or effort, an underlying dimension of cognitive tasks demonstrated to be associated with impaired performance in depression (cf. Cohen et al., 1982; Hartlage et al., 1993). Moreover, most studies have focused on quantitative measures of memory performance (i.e., How much do they learn?) to the neglect of qualitative measures (i.e., How do they learn?).

In an attempt to synthesize and clarify the literature on depression and memory performance, Burt et al. (1995) conducted a meta-analysis of 99 studies of recall and 48 studies of recognition in clinically depressed and nondepressed samples. They found a significant, stable association between depression and memory impairment overall, as well as evidence linking depression with specific aspects of memory and with specific subsets of depressed individuals. Consistent with other investigators, they found greater associations of depression with both recall and recognition memory in inpatient depressives as compared to outpatients. Unexpectedly, significant age effects revealed a greater association of depression with impaired performance in younger patients compared to older patients. The authors offer a number of possible reasons for this finding, including the possibilities that age may have overlapped with other variables that were not controlled in most of the studies and the possibility that age-associated heterogeneity due to factors such as increased medical illness may have overshadowed a depression effect. While providing some useful conclusions about the literature on depression and memory performance to date, the authors acknowledged the significant limitations of their database of studies (e.g., 70% did not specify the subtype of depression in terms of unipolar or bipolar) and joined with other prominent depression researchers in calling for more research that clearly characterizes the nature of study participants and tasks (Reynolds et al., 1993).

Despite the fact that both old age and depression are associated with increased levels of medical illness, very few investigators have examined directly the impact of overall medical burden on the performance of older depressives. LaRue and colleagues failed to find an association between physical health and memory performance in two studies of hospitalized depressed patients (LaRue, 1989, 1992) and one investigation of a large community cohort (LaRue et al., 1995). Similarly, we found minimal impact of medical illness on the neuropsychological performance of elderly inpatient depressives when illness was assessed with a standardized measure of general medical burden (King et al., 1995). However, we are unaware of studies that have used a standardized instrument to directly examine the possible confounding role of general medical illness in producing apparent age or depression effects on specific aspects of memory performance.

In light of the limitations of previous studies conducted by us and others, the present study had four distinct aims: (1) to confirm the presence of previously described deficits in a range of well-defined learning and memory functions using a carefully diagnosed sample of elderly inpatients with unipolar major depression while controlling for age, sex, educational level, and estimated verbal intelligence; (2) to compare qualitative aspects of the depressives' performance with that of controls, especially in regard to the verbal learning strategies employed and the frequency of errors of intrusion; (3) to discern the nature of the relationships between age and depression on verbal recall, recognition, and retention; and (4) to determine whether potential depression or age effects were attributable to increased medical illness by using a standardized illness rating scale.

We predicted that elderly depressives' verbal learning would be quantitatively and qualitatively distinct from that of nondepressed elderly controls. More specifically, in terms of quantitative measures, we predicted that elderly depressives would be impaired on immediate and delayed recall of a word list, and that the impairments would be evident even after cues were provided. We did not expect elderly depressives to have more difficulty retaining words once they had been learned. Qualitatively, we predicted that the depressives would be less likely to use an effortful processing strategy that required the active imposition of structure on the material to be learned (i.e., clustering words according to their semantic properties). In contrast, we expected that depressives would not be impaired in the use of a less intentional, more passive strategy (i.e., processing words according to their order of presentation). As well, we predicted that depressives would not differ from controls in terms of the frequency of errors of intrusion, a deficit associated with dementia (Hill et al., 1993).

Our previous work reported that the verbal recall performance of severely depressed inpatients declined with age in a fashion parallel to that of nondepressed controls, whereas depressives' performance on tasks such as naming and Trail Making B declined more precipitously than that of controls (King et al., 1991, 1993). These findings were recently reinforced by Lyness et al. (1994) with regard to the memory performance of an outpatient sample. Therefore, we predicted that there would be parallel (i.e., noninteractive) negative effects of age on verbal recall, recognition, and retention in this study. Consistent with previous findings, we expected to find little or no impact of general medical illness on memory performance.

#### METHODS

#### **Research Participants**

Fifty-seven inpatients (40 women, 17 men) diagnosed with unipolar major depression using the Structured Clinical Interview for DSM-III-R (SCID; Spitzer et al., 1987) gave informed consent to participate in a study of the neuropsychological effects of depression. Forty of these patients were included in a previous study of the general neuropsychological effects of depression that did not involve the verbal learning findings reported here (King et al., 1995). Consecutive admissions to the acute-care psychiatric units of Strong Memorial Hospital were approached for participation in the study if they were depressed, at least 50 years of age, had at least an eighth grade education, had no prior record of neurologic or schizophrenic illness, and had no hearing or vision impairments that would preclude neuropsychological testing. With these exceptions, the sample was believed to be representative of elderly patients requiring acute treatment of depression. Scores on the 24-item Hamilton Depression Rating Scale (HDRS; Hamilton, 1967) ranged from 16 to 46, with a mean score of 28.91 (SD = 8.21). The depressives ranged in age from 50 to 86, with a mean age of 65.77 (SD = 10.22). Age of onset of first major depressive episode (SCID) ranged from 16 to 86 years, with a mean of 51.81 (SD = 18.35). Years of education ranged from 8 to 18, with a mean educational level of 12.67 (SD = 2.39). Scores on the Vocabulary subtest of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) were used to estimate verbal intelligence. These ranged from 10 to 65, with a mean score of 43.04 (*SD* = 13.73).

All patients were assessed within the 1st week of hospitalization to maximize the likelihood that they were fully symptomatic at the time of testing. One of the patients had been treated with electroconvulsive therapy (ECT) 10

months prior to the study; 8 others had received ECT more than 1 year previously. There were no differences in age, education, or neuropsychological test performance between depressives with and without history of ECT (t tests for unequal variances, alpha = .05), except that depressives with previous ECT recalled fewer words from the end of the word list [t(12.4) = -2.59, p < .02]. Thirty-eight of the depressed patients were taking psychotropic medication in the form of benzodiazepines, antidepressants, antipsychotics, or some combination of the three. Those taking psychotropic medication were no different from the rest of the depressed group in terms of age, education or neuropsychological measures (*t* tests for unequal variances, alpha = .05). Eleven of the patients were diagnosed major depression with psychotic features; all were carefully assessed to insure that they could understand and respond to the test materials. Although these patients typically had somatic or depressive delusions that warranted a psychotic designation, they did not have formal thought disorder or thought processing problems such as blocking or loose associations that interfered with their cooperation during testing. Jeste et al. (1996) reported that a younger psychotically depressed sample performed like schizophrenic patients on a range of neuropsychological measures. However, in a previous study (King et al., 1995) we found no differences between eight of these delusional patients and the other elderly depressed patients on general neuropsychological measures (e.g., naming, copying, Trail Making B). In the present study, there were no differences between delusional and nonpsychotic patients in terms of age, education, or CVLT performance, except that the delusional patients had a smaller learning slope [t(27.4) = -2.61, p < .01], and recalled fewer words from the end of the word list [t(17.6) = 2.61, p < .02]. All analyses of these variables were repeated with and without psychotic patients to determine whether this possible confound influenced the results (see Results).

All potential participants were screened to exclude those with a history of clinically significant head injury as defined by trauma followed by sustained loss of consciousness or subsequent anterograde amnesia. Patients with clinical evidence of focal or progressive brain disease were also excluded. Such determinations were made by careful physician-investigator review of all available patient records, including the admission histories and physical examinations (including neurologic exam) routinely performed by consulting internists, and all laboratory and radiologic assessments obtained by the treating clinicians. In addition to the SCID interview, all subjects received a careful assessment of mental status that included the Mini-Mental State Exam in all but 6 cases that were included prior to the use of this measure in our protocol (MMSE; Folstein et al., 1975). MMSE scores ranged from 18 to 30, with a mean score of 27.25 (SD = 2.52). To provide further information regarding functional level, we administered the Instrumental Activities of Daily Living (IADL) and Physical Self-Maintenance Scales (PSMS) to all but two depressed participants (Lawton & Brody, 1969). Family members and/or nursing staff were consulted for corroborating reports of participants' functioning when the subjects themselves were vague or had difficulty responding to the clinical interview. Scores ranged from 0 to 24 on the IADL, with a mean of 6.09 (SD = 6.49), and from 0 to 12 on the PSMS, with a mean score of 1.30 (SD = 2.36). Using all of these assessments, subjects who met DSM–III–R criteria for dementia or delirium were excluded. As well, participants were excluded if they had clinically manifest CNS effects from nonpsychotropic medication.

Thirty nondepressed controls (21 women, 9 men) were recruited from senior citizen centers or from a list of University of Rochester alumni. They were screened for presence or history of psychiatric or neurologic illness using the same measures administered to depressives. HRSD scores ranged from 0 to 9, with a mean score of 3.27 (SD = 2.55). Controls ranged in age from 50 to 83 years, with a mean age of 69.07 (SD = 8.74). Years of education ranged from 8 to 18, with a mean educational level of 13.00 (SD = 2.91). Vocabulary scores ranged from 14 to 65, with a mean Vocabulary score of 47.23 (SD = 12.40). Using unequal variances t tests (alpha = .05), there were no significant differences in age, p = .12; education, p = 0.59 or Vocabulary score, p = .15, between the depressed patients and the control participants. As well, there was no significant difference in the composition by sex of the two groups (chi-square test, alpha = .05).

#### **Memory Measures**

Verbal learning and memory were assessed using The California Verbal Learning Test (CVLT; Delis et al., 1987). The CVLT was selected because of its extensive validation with aged populations, its sound psychometric properties that included an ability to accommodate a wide range of performance levels (our own and other previous studies were hampered by ceiling and floor effects), and its assessment of a broad range of both quantitative and qualitative aspects of learning and memory. The CVLT was administered as part of a full neuropsychological battery that included tests of attention, language ability, memory, and visuospatial processing administered in a fixed order (see King et al., 1995).

The CVLT involves reading a list (List A) of 16 words to the subject for five learning trials and asking for recall of as many words as possible after each trial. The five trials were followed sequentially by the immediate presentation of a second list (List B), immediate recall testing for List B, shortand long-delay recall testing of List A (with and without semantic cuing), and long-delay recognition testing of the 16 target words from List A versus 28 distractor items. For the purposes of this study, the following measures were of primary interest: *total recall* (the total number of words recalled on Trials 1–5), *learning slope* (an index of the increment in words recalled per trial over Trials 1–5), *long-delay recall* (free recall after 20 min), *long-delay cued recall* (delayed recall in response to semantic cues about list items; e.g., clothing, fruits), *verbal retention* (a "percent savings" measure derived from the long-delay recall score divided by the short-delay recall score, multiplied by 100), discriminability (a measure of delayed recognition that takes into account both hits and false positives), semantic clustering (the degree to which the participant recalls target items in the order of categorical groups), serial clustering (the degree to which the participant recalls items in the order that they were presented), and free and cued recall intrusions (words recalled that were not on the list under both free and cued recall). Secondary measures included the percent of words recalled that were from the beginning (% primacy), middle (% middle), and end (% recency) of the list; the number of perseverations (repetitions of words in the same learning trial); and response bias (a participant's tendency toward positive or negative response set). Response bias scores on the CVLT range from -1.00 (a conservative response style) to 1.00 (a liberal response style), with a score of 0.00 reflecting a neutral style of responding.

#### Assessment of Medical Illness

Each participant's medical history, physical examination at the time of admission to the hospital (or study), and all available laboratory tests were reviewed by a physician investigator. All current medical diagnoses were used to complete the Cumulative Illness Rating Scale (CIRS; Linn et al., 1968), a valid and reliable assessment of overall medical burden based on evidence of organ system pathology (Conwell et al., 1993). For our purposes, the psychiatric item of the CIRS was excluded.

### **Statistical Analysis**

In order to compare the memory performance of depressives and controls, analysis of covariance was performed using the raw Vocabulary score as a covariate. Vocabulary was included as a covariate in all analyses in order to further examine the possibility that subtle differences in intelligence might have accounted for obtained differences between the groups. This procedure would not be expected to contaminate findings regarding level of depression and Vocabulary score, as there was no significant correlation between Vocabulary score and HRSD score, either within the depressed group [Pearson Correlation Coefficient, r = -.14, p = .31], or in the entire sample [r = -.18, p = .09]. The same ANCOVA model was used to test for differences between depressed patients and control participants in terms of learning strategies, that is, semantic and serial clustering, with adjustment for Vocabulary score. To evaluate the effect of age on the performance of depressives and controls, ANCOVA was conducted using age as a covariate (in addition to Vocabulary score). This ANCOVA included a test for parallelism of the regressions in each of the two groups. That is, we compared the regression coefficients relating age to the test score within each participant group. In this way we were able to determine whether the effects of age were similar in each group.

#### Depression, learning, and memory

In all analyses described below, plots of standardized residuals were used to check the required assumptions of normally distributed errors with constant variance. Examination of the residuals for some of the variables (see Results) revealed outliers whose scores were inconsistent with the ANCOVA model (standardized residual > 3 or < -3). The demographic and clinical characteristics of outliers were reviewed in order to determine if there were discernible reasons why they would differ from the rest of the group. As well, we determined whether the same individuals were outliers across different neuropsychological variables. We were unable to find any pattern in the data to account for the outliers and no single individual was a consistent outlier across variables. To examine the effect of outlying observations, they were removed from the data and the analyses were repeated. In all analyses of primary variables, the results were the same with or without outliers included. However, in the case of one secondary variable, perseverations, the results changed after removal of outliers (see Results). Because outliers represent observations inconsistent with the statistical analysis, results are presented with outliers removed.

Because multiple analyses were performed, we tested the strength of the findings by conducting separate Bonferroni corrections on the analyses of our 10 primary variables (i.e., these p values were multiplied by 10), our five secondary variables (i.e., these p values were multiplied by 5), and the five measures selected to test for age effects (i.e., these p values were multiplied by 5).

### RESULTS

The means and standard deviations of depressed and control participants on each of the primary and secondary variables are displayed in Table 1, along with the p values corresponding to the group effect from the quantitative analyses reported below. Although outliers were removed from some individual analyses, they were not removed from the entire study. Therefore, the data are displayed in the table without removing outliers. For the purpose of illustration, we included the raw total of correct responses made during CVLT recognition testing. This variable was used to derive the measure of discriminability and, therefore, was not used

 Table 1. Scores of older depressed and control participants on selected measures from the

 California verbal learning test (CVLT) and other neuropsychological tasks

Measure	Depressives $(N = 57)$		Controls ( $N = 30$ )	
	М	( <i>SD</i> )	М	(SD)
Primary CVLT variables				
Total recall [80]	37.67	(11.10)****	49.83	(8.53)
Learning slope	0.91	(0.62)****	1.33	(0.32)
Long-delay recall [16]	7.23	(3.62)***	10.27	(2.63)
Long-delay cued recall [16]	8.20	(3.51)****	11.53	(1.87)
Recognition [44] <sup>a</sup>	37.66	(4.20)	41.07	(2.64)
Discriminability [100]	85.89	(9.08)***	93.33	(6.00)
Verbal retention <sup>b</sup>	128.18	(70.20)	100.69	(16.27)
Semantic clustering	1.60	(0.68)*	2.04	(0.88)
Serial clustering	2.05	(1.38)	2.01	(1.20)
Free recall intrusions	4.33	(4.85)	3.00	(2.63)
Cued recall intrusions	4.42	(4.83)	2.43	(2.72)
Secondary CVLT variables				
% Primary recall	28.79	(7.81)	27.80	(4.89)
% Middle recall	40.18	(10.58)**	42.23	(5.23)
% Recency recall	31.39	(10.87)	27.17	(5.70)
Perseverations	3.98	(3.87)*	5.10	(4.25)
Response bias	0.18	(0.40)*	-0.03	(0.35)
General neuropsychological measures				
Naming [30]	25.16	(4.36)	26.33	(2.77)
F-words	10.42	(4.58)	12.67	(4.35)
Rey Copy [36]	22.00	(8.02)	24.82	(5.35)
WMS-R <sup>c</sup> [140] attention/concentration	83.76	(22.34)**	99.23	(13.62)
Trail Making B (time)	141.87	(93.32)*	85.45	(28.96)

Notes. Numbers in brackets are total possible scores.

<sup>a</sup>This score represents total correct during recognition testing and is used to derive "discriminability."

<sup>b</sup>(Delayed Recall/Immediate Recall)  $\times$  100.

<sup>c</sup>WMS–R = Wechsler Memory Scale–Revised.

p < .05, p < .01, p < .01, p < .001, p < .001, p < .0001.

as another dependent measure in the study. For the purpose of comparison, Table 1 displays the means and standard deviations of the depressives and controls on general neuropsychological measures administered as part of a different study (King et al., 1995). These measures included a 30item test of confrontation naming (Kaplan et al., 1978), generation of "f-words" (Benton & Hamsher, 1978), untimed copy of the Rey–Osterrieth Complex Figure (Corwin & Bylsma, 1993), the Attention/Concentration weighted sum from the Wechsler Memory Scale–Revised (WMS–R; Wechsler, 1987) and Trail Making B Time (Reitan, 1958).

## **Quantitative Features**

Controlling for Vocabulary score, the depressives had significantly lower scores on total recall [F(1,84) = 24.66, p < .0001], learning slope [F(1,83) = 15.36, p < .0002], long-delay cued recall [F(1,84) = 21.62, p < .0001], and discriminability [F(1,82) = 13.34, p < .0005]. Analysis of the long-delay recall data revealed an interaction of Vocabulary Score × Group [F(1,83) = 4.12, p < .05], such that depressives performed worse than controls as Vocabulary score increased. There was a trend toward greater verbal retention in the depressed group [F(1,78) = 3.48, p = .07]. Inspection of the learning slope, discriminability and verbal retention data revealed the existence of outliers who were removed from the analyses without changing the results. (There were two outliers for each of the three variables. All were depressed subjects; none were outliers across analyses.) Repeated analysis of learning slope data without psychotic patients did not change the results [F(1,72) = 9.87, p < .002].

Figure 1 depicts the average recall performance (with corresponding standard errors) of each group across the five learning trials, after a 20-min delay, and after the delayed provision of semantic cues. It can be seen that the depressives started off learning fewer words than controls and acquired new words less efficiently. The depressives maintained their lower level of performance after the delay period and after semantic cues were provided.

### **Qualitative Features**

The depressed participants were less likely to process words according to their semantic properties [F(1,84) = 5.47,

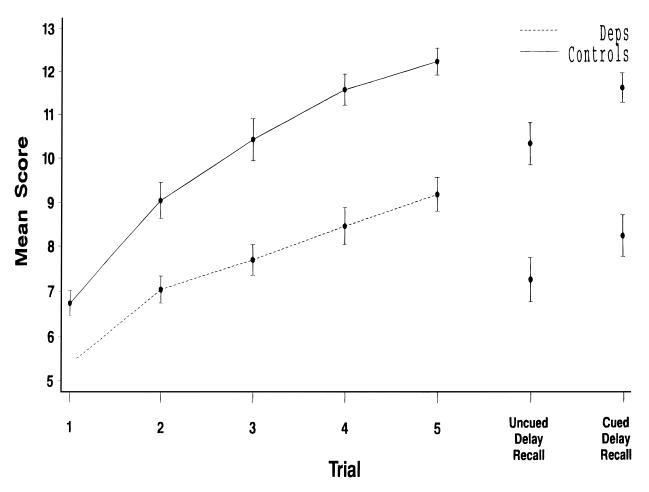


Fig. 1. Mean (standard error) recall performance of depressed and normal control participants across five presentations of the 16-word list and after 20-min delay.

p < .02]. As predicted, depressives and controls did not differ in the use of serial processing [F(1,84) = .04, p = .84], the number of free recall intrusions [F(1,82) = .46, p = .50], or the number of cued recall intrusions [F(1,82) = 2.12, p = .15]. (There were two depressed outliers removed from the analysis of free recall intrusions, and two depressed outliers removed from the analysis of cued recall intrusions. Removal of these outliers did not change the nature of the findings.)

#### **Bonferroni Corrections of Primary Analyses**

Given that 10 primary analyses were performed, we used the stringent Bonferroni correction to test the strength of these findings. All of the findings remained significant except for the interaction of Long-delay Recall × Vocabulary × Group, p = .45, and the difference between depressives and controls in the use of semantic processing, p = .20. Therefore, these findings should be interpreted cautiously. Regarding long-delay recall, it is important to note that lack of an interaction of Group × Vocabulary Score would allow one to interpret the main effect between groups [F(1,84) =14.96, p < .0002], a finding that remained highly significant after Bonferroni correction, p < .002, indicating that depressives performed more poorly on this task.

#### **Analyses of Secondary Measures**

In terms of secondary measures, the depressed participants recalled fewer words from the middle of the list [F(1, 84) =6.64, p < .01]. There was a trend for depressives to recall more words from the end of the list than control participants [F(1,83) = 3.46, p = .07; one depressive outlier removed]. Reanalysis of the recency data without depressives who had previously received ECT did not change the nature of the findings. However, when psychotic depressives were removed from the analysis, there was no longer a trend suggestive of difference between groups in terms of recency recall [F(1,73) = 1.66, p = .20]. There was no difference between groups with respect to primacy recall [F(1,83) =.18, p = .67; one depressive outlier removed]. Depressives made fewer perseverative errors than controls when two depressive outliers were removed [F(1,82) = 4.79, p < .03]. Also, depressives were inclined toward a more liberal response style than control participants [F(1, 84) = 3.98, p <.05]. Using the Bonferroni correction to guard against chance findings in these five secondary analyses, only percent middle recall maintained significance, p < .05.

#### **Analyses of Age Effects**

With respect to the relationships between depression, age, and cognitive functioning, there was an interaction of Age  $\times$  Group on total recall [F(1,81) = 15.00, p < .0002]. As shown in Figure 2, when Vocabulary score was held constant in the analysis, the performance of depressives declined with age whereas the performance of control participants did not change. Pearson product-moment correlations between age and total recall score in the depressed and normal control groups were, respectively, -.69 and -.06.

There was a similar interaction of Age  $\times$  Group on longdelay cued recall [F(1,81) = 7.68, p < .007], such that the performance of depressives declined with age faster than that of control participants (see Figure 3). Pearson productmoment correlations between age and long-delay recall in the depressed and normal control groups were, respectively, -.62 and -.44.

There were parallel negative effects of age on both groups' performance of long-delay free recall [F(1,81) = 12.46, p < .001], and discriminability [F(1,81) = 13.86, p < .001]. There were no effects of age on verbal retention score [F(1,75) = .09, p = .76]. (Two depressive outliers were removed from the verbal retention analysis without changing the results.)

#### **Bonferroni Corrections of Age Analyses**

The interaction between age and participant group remained significant for total recall, p = .001 and long-delay cued recall, p = .035. The parallel effects of age on depressed and control participants' performance of longdelay free recall and discriminability remained significant, p = .005.

#### **Effects of Medical Illness**

Reanalyses of the data controlling for the effect of medical illness did not change the findings reported above. Significant differences between the performance of depressed and control participants remained; CIRS score was not significantly associated with performance on any neuropsychological variable. Similarly, the results regarding age and depression effects were unchanged.

#### **Effects of Age of Onset**

In light of finding parallel or interactive age effects on most verbal learning measures, we performed supplementary ANCOVAs on the data from depressed participants that included age of onset as a covariate, in addition to age. These analyses failed to reveal significant relationships between age of onset and verbal learning. In every case, the analyses confirmed the effects of age as reported above.

## DISCUSSION

The results of this study must be interpreted in the context of its methodological limitations. Most of the depressed subjects were taking psychoactive medication and some had undergone ECT as treatment for past depressive episodes. However, only 1 participant had been treated with ECT within the previous year and history of ECT was not associated with any of the dependent variables except for one

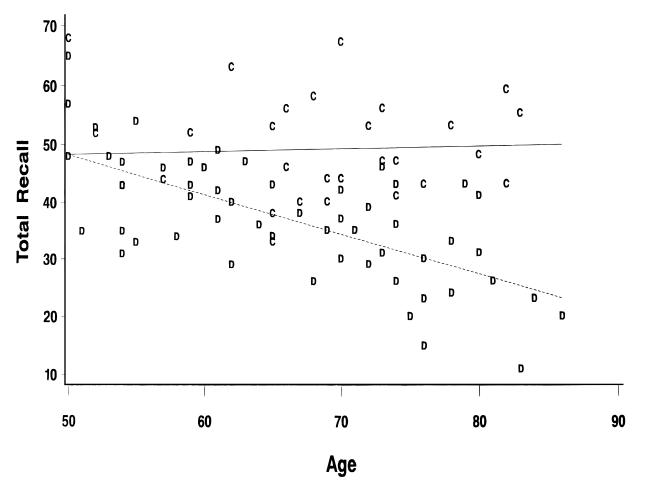


Fig. 2. The effects of age on total recall in depressed and normal control participants.

secondary measure, recency recall. Similarly, the use of medication was not associated with differences in participant characteristics or performance on any of the neuropsychological measures. Eleven of the 62 depressed patients had a diagnosis of major depression with psychotic features. These individuals had somatic or depressive delusions but were not suffering from formal thought disorder that impeded or interfered with their cooperation during testing. Jeste et al. (1996) reported that patients with psychotic depression performed like schizophrenic patients on neuropsychological measures. However, the psychotic patients in our study were no different from other depressed participants in terms of demographic characteristics, general neuropsychological measures (e.g., naming, copying, Trail Making assessed during a previous study) or CVLT measures, except that they had a smaller learning slope and a greater percentage of words recalled from the end of the word list. Results pertaining to learning slope were no different when psychotic patients were excluded from the sample. However, there was no longer a trend for depressives to recall more words from the end of the list when psychotic patients were excluded. For this reason and because history of ECT also affected performance on this variable, the results pertaining to recency recall should be considered inconclusive. It would be useful to conduct future studies with larger proportions of psychotically depressed patients so that their results could be compared to nonpsychotic patients. Our findings tentatively indicate that depressed patients with delusions learn less efficiently than other depressed patients, although they do not appear to perform differently in terms of total immediate recall, delayed recall, retention, or discriminability.

Another limitation was the lack of a nondepressed patient group to control for the effects of hospitalization or nondepressive psychopathology. As discussed previously, we tested for effects of some of the potentially contaminating variables associated with hospitalization, such as previous ECT or the use of psychotropic medications. There was not a discernible effect of any of these factors on the variables of primary interest. However, we did not test the specificity of these findings by comparing the depressed group to a patient group representative of a different disorder. There is evidence of memory impairments associated with schizophrenia and with mixed groups of patients with substance abuse, anxiety, and personality disorders (Burt et al., 1995). More research is needed to clarify the diagnostic specificity of the learning and memory deficits reported here.

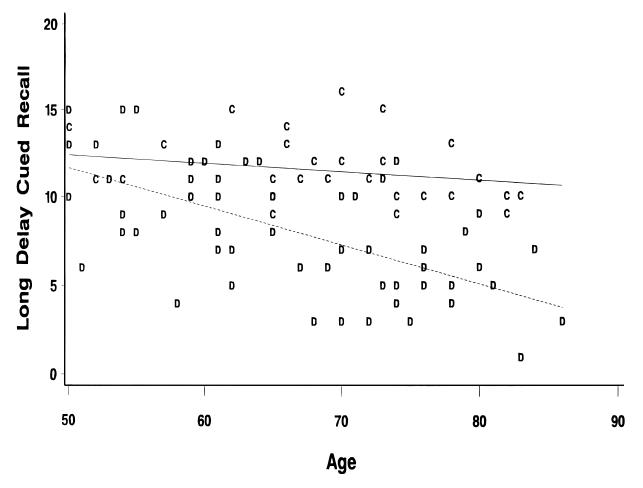


Fig. 3. The effects of age on long-delay cued recall in depressed and normal control participants.

In regard to the primary variables of interest, this study confirmed the presence of significant verbal learning deficits in carefully diagnosed elderly inpatients with major depression. Unlike many previous studies, this investigation used participant groups that were equivalent in terms of average educational level, intellectual level, and gender composition, while at the same time controlling in the analyses for subtle group differences in intelligence. We found the depressed patients to have impaired acquisition of verbal information in terms of both quantity (total recall) and efficiency (learning slope). As well, they recalled less information after a delay period with and without the provision of semantic cues, and they were less able to discriminate previously presented words from new words after the delay period. There was a trend toward higher verbal retention in the depressed group (percent of previously learned information that was recalled after the long delay interval), suggesting that the depressives benefited more than controls from the delay interval itself and/or the cued recall trial that occurred during the delay interval. However, this finding should be interpreted most cautiously as it was no longer significant after stringent correction for Type I error was applied. Taken together, these findings suggest that elderly

depressives suffer from impairments in the encoding or acquisition of information, rather than deficits in the retention or retrieval of information. Our own previous work and that of others has revealed that these deficiencies exist as part of a broad base of neuropsychological impairments associate with severe late-life depression (cf. King et al., 1995; LaRue, 1992; Speedie et al., 1990).

In terms of qualitative aspects of performance, our predictions were only partially confirmed. Our initial analyses revealed that these older depressives were less likely to actively impose a conceptual structure on the material to facilitate learning than were nondepressed controls, although this finding was no longer significant after we applied a stringent correction for Type I error. Previous investigators studying both older and younger patients reported a depressionassociated difficulty in using an active or elaborative strategy to enhance learning (Bäckman & Forsell, 1994; Hart et al., 1987a; Weingartner et al., 1981). As predicted, these older depressives did not make more errors of intrusion than controls, consistent with previous studies contrasting depression and dementia (Hill et al., 1993). In terms of the serial position of words that were learned, the elderly depressives were less likely to learn words from the middle of the list, a finding that suggests a susceptibility to interference effects from words presented previously and subsequently in the list.

Our findings stand in direct contrast to the negative findings of other investigators who studied less severely depressed, typically outpatient or community samples (Bieliauskas, 1993; Lyness et al., 1994; Miller & Lewis, 1977; Niederehe, 1986) or samples that had been screened to exclude patients with severe cognitive deficits (Poon, 1992). Reviewing the literature on depression and cognitive functioning, LaRue (1992) concluded that significant cognitive impairments were for the most part only to be found in severely depressed inpatients. Burt et al. (1995) supported this conclusion specifically with regard to memory functioning. However, Bäckman and Forsell (1994) found verbal recall and recognition deficits in a carefully controlled study of community-dwelling elders with major depression, suggesting that the question of memory functioning in moderately depressed outpatients is as of yet unresolved. With regard to the exclusion of patients with deficits on cognitive screening exams, this procedure confounds dependent and independent variables of uncertain validity. While we agree that efforts must be made to remove participants who suffer diagnosable dementia, we have attempted to make this differentiation on the basis of history, laboratory, and physical examination findings, rather than arbitrarily restricting the variance of the variables being studied.

Considering the relationship between depression and age effects, we found parallel negative effects of age on depressed and control participants' performance of longdelay free recall and discriminability, but interactive effects of age and depression on immediate total recall and longdelay cued recall. Cross-sectional findings such as these must be interpreted cautiously, given that they do not allow for the control of possible cohort effects (Schaie, 1989). As well, the interaction of age and group on some variables could reflect either an increased susceptibility to age effects among depressives, or an increased susceptibility to depression effects among the aged. With these cautions in mind, it is important to consider the findings in the context of previous work. Whereas the parallel effects were consistent with our own and others' previous findings (King et al., 1991, 1993, 1995; Lyness et al., 1994), the interactive effects on total recall and long-delay cued recall were unexpected and seemingly contrary to these previous findings. However, there were significant task differences across studies that could account for these discrepancies. First, some of our previous work used a 10-word list that was subject to ceiling effects in the control group and could have obscured the detection of interactive differences between the groups (King et al., 1991, 1993). As well, the increased list length of the CVLT might have had a selective negative effect on depressive performance as it has been shown that depressives do especially poorly when long memory lists are employed (Henry et al., 1973). These factors would not explain the discrepancy between the present findings and those of Lyness et al.

(1994), who also used the CVLT, although other methodologic variations such as the population sampled (inpatient *vs.* outpatient) and the method of testing for age effects (age analyzed as a continuous *vs.* categorical variable) could account for the different findings.

What might account for the finding of interactive age effects on some tasks and additive, parallel effects on others? Previously, we reported interactive effects of depression and age on tasks that involved complex visuoperceptual processing, such as Trail Making B, the Rey Complex Figure, and the Hooper Test of Visual Organization (King et al., 1995), as well as on Confrontation Naming, a task requiring retrieval of specific information from the semantic memory store (King et al., 1991, 1993, 1995). Given reports of age-associated impairment of retrieval of specific information (words rather than concepts) from the memory store (cf. Bowles & Poon, 1985), and depression-associated impairment on any task requiring more active or effortful processing (cf. Roy-Byrne et al., 1986; Tancer et al., 1989), one might expect synergistic, negative effects of both age and depression on verbal memory tasks that require greater self-initiated search. This pattern fits our findings in that there was an interaction effect of depression and age on total recall, a task that might be considered the most effortful in terms of immediate, uncued processing, but no interaction on discriminability, a delayed recognition task that would be considered the least effortful. However, we did not find interactive effects on long-delay free recall, a task that requires some degree of active spontaneous retrieval and we did find an interactive effect on long-delay cued recall, a more structured task. The inconsistency in this pattern of findings points to the need for further research, especially studies using methodologies that more directly determine the amount of attention or effort involved in a task (e.g., dual-task methods; Tyler et al., 1979).

It is tempting to propose that there is a specific neurobiological substrate for the pattern of memory findings in elderly individuals with major depression. Recently, Massman et al. (1992) reported that a subgroup of moderately depressed patients (M age = 46.1 years) performed like patients with Huntington's disease (HD) in terms of significant recall deficits, inefficient learning across trials, mild recognition deficits, and reduced use of elaborative learning strategies. Caine et al. (1986) reported similar deficits in HD patients, and elsewhere suggested that depressives' deficits may be similar to those of patients with so-called subcortical neurologic disease (Caine, 1986; King & Caine, 1990). The depressives in the current study also performed in a qualitatively similar fashion to HD patients, in terms of their significant recall and recognition deficits, and their tendency to acquire relatively fewer words from the middle of the stimulus word list. However, a broader view of depressives' neuropsychological performance suggests that other processes (and by implication, brain regions) are involved (King & Caine, 1996; Speedie et al., 1990). While our current data point to the possible involvement of the same cortex-to-striatum-to-thalamus circuitry that is disrupted in Huntington's disease, we caution that such a phenotypic inferential model of presumed neurobiological substrates must be viewed tentatively, and perhaps is most useful to establish hypotheses for future studies utilizing functional neuroimaging.

Despite the acknowledged heterogeneity in depressives' cognitive performance illustrated by Massman et al. (1992) and described elsewhere (Caine et al., 1993), efforts to identify a definable subgroup of patients in our sample according to cognitive performance were not successful. We were unable to identify consistent outliers across tasks or to characterize outliers in any consistent fashion in terms of age, education, Vocabulary, or MMSE scores. Moreover, as suggested by an anonymous reviewer of this paper, we attempted to identify a subgroup of more severely cognitively impaired depressed patients who scored 1 standard deviation below the control mean on three key tasks: total recall, discriminability, and verbal retention. Six individuals scored below 1 standard deviation on all three tasks. There were no differences between these 6 depressed patients and the other 51 patients in terms of age, Hamilton, or CIRS score, although they did have significantly higher Vocabulary scores (*M* Vocabulary score = 52.33, SD = 6.41) than the other depressed participants [*M* Vocabulary score = 41.94, *SD* = 13.98; t(11.8) = -3.18, p < .01]. When we reanalyzed the data without these 10 participants, there was no substantive change in the findings. The remaining 51 depressed patients were still impaired relative to controls on all aspects of performance reported previously. Therefore, we concluded that our findings were indeed descriptive of this entire sample of depressed elderly inpatients, rather than reflecting a misleading "average" of different groups.

In light of the present findings and the reports of others reviewed here, there is little question that severely depressed elderly inpatients have significant deficits in a range of explicit verbal learning functions. Moreover, there is strong evidence to suggest a synergy between depressionassociated and age-associated vulnerabilities in elaborative, intentional memory processes. It remains to be resolved whether or not these impairments extend to automatic or implicit processing (see Hartlage et al., 1993; King & Caine, 1996, for reviews) and whether the explicit deficits reported here are present in less severely depressed elderly individuals. Nevertheless, the etiology of the established memory deficits of severely depressed inpatients is worthy of definition. Future studies of this population should move beyond the simple description of such impairments and focus directly on potential neurophysiological and neuroanatomical mechanisms.

## ACKNOWLEDGMENTS

This study was supported by NIMH Grants MH-40381 (E.D.C.) and MH-51201 (Y.C.). The authors thank Kim Finch, Diane Gill, Jack Herrmann, Christine Brower, and Carrie Irvine for assistance with the data collection and/or management and Janet Werkheiser and Priscilla Field for preparation of this manuscript.

## REFERENCES

- Bäckman, L. & Forsell, Y. (1994). Episodic memory functioning in a community-based sample of old adults with major depression: Utilization of cognitive support. *Journal of Abnormal Psychology*, 103, 361–370.
- Benton, M.A. & Hamsher, K. (1978). *Multilingual Aphasia Examination* (Rev. ed.). Iowa City: University of Iowa.
- Bieliauskas, L.A. (1993). Depressed or not depressed? That is the question. *Journal of Clinical and Experimental Neuropsychology*, *15*, 119–134.
- Boone, K.B., Lesser, I., Miller, B., Wohl, M., Berman, N., Lee, A., & Palmer, B. (1994). Cognitive functioning in a mildly to moderately depressed geriatric sample: Relationship to chronological age. *Journal of Neuropsychiatry*, 6, 267–272.
- Bowles, N.L. & Poon, L.W. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology*, 40, 71–77.
- Burt, D.B., Zembar, M.J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin*, 117, 285– 305.
- Caine, E.D. (1986). The neuropsychology of depression: The pseudodementia syndrome. In I. Grant, & K.A. Adams (Eds.), *Neuropsychological assessment and neuropsychiatric disorders* (pp. 221–243). New York: Oxford University Press.
- Caine, E.D., Bamford, K.A., Schiffer, R.B., Shoulson, I., & Levy, S. (1986). A controlled neuropsychological comparison of Huntington's disease and multiple sclerosis. *Archives of Neu*rology, 43, 249–254.
- Caine, E.D., Lyness, J.M., & King, D.A. (1993). Reconsidering depression in the elderly. *American Journal of Geriatric Psychiatry*, *1*, 4–20.
- Cohen, R.M., Weingartner, H., Smallberg, S.A., Pickar, D., & Murphy, D.L. (1982). Effort and cognition in depression. *Archives* of General Psychiatry, 39, 593–597.
- Conwell, Y., Forbes, N.T., Cox, C., & Caine, E.D. (1993). Validation of a measure of physical illness burden at autopsy: The cumulative illness rating scale. *Journal of the American Geriatrics Society*, *41*, 38–41.
- Corwin, J. & Bylsma, F.W. (1993). Translations of excerpts from Andre Rey's *Psychological examination of traumatic encephalopathy* and P.A. Osterrieth's *The Complex Figure Copy Test*. *Neuropsychologist*, 7, 3–15.
- Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *The California Verbal Learning Test: Research Edition*. San Antonio, TX: The Psychological Corporation.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.
- Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, 6, 278–296.
- Hart, R.P., Kwentus, J.A., Hamer, R.M., & Taylor, J.R. (1987a). Selective reminding procedure in depression and dementia. *Psychology and Aging*, 2, 111–115.
- Hart, R.P., Kwentus, J.A., Taylor, J.R., & Harkins, S.W. (1987b). Rate of forgetting in depression and dementia. *Journal of Consulting and Clinical Psychology*, 55, 101–105.
- Hartlage, S., Alloy, L.B., Vazquez, C., & Dykman, B. (1993). Automatic and effortful processing in depression. *Psychological Bulletin*, 113, 247–278.

- Henry, G., Weingartner, H., & Murphy, D.L. (1973). Influence of affective states and psychoactive drugs on verbal learning and memory. *American Journal of Psychiatry*, 130, 966–971.
- Hill, C.D., Stoudemire, A., Morris, R., Martino-Saltzman, D., Markwalter, H.R., & Lewison, B.J. (1993). Similarities and differences in memory deficits in patients with primary dementia and depression-related cognitive dysfunction. *Journal of Neuropsychiatry and Clinical Neurosciences*, 5, 277–282.
- Jeste, D.V., Heaton, S.C., Paulsen, J.S., Ercoli, L., Harris, M.J., & Heaton, R.K. (1996). Clinical and neuropsychological comparison of psychotic depression with nonpsychotic depression and schizophrenia. *American Journal of Psychiatry*, 153, 490– 496.
- Kaplan, E.F., Goodglass, H., & Weintraub, S. (1978). *The Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.
- King, D.A. & Caine, E.D. (1990). Depression. In J.L. Cummings (Ed.), *Subcortical dementia* (pp. 218–230). New York: Oxford University Press.
- King, D.A. & Caine, E.D. (1996). Cognitive impairment in major depression: Beyond the pseudodementia syndrome. In I. Grant & K.A. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric syndromes* (2nd ed., pp. 200–217). New York: Oxford University Press.
- King, D.A., Caine, E.D., & Cox, C. (1993). Influence of depression and age on selected cognitive functions. *Clinical Neuropsychologist*, 7, 443–453.
- King, D.A., Caine, E.D., & Cox, C. (1995). Neuropsychological effects of depression and age in an elderly sample: A confirmatory study. *Neuropsychology*, 9, 399–408.
- King, D.A., Caine, E.D., Conwell, Y., & Cox, C. (1991). The neuropsychology of depression in the elderly: A comparative study with normal aging and Alzheimer's disease. *Journal of Neuropsychiatry and Clinical Neurosciences*, *3*, 163–168.
- LaRue, A. (1989). Patterns of performance in the Fuld Object Memory Evaluation in elderly inpatients with depression or dementia. *Journal of Clinical and Experimental Neuropsychology*, *11*, 409–422.
- LaRue, A. (1992). *Aging and neuropsychological assessment*. New York: Plenum Press.
- LaRue, A., Goodman, S., & Spar, J.E. (1992). Risk factors for memory impairment in geriatric depression. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *5*, 178–184.
- LaRue, A., Swan, G.E., & Carmelli, D. (1995). Cognition and depression in a cohort of aging men: Results from the western collaborative group study. *Psychology and Aging*, 10, 30–33.
- Lawton, M.P. & Brody, E.M. (1969). Assessment of older people: Self-maintenance and instrumental activities of daily living. *Gerontologist*, 9, 179–186.
- Linn, B.S., Linn, M.W., & Gurel, L. (1968). Cumulative Illness Rating Scale. *Journal of the American Geriatric Society*, 16, 622–626.
- Lyness, S.A., Eaton, E.M., & Schneider, L.S. (1994). Cognitive performance in older and middle-aged depressed outpatients and controls. *Journal of Gerontology: Psychological Sciences*, 49, 129–136.
- Massman, P.J., Delis, D.C., Butters, N., Dupont, R.M., & Gillin,

C. (1992). The subcortical dysfunction hypothesis of depression: Neuropsychological validation in a subgroup of patients. *Journal of Clinical and Experimental Neuropsychology*, *14*, 687–706.

- Miller, E. & Lewis, P. (1977). Recognition memory in elderly patients with depression and dementia: A signal detection analysis. *Journal of Abnormal Psychology*, 86, 84–86.
- Niederehe, G. (1986). Depression and memory impairment in the aged. In L.W. Poon (Ed.), *Clinical memory assessment of older adults* (pp. 226–237). Washington, DC: American Psychological Association.
- Niederehe, G. & Camp, C.J. (1985). Signal detection analysis of recognition memory in depressed elderly. *Experimental Aging Research*, *11*, 207–213.
- Poon, L.W. (1992). Toward an understanding of cognitive functioning in geriatric depression. *International Psychogeriatrics*, 4, 241–266.
- Raskin, A. (1986). Partialing out the effects of depression and age on cognitive functions: Experimental data and methodologic issues. In L.W. Poon (Ed.), *Clinical memory assessment of older adults* (pp. 244–256). Washington, DC: American Psychological Association.
- Reitan, R.M. (1958). Validity of the Trail Making Test as an indication of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Reynolds, C.F., III, Lebowitz, B.D., & Kupfer, D.J. (1993). Recommendations for scientific reports on depression in late life. *American Journal of Geriatric Psychiatry*, 1, 231–233.
- Roy-Byrne, P.P., Weingartner, H., Bierer, L.M., Thompson, K., & Post, R.M. (1986). Effortful and automatic cognitive processes in depression. *Archives of General Psychiatry*, 43, 265–267.
- Schaie, K.W. (1989). The hazards of cognitive aging. *Gerontologist*, 29, 484–493.
- Speedie, L.J., Rabins, P.V., & Pearlson, G.D. (1990). Confrontation naming deficit in dementia of depression. *Journal of Neuropsychiatry and Clinical Neurosciences*, 2, 59–63.
- Spitzer, R.L., Williams, J.B.W., & Gibbon, M. (1987). Structured Clinical Interview for DSM–III–R (Rev. ed.). New York: New York State Psychiatric Institute, Biometrics Research Department.
- Tancer, M.E., Brown, T.M., Evans, D.L., Ekstrom, D., Haggerty, J.J., Pedersen C., & Golden, R.N. (1989). Impaired effortful cognition in depression. *Psychiatry Research*, 31, 161–168.
- Tyler, S.W., Hertel, P.T., McCallum, M.C., & Ellis, H.C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 607–617.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–Revised manual*. New York: The Psychological Corporation.
- Wechsler, D. (1987). *Wechsler Memory Scale–Revised manual*. New York: The Psychological Corporation.
- Weingartner, H., Cohen, R.M., Murphy, D.L., Martello, J., & Gerdt, C. (1981). Cognitive processes in depression. Archives of General Psychiatry, 38, 42–47.
- Weingartner, H. & Silberman, E. (1982). Models of cognitive impairment: Cognitive changes in depression. *Psychopharmacol*ogy Bulletin, 18, 27–42.