Psychological Medicine, 1994, **24**, 829–847.

Cognitive function in depression: its relationship to the presence and severity of intellectual decline

R. G. BROWN,[1] L. C. SCOTT, C. J. BENCH AND R. J. DOLAN

From the MRC Human Movement and Balance Unit, Institute of Neurology; Academic Department of Psychiatry, Royal Free School of Medicine; MRC Cyclotron Unit, Hammersmith Hospital; and National Hospital for Neurology and Neurosurgery, London

SYNOPSIS Cognitive dysfunction is an integral feature of depression, in some cases of sufficient severity to warrant a diagnosis of dementia. There has been little systematic investigation of whether cognitive dysfunction is an inevitable consequence of depression, or is specific to a subgroup of depressed patients. Related to this is the distribution of cognitive dysfunction, whether there is a continuum of impairment or a distinct demented subgroup. Finally, there is the question of which aspects of cognitive function are most sensitive to the intellectual decline seen in depression. A study is described which addresses these issues. The distribution of global cognition was found to be normally distributed in the sample of 29 patients assessed. Based on this distribution and the scores of a control sample, the patients were classified as unimpaired, borderline or impaired. Two sets of independent comparisons were carried out. First, the unimpaired depressed patients were compared to matched non-depressed controls. Significant deficits were found on a range of neuropsychological measures covering aspects of language function, memory, both recall and recognition, attention and behavioural regulation. These same patients were also compared with two groups of matched depressed patients, with varying degrees of global cognitive impairment. In general, the cognitive measures showed a gradient of dysfunction across the three patient groups. Significant differences between the depressed groups were shown on measures of immediate recall, attention and behavioural regulation. The possible significance of attentional factors for the observed memory dysfunction is discussed.

INTRODUCTION

Depressed individuals frequently complain of poor memory and concentration. Much effort has gone into assessing the nature and severity of this impairment. Although focused almost exclusively on memory, a few studies have shown deficits in areas such as abstract reasoning (Braff & Beck, 1974), simple perceptual discrimination (Cornell *et al.* 1984) and verbal fluency (Robertson & Taylor, 1985). A simple explanation of the observed deficits is that they reflect poor motivation, or distraction from depressive thoughts (see Jorm, 1986). However, there is increasing consensus that cognitive dysfunction is intrinsic to depression and directly related to the neurobiology of the illness.

[1] Address for correspondence: Dr R. G. Brown, MRC Human Movement and Balance Unit, Institute of Neurology, Queen Square, London WC1N 3BG.

Neuropsychological research has addressed a number of issues relating to cognitive dysfunction in depression. Some has been concerned primarily with clinical questions: whether there are differences between clinical subgroups (e.g. Savard *et al.* 1980); the effects of treatment (Sternberg & Jarvik, 1976; Frith *et al.* 1983) and the changes with remission of the depression (Savard *et al.* 1980). Other research has addressed the nature of the processes underlying the deficits. This research has employed various explanatory frameworks, such as information processing (Hasher & Zaks, 1979), 'effort' (Cohen *et al.* 1982), encoding strategy (Miller & Lewis, 1977), processing resources (Watts *et al.* 1990) or arousal and activation (Weingartner *et al.* 1981). To date, however, no single framework, 'theory' or explanation, has proven to be of general value in understanding the nature of cognitive impairment in depression. The reader

is referred to a number of reviews (Miller, 1975; McAllister, 1981; Willner, 1984; Jorm, 1986; Widlocher & Hardy-Bayle, 1989; Newman & Sweet, 1992).

The literature on the neuropsychology of depression has remained separate from research concerned with the relationship between depression and dementia (Cummings & Benson, 1984). Some patients with depression suffer cognitive impairment of sufficient severity to warrant a diagnosis of dementia. The apparent reversibility of the dementia with remission of the depressive illness has led to the label 'pseudodementia' (Kiloh, 1961) to distinguish it from progressive dementias such as Alzheimer's disease and Pick's disease. More recently the reversibility, or otherwise, of cognitive impairment has become less important in the diagnosis of dementia (Jorm, 1986). As a result the terms 'depressive dementia' or 'dementia syndrome of depression' (Cummings, 1989) have become the accepted diagnostic labels.

There are thus two broad strands of research concerned with, first, the nature of the impairment in depressed patients *per se*, and secondly, the nature of depressive dementia, and its relationship to other dementing disorders and its neurobiological substrate. To date, little attempt has been made to combine these two issues. However, some important questions arise when considering depression and cognitive function in a broader context. Are we dealing with a single entity of cognitive impairment in depression, with dementia being the extreme case? Is cognitive deficit, whether mild or severe, an inevitable consequence (or concomitant) of depression?

Assessing such questions on the basis of existing studies is difficult. In the majority of studies, patients have been selected on the basis of their affective disturbance alone. In such a sample it is possible that only a proportion will exhibit cognitive deficits. Other patients, with equally severe depression, may have no obvious cognitive impairment. The effect of combining samples of unimpaired depressed with demented depressed patients would simply be to show an average effect which fails to reflect the true nature or severity of cognitive disturbance in either of the subgroups. One approach is to assess a selected group of patients who show normal global cognitive function. Specific deficits in this group could confidently be attributed to depression, and not the influence of a subgroup of patients with a possibly independent dementing disorder. Conversely, if such patients *fail* to show any deficits, this would call into question models which suggest that cognitive dysfunction is an epiphenomenon of the depressive symptoms.

In the present paper, we examine the neuropsychological profiles of a sample of depressed patients classified on the basis of overall intellectual function. Because of the shortage of data relating to this issue we chose an exploratory approach employing a broad assessment of cognitive function. Our aim was to address three main questions. First, what is the distribution of cognitive dysfunction in a sample of depressed patients? Is there a continuous distribution, or is there evidence for a distinct subgroup of demented patients? Secondly, is there evidence of specific cognitive dysfunction even in a sample of depressed patients selected for the absence of global cognitive impairment? Thirdly, within a depressed sample, which specific aspects of cognitive function are most susceptible to impairment?

METHOD

Subjects

Patients were recruited from district psychiatric services in north London and the National Hospital for Neurology and Neurosurgery. Potential patients were administered the Schedule for Affective Disorders and Schizophrenia (SADS) (Endicott & Spitzer, 1978). Subjects who satisfied Research Diagnostic Criteria (Spitzer *et al.* 1977) (RDC) for Major Depressive Disorder were further screened with the following exclusion criteria: past or present history of neurological disease, drug or alcohol abuse, any significant past medical illness, a score of more than 4 on the Hachinski ischemia scale (Hachinski *et al.* 1975). In total, 29 patients entered into the study. These had moderate to severe depression as rated on the 17-item Hamilton Rating Scale for Depression (Hamilton, 1960) (mean = 25·0, S.D. = 4·2, range = 17–34). The mean age of the sample was 58·0 years (S.D. = 13·0). The sample comprised 19 males and 10 females. Their mean number of years of education was 11·3 years (S.D. = 3·0). Of

the 29 patients, 14 were taking antidepressant medication at the time of assessment. Most of the patients (26) were classified as having unipolar depression, with only three having bipolar depression.

Twenty healthy control subjects were also assessed. They comprised 6 males and 14 females, with a mean age of 58·4 years (S.D. = 14·5), and mean of 12·4 years of education (S.D. = 1·8). The same exclusion criteria were applied as with the patients. All subjects gave informed consent.

Psychiatric assessment

Previously (Bench *et al.* 1993), the SADS results of a larger sample, including the present subjects, were subjected to a principal component analysis with varimax rotation. The results revealed one factor 'Anxiety/Somatism', and another labelled 'Mood/Retardation'. For each individual in the present study two factor scores were derived and used in subsequent correlational analyses.

Neuropsychological assessment

The neuropsychological assessment comprised two sections. The first was a basic screening of cognitive function employing the CAMCOG (Roth *et al.* 1988). This measure comprises an expanded version of the Mini-Mental State Examination (MMSE) (Folstein *et al.* 1975), and provides a total score (maximum 107) as well as subscale scores for different aspects of cognition (see Table 1). A MMSE score can also be calculated (maximum score 30). A cut-off of 80 on the CAMCOG was found by Roth *et al.* (1986) to be useful to indicate the presence of significant cognitive impairment. The second part of the investigation comprised a battery of neuropsychological tests, administered over two sessions 24 h apart. Each subject was assessed at the same time of day on the two sessions. The order of test presentation was standardized as far as possible. Subjects were administered 5 subtests of the Wechsler Adult Intelligence Scale (WAIS-R) (Wechsler, 1986): Vocabulary, Similarities, Comprehension, Arithmetic and Digit Span, from which a pro-rated Verbal IQ was calculated. Reading ability was assessed using the Schonell Graded Word Reading Test (Schonell, 1942; Nelson & McKenna, 1975). Verbal memory and learning were assessed using: Wechsler Logical Memory (LM) and Word-Pair Associate Learning Test (PALT) (Wechsler, 1945); Rey Auditory Verbal Learning Test (RAVLT) (Taylor, 1959). Recall was assessed immediately and after delays of 1 and 24 h. The effects of list organization was assessed using three, 12-word lists comprising (A) unrelated words, (B) words from 3 categories randomly presented and (C) words from 3 categories clustered by category. Immediate free recall was assessed, followed, for lists B and C, by recall in which the subject was cued with the category names. After cued recall, subjects were given a recognition test in which the 12 words in each list were randomly mixed with 12 new items. For the lists B and C, the distractors belonged to the same semantic categories as the original words. Subjects were shown the words one at a time and asked whether or not they recognized them. Short-term memory was assessed using the Brown-Peterson (BP) test (Brown, 1958; Peterson & Peterson, 1959). A three-consonant trigram was read out to the subject which they had to repeat immediately or after a delay of 5, 10, 20 or 30 s. Rehearsal was prevented by asking the subjects to count backwards from a 3-digit number, presented immediately after the trigram. A number of tests related to language and 'executive' function were also administered. Verbal fluency was assessed using 3 conditions: 'free' (any word), 'category' (boys' names) and 'letter' (any word beginning with the letter 's'). Subjects performed each task for 60 s. Separate scores were given for the number of words generated in the first and second 30 s of each test. Language comprehension was assessed using an abbreviated version of the Token Test (Spreen & Benton, 1969). Subjects were given 16 commands to carry out. A 'strict' score was derived from the subjects performance on a single reading of the command. If the subject made an error, the command was given again, and a separate, 'lenient' score derived. Conceptual ability was assessed using the Weigl Test (Weigl, 1941).

Statistics and organization of results

The main methods were multivariate analysis of variance (MANOVA) and repeated measures analysis of variance (ANOVA). Planned contrasts were employed where specific effects were to be tested on an *a priori* basis. Where significant omnibus effects and interactions were

Table 1. *Details (mean and standard deviation) of the control and three depressed groups, for age, years of education, Hamilton score, Schonnel score, and results from the CAMCOG and WAIS (with scaled scores). The bracketed figures in italics following the measure name refer to the maximum score on the test or subtest (see also subsequent tables)*

Measure	Control group ($N = 16$) Mean (S.D.)	Depressed groups: Unimpaired ($N = 10$) Mean (S.D.)	Borderline ($N = 10$) Mean (S.D.)	Impaired ($N = 9$) Mean (S.D.)
Age	62·6 (14·1)	61·9 (13·4)	53·5 (31·5)	58·7 (11·1)
Years of education	12·4 (1·8)	11·9 (3·4)	11·2 (3·7)	10·7 (1·9)
Hamilton Depression Scale score	—	24·4 (4·2)	25·6 (5·8)	25·1 (3·8)
Schonnel Reading Test score (*100*)	94·5 (6·5)	92·1 (13·8)	80·5 (18·4)	63·2 (25·2)
CAMCOG total score (*107*)	97·7 (2·3)	96·1 (3·8)	86·7 (2·5)	71·1 (10·7)
Orientation (*10*)	9·9 (0·3)	9·6 (1·9)	9·3 (0·5)	6·3 (2·2)
Language (*30*)	28·6 (1·1)	26·3 (1·8)	25·0 (1·4)	22·9 (2·3)
Memory (*27*)	22·4 (1·9)	22·8 (1·8)	19·9 (2·8)	15·6 (4·8)
Attention (*7*)	6·8 (0·5)	6·6 (1·0)	5·1 (1·4)	36·0 (2·5)
Praxis (*12*)	11·2 (0·8)	11·4 (0·8)	10·2 (1·2)	8·4 (2·2)
Calculation (*2*)	2·0 (0·0)	2·0 (0·0)	2·0 (0·0)	1·4 (0·5)
Abstraction (*8*)	7·7 (0·9)	7·3 (0·8)	6·4 (1·4)	5·3 (2·3)
Perception (*11*)	10·3 (0·8)	10·2 (0·8)	8·9 (1·6)	7·8 (1·8)
MMSE total score (*30*)	29·2 (1·0)	28·9 (1·2)	26·2 (2·6)	19·7 (4·2)
WAIS Verbal IQ	113·6 (12·5)	110·3 (19·3)	90·1 (7·2)	86·7 (16·2)
Vocabulary	12·3 (3·0)	12·0 (4·1)	7·9 (1·7)	6·7 (2·6)
Comprehension	12·5 (2·5)	11·4 (2·9)	9·1 (2·1)	7·7 (2·6)
Similarities	10·9 (1·9)	9·7 (3·3)	6·6 (1·4)	6·3 (1·5)
Digit Span	11·1 (3·4)	8·9 (3·8)	8·5 (3·2)	6·3 (1·5)
Arithmetic	11·2 (2·1)	10·6 (3·1)	6·9 (1·7)	5·6 (2·4)

obtained, further comparisons were carried out. Conventional significance levels $< 0{\cdot}05$, $< 0{\cdot}01$ and $< 0{\cdot}001$ are adopted. P values of $> 0{\cdot}10$ are considered 'not significant'. Those between 0·05 and 0·10 ($P < 0{\cdot}10$) are reported as 'approaching significance'.

The first part of the results section is a classification of the depressed sample into three subgroups based on their total CAMCOG scores in relation to the control sample. Analysis of the four groups (one control and three depressed) were then carried out in two independent sets of comparisons. First, between a group of patients defined *a priori* as 'unimpaired' and the control group. Second, between the three subgroups of depressed patients, classified according to the total CAMCOG scores.

RESULTS

Total CAMCOG scores – classification of patients

Fig. 1 shows the distributions of total CAMCOG scores. The scores of the 20 control subjects ranged from 92–104 out of a maximum of 107. Of the patients, 10 (34·5%) scored 92 or more, with an upper score of 101. These depressed subjects will be considered 'unimpaired' depressed (UD). The remaining 19, divide into two groups: 10 (34·5%) 'borderline impaired' (BD) scored less than 92 but more than 81, while the remaining 9 (31%) scored 81 or less and formed the 'impaired' group (ID). Of the 20 control subjects, 4 achieved higher CAMCOG scores than any depressed patient. To improve matching, these 4 subjects were excluded from subsequent analysis, leaving a final control group of 16 subjects.

Comparison of unimpaired depressed and control groups

Age, sex, years of education, reading ability and CAMCOG score (Table 1)

Age, years of education, Schonell reading test score and total CAMCOG score, were entered into a MANOVA. The two groups did not differ on the set of variables. Univariate comparisons confirmed that the two groups were matched on each measure. There was a significant difference in the sex ratios of the two groups ($P < 0{\cdot}05$).

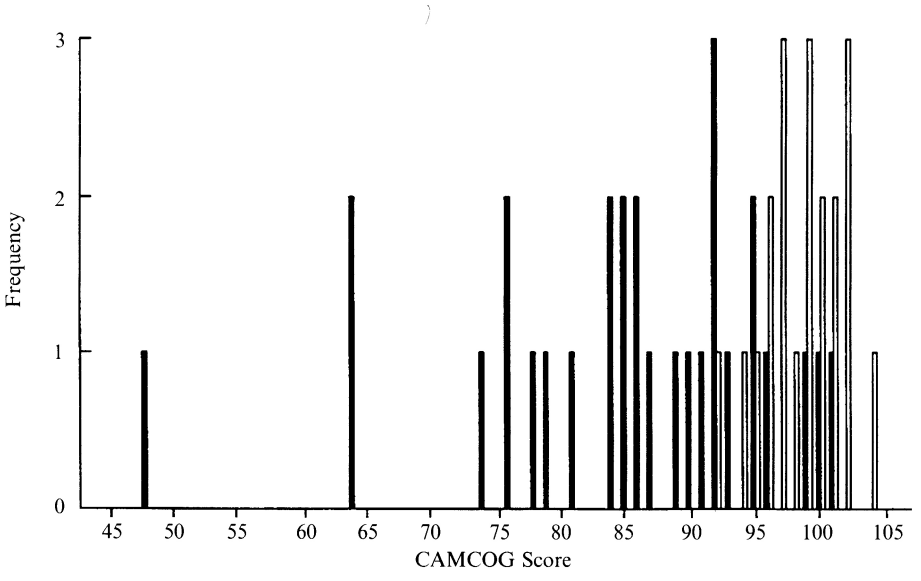

FIG. 1. Distribution of total CAMCOG scores in the controls group ($N = 20$) (▯) and depressed group ($N = 29$) (▮).

However, comparison of the male and female subjects in the two groups failed to reveal any consistent difference in performance on the various neuropsychological tests between the two sexes. Consequently, subject's sex was not considered further.

CAMCOG subscales (Table 1)

The MANOVA of subscale scores failed to reveal any significant overall group difference (the Arithmetic subscale score was excluded from this analysis as all subjects scored at ceiling). Despite the non-significant omnibus-F, univariate F tests demonstrated significant differences between the two groups on the two language tests, Comprehension ($P < 0{\cdot}05$) and Expression ($P < 0{\cdot}001$), with the UD group performing less well than the controls. In addition to the subscale analysis, one test item, serial sevens, was analysed individually because of its utility as a clinical measure of attention. There was no difference between the performance of the two groups (UD group, mean = 4·8, S.D. = 0·4; control group, mean = 4·9, S.D. = 0·3).

WAIS – Verbal IQ and Verbal Subscales (Table 1)

There was no significant difference between the groups on Verbal IQ pro-rated from the 5 subscales ($P > 0{\cdot}10$). A MANOVA with all of the subscale scores, revealed no significant overall difference between the control and UD groups. Univariate F tests revealed that the two groups were matched on all subscales.

Memory

Immediate free recall (Table 2)

The UD group showed a mixed pattern of performance on the measures of immediate free recall. For the prose passage of the LM test, their performance was lower than controls but the difference was not significant. Similarly, the group difference for recall of the first presentation of the 15 word list from the RAVLT only approached significance ($P < 0{\cdot}10$). Significant differences, however, were found for recall of the three word lists ($P < 0{\cdot}01$). Across subjects recall of the unrelated list (A) was inferior to the two categorized lists (B and C). Of these, recall of the unclustered list (B) was superior to that for the clustered list (C). This pattern of recall by list type was the same in the two groups. All of the above measures involved supraspan material. Mean memory span as measured by the digit span subtest of the WAIS-R was 6·6 (S.D. = 2·1) in the UD group and 7·1 (S.D. = 1·6) in the controls. The difference was not significant.

Overall, therefore, immediate recall of supraspan verbal material appeared to be inferior in

Table 2. *Results (mean and standard deviation) for control and the three depressed groups on Logical Memory, Paired Associate Learning and Rey Auditory Verbal Learning Test*

Measure		Control group ($N = 16$) Mean (S.D.)	Depressed groups: Unimpaired ($N = 10$) Mean (S.D.)	Depressed groups: Borderline ($N = 10$) Mean (S.D.)	Depressed groups: Impaired ($N = 9$) Mean (S.D.)
Wechsler Logical Memory					
Immediate recall score (*46*)		24·4 (5·1)	20·9 (8·8)	14·6 (4·9)	10·4 (5·6)
1 h delayed recall score		20·6 (5·6)	14·4 (9·6)	8·1 (5·3)	5·4 (6·4)
1 h delayed recall (% immediate)		83·9 (11·2)	63·6 (28·1)	41·8 (32·0)	40·7 (39·8)
24 h delayed recall score		19·4 (6·2)	11·8 (9·9)	7·1 (5·2)	5·8 (6·5)
24 h delayed recall (% immediate)		82·4 (14·1)	50·5 (32·7)	44·1 (34·9)	47·2 (46·9)
Paired Associate Learning Task					
Easy items	Trial 1 (*4*)	3·4 (0·7)	3·3 (0·8)	3·4 (0·7)	2·7 (1·2)
	Trial 2	3·8 (0·6)	3·8 (0·6)	3·4 (0·7)	2·8 (1·3)
	Trial 3	4·0 (0·0)	3·9 (0·3)	3·6 (0·5)	2·8 (1·5)
	Trial 4	4·0 (0·0)	3·9 (0·3)	3·8 (0·4)	2·7 (1·6)
	Trial 5	4·0 (0·0)	3·9 (0·3)	3·8 (0·4)	2·6 (1·7)
	Trial 6	4·0 (0·0)	3·8 (0·6)	3·9 (0·3)	3·0 (1·3)
	Total (*24*)	23·2 (1·0)	22·6 (1·2)	22·0 (1·3)	16·7 (8·1)
Hard items	Trial 1 (*4*)	1·0 (0·9)	1·0 (1·3)	0·5 (0·7)	0·1 (0·3)
	Trial 2	2·6 (1·3)	1·5 (1·2)	1·3 (0·8)	0·3 (0·5)
	Trial 3	3·3 (1·0)	2·4 (1·1)	1·8 (1·2)	0·7 (0·9)
	Trial 4	3·6 (0·7)	3·0 (0·9)	2·5 (0·5)	1·8 (1·1)
	Trial 5	3·6 (0·9)	2·8 (1·4)	2·5 (1·2)	1·4 (1·4)
	Trial 6	3·8 (0·8)	3·1 (1·2)	2·9 (1·5)	1·1 (1·3)
	Total (*24*)	17·9 (4·3)	13·8 (6·3)	11·2 (3·9)	5·2 (4·6)
1 h delayed recall					
Easy Items		4·0 (0·0)	3·8 (0·4)	4·0 (0·0)	2·4 (1·3)
Hard Items		2·0 (1·2)	1·9 (1·3)	1·8 (1·4)	0·7 (0·7)
24 h delayed recall					
Easy Items		4·0 (0·0)	3·8 (0·4)	3·7 (0·5)	2·4 (1·2)
Hard Items		1·9 (1·3)	0·9 (1·0)	1·1 (1·3)	0·2 (0·4)
Rey Auditory Verbal Learning Test					
List A	Trial 1 (*15*)	5·5 (1·5)	4·1 (2·1)	4·4 (1·1)	2·9 (1·5)
	Trial 2	9·1 (1·9)	7·0 (2·5)	6·5 (1·3)	4·9 (2·1)
	Trial 3	10·8 (3·0)	8·7 (3·6)	7·2 (2·0)	5·7 (2·9)
	Trial 4	12·3 (2·1)	10·4 (2·2)	9·7 (3·8)	5·9 (3·1)
	Trial 5	13·1 (1·7)	11·8 (2·6)	10·2 (3·3)	7·0 (2·7)
List B (*15*)		5·6 (1·9)	3·8 (1·4)	4·2 (2·4)	2·7 (1·0)
List A	Trial 6	10·8 (1·9)	8·5 (5·1)	7·2 (3·8)	3·1 (2·9)
1 h delayed recall		7·8 (4·4)	3·9 (4·1)	3·7 (3·4)	1·0 (2·4)
24 h delayed recall		8·3 (2·9)	2·8 (4·0)	4·2 (3·5)	0·7 (2·0)

the UD group for a variety of materials. The most sensitive measures, however, tended to be recall of word lists, whether organized or not. Recall of structured prose material, in contrast, was relatively intact.

Learning supraspan material (Table 2)

The ability to learn supraspan material was assessed by the RAVLT and the PALT. In both cases, the two groups showed clear evidence of learning ($P < 0{\cdot}001$) with the exception of the PALT easy items which approached ceiling in both groups by trial 2. The rate of learning, as indicated by the polynomial trends in the trial data, did not differ between the two groups. On the RAVLT, list B could potentially interfere with the list A material learned over the preceding 5 trials. Subsequent free recall of list A revealed significant interference ($P < 0{\cdot}001$). The size of the effect, however, did not differ in the two groups.

Delayed recall (Table 2)

Delayed recall after 1 h and 24 h was assessed

Table 3. *Results (mean and standard deviation) for the control and three depressed groups on immediate recall, cued recall and recognition of the three word lists A (uncategorized), B (categorized – unorganized) and C (categorized – organized)*

Measure	Control group ($N = 16$) Mean (S.D.)	Depressed groups: Unimpaired ($N = 10$) Mean (S.D.)	Borderline ($N = 10$) Mean (S.D.)	Impaired ($N = 8$) Mean (S.D.)
List A (uncategorized word list)				
Free Recall score (*12*)	6·6 (1·9)	4·1 (1·5)	4·3 (1·5)	3·2 (1·7)
Recognition				
Total correct (*24*)	21·6 (1·6)	19·2 (1·8)	19·9 (1·9)	16·9 (2·1)
True positive responses (*12*)	10·6 (1·2)	7·7 (1·9)	9·2 (1·4)	7·8 (2·4)
True negative responses (*12*)	11·6 (0·6)	11·5 (1·0)	10·7 (1·3)	9·1 (2·3)
False positive response (*12*)	0·4 (0·6)	0·5 (1·1)	1·3 (1·3)	2·8 (2·2)
False negative response (*12*)	1·9 (1·2)	4·3 (1·9)	2·8 (1·4)	4·2 (2·4)
List B (categorized unclustered word list)				
Free Recall score (*12*)	7·2 (2·3)	5·8 (2·3)	5·1 (1·6)	3·5 (1·3)
Cued Recall Score (*12*)	7·9 (1·2)	7·3 (2·5)	6·0 (2·0)	4·1 (2·6)
Recognition				
Total correct (*24*)	19·6 (1·7)	19·0 (2·4)	17·1 (2·9)	16·0 (2·6)
True positive responses (*12*)	10·1 (1·2)	10·1 (1·5)	9·0 (1·5)	8·1 (2·5)
True negative responses (*12*)	9·5 (1·6)	8·9 (1·7)	8·1 (2·4)	7·9 (1·9)
False negative response (*12*)	2·5 (1·6)	3·1 (1·7)	3·9 (2·4)	4·1 (1·9)
False negative response (*12*)	1·9 (1·2)	1·9 (1·5)	3·0 (1·5)	3·9 (2·5)
List C (categorized clustered word list)				
Free Recall score (*12*)	6·6 (1·5)	4·8 (1·6)	5·5 (1·8)	3·8 (1·4)
Cued Recall Score (*12*)	7·1 (1·3)	5·8 (1·2)	6·0 (1·8)	4·4 (2·3)
Recognition				
Total correct (*24*)	21·0 (1·5)	19·2 (1·9)	18·6 (1·4)	17·9 (1·6)
True positive responses (*12*)	10·5 (1·3)	9·3 (1·5)	9·0 (1·7)	8·5 (2·3)
True negative responses (*12*)	10·5 (1·2)	9·9 (1·6)	9·6 (1·9)	9·4 (1·3)
False positive response (*12*)	1·5 (1·4)	2·1 (1·6)	2·4 (1·9)	2·6 (1·3)
False negative response (*12*)	1·5 (1·3)	2·7 (1·5)	3·0 (1·7)	3·5 (2·8)

for the LM, RAVLT and PALT. Because of differences in overall performance of the groups prior to the delay, delayed recall was considered in relation to immediate recall (LM) or recall after trial 5 of the two learning tests (RAVLT and PALT). With the exception of the PALT easy items both groups recalled significantly less material after delay. For the PALT Hard items, the control group showed no further forgetting between 1 h and 24 h. The patients, however, tended to continue forgetting with increasing delay ($P < 0{\cdot}10$). A similar pattern was seen with the LM test. Although the differential effect of delay on the raw recall scores failed to reach significance, the groups showed a clear effect when percentage recall scores were considered. The delayed recall of the controls remained at over 80% of immediate recall levels, even after 24 h. The UD group's performance, however, dropped to 63·6% after 1 h and 50·5% after 24 h ($P < 0{\cdot}05$). Finally, for the RAVLT, the UD group again showed differential effect of delay ($P < 0{\cdot}01$) with a tendency for patients but not the controls to continue to forget items over the 24 h delay period.

Cued recall and recognition (Table 3 and Figs 2a, b)

Cued recall was assessed for the categorized lists A and B immediately after free recall. Overall cueing led to a significant increase in the number of items recalled, regardless of list and in both groups. There were no differential effects of either list or group on the cueing effect.

Recognition performance was assessed for all three lists, A, B and C. Analysis of the recognition data was restricted to the 'hits' (true positive responses) and 'false alarms' (false negative responses) as, with an equal number of target items and distracters, these provide a

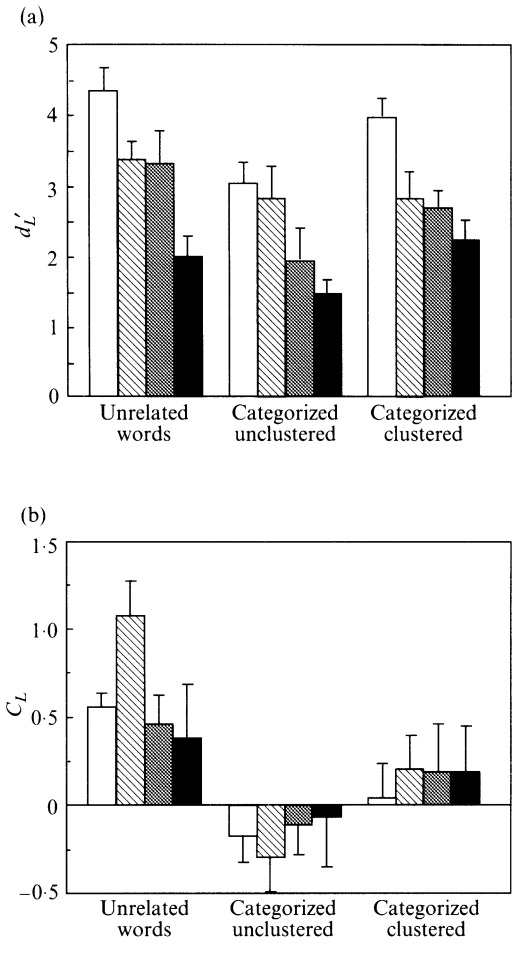

FIG. 2. Mean (and S.E.) signal detection parameters d_L' (a) and C_L (b) for the three word lists, measured in the control group (□) and the three depressed groups (▨, UD; ▩, BD; ■, ID).

complete description of recognition performance. Considering the two response types separately revealed a significant difference between the groups for hits ($P < 0{\cdot}01$) but not for false alarms, with UD patients identifying fewer items. In addition, for the hits, there was a significant group by list interaction ($P < 0{\cdot}05$). Planned contrasts revealed that the source of the interaction lay in the difference between the groups for unrelated versus categorized lists ($P < 0{\cdot}05$), rather than between the clustered and unclustered categorized lists.

Employing the procedure recommended by Corwin *et al.* (1990), two independent parameters d_L' and C_L were derived from the hit and false alarm results interpretable as sensitivity and response bias. A d_L' of zero is equivalent to chance responding. Increasingly positive values correspond to increasing ability to discriminate previous items from novel items, whereas negative values correspond to worse than chance performance. For C_L, a value of zero implies no bias. Increasing positive values imply an increasingly conservative response bias (saying that the item is novel) while increasing negative values imply an increasingly liberal bias (saying that the item is one presented previously).

The UD group showed a significantly lower d_L' (Fig. 2a) ($P < 0{\cdot}05$). Across the two groups, d_L' differed significantly between lists ($P < 0{\cdot}01$), with a significantly higher sensitivity for the unrelated list than for the categorized lists ($< 0{\cdot}01$). The group by list interaction was not significant. In contrast to the effect for sensitivity, response bias (C_L) (Fig. 2b) did not differ between the two groups, although the data in Fig. 2b suggests an increase in bias for the unrelated list A. Across groups, there was a strong effect of list on bias ($P < 0{\cdot}001$). Univariate analyses revealed that the bias was significant and positive for List A ($P > 0{\cdot}001$), non-significant (i.e. zero bias) for list C, and with a trend towards a negative bias for List B ($P < 0{\cdot}10$). However, this pattern of bias was the same in the two groups. Thus, the recognition performance of the UD patients was characterized by a significant decrease in sensitivity (i.e. ability to discriminate previously presented items from novel items), but with no abnormality in response bias.

Short-term memory (Brown-Peterson task) (Fig. 3)

The UD group recalled significantly fewer trigrams overall ($P < 0{\cdot}05$), with a significant group by delay interval interaction ($P < 0{\cdot}05$). *Post hoc* analysis revealed that the two groups did not differ for immediate recall (0 s delay), with both groups achieving near perfect performance (maximum score = 15). Significant differences ($P < 0{\cdot}001$) between the groups, however, were found at all other delay intervals. Analysis of intervals 5–30 s confirmed the significant main effects of group and interval. The interaction between these two effects, however, was no longer significant. Thus, the patients showed a deficit in recall after delay

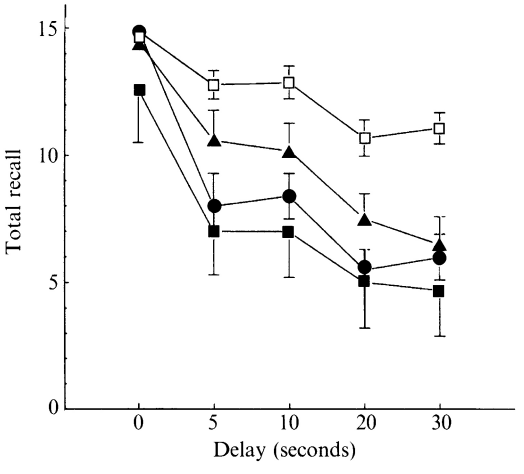

FIG. 3. Mean (and S.E.) recall score (maximum = 15) for the 5 delay intervals of the Brown–Peterson test, measured in the control group (□) and the three depressed groups (▲, UD; ●, BD; ■, ID).

compared to the no-delay condition, but the degree of impairment was unrelated to the duration of the delay interval.

Language and executive function (Table 4)

In the verbal fluency tests the patients generated fewer words overall ($P < 0{\cdot}01$). In both groups and to a similar extent, most words were generated in the free condition and least in the letter 's' condition. All subjects generated more words in the first 30 s of each condition than in the second ($P < 0{\cdot}001$). However, this effect was the same in the two groups.

On the Token test, the control subjects made relatively few errors, even with the strict scoring (one attempt). Half (8/16) made 1 error, and 1 subject made 2 errors. The UD group made significantly more errors (Mann–Whitney U test, $P < 0{\cdot}01$) with 5/10 making 1 or 2 errors and 3/10 making 3 or more errors (maximum 9). With the lenient scoring criterion 2/16 controls made a single error, while 1 patient made a single error and 2 made 2 errors. Performance by this criterion did not distinguish the groups.

No significant deficit was found in the UD group on the Weigl sorting test. All of the control subjects and 9/10 of the patients were able to sort the tokens successfully according to both methods of classification (shape and colour).

Comparison of three depressed samples

Age, sex, years of education and reading ability and level of depression (Table 1)

The three depressed groups (UD, BD and ID) did not differ in mean age, years of education, Hamilton score or sex ratios. Two of the 3 bipolar patients were in the UD group and 1 in the BD group. A significant difference was found between the groups for the Schonell reading test ($P < 0{\cdot}05$). *Post hoc* comparisons revealed only that the UD group scored higher than the ID group.

CAMCOG and MMSE (Table 1)

There was a highly significant difference between

Table 4. *Results (mean and standard deviation) for the control and three depressed groups on the verbal fluency tasks*

Measure	Control group ($N = 16$) Mean (S.D.)	Depressed groups: Unimpaired ($N = 10$) Mean (S.D.)	Borderline ($N = 10$) Mean (S.D.)	Impaired ($N = 9$) Mean (S.D.)
Free condition				
First 30 s	15·1 (2·2)	14·5 (2·2)	13·4 (5·5)	10·3 (3·5)
Second 30 s	12·9 (3·0)	11·2 (2·1)	10·9 (2·8)	6·8 (3·0)
Total (60 s)	28·0 (3·5)	25·7 (3·3)	24·3 (6·9)	17·1 (5·3)
Boy's names				
First 30 s	12·2 (2·6)	9·9 (2·4)	10·0 (3·0)	6·9 (2·0)
Second 30 s	8·3 (3·2)	6·3 (2·0)	5·9 (1·4)	3·7 (2·2)
Total (60 s)	20·5 (5·2)	16·2 (3·2)	15·9 (3·9)	10·6 (2·8)
Letter S				
First 30 s	11·5 (2·5)	9·4 (2·5)	6·9 (3·4)	5·7 (2·7)
Second 30 s	7·9 (3·0)	6·4 (1·6)	5·5 (4·4)	3·0 (1·9)
Total (60 s)	19·4 (4·9)	15·8 (3·6)	12·4 (7·4)	8·7 (4·2)

the groups in total CAMCOG score ($P < 0{\cdot}001$), with each group differing significantly from each other. For the MMSE total, however, the UD and BD groups did not differ significantly, a reflection, perhaps, of the lower ceiling of the MMSE and decreased sensitivity to identify mild levels of impairment. Univariate statistics revealed that the three groups differed significantly on all CAMCOG subscales. *Post hoc* comparisons revealed that the ID group performed significantly worse than the UD group on all scales, and worse than the BD group for Orientation, Memory and Calculation. In contrast, the UD and BD groups did not differ significantly on any scale. The groups differed on the serial sevens item (UD group mean = 4·8, S.D. = 0·4; BD group mean = 3·5, S.D. = 1·1; ID group mean = 2·2, S.D. = 2·0) ($P < 0{\cdot}01$), with the ID group significantly worse than the UD group.

WAIS – Verbal IQ and subscale scores (Table 1)

There was a significant difference between the pro-rated Verbal IQ of the three groups ($P < 0{\cdot}01$), with the UD group having a higher verbal IQ than the others. The BD and ID groups did not differ. MANOVA with all subscale scores, revealed a significant overall effect of group ($P < 0{\cdot}01$). Univariate tests revealed that the groups differed significantly on all subscales ($P < 0{\cdot}01$) except digit span. *Post hoc* comparisons between the groups on the scales with significant F statistics revealed that the UD and ID groups differed significantly in each case. Furthermore, the UD group performed significantly better than the BD group on three of the tests: Vocabulary, Similarities and Arithmetic, but not on Comprehension.

Memory

Immediate recall (Table 2)

On the recall of the LM passages, a significant difference was shown between the groups ($P < 0{\cdot}05$), with the ID group recalling significantly less than the UD group. On the first trial of the RAVLT, no clear or significant group differences were found, although in this instance, there was a tendency for the BD group to recall more than the UD group. A similar pattern was shown for digit span (UD mean = 8·2, S.D. = 3·7; BD mean = 8·7, S.D. = 2·9; ID mean = 6·0, S.D. = 2·1), and the recall of two of the separate word lists (A and C). No significant group differences were found, however, except for the List (B) with the UD group recalling more items than the ID group. Thus, the recall of prose material but not word lists showed a clear gradient of performance with increasing overall cognitive impairment.

Learning supraspan material (Table 2)

In general, the three groups could not be distinguished statistically for rate of learning on either the RAVLT and PALT tests. Clear evidence of learning was demonstrated in all cases apart from the easy items on the PALT. Although not significant, the UD group tended to improve performance most with practice in the learning tests and the ID group least.

Delayed recall (Table 2)

Overall the tests failed to reveal any differential effect between the groups of delay recall performance relative to immediate recall. For the LM test, all groups showed a decrease in recall from immediate to 1 h delay. Overall, however, there was no significant decrease over the next 24 h. The effect of delay on recall did not differ between the groups. Considering delayed recall as a percentage of immediate recall, there was no significant differences in performance between groups. For the PALT the effect of delay on recall was examined for the easy and hard items together. Across trial type (easy or hard) and recall delay, there was a significant difference between the groups ($P < 0{\cdot}001$), with the ID patients recalling less than the other two groups. However, none of the two- or three-way interaction involving group were significant. Finally, for the RAVLT there was a significant decrease in recall from trial 5 (immediate recall) to 1 h delayed recall ($P < 0{\cdot}001$), but not from 1 h to 24 h ($P > 0{\cdot}10$). Once again there was no significant difference in the performance of the three groups across the two delay intervals.

Cued recall and recognition (Table 3 and Figs 2a, b)

Comparing the free and cued recall performance of lists B and C for the three groups showed a significant effect of cueing ($P < 0{\cdot}05$) with more

being recalled in the cued condition than in the free condition. However, the effect of cueing was similar in the three groups and for the two lists.

For the recognition data attention was limited to hits and false alarms. For the hits, there was no significant difference, overall, between the three groups. The effect of list (A, B, C) approached significance ($P < 0{\cdot}10$) with a lower hit rate for the unrelated list (A) than for the categorized lists (B and C) ($P < 0{\cdot}05$), but with no difference between the latter two lists. The group by list interaction also approached significance ($P < 0{\cdot}10$). Subsequent one-way ANOVAs revealed no significant differences between the three groups for lists A or C, while the result for list B approached significance ($P < 0{\cdot}10$) with the UD group tending to perform best and the ID group worse. Of the three groups only the UD group showed a difference in hits rates between the three lists ($P < 0{\cdot}10$) with the effect attributable to a lower hit rate in the unrelated list A than in the two categorized lists ($P < 0{\cdot}05$). For the false alarms there was a trend, across all three lists for the UD group to make the fewest number of false positive responses and the ID group the most. However, neither the group effect nor the group by list interaction were significant.

Of the signal detection parameters, there was a significant difference in the sensitivity (d_L') between the groups ($P < 0{\cdot}01$) (Fig. 2a), with the UD and ID groups differing significantly ($P < 0{\cdot}001$). Overall there was a significant effect of list on sensitivity ($P < 0{\cdot}05$), but this list effect were similar in the three groups. In contrast to sensitivity, bias (C_L) did not differ between the groups when averaging across the three lists (Fig. 2b). A highly significant effect of list was shown ($P < 0{\cdot}001$), with significant differences between both random and categorized lists ($P < 0{\cdot}001$) and between the two categorized lists ($P < 0{\cdot}05$). In addition, the group by list interaction was significant ($P < 0{\cdot}05$). Planned contrasts revealed that this interaction was due to the pattern of response by the groups to the random *v.* categorized lists ($P < 0{\cdot}05$) rather than between the two categorized lists ($P < 0{\cdot}10$). The main finding, seen in Fig. 2b, was the higher positive bias of the UD group for the random list, while the three groups did not differ for lists B and C.

Short-term memory (Brown-Peterson task) (Fig. 3)

While there was a significant effect of delay on recall ($P < 0{\cdot}001$), there was no overall difference between the groups and no group by delay interaction. Thus, all subjects recalled less with increasing recall delay.

Language and executive function (Table 4)

Across the three verbal fluency conditions, there was a significant difference between the groups ($P < 0{\cdot}01$), with the main difference between the ID and the other two groups. However, this was the only significant difference between the groups. All groups generated most words in the 'free' condition relative to the other two ($P < 0{\cdot}001$), and in the first 30 s of each condition relative to the second ($P < 0{\cdot}001$).

On the token test only one subject (in the impaired group), made more than two errors with the lenient scoring criterion. Errors with the strict criterion, however, were shown by subjects in each of the groups. Adopting the definition of impairment derived from the control group data (2 or more errors in the strict scoring criterion), three of the UD depressed and BD depressed groups were impaired on the token test. In contrast, almost all (8/9) of the ID group, had an abnormal performance on this test ($P < 0{\cdot}05$).

The ID groups also tended to have problems on the Weigl test. Of the UD group, 9/10 were able to sort by two categories successfully, as were 8/10 of the BD group. However, only 4/9 of the ID group were able to sort the shapes by more than one category. This difference approached significance ($P < 0{\cdot}10$).

Associations between clinical measures and neuropsychological measures

Because of the large number of neuropsychological measures, the data were first subjected to a principle components analysis with varimax rotation. A set of variables were chosen after examining the correlation matrix to eliminate redundant variables. The final set of variables produced a 5 factor solution accounting for 72% of the variance. Factor scores were correlated with total Hamilton score, the two factor scores derived from the SADS, and the

Table 5. *Summary of previous studies on cognitive function in depressed patients*

Task	Significant impairment shown	No significant impairment shown	Note
Immediate free recall			
Prose passages	Breslow *et al.* 1980 Hart *et al.* 1987*a*, *b* Watts & Cooper, 1989		
Word lists, paired associates	Sternberg & Jarvik, 1976 Weingartner *et al.* 1981 Calev *et al.* 1986 Roy-Byrne *et al.* 1986 Wolfe *et al.* 1987 Golinkoff & Sweeney, 1989 Watts *et al.* 1990		
Effect of semantic organization		Watts *et al.* 1990 Weingartner *et al.* 1981	
Learning supraspan material	Wolfe *et al.* 1987 Golinkoff & Sweeney, 1989		
Delayed recall	Sternberg & Jarvik, 1976	Weingartner & Silberman, 1982 Kopelman, 1986 Wolfe *et al.* 1987	
Recognition memory			
Hit rate (true +ve)	Miller & Lewis, 1977 Dunbar & Lishman, 1984* Cole & Zarit, 1984 Calev & Erwin, 1985 Watts *et al.* 1987 Wolfe *et al.* 1987 Golinkoff & Sweeney, 1989	Dunbar & Lishman, 1984*	*Impairment only for words of positive hedonic tone
False alarm rate (false -ve)	Frith *et al.* 1983 Watts *et al.* 1987* Wolfe *et al.* 1987†	Miller & Lewis, 1977 Dunbar & Lishman, 1984 Calev & Erwin, 1985 Watts *et al.* 1987* Wolfe *et al.* 1987†	*Increase in vocalization condition, decrease in silent reading condition †Impairment in unipolar patients but not in bipolar
Sensitivity	Dunbar & Lishman 1984* Watts *et al.* 1987 Corwin *et al.* 1990†	Miller & Lewis, 1977 Dunbar & Lishman, 1984* Corwin *et al.* 1990†	*Normal for neutral words, increased for negative words and decreased for positive words †Impaired only in more severely depressed patients
Response bias	Watts *et al.* 1987	Miller & Lewis, 1977* Dunbar & Lishman, 1984* Corwin *et al.* 1990*	*All showed more positive (conservative) bias
Short-term memory			
Digit span	Breslow *et al.* 1980		
Brown–Peterson paradigm	Cohen *et al.* 1982*	Cohen *et al.* 1982*	*Impairment only in 'severely depressed' patients
Language and executive function			
Conceptual ability	Savard *et al.* 1980*	Savard *et al.* 1980* Hart *et al.* 1987*a*†	*Halstead Categories Test – impaired only in bipolar patients †Wisconsin Card Sorting Test
Verbal fluency	Robertson & Taylor, 1985 Hart *et al.* 1987*a* Wolfe *et al.* 1987*	Wolfe *et al.* 1987*	*Impairment only in bipolar patients

individual SADS subscale scores. Of the 70 paired associations, only one was significant at the 0·01 level, close to chance level.

Effect of antidepressant medication on cognitive function

Of the total sample of depressed patients, almost half (14/29) were taking antidepressant medication at the time of assessment. This same ratio was represented in each of the three depressed subgroups (UD: 5/10; BD: 5/10; ID: 4/10). Overall the medicated ($N = 14$) and unmedicated patients ($N = 15$) did not differ on the set of variables age, years of education, Hamilton

score or CAMCOG total. To determine whether the presence or absence of medication had any effect on neuropsychological function, independent of overall level of cognitive impairment, a representative set of 21 neuropsychological measures were entered into a MANOVA (Logical memory: immediate recall and 1 h delayed recall as % of immediate recall. Rey AVLT: total score and 1 h delayed recall; PALT total easy and total hard; Brown-Peterson recall after 0 and 30 s; free recall and recognition score for word lists A, B and C and cued recall for lists B and C; total Verbal fluency score; Token test: strict and lenient score). The results revealed no significant difference between the medicated and unmedicated depressed groups. Independent univariate F tests revealed no significant differences between the groups for any measure, with the F score in all but two instances being less than 1. Thus, the medication status of the patients appeared to have no impact on their overall level of cognitive function, or on their performance on the specific neuropsychological tests.

DISCUSSION

The first empirical question addressed by this study was the distribution of overall intellectual function in a sample of depressed patients. Specifically, was there any evidence for a distinct subgroup of demented patients. The data revealed that CAMCOG scores were unimodally distributed with an extended 'tail' towards the impaired end of the continuum (Fig. 1). The classification of patients as 'impaired' or 'unimpaired', therefore, must be based on cut-offs derived from a normative sample.

Methodologically, the description of the 'unimpaired' group was simply a matter of taking the range of values obtained from a normal control group. Less straightforward was the classification of the remaining patients. As a major concern of the study was the issue of dementia in depression it was not sufficient simply to take the cut-off provided by the minimum control group score on the CAMCOG. Although all patients scoring less than this would lie outside of the 'normal' range, this is not the same as saying that they had a cognitive impairment of sufficient severity to warrant a diagnosis of dementia. Roth *et al.* (1986) recommended a cut-off of 80 as providing acceptable levels of both sensitivity and specificity for their mixed population of patients. In the present study a cut-off of 81 was employed instead as this had the practical advantage of dividing the remaining patients into two approximately equal halves. Thus, we obtained a sample of patients classified as 'impaired' who constituted 31% of the total depressed sample. Although one might label such patients as 'demented', we will continue with the label 'impaired' to avoid the implication that we have provided a prevalence figure for dementia in depression.

Having identified three subgroups of patients we could then turn to the second question. How normal were the unimpaired (UD) patients? Were there deficits in specific aspects of cognition, despite globally defined intellectual function within the normal range? The results revealed clearly that the UD group were impaired on a wide range of measures. Those most sensitive to the presence of depression included deficits in recall memory, particularly after a delay, aspects of recognition memory, short-term memory, verbal fluency and language comprehension. How does this pattern of results compare with that reported elsewhere in the literature? The aspects of cognitive function under consideration are shown in Table 5, together with the results of the main published studies.

Considering first long-term memory and learning. The UD patients showed no significant impairment on the immediate recall of the 'easy' items of the PALT or for the prose passages of the LM test. While the absence of impairment for the PALT may be due to ceiling effects, the same explanation cannot account for the results for the LM test. The lack of any significant impairment on this test contrasts with other studies which have found impaired immediate memory for prose passages in depressed samples. Inadequate power of the statistical test may partly account for the lack of a significant result in the present case. In addition, however, the previous studies all used samples unselected for the severity of their overall level of cognitive function. Our results suggest that immediate recall of prose material may be a relatively

insensitive measure of memory dysfunction in intellectually intact depressed patients. However, other measures appear to be more sensitive, and showed significant effects even with the present sample size and in globally intact subjects. Impaired performance was observed in the recall of supra-span word lists (RAVLT and the various random, organised and categorized word lists) and for 'hard' paired associates. In these cases, the findings concur with the majority of previous studies (see Table 5).

Also assessed was the effects on recall of semantic categorization and clustering of the stimulus material. It was shown that these effects were normal, a similar result to that shown by Watts *et al.* (1990). Weingartner *et al.* (1981) found categorization and clustering of the material at input aided patients to the point that their performance was no longer significantly impaired. In the present study, however, performance remained impaired, even with organisation of the material.

Distinct from the ability to recall supra-span material after a single presentation, is the ability to learn that material with repetition. Verbal learning was assessed by the RAVLT and the PALT. In both tests, although overall performance was poorer (with the exception of the 'easy' pairs), the *rate* of learning was relatively normal in the UD group. Surprisingly little data exists on this aspect of memory in the literature, and that which exists suggests impairment. Although it is difficult to generalize from such a small set of findings, the present study suggests that this impairment in verbal learning with repetition (as opposed to immediate recall), may not be typical of all depressed patients.

Although many studies have examined the immediate recall of verbal material, surprisingly few have assessed recall after a delay. Those that have provide no consistent pattern. In the present study, however delayed recall, over both 1 h and 24 h, was impaired for all material with the exception of the 'easy' paired associates. Even LM which showed no significant deficit in immediate recall, showed an impairment in the UD patients after a delay. This suggests that delayed recall may provide a more sensitive index of mnestic capacity in depressed subjects.

The final aspect of long-term memory examined in the present study was recognition. Evaluation of the various studies employing recognition paradigms in depressed samples is complicated by differences not only in the basic paradigm, but also in the indices of performance employed, all of which assess a different aspect of recognition memory. However, a majority of studies report the hit rate (true positive responses), and almost all indicate impaired performance in depressed subjects. The present study suggests that the level of structure may be important, with a significant decrease in hit rate being observed only for a list of unrelated words. Where the lists were made up of items from a small number of categories, no impairment in hit rate was observed.

The consensus in the literature on hit rate is not shared by other recognition memory parameters. Of those studies that report the false alarm rate (false positive responses) some find a significant increase, some show a decrease, while others, including the present study, find no significant difference. Only a few studies have employed signal detection analysis as a way of quantifying recognition performance. Sensitivity provides the main index of mnestic ability for recognition performance. Across studies the results appear to depend upon the nature of the material or the level of depression of the sample. The present study, however, provided clear evidence of impairment, unrelated to depression severity and in all lists assessed. An equally important aspect of memory function is response bias, i.e. their willingness to commit themselves to a decision about whether they have seen a test item before. A more conservative response bias, if a general characteristic of performance, might cause subjects to perform less well on all memory tasks not because the memories are less accessible but because the subjects lack confidence in them. Several studies have shown a more significant positive (i.e. conservative) bias in their depressed samples. These findings contrast with those of the present study, and those of Watts *et al.* (1987) which found no significant change. However, from Fig. 1 it can be seen that there was at least a trend for the UD group to show a greater level of positive bias for the list made up of unrelated words.

It seems likely that, whatever their theoretical value, the two signal detection parameters of recognition memory performance are highly sensitive to differences in patient sample, material and possibly method of testing. Without

being able to generalize, therefore, we can say only that our sample of 'unimpaired' depressed subjects showed a decreased sensitivity, with increasing impairment across the patient groups with increasing overall cognitive dysfunction. In contrast, response bias was generally normal, at least for lists comprized of categorized words. This finding on bias is important if the results of the other memory tests are to be interpreted in terms of mnestic ability rather than as an artefact of a more conservative strategy in the sample being studied.

The discussion, to this point, has concentrated on aspects of long-term memory function. In contrast to the large number of studies on this facet of memory there has been virtually no systematic attention paid to short-term memory, i.e. the retention of small amounts of information over durations measurable in seconds. At the simplest level, it has been assessed using measures of memory span for digits. Breslow *et al.* (1980) reports a significant decrease in span, a result which was not replicated in the present sample in the UD group.

A second approach to assessing short-term memory has been to assess recall after short delays during which time rehearsal is prevented using the Brown-Peterson paradigm. In the present study the UD group showed an impairment in recall after a delay relative to immediate recall. The degree of deficit, however, was not delay dependent suggesting that the main factor was the introduction of a delay *per se* or the presence of the backward counting task employed during the delay period. This latter effect might suggest that the patients have an increased susceptibility to interference from a second task. The absence of any differential interference effect on the RAVLT test seems to argue against this possibility. However, a critical difference between the two tasks is that the Rey material had been learned, through repetition, prior to the interference task, whereas in the short-term memory task the interference task occurred after a single presentation of the stimulus material and prevented any rehearsal. In this respect, the deficit on short-term memory task may be a better indicator of the types of difficulties that patients may have with registering new information in every-day life. Only one other study has utilized the Brown-Peterson paradigm with a depressed sample. Cohen *et al.* (1982) found normal performance, both immediately and after delay, for a group of 3 'moderately depressed' patients (mean Hamilton Depression Rating 21·9). A group of 5 'severely depressed' patients (mean Hamilton Depression Rating 44·6) showed impaired performance with greater deterioration with increasing delay interval. These findings contrast with our own, at least with regard to the performance of the moderately depressed group with depression levels similar to our own UD sample. However, the sample size of only 3 patients makes the results of Cohen and colleagues difficult to evaluate.

The final aspects of cognitive function assessed were language and executive function. Conceptual ability was assessed by the Weigl test, and showed no significant impairment in the UD group. Although there are few studies on this aspect of cognition, problems with conceptual ability does not appear to be characteristic of patients with depression. However, a different picture is observed for language function and particularly verbal fluency. In the present study, despite the UD and control groups being matched for overall CAMCOG score, the patients showed significantly impaired function on both the expressive language and comprehension subscales. Both subscales include a number of different aspects of language related tests. The comprehension subscale involves carrying out motor responses to command (e.g. 'Tap each shoulder twice with two fingers keeping your eyes shut') and answering questions (e.g. 'Was there radio in this country before television was invented?'). The deficit on the comprehension subscale is consistent with the impairment shown by the UD group on the Token test. An important result on this latter test, however, was that the patients were impaired only when attempting to carry out the command on the first attempt. After repetition of the command, no significant deficit was apparent. Such a pattern of performance seems to argue against any basic deficit in linguistic function. Rather, the poor performance may be more reasonably attributed to difficulty with registering the command rather than comprehending it. The deficit, therefore, may be one of attention or short-term memory.

Expressive language on the CAMCOG was assessed by defining words, naming pictures,

verbal fluency (animals in 60 s), sentence repetition ('no ifs ands or buts') and writing name and address. Once again, the task of sentence repetition may be sensitive to deficits in attention. One of the main reasons for the overall deficit on the Expressive language subscale, however, was poor performance of the UD group on the verbal fluency task, which contributes almost half of the total subscale score. A deficit in verbal fluency was found not only in the CAMCOG, but also when tested separately for the free generation of words, and for words beginning with the letter 's', and boy's names. Similar findings are reported in other studies. The cause of this fluency deficit, however, is less clear. If it represents a deficit in access to semantic memory one might expect to find some variation in the level of impairment dependent upon the particular demands of the task. In the present study, however, the patients appeared to show a constant deficit regardless of the precise nature of the fluency condition. Alternatively, the slowness in word generation may be a non-specific feature of the slowness found in patients with depression in their movement and spontaneous speech and gesture ('psychomotor retardation'). Unfortunately, no other time-dependent tasks were administered in the present study, making it difficult to determine the degree to which motor or 'cognitive' slowness is a general feature of the sample of patients studied.

To summarize, we have demonstrated cognitive dysfunction on a range of measures in a sample of depressed patients who show no global impairment. This provides clear evidence that cognitive dysfunction is a real feature of depressive illness, and that previous studies have not been reporting averaged group effects biased by the influence of a proportion of patients with more severe global deterioration. The deficits on which the UD group were impaired were, with only a few exceptions, consistent with the existing literature on cognitive function in depression. In addition, however, the results point to some important aspects of impairment not generally investigated, aspects which may relate more closely to 'real-life' memory problems reported by patients. These relate to poor performance on language comprehension tests (following a complex set of instructions), and poor short-term memory when rehearsal is prevented. It would be important for future research to determine the nature of the processes underlying these deficits and their relationship to the more pervasive impairments in long-term memory.

While it is important to be able to identify cognitive dysfunction in depressed patients relative to controls, it is equally important to distinguish degrees of dysfunction within a depressed sample. Specifically, which measures of cognitive function are most sensitive to cognitive decline and which remain relatively constant? An answer to this question may have important clinical implications for the choice of tests to identify demented depressed patients or to monitor change with treatment or progression of the impairment. Comparisons between the three depressed groups was made easier by the fact that they were matched for age and years of education, Hamilton score and the proportions receiving antidepressant medication. A significant difference was found, however, for reading ability as measured by the Schonell reading test. The UD group performed best, the ID group worse, and the BD group intermediate. This finding raises the possibility that the global differences between the groups represent simply a difference in pre-morbid cognitive function. A number of facts, however, argue against this possibility, and against the idea that reading age is an unbiased indicator only of premorbid intelligence in a patient group. First, even using the Schonell score as a covariate, there was a highly significant difference between the total CAMCOG score of the three groups ($P < 0{\cdot}001$). This probably reflects the fact that the CAMCOG is not an IQ test, and that even normal subjects with low IQ's will still score outside the range characteristic of patients with cognitive impairment. Secondly, total CAMCOG score is significantly correlated with reading ability only in the patient group ($r = 0{\cdot}79$, $P < 0{\cdot}001$). No such association exists in the control group ($r = 0{\cdot}27$, $P > 0{\cdot}10$) suggesting that the association in the patients is due to the presence of some additional factor in the patient group to which both measures are related. Thirdly, one reason for a low reading score on the Schonell might be a failure of the subjects to attend to the precise spelling of the more difficult words. As most of these are phonetically regular, an attentive and motivated subject could produce an accurate pronunciation. Consistent with

such an explanation, reading ability loaded on the same factor in the principal components analysis as other attention related tasks such as serial sevens, the Token Test as scored by the strict criteria and delayed recall on the Brown-Peterson task. Even if the influence of Verbal IQ is taken into account, reading ability and serial sevens still have a partial correlation of 0·60 ($P < 0·001$). Finally, although the data are not presented here, a follow-up of the patients after recovery of their depression showed a significant improvement not only in overall cognitive function but also in reading ability. Together these findings indicate that the deficit in reading ability in the ID patients may be considered a facet of their cognitive impairment, probably related to attention, and not a trait indicator of lower pre-morbid intelligence.

The data reported here suggests that the large majority of tests revealed a gradient of function with the UD group performing best, the ID group worst, and the BD group intermediate. This pattern confirms the conclusion that cognitive dysfunction in depression is continuously distributed with no evidence for a discretely defined demented subgroup. There was considerable variation between the tests, however, in the degree to which the three groups could be distinguished statistically. This may relate to the tests themselves (e.g. ceiling effects) or the variability in performance within the depressed sample. In any event, it is probably unwise to draw firm conclusions from the specific pattern of tests which show significant group differences and those which do not.

With the present sample sizes the three patient groups could be distinguished on the immediate recall of supraspan material, particularly the prose passage of the LM test. In contrast, on this and other tests, the results failed to reveal any systematic and significant difference between the three groups for the rate of new learning, the effect of cueing on recall, and effect of delay on recall. Of the various recognition memory parameters, sensitivity showed the clearest group differences. Importantly, significant group effects were non limited to memory function. The depressed groups could also be distinguished on a number of tests which have a large attentional component. These included serial sevens, the ability accurately to follow a complex verbal command, and the Weigl sorting test. In addition, a significant gradient of impairment was found for verbal fluency. An important question is the degree to which these various aspects of cognitive impairment represent the effect of a single common factor in contrast to a number of independent factors. In particular, a primary attentional deficit may have important implications for a wide range of tests, including tests of memory, particulary the immediate recall of supraspan material. In contrast, attentional factors may be expected to have less impact on, for example, delayed recall where memory function *per se* may be more important. Unfortunately, the role of attention in cognitive function in depression has not been the subject detailed investigation in the literature to date. The present findings, with the relatively crude measures employed, can provide only a preliminary suggestion of its importance. However, these behavioural data are consistent with the evidence from Positron Emission Tomography (PET) on the cohort of patients from which the present sample was drawn (Bench *et al.* 1992). Regional cerebral blood flow (rCBF) in the depressed patients was significantly reduced in the regions of the anterior cingulate cortex and left dorsolateral prefrontal cortex compared to controls, brain areas implicated more with attention and behavioural regulation than with memory function. These findings suggest that a thorough investigation of attentional function in depression is justified, both in its own right and to determine its relationship to memory impairment.

Finally, one unexpected finding from the present study was the absence of significant relationship between cognitive function and any index of depression relating to severity, symptomatology or treatment. The three depressed groups, with global levels of impairment ranging from normal to 'demented', had identical mean Hamilton scores. This provides powerful evidence that cognitive dysfunction cannot be considered as an epiphenomenon of depressive symptomatology as has been suggested by some investigators. It further suggests that we need to look for different neurobiological explanations for the affective and the cognitive features of depression (Dolan *et al.* 1992).

L.C.S. and C.J.B. were supported by a grant from the Wellcome Trust.

REFERENCES

Bench, C. J., Friston, K. J., Brown, R. G., Scott, L. C., Frackowiak, R. S. J. & Dolan, R. J. (1992). The anatomy of melancholia – focal abnormalities of cerebral blood flow in major depression. *Psychological Medicine* **22**, 607–615.

Bench, C. J., Friston, K. J., Brown, R. G., Frackowiak, R. S. J. & Dolan, R. J. (1993). Regional cerebral blood flow in depression measured by Positron Emission Tomography: the relationship with clinical dimensions. *Psychological Medicine* **23**, 579–590.

Braff, D. L. & Beck, A. T. (1974). Thinking disorder in depression. *Archives of General Psychiatry* **31**, 456–459.

Breslow, R., Kocsis, J. & Belkin, B. (1980). Memory deficits in depression: evidence utilizing the Wechsler Memory Scale. *Perceptual and Motor Skills* **51**, 541–542.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology* **10**, 12–21.

Calev, A. & Erwin, P. G. (1985). Recall and recognition in depressives: use of matched tests. *British Journal of Clinical Psychology* **24**, 127–128.

Calev, A., Korin, Y, Shapira, B., Kugelmass, S. & Lerer, B. (1986). Verbal and non-verbal recall by depressed and euthymic affective patients. *Psychological Medicine* **16**, 789–794.

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.

Cole, K. D. & Zarit, S. H. (1984). Psychological deficits in depressed medical patients. *Journal of Nervous and Mental Diseases* **172**, 150–155.

Cornell, D. G., Suarez, R. & Berent, S. (1984). Psychomotor retardation in melancholic and non-melancholic depression: cognitive and motor components. *Journal of Abnormal Psychology* **93**, 152–157.

Corwin, J., Peselow, E., Feenan, K., Rotrosen, J. & Fieve, R. (1990). Disorders of decision in affective disease: an effect of β-adrenergic dysfunction. *Biological Psychiatry* **27**, 813–833.

Cummings, J. L. (1986). Subcortical dementia. Neuropsychology, neuropsychiatry and pathophysiology. *British Journal of Psychiatry* **149**, 682–697.

Cummings, J. L. (1989). Dementia and depression: an evolving enigma. *Journal of Neuropsychiatry* **1**, 236–242.

Cummings, J. L. & Benson, D. F. (1984). Subcortical dementia. Review of an emerging concept. *Archives of Neurology* **41**, 874–879.

Dolan, R. J., Bench, C. J., Brown, R. G., Scott, L. C., Friston, K. J. & Frackowiak, R. S. J. (1992). Regional cerebral blood flow abnormalities in depressed patients with cognitive impairment. *Journal of Neurology, Neurosurgery and Psychiatry* **55**, 768–773.

Dunbar, G. C. & Lishman, W. A. (1984). Depression, recognition-memory and hedonic tone. A signal detection analysis. *British Journal of Psychiatry* **144**, 376–382.

Endicott, J. & Spitzer, R. L. (1978). A diagnostic interview. The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry* **35**, 837–844.

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.

Frith, C. D., Stevens, M., Johnstone, E. V., Deakin, J. F. W., Lawler, P. & Crow, T. J. (1983). Effects of ECT and depression on various aspects of memory. *British Journal of Psychiatry* **142**, 610–617.

Golinkoff, M. & Sweeney, J. A. (1989). Cognitive impairments in depression. *Journal of Affective Disorders* **17**, 105–112.

Hachinski, V. C., Iliff, L. D., Zilkha, E., Du Boulay, G. H., McAllister, V. L., Marshall, J., Ross Rissell, R. W. & Symon, L. (1975). Cerebral blood flow in dementia. *Archives of Neurology* **32**, 632–637.

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

Hart, R. P., Kwentus, J. A., Taylor, J. R. & Harkins, S. W. (1987*a*). Rate of forgetting in dementia and depression. *Journal of Consulting and Clinical Psychology* **55**, 101–105.

Hart, R. P., Kwentus, J. A., Wade, J. B. & Hamer, R. M. (1987*b*). Digit symbol performance in mild dementia and depression. *Journal of Consulting and Clinical Psychology* **55**, 236–238.

Hasher, L. & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology* **108**, 356–388.

Jorm, A. F. (1986). Cognitive deficits in the depressed elderly: a review of some basic unresolved issues. *Australian and New Zealand Journal of Psychiatry* **20**, 11–22.

Kiloh, L. G. (1961). Pseudo-dementia. *Acta Psychiatrica Scandinavica* **37**, 336–351.

Kopelman, M. D. (1986). Clinical tests of memory. *British Journal of Psychiatry* **148**, 517–525.

McAllister, T. W. (1981). Cognitive functioning in the affective disorders. *Comprehensive Psychiatry* **22**, 572–586.

Miller, W. R. (1975). Psychological deficits in depression. *Psychological Bulletin* **82**, 238–260.

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.

Nelson, H. E. & McKenna, P. (1975). The use of current reading ability in the assessment of dementia. *British Journal of Social and Clinical Psychology* **14**, 259–267.

Newman, P. J. & Sweet, J. J. (1992). Depressive Disorders. In *Handbook of Neuropsychological Assessment: A Biopsychosocial Perspective* (ed. A. E. Puente and R. J. McCaffrey), pp. 263–307. Plenum Press: New York.

Peterson, L. R. & Peterson, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology* **58**, 193–198.

Robertson, G. & Taylor, P. J. (1985). Some cognitive correlates of affective disorders. *Psychological Medicine* **15**, 297–309.

Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S. & Goddard, R. (1986). CAMDEX: a standardised instrument for the diagnosis in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry* **149**, 698–709.

Roth, M., Huppert, F. A., Tym, E. & Mountjoy, C. Q. (1988). *CAMDEX: the Cambridge Examination for Mental Disorders of the Elderly*. Cambridge University Press: Cambridge.

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.

Savard, R. J., Rey, A. C. & Post, R. M. (1980). Halstead-Reitan Category Test in bipolar and unipolar affective disorders: relationship to age and phase of illness. *Journal of Nervous and Mental Diseases* **168**, 297–304.

Schonell, F. (1942). *Backwardness in the Basic Subjects*. Oliver & Boyd: London.

Spitzer, R. L., Endicott, J. & Robins, E. (1977). *Research Diagnostic Criteria for a Selected Group of Functional Disorders*. Biometrics Research Division, New York State Psychiatric Institute: New York.

Spreen, O. & Benton, A. L. (1969). Neurosensory Centre Comprehensive Examination for Aphasia. Neuropsychology Laboratory, Department of Psychology, University of Victoria: Victoria, BC, Canada.

Sternberg, D. E. & Jarvik, M. E. (1976). Memory function in depression: improvement with antidepressant medication. *Archives of General Psychiatry* **33**, 219–224.

Taylor, E. M. (1959). *Psychological Appraisal of Children with Cerebral Defects*. Harvard University Press: Cambridge, Mass.

Watts, F. N. & Cooper, Z. (1989). The effects of depression on structural aspects of the recall of prose. *Journal of Abnormal Psychology* **98**, 150–153.

Watts, F. N., Morris, L. & MacLeod, A. K. (1987). Recognition memory in depression. *Journal of Abnormal Psychology* **96**, 273–275.

Watts, F. N., Dalgleish, T., Bourke, P. & Healy, D. (1990). Memory deficits in clinical depression: processing resources and the structure of materials. *Psychological Medicine* **20**, 345–349.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology* **19**, 87–95.

Wechsler, D. (1986). *The Wechsler Adult Intelligence Scale – Revised. Incorporating British Amendments to the Test Administration Section*. The Psychological Corporation Ltd.: New York.

Weigl, E. (1941). On the psychology of so-called processes of abstraction. *Journal of Abnormal and Social Psychology* **36**, 3–33.

Weingartner, H. & Silberman, E. (1982). Models of cognitive impairment: cognitive changes in depression. *Psychopharmacological Bulletin* **18**, 27–42.

Weingartner, H., Cohen, R. M., Murphy, D. L., Martello, J. & Gerdt, C. (1981). Cognitive processes in memory. *Archives of General Psychiatry* **38**, 42–47.

Widlocher, D. & Hardy-Bayle, M-C. (1989). Cognition and control of action in psychopathology. *European Bulletin of Cognitive Psychology* **9**, 1–33.

Willner, P. (1984). Cognitive functioning in depression: a review of theory and research. *Psychological Medicine* **14**, 807–823.

Wolfe, J., Granholm, E., Butters, N., Saunders, E. & Janowsky, D. (1987). Verbal memory deficits associated with major affective disorders: a comparison of unipolar and bipolar patients. *Journal of Affective Disorders* **13**, 83–92.