This is to certify that the

thesis entitled

The Impact of Depressive Symptoms
on Explicit Memory Performance in the Elderly

presented by

Timothy L. Gannon

has been accepted towards fulfillment
of the requirements for

M.A. degree in Psychology

Major professor

Date March 29, 1994

THE IMPACT OF DEPRESSIVE SYMPTOMS ON EXPLICIT MEMORY PERFORMANCE IN THE ELDERLY

By

Timothy Leo Gannon

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1994

ABSTRACT

THE IMPACT OF DEPRESSIVE SYMPTOMS
ON EXPLICIT MEMORY PERFORMANCE
IN THE ELDERLY

By

Timothy Leo Gannon

This study investigated whether depressive symptoms directly affect explicit memory performance. Participants (n = 45) for this study were community dwelling elderly (Mean age = 68) who were offered periodic assessments of their mood and memory, in addition to, a 7-session workshop that focused on relaxation or cognitive strategies for alleviating depression and memory problems. Level of depressive symptoms were assessed with the Hamilton Depression Rating Scale, the Beck Depression Inventory, and the Geriatric Depression Scale while total recall was assessed with the Selective Reminding Test and the Logical Memory Test. These test scores were then combined to produce a depression and memory factor. The average correlation between these factors was -.35 ($p<.05$) within the same test period and -.38 ($p<.05$) across test periods. A relationship between symptoms and memory performance was found. Implications, as well as, limitations of these findings were discussed.

ACKNOWLEDGMENTS

I would like to take a moment to express my appreciation to the members of my thesis committee, as well as, Dr. John Hunter. First, I would like to thank Dr. Ralph Levine, who served as my statistical advisor. Dr. Levine spent an exceptional amount of time and patience broadening my understanding of the theory and mechanics of the statistical procedures used in this study. In addition, I would also like to thank Dr. Hunter, my former statistics mentor, whose statistical software programs, expertise, and availability also helped make this study possible. Next, I would like to thank Dr. John Hurley for his insightful comments on my thesis proposal and final manuscript. Dr. Hurley put considerable effort into increasing my understanding of what is required for a "journal quality" manuscript. Finally, I would like to thank Dr. Norman Abeles, my mentor and committee chairperson. Through the effort of Dr. Abeles, I have always had funding to ease the financial burden of supporting a family while in graduate school. In addition, Dr. Abeles has always been readily accessible for consultations regarding clients, research, and coursework issues. It has also been through his ongoing efforts that I have obtained valuable experience in memory and mood research, neuropsychological and learning characteristic assessments, and specific training in clinical neuropsychology. I am indebted to each of you gentlemen.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D)

LIST OF TABLES

INTRODUCTION

Geriatric memory problems are one of the primary reasons for family referrals (Heath, Grant, Kamps, & Margolin, 1991). Craik (1991) reported that age-related declines in memory ability have been observed in explicit memory, working memory, episodic memory, semantic memory, and prospective memory. Hypotheses that have stated to account for these declines include differences in the ability to self-initiate encoding and retrieval processes, fewer processing resources, additional attentional demands competing for limited attentional resources, less mental energy, and brain structures differentially susceptible to the aging process (Craik, 1991). Lowenthal et al. (1967) observed that self-reports of declining memory increased from 31 percent for those 60 to 64 years of age to 65 percent for those over 75 years of age. However, since complaints about memory deficits are not always correlated with actual memory performance (N = 153; r = .05; Kahn, Zarit, Hilbert, & Niederehe, 1975) and memory complaints have been positively associated with depression (N = 67, r = .25, West, Boatwright, & Schleser, 1984; N = 120, r = .52, Shelton & Parsons, 1987; N = 144, r = .49, Tun, Perlmutter, Russo, & Nathan, 1987), an individual's memory complaints may not always be symptomatic of memory impairment.

Due to the essential role that memory plays in learning and functioning in day-to-day life, memory impairment can have a devastating impact on the individual, family, and society (e.g., health care system). Presently, researchers are attempting to operationalize the characteristics of age-related memory decline further, as well as, develop effective treatment regimens. For example, the National Institute of Mental Health (NIMH) has

introduced the term age-associated memory impairment (AAMI; Crook, Bartus, Ferris, Whitehouse, Cohen, & Gershon, 1986) to characterize healthy individuals who are at least 50 years of age who have both objective declines in memory (i.e., ≥1 SD below young adult norms), as well as, subjective memory complaints (Crook, 1989). So far, research on AAMI has focused on improving its inclusionary criteria and terminology (Smith, Ivnik, Petersen, Malec, Kokmen, & Tangalos, 1991). For example, Blackford and LaRue (1989) have suggested that AAMI be further subdivided into age-associated memory decline (i.e., ≥1 SD below young adult norms on 25% of the tests administered) and late-life forgetfulness (i.e., ≥1 SD below young adult norms on 50% or more of the tests administered).

Memory Training Studies

Non-invasive treatment regimens for age-related memory decline have centered on teaching the elderly memory training strategies. The typical paradigm used to assess their effectiveness consists of (a) assessment on a memory task (e.g., word list, prose recall), (b) instruction in one or more memory strategies (e.g., method of loci) across one or multiple sessions, and (c) re-testing on an equivalent form of the original memory task. Across 14 explicit memory training studies conducted with the elderly that used word list tasks and/or prose recall to assess memory ability, all but one (Yesavage, Rose, & Spiegel, 1982) reported significant improvements in total recall when compared to pre-training scores or to a control group (Hulicka & Grossman, 1967; Robertson-Tchabo, Hausman, & Arenberg, 1976; Schmitt, Murphy, & Sanders, 1981; Zarit, Cole, & Guider, 1981; Schaffer & Poon, 1982; Yesavage, Rose, & Spiegel, 1982; Rose & Yesavage, 1983; Yesavage & Rose, 1983; Anschutz, Camp, Markley, & Kramer, 1985; Scogin, Storandt, & Lott, 1985; Johnston & Gueldner, 1989; Rebok & Balcerak, 1989; Hill, Storandt, & Simeone, 1990; Yesavage, Sheikh, Friedman, & Tanke, 1990). These training studies used a wide range of techniques, with mnemonics (e.g., method of loci) being the most common. Other techniques included word associations (i.e., mediators), affective

judgments, organizational relationships, relaxation training, concentration training, and group discussions on memory issues (e.g., difference between vital and unnecessary remembering; memory loss experiences). The designs used in some of these studies enabled the authors to demonstrate that improvement in recall was not a result of practice effects (Robertson-Tchabo et al. 1976; Yesavage & Rose, 1983; Hill et al. 1990), motivational influences (Hill et al. 1990), or rehearsal length (Schmitt et al. 1981).

Although these studies provide evidence that improvements in explicit memory are possible, their results were based on changes in group means. Schaffer and Poon (1982) point out that, at the individual level, effects of memory training can be quite variable. In order for the rehabilitation psychologist to effectively implement memory training interventions among the elderly, factors that impede memory training need to be identified. Johnson and Magaro (1987) have suggested that one such factor may be the degree of depressive symptomatology. They have hypothesized that the degree of memory impairment in an individual may be partly a result of the severity of the individual's depressive symptomatology (e.g., greater depressive symptomatology results in greater memory deficits). Here, the severity of depressive symptoms denotes a syndrome rather than a nosological category. For example, depression as a syndrome indicates a group of symptoms (e.g., depressed mood, diminished interest in activities, significant weight loss or gain, fatigue) that may be primary to a diagnosis of Major Depression (i.e., a nosological category) or be secondary to another psychopathological disorder (e.g., Schizophrenia). These disorders (e.g., Major Depression, Schizophrenia) are examples of nosological categories that include the depressive syndrome as part of their diagnostic criteria (DSM-III-R, 1987).

Meta-Analysis

In order to examine the hypothesis of Johnson and Magaro (1987), a meta-analysis was first conducted (Hunter & Schmidt, 1990). The criteria used for this procedure are described in detail in Appendix A. A meta-analysis was chosen over a narrative review

due to the inherent biases in summarizing the results of a group of studies in a narrative format (Smith, Glass, & Miller, 1980). A search of the literature resulted in 9 studies from which 13 effect-sizes were calculated (Gibson, 1981; Zarit, 1982, O'Hara, Hinrichs, Wallace, Lemke, & Kohout, 1986; Hart, Kwentus, Taylor, & Hamer, 1987; Hart, Kwentus, Taylor, & Harkins, 1987; Poitrenaud, Moy, Girousse, Wolmark, & Piette, 1989; Feehan, Knight, & Partridge, 1991; King, Caine, Conwell, & Cox, 1991; Lichtenberg, Manning, & Turkheimer, 1992). The additional effect-sizes were due to the inclusion of multiple groups or multiples measures of depression and memory within some of the studies. With the exception of Zarit (1982) and Lichtenberg et al. (1992), in which subjects were not assessed for depression (i.e., nosological category), and one group in O'Hara et al. (1986) that did not meet the criteria for depression, the remaining studies used individuals who met the criteria for depression as defined in the Diagnostic and Statistical Manual of Mental Disorders III, (DSM-III; American Psychiatric Association, 1980), its revision (DSM-III-R; American Psychiatric Association, 1987), or the Research and Diagnostic Criteria (RDC; Spitzer, Endicott, & Robins, 1978). Other than Zarit (1982), who did not use a control group, each of these studies compared groups with elevated depressive symptoms with groups of healthy elderly free of psychiatric diagnoses and elevated levels of depressive symptoms on a word list or prose recall task.

For each of these studies, "d," an effect-size statistic and "r," the point-biserial correlation, were computed (Hunter & Schmidt, 1990; pp. 272-273). The d statistic expresses the difference between the depressed and non-depressed groups in standard deviation units while the point biserial correlation describes the amount of relationship that exists between the group (i.e., depressed versus control) and the level of explicit memory. For more detailed characterizations of these measures and the depression and memory tests used in these studies, please consult Appendix A. Due to the small number of studies, the effects of moderator variables could not be assessed. Hence, the overall findings are limited to the weighted average and weighted standard deviation for the d and

$\underline{r}$ statistics (Table 1). As can be seen, the point biserial correlations across each of these studies suggest a negative relationship ($\underline{r}_{\underline{avg}}$ = -.360) between group and explicit memory ability. In other words, greater memory ability was associated with the healthy control groups while the depressed groups had relatively poorer explicit memory. This finding does appear to be congruent with the relationship hypothesized by Johnson and Magaro (1987). However, since the majority of these studies used depressed patients, another interpretation of these findings is that memory impairment is related to some other facet of depression (i.e., nosological category) and that the relationship with the severity of depressive symptoms is merely spurious. For example, O'Hara et al.'s (1986) group that did not meet the criteria for depression yet reported high levels of depressive symptoms had a very low correlation with memory recall ($\underline{r}$ = -.111). Such a low correlation does not suggest a relationship between depressive symptom severity and memory impairment. In addition, these nine studies did not address the issue of causality which is hypothesized to exist (Johnson & Magaro, 1987). Specifically, depressive symptoms are hypothesized to directly impair memory ability. Hence, the purpose of this study is to determine if the severity of depressive symptoms directly affects explicit memory performance as measured through total recall scores on word list and prose recall tasks. Based on the hypothesis of Johnson and Magaro (1987), path coefficients from the depression factor to the memory factor should be the largest.

Table 1 Meta-Analysis Findings

Author(s) Sample	Depressed n	Depressed M	Depressed SD	Control n	Control M	Control SD	d	r
1. Gibson (1981)								
I & O	20	25.65	6.16	20	34.70	6.63	-1.41	-.59
2. Zarit (1982)								
CD?				79			-.69	-.33
CD?				79			-.41	-.21
CD?				79			-.64	-.31
CD?				79			-.47	-.23
3. O'Hara et al. (1986)								
CD	22	5.00	2.09	25	6.08	1.85	-.55	-.27
C	23	5.57	2.78	25	6.08	1.85	-.22	-.11
4. Hart, Kwentus, Taylor, and Hamer (1987)								
O	14	70.70	16.50	16	83.60	12.40	-.89	-.42
5. Hart, Kwentus, Taylor, and Harkins (1987)								
?	10	12.20	1.50	14	17.70	1.40	-3.81	-.90
6. Poitrenaud et al. (1989)								
I	24	14.29	3.67	33	19.03	4.31	-1.17	-.52
7. Feehan et al. (1991)								
O	10	5.50	2.43	10	7.20	2.46	-.70	-.34
8. King et al. (1991)								
I	23	6.40	1.50	23	7.50	1.40	-.76	-.36
9. Lichtenberg et al. (1992)								
CD?	16	7.30	3.50	19	11.50	2.20	-1.63	-.71
				Weighted Mean:			-.84	-.36
				Weighted SD:			-.68	-.18

?: Patients diagnosed depressed, did not specify if inpatient or outpatient
C: Community Dwelling, did not meet criteria for depression
CD: Community Dwelling, met depression criteria in epidemiological study, never treated
CD?: Community Dwelling, were not assessed for depression
I/O: Inpatient/Outpatient diagnosed depressed

METHOD

Participants

The participants were a subset of community-dwelling elderly (N=208) recruited through newspaper and radio ads, circulars, and presentations at local senior citizen centers. Each participant was offered multiple assessments of their mood and memory, as well as, a 7-session workshop that taught primarily relaxation or cognitive strategies for alleviating depression and memory problems. Forty-five participants (i.e., "stayers") who agreed to be assessed across three occasions and had a Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) score of 24 or better were included in the present study. These participants were composed of both men (n = 13) and women (n = 32) whose mean age was 68 (SD = 10y, Range = 52 to 92y) and who had, on average, 13 years of schooling (M = 13y, SD = 3y, Range = 5 to 20y).

Test Instruments for Severity of Depression

Since multiple measures of depression could more fully capture any depressive symptomatology being experienced by the participants (Kendall et al., 1987), three instruments were chosen. The first instrument was an observer-rating scale, the Hamilton Depression Rating Scale (Hamilton, 1960). The remaining two, the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and the Geriatric Depression Scale (Yesavage et al., 1983), were self-report measures. Together, these three measures more fully address the DSM-III-R criteria for depression and criteria characteristic of depression in the elderly population (Weiss, Nagel, & Aronson, 1986).

The Hamilton Rating Scale for Depression (HAM), is a 17-item scale that was originally designed to measure the severity of depression in individuals already diagnosed depressed. However, since its inception, it has been extensively used to assess depressive symptoms in other populations as well (Hedlund & Vieweg, 1979). Test-retest stability coefficients for the HAM ranged from .61 (Mean interval = 15 days; Lyons, Strain, Hammer, Ackerman, & Fulop, 1989) to .70 (3 week interval; Maier, 1990) while Cronbach's (1951) alpha was .76 (Rehm & O'Hara, 1985). Sensitivity to change, which is a scale's ability to detect changes (e.g., in depressive symptoms) in individuals across time, was .65 for the HAM (Maier, 1990). Correlations of the HAM and psychiatrist's ratings of global severity of depressed patients ranged from a mean of .88 (range = .84 to .90) for three studies reviewed by Hedlund and Vieweg (1979) to .67 (Maier, 1990). Stukenberg, Dura, and Kiecolt-Glaser (1990) observed that the area under the ROC (i.e., Receiver Operating Characteristics) curve, which represents the probability of a measure correctly categorizing subjects (e.g., depressed versus nondepressed), was .85 for the HAM. In a sample of psychiatric inpatients (Lichtenberg, Steiner, Marcopulos, & Tabscott, 1992), the HAM's sensitivity and specificity was 9% and 92%, respectively. These indices were determined by comparing those subjects who fell above and below a cut-off criterion. Sensitivity is the proportion of individuals correctly diagnosed depressed (i.e., proportion of true positives) while specificity is the proportion of individuals correctly diagnosed affectively healthy (i.e., proportion of true negatives; Masur et al. 1989). The authors mentioned that the low sensitivity of 9% is at odds with what is reported in the literature and suggested that it may have been due to their study's large number of demented subjects.

Concurrent validity for the HAM was determined based on its correlation with other measures of depression. The HAM had an average correlation of .77 (range = .73 to .80) with the Beck Depression Inventory across two studies that examined non psychiatric samples (Beck, Steer, & Garbin, 1988). This correlation with the Beck

Depression Inventory was lower, however, in a sample of mixed psychiatric and medical inpatients (r = .52; Fitzgibbon, Cella, & Sweeney, 1988) and across seven studies assessing psychiatric inpatients (r_{avg} = .58; range = .31 to .82; Hedlund & Vieweg, 1979). The HAM also had a correlation of .60 with the Brief Symptom Inventory Depression scale (Derogatis & Spencer, 1982) in a sample of community-dwelling elderly (Stukenberg, Dura, & Kiecolt-Glaser, 1990). In a review of seven factor analyses, Hedlund and Vieweg (1979) concluded that the first of two factors generally evident across the studies appeared to measure the severity of depressive symptoms. However, in a more recent study, O'Brien and Glaudin (1988) observed four factors that they labeled somatic complaints, anorexia, sleep disturbance, and agitation/retardation. The 23-item version of the HAM was used for the present study. Scores on this version can range from 0 to 75. Zero denotes a lack of depressive symptoms while 75 suggests severe depression.

The Beck Depression Inventory (BDI) is a 21-item, multiple-choice inventory that rates the intensity of depressive symptoms, especially those of a somatic nature. Based on a psychometric review of the BDI (Beck, Steer, and Garbin, 1988), Cronbach's (1951) alpha had a mean of .81 (range = .73 to .92) based on 15 samples of non psychiatric subjects while test-retest stability coefficients had a mean of .76 (1 hour to 4 month interval; range = .60 to .90) based on 5 samples of non psychiatric subjects. Correlations of the BDI and clinical ratings of depression for three samples of non psychiatric patients, had a mean of .60 (range = .55 to .73). The concurrent validity of the BDI, when correlated with the Geriatric Depression Scale, was .85 with medical outpatients (Norris, Gallagher, Wilson, & Winograd, 1987). The value with the HAM across two non psychiatric samples, however, was slightly lower (r_{avg} = .77, range = .73 to .80; Beck et al., 1988). In a review by Beck et al. (1988), the authors concluded that the most recent factor studies suggest one depressive factor which measures depressive severity. Scores

on the BDI can range form 0, which indicates that depressive statements were not endorsed, to a maximum value of 36 which indicates severe depression.

The Geriatric Depression Scale (GDS) is composed of 30 yes-no questions that primarily examine mood and psychological symptoms. Test-retest stability coefficients of .98 (Mean interval = 15 days; Lyons, Strain, Hammer, Ackerman, & Fulop, 1989) and .85 (1 month interval; Parmalee, Lawton, & Katz, 1989) have been reported. The mean value for Cronbach's (1951) alpha was .91 based on studies by Yesavage et al. (r = .94; 1983), Parmalee et al. (r = .91; 1989), and Salamero and Marcos (r = .87, 1992). A study by Harper, Kotik-Harper, and Kirby (1990) yielded a sensitivity of 85% for the GDS in a sample of depressed elderly, while in a study by Koenig, Meador, Cohen, and Blazer (1988), a sensitivity of 92% and a specificity of 89% were obtained. Concurrent validities of .62 and .81 were obtained when the GDS was correlated with the HAM at two different assessment periods with medical inpatients (Lyons, Strain, Hammer, Ackerman, & Fulop, 1989). This value was slightly higher (r = .83) when medical outpatients were assessed (Norris, Gallagher, Wilson, & Winograd, 1987). Correlations were also slightly higher between the GDS and BDI (r = .85; Norris et al., 1987). Two factor analyses of the GDS found six (Parmalee et al., 1989) and nine-factor (Salamero & Marcos, 1992) solutions although both studies mentioned that the GDS was unidimensional (i.e., general depression factor) based on Cattell's (1966) scree criterion. Scores on the GDS can range from a score of 0, which indicates that no depressive statements were endorsed, to a maximum value of 30 which indicates severe depression.

Test Instruments for Explicit Memory

Logical Memory (LM) Form I of the Wechsler Memory Scale (Wechsler & Stone, 1945) served as a measure of prose recall. It has an average test-retest stability of .79 (time interval not mentioned; Bowden & Bell, 1992) and was found to be significantly correlated (r = -.375) with the Halstead-Reitan Average Impairment Scale in a sample of brain-damaged subjects (Russell, 1975). LM is one of four tests (i.e., Storandt Brief

Dementia Battery) used in a discriminant function for detecting Alzheimer's disease (Storandt, Botwinick, Danziger, Berg, and Hughes, 1984). This function had a sensitivity of 98% on a cross-validation sample. Across several factor analyses reviewed in Erickson and Scott (1977) and Prigatano (1978), LM consistently loaded on the memory factor. Total recall scores for LM can range from 0 to 22. The higher the score, the greater number of information units are recalled by the participant.

The Selective Reminding Test (SRT, Buschke, 1973) served as the word list task with lists of equivalent words (Masur et al., 1989) being used for each of the three assessments (Table 2). Each list contained 12 unrelated words (Hannay & Levin, 1985). Test-retest stability has been found to range from .73 (6 month interval; Ruff, Quayhagen, & Light, 1988) to .89 (2 hour interval; Masur et al., 1989). In a study by Masur et al. (1989), the SRT displayed a sensitivity and specificity in the detection of patients with mild Alzheimer's disease of 86 and 99%, respectively. However, in a second study (Masur, Fuld, Blau, Crystal, & Aronson,1990) with patients diagnosed with dementia based on DSM-III-R criteria, the SRT obtained a sensitivity of 47% and a specificity of 86%. Total recall scores were based on the sum of words recalled across 6 trials and could range from 0 to a maximum value of 72. The higher the number, the more words were recalled by the participant across the trials. This was chosen as a measure of long term explicit memory over other word list tasks since it controls for the influence of short term memory (Buschke, 1973; Buschke & Fuld, 1974).

To provide a measure of subjective memory complaints, two self-report items of memory problems were created. The first item (PROB) asked, "Are you currently experiencing any problems with your memory?" while the second item (#PROB) stated, "If (you answered) yes (to the 1st item), please tell us what kinds of problems you are experiencing. Remembering names ( ), faces ( ), date, month, or year ( ), appointments ( ), where you put things ( ), what you went into a room to do/get ( ), to take medication

( ), other ( )." The stability coefficient was .77 for the first item and .66 for the second item. These coefficients were based on 51 control subjects who took the first two assessments as part of the larger study. #PROB could range from 0 problems to a maximum value of 8 problems.

Table 2 SRT Word Lists

Test 1	Test 2	Test 3
Shine	Bowl	Throw
Disagree	Passion	Lily
Fat	Dawn	Film
Wealthy	Judgment	Discreet
Drunk	Grant	Loft
Pin	Bee	Beef
Grass	Plane	Street
Moon	County	Helmet
Prepare	Choice	Snake
Prize	Seed	Dug
Duck	Wool	Pack
Leaf	Meal	Tin

Test Instrument for Mental Status

The Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) is a 30-item form that served as a cognitive impairment (i.e., mental status) examination. This screening test provides a brief measure of orientation, registration, attention and calculation, recall, and language. Based on a review of the MMSE by Tombaugh and McIntyre (1992), the mean coefficient alpha across five studies was .72 (range = .54 to .96) while the mean stability coefficient across 25 samples was .82 (range = .80 to .97; time interval 1 day to 2 months). The sensitivity of the MMSE across 27

samples that examined demented subjects had a mean value of .79 (range = .20 to 1.00) while the specificity across 21 samples had a mean value of .83 (range = .46 to 1.00). Construct validity, based on 14 studies between the MMSE and other cognitive status measures, ranged from .70 to .90. The MMSE was included for two reasons. First, to identify any difference in cognitive status that may have been present between those participants who only agreed to a single testing session and those that took all three. Second, since cognitive impairment (i.e., a low MMSE) is also associated with poorer memory, it was used as a screening measure for the stayers to avoid any confounds. The 23/24 cutoff recommended by Folstein et al. (1975) was used for the present study. Scores could potentially range from a low of 0 indicating severe impairment to a maximum value of 30 indicating an unimpaired mental status.

Procedure

Subjects were assessed on the three depression scales, two objective memory scales, the 2-item subjective memory problem scale, and the test of cognitive status on three occasions at approximately three month intervals. Testing was done by clinicians enrolled in Michigan State University's clinical psychology program. Approximately seven hours of training were given to each of the clinicians in administering and scoring the tests. During their training phase, each clinician observed an assessment by another clinician with at least 1 year experience and was then subsequently observed by the same clinician during their first testing session. Furthermore, all tests were re-scored by the author to ensure accuracy. Whenever, differences in total score existed, totals were retabulated until the same score was produced on two successive occasions.

RESULTS

Data analyses were subdivided into four components. First, the psychometric properties of the depressive and memory indices were assessed. Second, the "stayers" were compared to the "decliners." Third, stayers who participated in the relaxation or memory workshops (n = 27) were compared to the stayers who did not participate (n = 18) in order to determine the presence of any treatment effects. Fourth, the stayers' levels of depressive symptoms and amount of memory ability were evaluated for the presence of any causal relationships.

Psychometric Properties

First, each of the stayer's total scores from the depression and memory measures were transformed into z scores and summed to produce a depression factor (DEP) and a memory factor (MEM). Construct validity was then empirically tested with confirmatory factor analyses (CFA). An initial CFA with the Hamilton Depression Scale (HAM) indicated that 4 test items had zero variance and 6 others had negative loadings on a factor comprising all the HAM test items.. It was assumed that the test items with negative factor loadings were measuring something different from the construct of depression being assessed by the other test items. Consequently, all ten of these test items were removed so the HAM total score did not reflect these items in all subsequent analyses. Next, CFA's were performed to test how each depression and memory test loaded on their respective factor. The results of these analyses can be found in Table 3.

Next, stability was assessed using Cronbach's (1951) alpha. The resulting stability coefficients can be found in Table 4. In summary, it appears that there was adequate

validity and reliability for the measures used in this analysis. A more detailed description of these psychometric analyses can be found in Appendix B, as well as, the stayers' raw scores for the depression and memory measures.

Table 3 DEP and MEM Test-Factor Correlations

	DEP1*	DEP2	DEP3
HAM	.71	.79	.83
BDI	.79	.72	.78
GDS	.75	.96	.71
	MEM1*	MEM2	MEM3
LM	.73	.74	.71
SRT	.73	.74	.71

* the number following the DEP and MEM indicates the test period from which the factors were calculated

Table 4 Coefficient Alphas for the DEP and MEM Factors

	Time 1	Time 2	Time 3
DEP	.79	.86	.82
MEM	.69	.71	.67

Stayers versus the Decliners

Since there was a large number of participants who were involved in the first mood and memory assessment but declined to take all three assessments, these analyses were

conducted to determine if any differences existed between the stayers and decliners that could affect the results generalizability. The following variables were examined for differences: general characteristics (i.e., age, gender, education), severity of depressive symptoms (i.e., HAM, BDI, GDS, DEP), explicit memory (i.e., LM, SRT, MEM), subjective memory complaints (i.e., PROB, #PROB), cognitive mental status (i.e., MMSE), and if they declined or chose to participate in one of the depression/memory workshops (i.e., C/W). Those who took part in the relaxation (n = 11) or memory workshop (n = 16) were combined due to small samples and since results from the larger study (N. Abeles, personal communication, April 15, 1992) indicated that the treatment effect was not significantly different between these two groups. In addition to these variables, a SR MEM factor, which was based on the summed z-scores for the two subjective memory complaint measures, was calculated. For this variable, as an individual's number of subjective memory complaints increases, their z score will become progressively more positive.

The means and standard deviations for each of these variables within the two groups (i.e., "stayers" and "decliners") were computed and are listed in Appendix C. From these, d statistics were calculated for each variable and then transformed to point biserial correlations (Appendix C). Stayers were found to be younger, have a higher ratio of men to women, have higher recall scores on the word list task, have more individuals with subjective feelings of memory problems and number of problems, be more cognitively unimpaired, and have a higher proportion of individuals who participated in the workshops. Another measure of variability between groups, the standard deviation ratio (V), was also calculated (Hunter, 1993b) and can be found in Appendix C. This measure compares the standard deviation of the above variables between the stayers and decliners. Results of this analysis indicated that the mental status of the decliners had greater variability then was found in the stayers.

Workshop versus Controls

Initially, control and workshop groups were analyzed separately for the presence of any changes in their DEP and MEM factors between time 1 and 2, time 2 and 3, and time 1 and 3 (Appendix D). These changes between test periods were then expressed as treatment correlations that had been corrected for attenuation (Hunter, 1994). Further information on correcting for attenuation can be found in Appendix J. Analysis of the control condition indicated that there was a high probability that depressive symptoms increased in severity between test period 2 and 3 (Odds ratio = 4.38). Calculations also indicated a high probability that memory recall increased between time 2 and 3 (Odds ratio = 7.87) and between time 1 and 3 (Odds ratio = 12.48). Analysis of the workshop condition indicated a high probability that depressive symptoms became less severe between time 1 and 2 (Odds ratio = 34.58). Results also indicated that it was highly probable that memory recall increased between time 1 and 2 (Odds ratio = 15.69) and time 1 and 3 (Odds ratio = 13.24). These analyses also indicated that there was no significant interaction for either the control or workshop condition ($p>.05$). In other words, once measurement error was controlled for, each of the participants within a condition (i.e., workshop or control) changed approximately the same amount in terms of depressive symptoms and memory recall between the test periods listed above. Ninety five percent confidence intervals with worst case, best case, and odds ratio that provided the basis for the above results are also tabulated in Appendix D for both the control and workshop conditions.

Next, treatment correlations for the control and workshop conditions were compared during each of the time periods. Gain scores (i.e., change scores) were calculated based on the differences in treatment correlations between the control and workshop conditions (Appendix D). Results of this analysis indicated that it is highly probable that there was a difference in the DEP treatment effect between the control and workshop conditions when time 1 was compared to time 2 (Odds ratio = 3.67). In this

case, it was due to the workshop group evincing a greater decrease in self-reported depressive symptoms during this time frame than was found in the control group.

Causality

First, a stability analysis (Heise, 1969) was conducted on the DEP and MEM factors over the three time periods to assess their temporal stability independent of measurement error. A value of 1.00 would indicated perfect stability across time. As can be seen in Table 5, the values for the DEP and MEM factors were very stable. In other words, the participants' self-reported depression and explicit memory ability changed very little across time. These changes were so minimal that any causal relationship that may have existed between the two factors could not be determined.

Table 5 Stability Analysis of the DEP and MEM Factors

	DEP Factor	MEM Factor
Stability between Time 1 and 2	.95	.99
Stability between Time 2 and 3	.96	.97
Stability between Time 1 and 3	.92	.96

Table 6 displays the correlations between the DEP and MEM factors. The time periods are designated by the number following the DEP and MEM factor. For example, DEP1 indicates the DEP factor at time 1. In each case, there is a negative correlation between the DEP and MEM factors. In addition, these negative correlations are slightly larger across time (r_{avg} = -.38) then within the same time period (r_{avg} = -.35). These correlations were corrected for attenuation. A path diagram of the DEP and MEM factors can be found in Appendix E.

Table 6 Correlations Between DEP and MEM Factors

	DEP1	DEP2	DEP3	MEM1	MEM2	MEM3
DEP1	-					
DEP2	1.00	-				
DEP3	1.00	1.00	-			
MEM1	-.38	-.46	-.39	-		
MEM2	-.44	-.35	-.33	1.00	-	
MEM3	-.46	-.38	-.32	1.00	1.00	-

DISCUSSION

The analyses for this study were subdivided into four components. First, the psychometric properties of the depressive and memory indices were determined. Second, the 45 "stayers" were compared to the 163 "decliners." Third, the 27 stayers who participated in the relaxation or memory workshops were compared to the 18 stayers who did not participate in order to assess the presence of any treatment effects. Fourth, depressive and memory indices among the stayers were examined for the existence of any causal relationships. This same format will be used to discuss the findings of these analyses.

Psychometric Properties

Results indicated that there was reasonable evidence of construct validity and reliability of the selected measures of depression and memory after test items were removed from the Hamilton Depression Rating Scale (HAM) due to a lack of variance and negative factor loadings. One possibility for the problematic test items encountered with the HAM may be due to its use on a population for which it was not designed for. Even though the HAM has been extensively used with other populations (Hedlund & Vieweg, 1979), it was originally designed for assessing the severity of depression in individuals already diagnosed depressed. That population is quite different from the present sample which reported relatively few depressive symptoms. Another possibility is that the clinicians did not give enough time for the HAM interview in order to bring out responses to each of the test items (Hamilton, 1960). However, each of the clinicians in this study had at least one year experience in supervised geropsychiatric assessments so this

explanation seems unlikely. One factor which apparently contributed was the 23-item version of the HAM used in this study. This version contains the original 17 test items plus an additional 6 test items. Within this study, three of these additional test items had zero variance while another 3 had negative factor loadings. It may be that the 23-item version was psychometrically unsound from its inception since the psychometric properties reported in the literature for the HAM appear to only be based on the original 17-item test.

Stayers versus Decliners

Although the stayers and decliners had a high probability of being different on several variables, it is doubtful whether these differences are clinically meaningful. For example, even though both groups differed on mean recall scores (i.e., Selective Reminding Test) and mental status scores (i.e., Mini-Mental State Examination), both group's mean scores are in the healthy range based on normative data (Masur et al., 1989; Tombauch & McIntyre, 1992). In addition, the difference between the average number of subjective memory problems being perceived between the two groups was less than half a problem. This also does not suggest a meaningful difference in perceived memory problems.

Workshop versus Controls

The first surprising finding was that, after correcting for measurement error, all of the individuals in the workshop changed approximately equal amounts in memory ability across the three test periods. Although others have reported high levels of individual variability in memory training treatment effects (Schaffer & Poon, 1982; Yesavage et al., 1990), no mention was made in these studies of correcting for the effects of measurement error. The unfortunate consequence of this practice is that misconceptions, such as high individual variability being typical among memory training programs, results. When measurement error is corrected for, the overall efficacy of these programs appears to be much better.

The second finding was that the workshop condition only had a minimal treatment effect. Based on the number of studies cited in the introduction that observed significant treatment effects, this came as a surprise. However, Schaffer and Poon (1982) contend that this is not atypical and instead hypothesize that their may be many unpublished memory training studies with relatively neutral or negative results. Looking back, we might have achieved greater changes in memory ability had we included memory tests more relevant to the memory strategies that were taught in the workshop. For example, one strategy that the elderly were taught was to group like-words (e.g., all meats or all vegetables) in order to increase their level of recall. If a shopping list task had been used instead of the Selective Reminding Test, it is possible that greater recall scores would have resulted. Also, although the elderly were assessed for prose recall using Logical Memory, no strategies for improving prose recall were given during the workshop. However, the workshops were purposely structured this way in order to assess the potential for transfer of training to other different but related memory tasks. Based on this study, proven memory training strategies do not appear to transfer to related memory tasks as were used in our memory assessments.

Another factor that may have negatively impacted the level of recall in the workshop condition was the presence of any participant anxiety. Anxiety has also been associated with decreased memory performance (Bhagia & Pal, 1986). However, since an anxiety measure was not part of the test battery and this relationship does not imply causality, this possibility remains only conjecture. Also, participants may have reverted to their former memory strategies to increase their "comfort zone" or, since they were not expressly asked to use their "new" memory strategies, they may not have chosen to use them.

Causality

A stability analysis revealed that our relatively well-educated elderly sample had a very stable level of self-reported depressive symptoms and explicit memory ability across

time. Because of these minimal changes, causality could not be determined. However, a negative relationship was found between depressive symptoms and explicit memory ability. In other words, as the level of depressive symptoms increased, total recall scores of the two measures of explicit memory decreased. The negative correlations encountered were similar to those observed by Feehan et al. (1991) and King et al. (1991). Unexpectedly, this relationship was found to be equally strong across time periods. One could hypothesize that an individual's present level of depressive symptoms influences future memory ability and that one's present memory ability influences future levels of depressive symptoms. However, due to the minimal variation across time, these hypotheses of causality cannot be determined from this sample. It was hoped that the results of this study could have been used to design future memory training programs and/or memory remediation interventions more effectively. If depression does decrease memory recall, the client could first be treated for depression before entering a memory training or rehabilitation program or, alternatively, the program could have a depression component within its design. However, it remains the task of future studies to determine if such a relationship exists.

Suggestions for Future Research

Ideally, examination of the Johnson and Magaro's (1987) hypothesis should be carried out concurrently on depressed and non-depressed psychiatric patients, patients with high levels of depressive symptoms but not meeting formal diagnositc criteria for depression, and those who are affectively healthy and do not meet the criteria for a psychiatric diagnosis. With these four groups, a study should be able to tease out the effects of depression, psychiatric hospitalization, and depressive symptoms on memory recall. As for the memory measures, ideally visual, as well as, verbal explicit memory should be assessed in order to more fully delineate any effects of depressive symptoms on explicit memory. These measures could be even further refined by having verbal explicit memory broken down into aural, visual, and mixed input. Also, adding some "everyday"

explicit memory measures like a shopping list task or a newspaper article may reveal something not observed in the more "removed-from-the-real-world" laboratory measures. As for the problems of limited change in memory and depression levels across time, having assessments over additional time periods may solve this problem. Using a time series analysis in which the participants are assessed for 5 or even 10 test periods may be able to tease out any causal relationships that do exist.

Work could also be focused on the effects of depressive symptoms on the other types of memory suffering age-related declines (e.g., episodic memory, semantic memory, prospective memory). Does Johnson and Magaro's (1987) hypothesis extend to these forms of memory also? Does this hypothesis extend to other forms of memory that do not appear to be susceptible to the effects of age (e.g., implicit memory)? Is there an interaction effect, where those forms of memory exhibiting age-related declines are differentially susceptible to the effects of depressive symptoms? Also, do depressive symptoms effect memory in the same temporal plane or do they have their greatest impact on some future time frame? In addition, does memory ability affect depressive symptoms? Furthermore, do all depressive symptoms exert these effects or are just certain symptoms associated with this occurrence? In summary, there are many questions that still need to be addressed in this area of mood and memory.

APPENDIX A

APPENDIX A

Characteristics of the Meta-Analysis

Medline and Psychlit were searched for relevant articles published from 1979 to 1993. Relevancy was based on the following criteria: (a) the sample needed to include elderly subjects (i.e., 55 years of age or older), (b) subjects had to be assessed for depressive symptomatology, (c) the article had to be published in English, and (d) explicit memory had to be assessed by either a word list or prose task. This analysis was limited to studies with elderly individuals for purposes of comparison with this study and due to the possibility that the relationship between memory impairment and depressive symptomatology may be moderated by age (Lichtenberg, Manning, & Turkheimer, 1992). In addition, studies were limited to those using word list and prose recall tasks since they are explicit memory measures that are analogous to many everyday situations (e.g., shopping list, to-do list, newspaper and magazine articles) that the elderly encounter. Relevant studies cited in the articles obtained through Medline and Psychlit were also acquired. The resulting studies, with the depressive and memory measures used in each, are found in Table 7.

Statistical analysis of these studies was done with the d and r statistics. The d statistic was used since it is analogous to the "t" statistic but is independent of sample size. For this analysis, d was based on the difference between the means of the two groups divided by the within-group standard deviation. The within-group standard deviation was used since it has approximately half the sampling error found in the control-group standard deviation (Hunter & Schmidt, 1990). The point biserial correlation (r) was used as

another measure of treatment effect and was calculated by transforming the d statistic (Hunter & Schmidt, 1990). This correlation describes the degree of relationship that exists between the group (i.e., depressed or control) and the level of explicit memory ability (i.e., total recall score). In addition, since the point-biserial correlation is affected by sample size, the rs for studies having unequal sized groups were corrected for attenuation (Hunter & Schmidt, 1990; p. 274). The larger correlation that results from this correction is what would have occurred had the samples been of equal size.

The overall results of statistics d and r were expressed as a weighted average and weighted standard deviation. These measures were used to take into account the variation in sample size across the studies. This average sample correlation (r_{avg}), which can be considered an unbiased estimate of the population correlation (Hunter & Schmidt, 1990), appears congruent with the relationship hypothesized by Johnson and Magaro (1987).

Table 7 Depression and Memory Measures

	Depression Measures	Memory Measures
Gibson (1982)		
	Psychiatrist rating	10-item word list
Zarit (1982)	Zung Depression Scale	15-item shopping list task
	Brief Symptom Inventory	15-item word list
O'Hara et al. (1986)		
	RDC diagnosis	20-item word list
	Schedule for Affective Disorders and Schizophrenia (SADS-L) diagnosis	
	Center for Epidemiological Studies Depression Scale (CES-D)	
Hart, Kwentus, Taylor, & Hamer (1987)		
	DSM-III diagnosis	Selective Reminding Test
	Hamilton Depression Scale	
Hart, Kwentus, Taylor, & Harkins (1987)		
	DSM-III diagnosis	Logical Memory Form I
	Hamilton Depression Scale	
Poitrenaud et al. (1989)		
	DSM-III diagnosis	20-item shopping list task
	Montgomery and Asberg Depression Rating Scale	
	Geriatric Depression Scale	
Feehan et al. (1991)		
	DSM-III diagnosis	Logical Memory Form I
	Geriatric Depression Scale	
King et al. (1991)		
	DSM-III-R diagnosis	10-item word list
	Hamilton Depression Scale	
Lichtenberg et al. (1992)		
	Geriatric Depression Scale	Logical Memory Form I

APPENDIX B

APPENDIX B

Psychometric Properties

First, psychometric analyses of the three depressive measures (i.e., HAM, BDI, GDS) and the two explicit memory measures (i.e., LM, SRT) were undertaken. In addition to these variables, a DEP factor and MEM factor were also analyzed. The DEP factor was based on the 45 participants' three depression scores. Each score was transformed into a z-score and these were aggregated to yield a DEP factor. Consequently, as the respondents endorsed fewer depressive symptoms across the three measures, their score on the DEP factor would become more negative or progressively more positive as more symptoms were acknowledged. The MEM factor was based on the summed z-scores of the individual's two explicit memory measures. As a result, lower recall scores would produce more negative z scores for the MEM factor while greater recall scores would result in more positive z scores. Z-scores were used as the basis of these factors since the total score ranges varied across each test measure.

Initially, the construct validity of each of the depressive measures (i.e., HAM, BDI, GDS) were empirically tested using confirmatory factor analysis. Confirmatory factor analyses (CFA; Hunter & Cohen, 1969) were performed on each of the depressive measures based on all participants who were assessed across each of the three time periods. For each CFA, all test items making up that depressive instrument were grouped under one factor. Test contamination was then defined as any test item which had a negative loading on its corresponding factor. Test items with negative factor loadings were assumed to be measuring something different from the construct of depression being

assessed by the other test items and were removed to increase the internal consistency of the measure. In each analysis, the correlations between the test items and their corresponding factor were corrected for attenuation (Hunter & Cohen, 1969). Also, CFA's were only performed on the depression measures since the scores for the two memory measures were not based on multiple test items (Table 8).

Table 8 Item-Factor Correlations for Depressive Measures

	HAM	BDI	GDS
Time 1 (n = 208)			
Range	.17 - .71	.26 - .63	.27 - .64
M	.44	.45	.46
Time 2 (n = 120)			
Range	.20 - .72	.20 - .71	.17 - .76
M	.48	.45	.45
Time 3 (n = 45)			
Range	.06 - .87	.12 - .63	.06 - .68
M	.41	.38	.37

Analysis of the CFA's for the BDI and GDS indicated that loadings between test items and their corresponding factors were all positive across each of the three time periods. The HAM, however, had 4 test items eliminated due to zero variance, and six removed due to negative test item-factor loadings. In all subsequent analyses, the HAM total score did not reflect these "contaminated" test items. Across each of the three depressive measures, the mean loading was lowest in the third test period indicating a lower level of internal consistency.

Next, stability was determined for the three depressive measures using Cronbach's (1951) alpha. The alphas for each of the depressive measures are found in Table 9. For each depression scale, the time period containing the highest mean test item-factor loading resulted in the highest coefficient alpha and the time period with the lowest mean test item-factor loading resulted in the lowest coefficient alpha.

Table 9 Coefficient Alphas for Depressive Measures

	HAM	BDI	GDS
Time 1	.75	.84	.89
Time 2	.77	.84	.88
Time 3	.70	.78	.83

Following this, construct validity of the DEP and MEM factor were empirically tested by determining the loadings of the HAM, BDI, and GDS with the DEP factor and the SRT and LM with the MEM factor. As Table 10 indicates, the factor loadings with the DEP factor ranged from .71 to .96. The HAM loaded highest with the DEP factor at time 3, the BDI loaded highest at time 1, and the GDS at time 2. The loadings of the LM and SRT with the MEM factor were consistently in the low .7 range across the three time periods with both memory tests having their highest loading with the MEM factor during the second test period. The stability of the stayers' DEP and MEM factors were then determined across the three time periods. The resulting coefficient alphas for the DEP and MEM factors are reported in Table 11. Here, the alphas are found to be highest in the second period while the lowest alpha varied with the measure. For the DEP factor, time 1 was associated with the lowest alpha, while for the MEM factor, time 3 had the lowest

alpha. As can be seen, reliability was moderately higher for the DEP factor than for the MEM factor. Also, for those readers that prefer the raw scores of the depression and memory measures, these are included in Table 12.

Table 10 Stayers Test-Factor Correlations

	DEP1	DEP2	DEP3
HAM	.71	.79	.83
BDI	.79	.72	.78
GDS	.75	.96	.71
	MEM1	MEM2	MEM3
LM	.73	.74	.71
SRT	.73	.74	.71

Table 11 Stayers Coefficient Alphas for the DEP and MEM Factors

	Time 1	Time 2	Time 3
DEP	.79	.86	.82
MEM	.69	.71	.67

Table 12 Stayers Raw Score Means and SDs

	TIME 1		TIME 2		TIME 3	
	M	SD	M	SD	M	SD
HAM	3.6	4.1	3.1	3.3	2.5	3.0
BDI	6.8	5.9	6.1	5.3	6.4	5.1
GDS	7.1	5.1	5.4	4.8	6.2	4.8
LM	6.2	2.4	6.7	2.5	7.0	2.6
SRT	38.4	12.9	38.7	11.7	39.0	11.7

APPENDIX C

APPENDIX C

Stayers versus Decliners

The means and standard deviations for the stayers and decliners were initially determined and are listed in Table 13. Secondly, point biserial treatment correlations were calculated. The point biserial correlations represent the degree of relationship between the group (i.e., stayers and decliners) and each of the above variables (e.g., age, HAM).For each point biserial correlation, a 1-sided, 95% confidence interval with worst case, best case, and odds ratio was calculated (Hunter, 1993a). A one-sided confidence interval was used since it was hypothesized that decliners would be more depressed and have a poorer memory than stayers. The 95% confidence interval gives the interval that 95 times out of 100 should contain the treatment correlation. It is bounded by the worst case, which represents the lowest possible correlation that would be expected in 95 times out of 100, and the best case, which represents the highest possible correlation that would be expected in 95 times out of 100. The true treatment correlation should lie somewhere in the middle of this interval. The odds ratio gives the probability that the correlation is positive. The greater the odds ratio, the greater the chance that the two groups are different on that variable. Variables with odds ratios that were equal to or greater than 3:1 were considered to have a high probability of being different between groups. A value of 3:1 was chosen since it results in a 75% chance that there is a positive correlation in the population. These results are presented in Table 14.

Finally, the standard deviation ration was calculated. The standard deviation ratio ($\underline{V}$) is formed by the larger standard deviation divided by the smaller standard deviation.

It provides another index of differences in variation that may have existed between the two groups on the variables assessed. For each of the standard deviation ratios, a 2-sided 95% confidence interval with worst and best case were calculated (Hunter, 1993c). This gives us the lowest (i.e., worst case) standard deviation ratio and the highest (i.e., best case) standard deviation ratio that would be expected in 95 times out of 100. The true standard deviation ratio is expected to occur somewhere in the middle of these two values. Intervals not containing 1.00 are considered to have a high probability of being different. As can be seen in Table 15, only the variable MMSE filled this criteria. It appears that the standard deviation is twice as big in the "decliners" as it is in the "stayers" (Decliners MMSE SD: 2.9; Stayers MMSE SD: 1.6).

Table 13 Comparison of Stayers and Decliners Means and SDs

	STAYERS n = 45		DECLINERS n = 163	
	M	SD	M	SD
AGE	68.0	09.8	70.5	09.4
GENDER	00.3	00.5	00.2	00.4
EDUCATION	13.3	02.9	13.2	02.8
HAM	03.6	04.1	03.5	03.7
BDI	06.8	05.9	07.3	06.0
GDS	07.1	05.1	07.0	05.8
DEP	-00.1	02.5	00.0	02.7
LM	06.2	02.4	05.8	02.9
SRT	38.4	12.8	34.9	12.8
MEM	00.3	01.6	-00.1	01.8
PROB	00.8	00.4	00.8	00.4
# PROB	02.5	02.0	02.2	01.8
SR MEM	00.2	01.8	-00.1	01.8
MMSE	28.0	01.6	27.3	02.9
C/W	00.600	00.5	00.4	00.5

Table 14 Stayers and Decliners Correlation Coefficients

	r	REL.	WORST CASE	BEST CASE	ODDS RATIO
AGE	0.11	1.00	0.00	0.22	16.62
GENDER	0.05	1.00	-0.06	0.17	03.55
EDUCATION	0.02	1.00	-0.10	0.13	01.51
HAM	0.01	0.75	-0.13	0.14	01.12
BDI	0.04	0.84	-0.09	0.16	02.27
GDS	0.01	0.89	-0.11	0.13	01.15
DEP	0.01	0.87	-0.11	0.13	01.23
LM	0.05	0.79	-0.07	0.19	03.64
SRT	0.11	0.81	0.00	0.25	18.29
MEM	0.09	0.76	-0.02	0.24	10.17
PROB	0.06	0.77	-0.06	0.20	04.18
# PROB	0.06	0.66	-0.07	0.21	04.20
SR MEM	0.07	0.80	-0.05	0.20	04.88
MMSE	0.11	0.82	0.00	0.25	17.21
C/W	0.20	1.00	0.10	0.31	925.2

Table 15 Stayers and Decliners Standard Deviation Ratios

	STANDARD DEVIATION RATIO ($\underline{V}$)	WORST CASE	BEST CASE
AGE	1.04	0.79	1.29
GENDER	1.08	0.83	1.34
EDUCATION	1.02	0.78	1.26
HAM	0.90	0.69	1.11
BDI	1.03	0.78	1.28
GDS	1.14	0.87	1.42
DEP	1.09	0.83	1.35
LM	1.19	0.91	1.48
SRT	1.00	0.76	1.24
MEM	1.13	0.86	1.40
PROB	1.11	0.84	1.37
# PROB	1.13	0.86	1.40
SR MEM	1.02	0.78	1.27
MMSE	1.86	1.41	2.30
C/W	1.02	0.78	1.27

APPENDIX D

APPENDIX D

Workshop versus Controls

The means and standard deviations for the control and workshop conditions were initially computed and are listed in Table 16. Next, treatment correlations that had been corrected for attenuation were determined for each condition and time period comparisons (Table 17). Here attenuation refers to the reduced size in the treatment correlation due to the effects of random error of measurement. Consequently, correction for attenuation restores, or increases the size of the correlation to what it would have been had perfect measurement been used (Hunter, 1990). These treatment correlations can be considered a measure of the size of the average treatment effect that occurred between the different time frames. Following this, the treatment correlations were determined for a 95% confidence interval for both the control (Table 18) and workshop (Table 19) condition.

Finally, the treatment correlations were compared between the control and workshop conditions (Table 20). From these, one-sided, 95% confidence intervals with worst case, best case, and odds ratios were calculated (Hunter, 1993a). One sided confidence intervals were chosen since it was hypothesized that the workshop condition should improve more than the control condition. Here, the 95% confidence interval reflects that region which captures the actual population treatment correlation 95% of the time. It is bounded by the worst and best case. The worst case is the lowest possible treatment correlation that should result in 95 times out of 100. The best case is the highest possible treatment correlation that should occur in 95 times out of 100. The odds

ratio, which indicates the chances of the treatment correlation being positive, was considered to have a high probability of being different if it was 3:1 or greater.

Table 16 Means and SDs for Control and Workshop Conditions.

	TIME 1	TIME 2	TIME 3
A. Control Condition			
DEP Factor			
M	0.185	-0.334	-0.096
SD	2.391	2.160	2.264
MEM Factor			
M	-0.060	-0.030	0.352
SD	1.596	1.741	1.502
B. Workshop Condition			
DEP Factor			
M	0.521	-0.118	-0.238
SD	3.018	2.897	2.488
MEM Factor			
M	-0.302	0.048	0.079
SD	1.918	1.776	1.888

Table 17 Treatment Correlations for Control and Workshop Conditions.

	TIME 1 & 2	TIME 2 & 3	TIME 1 & 3
CONTROL			
DEP	-0.13	0.06	-0.07
MEM	0.01	0.14	0.16
WORKSHOP			
DEP	0.12	-0.02	-0.15
MEM	0.11	0.01	0.12

Table 18 Control Treatment Correlations

	WORST CASE	BEST CASE	ODDS RATIO
DEP Factor			
TIME 1& 2	-0.26	0.01	0.07
TIME 2 & 3	-0.05	0.17	4.38
TIME 1 & 3	-0.21	0.08	0.28
MEM Factor			
TIME 1 & 2	-0.15	0.17	1.19
TIME 2 & 3	-0.05	0.33	7.87
TIME 1 & 3	-0.02	0.34	12.48

Table 19 Workshop Treatment Correlations

	WORST CASE	BEST CASE	ODDS RATIO
DEP Factor			
TIME 1& 2	0.02	0.22	34.58
TIME 2 & 3	-0.12	0.07	0.05
TIME 1 & 3	-0.25	-0.05	0.01
MEM Factor			
TIME 1 & 2	-0.01	0.23	15.69
TIME 2 & 3	-0.11	0.13	1.25
TIME 1 & 3	-0.01	0.25	13.24

Table 20 Comparisons of Treatment Correlations

	WORST CASE	BEST CASE	ODDS RATIO
DEP			
TIME 1& 2	-0.26	0.75	3.67
TIME 2 & 3	-0.59	0.43	0.65
TIME 1 & 3	-0.59	0.42	0.65
MEM			
TIME 1 & 2	-0.41	0.61	1.69
TIME 2 & 3	-0.64	0.38	0.51
TIME 1 & 3	-0.54	0.46	0.82

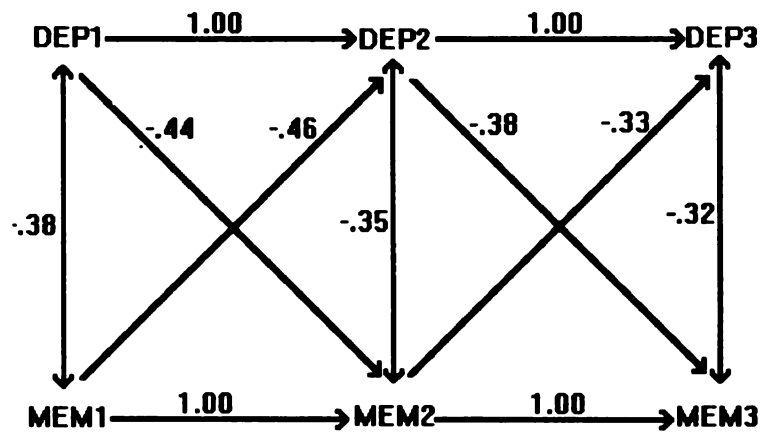
APPENDIX E

APPENDIX E

Causality

The path diagram was based on correlations that had been corrected for attenuation. As can be seen in Table 21, due to the minimal change, the stability coefficients were 1.00 between the depressive measures and between the memory measures across the three time periods.

Table 21 Path Diagram



LIST OF REFERENCES

LIST OF REFERENCES

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1987). Diagnostic and statistical manual of mental disorders. (Revised 3rd ed.). Washington, DC: American Psychiatric Association.

Anschutz, L., Camp, C., Markley, R., & Kramer, J. (1985). Maintenance and generalization of mnemonics for grocery shopping by older adults. Experimental Aging Research, 11, 157-160.

Beck, A., Steer, R., & Garbin, M. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review, 8, 77-100.

Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 53-63.

Bhagia, N., & Pal, H. (1986). The effect of anxiety on memorization. Perspectives in Psychological Researches, 9, 60-61.

Blackford, R., & LaRue, A. (1989). Criteria for diagnosing age-associated memory impairment: Proposed improvements from the field. Developmental Neuropsychology, 5, 295-306.

Bowden, S., & Bell, R. (1992). Relative usefulness of the WMS and WMS-R: A comment on D'Elia et al. (1989). Journal of Clinical and Experimental Neuropsychology, 14, 340-346.

Buschke, H. (1973). Selective reminding for analysis of memory and learning. Journal of Verbal Learning and Verbal Behavior, 12, 543-550.

Buschke, H., & Fuld, P. (1974). Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology, 24, 1019-1025.

Cattell, R. (1966). The scree test for number of factors. Multivariate Behavioral Research, 1, 245.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Craik, F. (1991). Memory functions in normal aging. In T. Yanagihara & R. Petersen (Eds.), Memory disorders research and clinical practice (pp. 347-367). New York: Marcel Dekker.

Crook, T. (1989). Diagnosis and treatment of normal and pathological memory impairment in later life. Seminars in Neurology, 9, 20-30.

Crook, T., Bartus, R., Ferris, S., Whitehouse, P., Cohen, G., & Gershon, S. (1986). Age associated memory impairment: Proposed diagnostic criteria and measures of clinical change-Report of a National Institute of Mental Health Work Group. Developmental Neuropsychology, 2, 261-276.

Derogatis, I., & Spencer, P. (1982). The Brief Symptom Inventory (BSI): I. Administration, scoring, and procedures manual. Baltimore, MD: Johns Hopkins University Press.

Erickson, R., & Scott, M. (1977). Clinical memory testing: A review. Psychological Bulletin, 84, 1130-1149.

Feehan, M., Knight, R., & Partridge, F. (1991). Cognitive complaint and test performance in elderly patients suffering depression or dementia. International Journal of Geriatric Psychiatry, 6, 287-293.

Fitzgibbon, M., Cella, D., & Sweeney, J. (1988). Redundancy in measures of depression. Journal of Clinical Psychology, 44, 372-374.

Folstein, M., Folstein, S., & McHugh, P. (1975). "Mini-Mental State" A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

Gibson, A. (1981). A further analysis of memory loss in dementia and depression in the elderly. British Journal of Clinical Psychology, 20, 179-185.

Hamilton, M. (1960). A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry, 23, 56-62.

Hannay, H., & Levin, H. (1985). Selective Reminding Test: An examination of the equivalence of four forms. Journal of Clinical and Experimental Neuropsychology, 7, 251-263.

Harper, R., Kotik-Harper, D., & Kirby, H. (1990). Psychometric assessment of depression in an elderly general medical population. Journal of Nervous and Mental Disease, 178, 113-119.

Hart, R., Kwentus, J., Taylor, J., & Hamer, R. (1987). Selective reminding procedure in depression and dementia. Psychology and Aging, 2, 111-115.

Hart, R., Kwentus, J., Taylor, J., & Harkins, S. (1987). Rate of forgetting in dementia and depression. Journal of Consulting and Clinical Psychology, 55, 101-105.

Heath, J., Grant, W., Kamps, C., & Margolin, E. (1991). Outpatient geriatric assessment: Associations between referral sources and assessment findings. Journal of the American Geriatrics Society, 39, 267-272.

Hedlund, J., & Vieweg, B. (1979). The Hamilton Rating Scale for Depression: A comprehensive review. Journal of Operational Psychiatry, 10, 149-162.

Heise, D. (1969). Separating reliability and stability in test-retest correlation. American Sociological Review, 34, 93-101.

Hill, R., Storandt, M., & Simeone, C. (1990). The effects of memory skills training and incentives on free recall in older learners. Journal of Gerontology, 45, 227-232.

Hulicka, I., & Grossman, J. (1967). Age-group comparisons for the use of mediators in paired-associate learning. Journal of Gerontology, 22, 46-51.

Hunter, J. (1990). Attenuation due to random error of measurement. Unpublished manuscript.

Hunter, J. (1993a). CORRECT: A program in Basica to correct correlations for various artifacts (1st revision, Sep. 1993). Unpublished manuscript.

Hunter, J. (1993b). SE: A program to compute standard errors. Unpublished manuscript.

Hunter, J. (1993c). CONFINT: A program in Basica to compute confidence intervals and statistical power analysis (2nd revision, Sep. 1993). Unpublished manuscript.

Hunter, J. (1994). WITHIN.BAS- The manual- A program to compute the analysis for a "within subjects" design. Unpublished manuscript.

Hunter, J., & Cohen, S. (1969). PACKAGE: A system of computer routines for the analyses of correlational data. Education and Psychological Measurement, 29, 697-700.

Hunter, J. & Schmidt, F. (1990). Methods of meta-analysis. Newbury Park: Sage Publication.

Johnson, M., & Magaro, P. (1987). Effects of mood and severity on memory processes in depression and mania. Psychological Bulletin, 101, 28-40.

Johnston, L., & Gueldner, S. (1989). Remember when? Using mnemonics to boost memory in the elderly. Journal of Gerontological Nursing, 15, 22-26.

Kahn, R., Zarit, S., Hilbert, N., & Niederehe, G. (1975). Memory complaint and impairment in the aged. Archives of General Psychiatry, 32, 1569-1573.

Kendall, P., Hollon, S., Beck, A., Hammen, C., & Ingram, R. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. Cognitive Therapy and Research, 11, 289-299.

King, D., Caine, E., Conwell, Y., & Cox, C. (1991). The neuropsychology of depression in the elderly: A comparative study of normal aging and Alzheimer's disease. The Journal of Neuropsychiatry and Clinical Neurosciences, 3, 163-168.

Koenig, H., Meador, K., Cohen, H., & Blazer, D. (1988). Self-rated depression scales and screening for major depression in the older hospitalized patient with medical illness. Journal of the American Geriatrics Society, 36, 699-706.

Lichtenberg, P., Manning, C., & Turkheimer, E. (1992). Memory dysfunction in depressed spousal caregivers. Clinical Gerontologist, 12, 77-80.

Lichtenberg, P., Steiner, D., Marcopulos, B., & Tabscott, J. (1992). Comparison of the Hamilton Depression Rating Scale and the Geriatric Depression Scale: Detection of depression in dementia patients. Psychological Reports, 70, 515-521.

Lowenthal, M., Berkman, P., Brisset, G., Buehler, J., Pierce, R., Robinson, B., & Trier, M. (1967). Aging and mental disorder in San Francisco. San Francisco: Jossey-Bass.

Lyons, J., Strain, J., Hammer, J., Ackerman, A., & Fulop, G. (1989). Reliability, validity, and temporal stability of the Geriatric Depression Scale in hospitalized elderly. International Journal of Psychiatry in Medicine, 19, 203-209.

Maier, W. (1990). The Hamilton Depression Scale and its alternatives: A comparison of their reliability and validity. In P. Bech & A. Coppen (Eds.), The Hamilton Scales (pp. 64-71). New York: Springer-Verlag.

Masur, D., Fuld, P., Blau, A., Crystal, H., & Aronson, M. (1990). Predicting development of dementia in the elderly with the Selective Reminding Test. Journal of Clinical and Experimental Neuropsychology, 12, 529-538.

Masur, D., Fuld, P., Blau, A., Thal, L., Levin, H., & Aronson, M. (1989). Distinguishing normal and demented elderly with the Selective Reminding Test. Journal of Clinical and Experimental Neuropsychology, 11, 615-630.

Norris, J., Gallagher, D., Wilson, A., & Winograd, C. (1987). Assessment of depression in geriatric medical outpatients: The validity of two screening measures. Journal of the American Geriatric Society, 35, 989-995.

O'Brien, K., & Glaudin, V. (1988). Factorial structure and factor reliability of the Hamilton Rating Scale for Depression. Acta Psychiatrica Scandinavica, 78, 113-120.

O'Hara, M., Hinrichs, J., Wallace, R., Lemke, J., & Kohout, F. (1986). Memory complaint and memory performance in the depressed elderly. Psychology and Aging, 1, 208-214.

Parmalee, P., Lawton, M., & Katz, I. (1989). Psychometric properties of the Geriatric Depression Scale among the institutionalized aged. Journal of Consulting and Clinical Psychology, 1, 331-338.

Poitrenaud, J., Moy, F., Girousse, A., Wolmark, Y., & Piette, F. (1989). Psychometric procedures for analysis of memory losses in the elderly. Archives of Gerontology and Geriatrics, (Suppl.1), 173-183.

Prigatano, G. (1978). Wechsler Memory Scale: A selective review of the literature. Journal of Clinical Psychology, 34, 816-832.

Rebok, G., & Balcerak, L. (1989). Memory self-efficacy and performance differences in young and old adults: The effect of mnemonic training. Developmental Psychology, 25, 714-721.

Rehm, L., & O'Hara, M. (1985). Item characteristics of the Hamilton Rating Scale for Depression. Journal of Psychiatric Research, 19, 31-41.

Robertson-Tchabo, E., Hausman, C., & Arenberg, D. (1976). A classical mnemonic for older learners: A trip that works! Educational Gerontology, 1, 215-226.

Rose, T., & Yesavage, J. (1983). Differential effects of a list-learning mnemonic in three age groups. Gerontology, 29, 293-298.

Ruff, R., Quayhagen, M., & Light, R. (1988). Selective Reminding Test: A normative study of verbal learning in adults. Journal of Clinical and Experimental Neuropsychology, 11, 539-550.

Russel, E. (1975). A multiple scoring method for the assessment of complex memory functions. Journal of Consulting and Clinical Psychology, 43, 800-809.

Salamero, M., & Marcos, T. (1992). Factor study of the Geriatric Depression Scale. Acta Psychiatrica Scandinavica, 86, 283-286.

Schaffer, G., & Poon, L. (1982). Individual variability in memory training with the elderly. Educational Gerontology, 8, 217-229.

Schmitt, F., Murphy, M., & Sanders, R. (1981). Training older adult free recall rehearsal strategies. Journal of Gerontology, 36, 329-337.

Scogin, F., Storandt, M., & Lott, L. (1985). Memory-skills training, memory complaints, and depression in older adults. Journal of Gerontology, 40, 562-568.

Shelton, M., & Parsons, O. (1987). Alcoholics' self-assessment of their neuropsychological functioning in everyday life. Journal of Clinical Psychology, 43, 395- 403.

Smith, M., Glass, G., & Miller. (1980). Benefits of psychotherapy. Baltimore, MD: John Hopkins University Press.

Smith, G., Ivnik, R., Petersen, R., Malec, J., Kokmen, E., & Tangalos, E. (1991). Age-associated memory impairment diagnoses: Problems of reliability and concerns for terminology. Psychology and Aging, 6, 551-558.

Spitzer. R., Endicott, J., & Robbins, E. (1978). Research Diagnostic Criteria: Rationale and reliability. Archives of General Psychiatry, 36, 773-782.

Storandt, M., Botwinick, J., Danziger, W., Berg, L., & Hughes, C. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer type. Archives of Neurology, 41, 497-499.

Stukenberg, K., Dura, J., & Kiecolt-Glaser, J. (1990). Depression screening scale validation in an elderly, community-dwelling population. Journal of Consulting and Clinical Psychology, 2, 134-138.

Tombaugh, T., & McIntyre, N. (1992). The Mini-Mental State Examination: A comprehensive review. Journal of the American Geriatrics Society, 40, 922-935.

Tun, P., Perlmutter, L., Russo, P., & Nathan, D. (1987). Memory self-assessment and performance in aged diabetics and non-diabetics. Experimental Aging Research, 13, 151-157.

Wechsler, D., & Stone, C. (1945). Wechsler memory Scale. New York: Psychological Corporation.

Weiss, I., Nagel, C., & Aronson, M. (1986). Applicability of depression scales to the old old person. Journal of the American Geriatric Society, 34, 215-218.

West, R., Boatwright, L., & Schleser, R. (1984). The link between memory performance, self-assessment, and affective status. Experimental Aging Research, 10, 197-200.

Williams, J., Little, M., Scates, S., & Blockman, N. (1987). Memory complaints and abilities among depressed older adults. Journal of Consulting and Clinical Psychology, 55, 595-598.

Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, V., Adey, M., & Leirer, V. (1983). Development and validation of a Geriatric Depression Screening Scale: A preliminary report. Journal of Psychiatric Research, 17, 37-49.

Yesavage, J., & Rose, T. (1983). Concentration and mnemonic training in elderly subjects with memory complaints: A study of combined therapy and order effects. Psychiatry Research, 9, 157-167.

Yesavage, J., Rose, T., & Spiegel, D. (1982). Relaxation training and memory improvement in elderly normals: Correlation of anxiety ratings and recall improvement. Experimental Aging Research, 8, 195-198.

Yesavage, J., Sheikh, J., Friedman, L., & Tanke, E. (1990). Learning mnemonics: Roles of aging and subtle cognitive impairment. Psychology and Aging, 5, 133-137.

Zarit, S. (1982). Affective correlates of self-reports about memory of older people. International Journal of Behavioral Geriatrics, 1, 25-35.

Zarit, S., Cole, K., & Guider, R. (1981). Memory training strategies and subjective complaints of memory in the aged. The Gerontologist, 21, 158-164.