REASONING PROCESSES IN DEPRESSION

A thesis submitted for the degree of

Doctor of Philosophy

by

Jane Elizabeth Baker

Department of Psychology

University College London

July, 1999

ACKNOWLEGEMENTS

The research reported in this thesis would not have been possible without the help and assistance of a number of people. I would therefore like to take this opportunity to thank them.

I am very grateful to Dr. Shelley Channon for supervising this project. I learnt a great many valuable lessons whilst working with her.

I would like to thank all the students who participated in the studies described in this thesis. The time and trouble they took was greatly appreciated.

I would also like to thank my parents, Howard and Christine Baker, my brothers, Michael and David Baker, and all my friends for their support and encouragement. Finally, a special thank you to my partner, Jon Furley, for all his support and patience.

ABSTRACT

Clinical accounts of depression suggest that it is associated with cognitive deficits in concentration, memory and 'thinking'. From a cognitive therapy perspective. it has been postulated that depression may be maintained by 'thinking errors' characterised by negative and biased interpretations of events. Research carried out to investigate cognitive function in depression has used both emotional and neutral stimuli, and has predominantly focused on memory and learning. With regard to neutral stimuli, a range of memory and learning tasks has been used, and depression has often, but not always, been associated with deficits. Current models suggest that the pattern of findings is consistent with a reduction in the cognitive capacity available to perform the tasks. In the light of these models it can be predicted that reasoning ability may also be affected.

In comparison to the large volume of research on memory for neutral materials, only a small body of work exists on reasoning processes in depression. It is difficult to draw any conclusions as to the likely nature and extent of any deficits on the basis of the existing work. Assessing reasoning in depression using neutral tasks is of interest from a clinical perspective since it can be established whether the 'thinking errors' which have been observed in clinical situations in relation to the personal concerns of the patient reflect more widespread difficulties with logical thought.

This thesis describes a series of experimental studies investigating reasoning in depression. The subjects in each study were dysphoric and nondysphoric undergraduate students. The initial studies compared dysphoric and nondysphoric subjects on tasks that have been widely used in the cognitive and neuropsychological literature. The subsequent studies were devised to explore further the nature and extent of reaoning deficits in the dysphoric subjects. Overall the pattern of deficits shown by the dysphoric subjects appeared to be consistent, and related to the difficulty of the task. The results are considered in the light of current models of cognitive function in depression and the implications of the findings are discussed.

PUBLICATIONS

The experiments described in chapter IV and chapter VIII have been published. The reference is Baker, J.E. & Channon, S. (1995). Reasoning in depression: impairment on a concept discrimination learning task. Cognition and Emotion, 9, 579-97.

The experiment described in chapter VI has been published. The reference is Channon, S. & Baker, J.E. (1996). Depression and problem-solving performance on a fault diagnosis task. Applied Cognitive Psychology, 10, 327-36.

The studies described in this thesis, including any published work, were designed and analysed by Jane Baker, in collaboration with myself as supervisor. Jane Baker was solely responsible for all the recruitment and data collection.

Shelley Channon

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 DEPRESSION

Depression has been labelled as the common cold of psychiatry, suggesting a widespread and relatively trivial disorder. This description does capture one of the essential elements of depression. Depression is common, and everyone is likely to experience a feeling they will label as depression at some time in their lives, although, at its shortest, it may last only minutes or hours. However, depression also has a number of other characteristics that make it far from trivial. Depression, unlike the common cold, has a high human cost. This can be measured in terms of the number of suicides carried out as a result of depression; the reduction in quality of life for both sufferers and their families; and the economic costs in terms of lost working days and expenditure on treatment.

As a first step in trying to increase understanding of a concept such as depression, it is important to establish a definition. As a term, 'depression' has come to be used in a number of different ways, and it is essential to avoid the resulting ambiguities. In lay-terms, 'depression' refers to feeling 'fed up', a feeling that is often short-lived. In medical terms 'depression' may be used at a symptom level to describe feelings of sadness, or it may be used to refer to a syndrome or a nosologic disorder. A depressive syndrome goes beyond the 'normal' experience of just feeling low or sad. It refers to a depressed mood, accompanied by a cluster of symptoms, that persists over time, and causes disruption and impairment of functioning. For depression to be a nosologic category then diagnostic procedures are required to exclude other diagnostic categories. For example, it is possible for two individuals to present with very similar syndromal depressive features, but for one also to meet diagnostic criteria for schizophrenia, while the other does not.

During an episode of clinically significant depression the mood component may be predominantly one of sadness, or it may be a diminished ability to experience pleasure, termed *anhedonia.* The mood symptoms are accompanied by symptoms in the domains of bodily functioning, behaviour and cognition. The depressed person may experience certain physical changes, such as feelings of listlessness or fatigue. There are often

changes in sleeping and eating patterns, and the latter may manifest themselves in weight changes and constipation. Changes in mood and physical well-being often lead to a loss of motivation and changes in behaviour, such as a withdrawal from normal activities. Depression is also evident in certain cognitive changes. The depressed person is more likely to report thoughts associated with self-doubt, self-blame, guilt and a belief in their own worthlessness. Depression may also impair intellectual functioning, with patients frequently reporting difficulties with concentration, memory and 'thinking'.

The different symptoms of depression may interact with each other in ways that perpetuate the disorder. For example, the belief that intellectual functioning is impaired may cause great distress to the individual, who may interpret their symptoms as evidence of a dementing process. It seems likely that perceived inadequacies in cognitive function are likely to increase feelings of worthlessness. If cognitive deficits actually exist, then they may exacerbate the individual's problems by causing difficulties in performing tasks necessary for everyday functioning. This may result in negative feedback about performance from others, and thereby give evidence to support the depressed person's negative view of themselves.

1.2 COGNITIVE FUNCTION IN DEPRESSION

The role of cognitive function in depression has been the subject of a large body of work. The negative cognitive content frequently reported by depressed patients has led many researchers to an interest in the processing of emotionally-salient or personally-relevant material by depressed individuals. This work has resulted in a number of theories concerned with the way depressive processing of emotional material might play a role in causing and/or maintaining a depressive episode, which in turn has led to the development of treatment strategies. This work is outlined in section 2.1.3.3 of the current thesis. In contrast, the difficulties with concentration, memory and 'thinking' often reported by depressed patients has led other researchers to explore depressive processing of neutral materials, with the aim of identifying the nature and extent of any depressive deficits so that methods for remediation can be developed. This work has used standard measures of cognitive function taken from the fields of cognitive psychology and neuropsychology and is reviewed in section 2.3.

Work on depressive processing of neutral materials has focused predominantly on memory

and learning, with other cognitive abilities receiving little attention. One ability that seems under-researched in comparison with memory and learning is reasoning. Reasoning refers to the ability to draw inferences, which is the means by which humans apply their knowledge and experience to specific situations. For example, if it is known it only snows when it is cold outside, then on seeing it is snowing the inference "it must be cold outside today" can be drawn. While reasoning has long been a subject of concern to philosophers, it has only expanded as a topic of interest in the field of psychology during the past thirty years, where it has been studied from both cognitive and neuropsychological perspectives. Only a small body of work exists on reasoning processes in depression, and it is difficult to draw any conclusions as to the likely nature and extent of any deficits on the basis of the existing work (see section 2.3.8).

While few studies have assessed reasoning processes in depression by examining performance on neutral tasks taken from the fields of cognitive psychology and neuropsychology, the work on processing emotionally-salient or personally-relevant stimuli has been concerned with reasoning in depression, although it has not usually been labelled as such. For example, the highly influential cognitive model of depression put forward by A.T. Beck (e.g. 1967; 1976; Beck, Rush, Shaw & Emery, 1979; see section 2.1.3.3.1) postulates depression is associated with systematic errors in thinking that help to maintain negative beliefs in the face of contradictory evidence. These 'thinking errors' can be interpreted as reasoning deficits in the sense that they are concerned with the processes by which individuals evaluate evidence and draw inferences. Similarly, the reformulated learned helplessness model of depression (Abramson, Seligman & Teasdale, 1978; see section 2.1.3.3.3) is concerned with the attributions depressed individuals make. An 'attribution' is the causal explanation an individual gives for an event. For example, an exam failure might be attributed to personal stupidity, lack of preparation, bad luck, etc. Because attributions require the individual to weigh up evidence and draw inferences, they are closely linked to the individual's capacity to reason.

The cognitive models of depression put forward by Beck (e.g. 1967; 1976; Beck, Rush, Shaw & Emery, 1979) and Abramson, Seligman and Teasdale (1978) have increased understanding of the process and treatment of depression, although more recently there have been challenges to the assumptions underpinning these theories. These are reviewed in section 2.1.3.3.5. However, it is important to consider whether the biases thought to

characterise depressive processing of emotionally-salient and personally-relevant stimuli reflect more widespread depressive difficulties with logical thought. Understanding the effect of depression on reasoning with neutral materials has implications for both models of emotional processing in depression and models developed to account for patterns of depressive performance on neutral tasks. By providing a cross-over between these fields, it should also act to strengthen the links between these two bodies of work which may in future prove beneficial in increasing our understanding of cognitive function in depression, and in developing ways to remediate any deficits.

1.3 THE AIMS OF THIS STUDY

The aim of this study is to investigate the nature and extent of any reasoning deficits associated with depression, and to explore the mechanisms that might underlie any deficits. Dysphoric and nondysphoric students will be compared on reasoning tasks widely reported in the fields of cognitive psychology and neuropsychology, and consisting of neutral stimuli. The findings will be considered in the light of relevant models taken from the fields of clinical psychology, cognitive psychology, and neuropsychology.

CHAPTER II

LITERATURE REVIEW

2.1 DEPRESSION

In section 1.1, it was noted depression is a widespread disorder, with serious implications in terms of cost for sufferers, their families, and society as a whole. It was also noted that the term 'depression' has a variety of meanings, and the importance of using stringent definitions that permit discrimination between these was highlighted. An outline was given of the symptoms associated with clinically relevant depression. In this section other aspects of depression relevant to the current thesis will be explored, including issues relating to the diagnosis, classification, epidemiology and aetiology of depression.

2.1.1 DIAGNOSTIC ISSUES

In section 1.1 syndromal depression was defined as a depressed mood accompanied by a cluster of symptoms, that persists over time, and causes disruption and impairment of functioning. Section 1.1 included a list of symptoms relating to the physical, behavioural and cognitive changes associated with a clinically relevant depressive syndrome. However, these symptoms may appear in almost any combination, often making diagnosis a highly subjective process. In section 1.1 it was noted that there is a crucial difference between diagnosing syndromal depression and establishing a nosologic category, the difference being the exclusion of other diagnostic categories in the latter. However, in practice this is often difficult to do, and may be subject to inconsistency.

During the past thirty years, the importance of making reliable diagnoses has been recognised and addressed. Several diagnostic systems have been developed (e.g. Diagnostic and Statistical Manual of Mental Disorders [DSM-III and DSM-III-R], American Psychiatric Association {APA} 1980, 1987), and these are often used in conjunction with standardised psychiatric interviews (e.g. Structured Clinical Interview [SCID], Spitzer, Williams, Gibbon & First, 1992). In order to meet diagnostic criteria for a depressive episode, symptoms must have been present for a specified time, and be of sufficient severity. A diagnosis of dysthymia is made if the symptoms are less severe, but of a more chronic nature. The DSM-III-R criteria for a Major Depressive Episode are shown in Table 2.1. However, in a review of the evidence, Costello (1993) concluded that the reliability and validity of three of these diagnostic systems was "not very good."

Table 2.1 DSM-III-R Diagnostic Criteria for Major Depressive Episode

A. At least five of the following symptoms have been present during the same two-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood, or (2) loss of interest or pleasure. (Do not include symptoms that are clearly due to a physical condition, mood-incongruent delusions or hallucinations, incoherence, or marked loosening of associations).

(1) Depressed mood (or can be irritable mood in children and adolescents) most of the day, nearly every day, as indicated either by subjective account or observation by others.

(2) Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated either by subjective account or observation by others of apathy most of the time).

(3) Significant weight loss or weight gain when not dieting (e.g. more than 5% of body weight in a month), or decrease or increase in appetite nearly every day (in children, consider failure to make expected weight gains).

(4) Insomnia or hypersomnia nearly every day.

(5) Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).

(6) Fatigue or loss of energy nearly every day.

(7) Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).

(8) Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).

(9) Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

B. (1) It cannot be established that an organic factor initiated and maintained the disturbance.

(2) The disturbance is not a normal reaction to the death of a loved one (Uncomplicated Bereavement).

C. At no time during the disturbance have there been delusions or hallucinations for as long as two weeks in the absence of prominent mood symptoms (i.e., before the mood symptoms developed or after they have remitted).

D. Not superimposed on Schizophrenia, Schizophreniform Disorder, Delusional Disorder, or Psychotic Disorder Not Otherwise Specified.

The concept of depression is associated with several long-standing controversies which remain unresolved. These fall loosely into two categories (Grove & Andreasen, 1992). The first includes issues relating to the boundaries of depression, both with regard to 'normality', and with regard to other psychiatric disorders. The second includes issues associated with the way in which depression may be classified into sub-types. These issues are considered in sections 2.1.1.1 and 2.1.1.2 below.

2.1.1.1 The Boundaries of Depression

The relationship between depression and normality has provoked a great deal of debate, and theorists remain divided. There are several intertwining and complex issues at stake in this debate. One issue is whether the depressive syndrome lies on a continuum with normal experience, or whether it represents a distinct illness. A related debate is whether mild, transient or situational depressive symptoms should be labelled as depression, or whether the term 'depression' should be reserved for disorders that meet diagnostic criteria of the type outlined above, with milder forms of symptomatology given another label (e.g. 'unhappiness'). The position adopted by individual theorists is likely both to reflect and inform their theoretical stance. For example, theorists interested in delineating the prognosis, course or aetiology of the disorder are likely to adopt a relatively narrow conceptualisation of depression as a means of identifying a homogeneous group for study. In contrast, those who conceptualise depression as being on a continuum with normal experience (e.g. Beck, 1976) may use models of normal function to inform models of depression, and vice versa. As noted above, this issue remains unresolved, but it influences other issues relating to depression, such as the diagnostic issues discussed above, and issues relating to classification and epidemiology discussed below.

When discussing the boundaries of depression, the ongoing debate regarding the relationship between depression and anxiety (see e.g. de Silva, 1994; Judd & Burrows, 1992) needs to be considered. While depression and anxiety are clearly distinct from one another at a conceptual level, with depression being associated with sadness and anxiety associated with fear, separating the two at an empirical level has proved difficult in both clinical and nonclinical samples (e.g. Wittchen & Essau, 1993). Clark and Watson (1991) reviewed the literature and identified five hypotheses of the relationship between depression and anxiety: a) that they are different points along the same continuum; b) that they are alternative manifestations of a common underlying diathesis; c) that they are

heterogeneous syndromes associated because of shared sub-types; d) that they are separate phenomena, each of which may develop into the other over the course of time, and e) that they are conceptually and empirically distinct phenomena. They also suggested a sixth possibility, that depression and anxiety have both common and specific features (Clark & Watson, 1991). Clark, Beck and Beck (1994) suggest difficulties in distinguishing between depression and anxiety have arisen because they share symptoms, have high co-morbidity rates, and there is item overlap on some standard depression and anxiety scales. It is not clear how the debate is likely to be resolved in the long-term, with evidence availabe to support all the alternative hypotheses.

2.1.1.2 Ways of Classifying Depression

As noted above, there is controversy regarding the way depression is classified into sub-types. This is related to the question of whether depression is a single, normally distributed dimension or whether it consists of discrete categories. Attempts to classify depression have often resulted in dichotomies, the most notable being neurotic versus psychotic, primary versus secondary, unipolar versus bipolar, and endogenous versus reactive.

The neurotic-psychotic distinction is an old one, but in general it denotes severity. The presence of delusional thinking, which may be either mood-congruent or non-congruent, triggers the diagnosis of psychotic. With regard to the primary-secondary distinction, primary depression refers to a first episode of depression not preceded by any other psychiatric or physical disorder. All remaining depressions that are preceded by, or accompany, other psychiatric or physical disorders are labelled as secondary depression. While this seems relatively clear, Grove and Andreasen (1992 p. 33) note that several problems have emerged regarding the identification of what should be included as a relevant antecedent diagnosis. They conclude that these problems have resulted in a decrease in the popularity of this distinction.

The most widely accepted distinction is between bipolar and unipolar depression, with bipolar disorders characterised by periods of both depression and mania, while unipolar disorders are associated with only one type of affect. In their review, Farmer and McGuffin (1989) concluded that this distinction has been accepted as real by neurobiologist researchers of depression, while other reviewers have concluded that the

evidence of differences between bipolar and unipolar depressive disorders in their course and epidemiology supports the distinction (e.g. Hirschfeld & Cross, 1982; Perris, 1992). The bipolar versus unipolar distinction is now incorporated in all the major classificatory systems (e.g. DSM-III-R, APA, 1987).

Unipolar depression subsumes a heterogeneous collection of disorders, and it is generally accepted there are likely to be one or more sub-types within this. The endogenous-reactive distinction has received most attention, but this concept remains controversial. This is due to confusion regarding the meanings of the terms which have changed over time. Originally the distinction was concerned with the aetiology of the disorder, with reactive depression thought to result from a life event, while endogenous depression was thought to result from biological processes, without an obvious external precipitant. There has been little support for a distinction based on these premises. Depressions which seem to follow life events do not necessarily differ in symptoms from those that do not, while depressions with 'endogenous' symptoms have often been found to follow stressful life events (e.g. Hirschfeld, 1981).

Over time the endogenous-reactive distinction has become associated with sub-types defined by symptom profiles. Endogenous depression, or melancholic depression as it is called in DSM-III-R (APA, 1987), is associated with the presence of symptoms thought of as predominantly biological. For example, the melancholic sub-type of depression in DSM-III-R is diagnosed if at least five of the following are present: loss of interest or pleasure in all, or almost all, activities; lack of reactivity to usually pleasurable stimuli; depression regularly worse in the morning; early morning awakening; psychomotor retardation or agitation; significant anorexia or weight loss; no significant personality disturbance before the first major depression; one or more previous major depressive episodes followed by complete, or nearly complete, recovery; previous good response to specific and adequate somatic anti-depressant therapy (DSM-III-R, APA, 1987). The remaining, and probably heterogeneous, depressions are subsumed by the reactive label. Gotlib and Hammen (1992) note that more work needs to be done to delineate meaningful sub-types in this residual (reactive) group, and that establishing the existence of any sub-types is only the first step in reaching an understanding of their meaning and implications.

Depression clearly varies in the degree of severity, and measures of this have been

developed. These take into account a range of symptoms commonly associated with depression, with each symptom given a score relating to its severity, and a total score calculated. The measures vary as to whether symptoms are rated by a clinician (e.g Hamilton Depression Rating Scale, Hamilton, 1967), or self-reported (e.g. Beck Depression Inventory [BDI], Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), but they have in common the notion that cut-off scores can be used to distinguish between different levels of symptomatology.

2.1.2 EPIDEMIOLOGY OF DEPRESSION

Information about the epidemiology of a disorder is important both in terms of informing theories of aetiology, and in terms of planning and targeting treatment resources. Researchers have used several ways of estimating the frequency of disorders. Point or period prevalence refers to the proportion of the population who have the disorder at a given time. Incidence refers to the number of new cases which emerge during a time period. Life-time risk refers to the likelihood an individual will develop the disorder at some time during their life, and can be calculated in different ways.

A large number of studies have been carried out to investigate the epidemiology of depression. Before considering some of these, it is worth quoting the very neat summary relating to the situation in this country provided by Paykel (1989a) who concluded:

"About 1 per 1000 of the general population are admitted to hospital annually with depression; about 3 per 1000 are referred to psychiatrists, of whom two are treated as out-patients. However, around 3 percent of the general population are treated in this country by general practitioners and an equal number probably consult and are not recognised. The prevalence rate in the population is about 5 percent although estimates vary considerably. The frequencies of the disorder will depend very much on how one defines it and where one studies it." (Paykel, 1989a, p.3).

Many early studies of the epidemiology of depression compared the rates of different disorders found in hospitals or other treatment settings. Although this approach can provide valuable information about treatment provision, it is flawed as a method of estimating the rates within the general population since not all those suffering from a disorder will seek treatment. Community studies, which select a random sample from the population and then identify the rates of different types of disorder within the sample, are

more methodologically sound. However, the lack of standardised diagnostic criteria in the very early studies made them difficult to interpret. Boyd and Weissman (1981) reviewed the findings of community studies that *had* used standard diagnostic techniques. They concluded that the point prevalence of depressive symptoms ranges between nine and twenty cases per 100 and that the point prevalence of unipolar depression is 3% for men and 4-9% for women. The lifetime risk of unipolar depression showed greater variation, and ranged from 2-12% for men and 5-26% for women (Boyd & Weissman, 1981).

More recently, the findings of large-scale community studies using standardised diagnostic techniques have become available, such as the Epidemiologic Catchment Area (ECA) research programme. This is a multi-centred programme carried out in the USA, with the results reported by a number of authors. For example, Myers, Weissman, Tischler, Holzer, Leaf et al. (1984) found the six-month prevalence of major depression at three of the sites was 2% to 3.5%, with the figure generally higher for women than for men, and for those under 65 years old. Surtees and Sashidaran (1986) compared a sample from Edinburgh with one of the ECA community samples, and found comparable rates for one-month prevalence. Karno, Hough, Burnam, Escobar, Timbers et al. (1987) calculated the lifetime prevalence of major depression at four of the ECA sites as ranging from 4% to 8.4%. However, it is important to note that while studies using the same methodology in different geographical areas may produce very similar estimates of depression rates, findings can differ when studies using different methodologies are compared. For example, Bebbington, Katz, McGuffin, Tennant and Hurry (1989) estimated lifetime (before age 65) risk of a minor depressive episode in Camberwell, London as 62%. The differences in chosen methodology are likely to reflect the many unresolved questions regarding the nature of depression, as discussed in section 2.1.1. For example, the selection of a cut-off point for caseness clearly has strong implications for the number of cases that will be subsequently identified, and this is likely to be influenced by whether clinical depression is perceived as on a continuum with 'normal' feelings of depression, or whether it is conceptualised as one or more discrete disorders.

Hirschfeld and Cross (1982) reviewed epidemiological studies that had considered psychosocial risk factors. They found studies consistently reported sex differences, with women reporting depressive symptoms and being diagnosed with unipolar depressive syndromes more frequently than men. A similar conclusion was reached by a later review

carried out by Nolen-Hoeksema (1987), with both reviews concluding that women suffer depression at a ratio of approximately 2:1 compared to men. Hirschfeld and Cross (1982) also identified other consistent findings. Depression was more prevalent in younger (aged 18-40 years) than older adults; in unmarried relative to married people; and depressive symptoms were more prevalent in those of lower socio-economic status (SES) compared to those of higher SES. For other factors the findings were not so clear. There was no conclusive evidence on urban versus rural areas, and any differences in terms of race seemed to be mediated by differences in SES.

This brief review indicates how problematic it has been to establish an estimate of the rate of depression which is universally applicable and acceptable. While the availability of standard diagnostic criteria has improved the reliability and validity of many studies, the ongoing debates about the nature and classification of depression, as outlined in section 2.1.1, still create difficulties when interpreting the findings of this research. Nevertheless, it can be seen that even the most conservative of estimates indicate depression is a widespread problem, and this is particularly true for certain sections of the population, such as women, the young, and those of lower SES.

2.1.3 THEORIES OF DEPRESSION

As a prevalent, distressing, potentially life-threatening disorder, many attempts have been made to establish the causal and maintaining mechanisms of depression, and the most effective methods of treatment. While later theories, such as Gilbert's (1992) biopsychosocial model, bridge theoretical boundaries, early theories focused on one aspect of the individual or environment. Constraints of space mean it is not possible to give an exhaustive review of all existing theories of depression. Instead, a brief outline will be given of the main theories, with more detail on those most pertinent to the current thesis.

2.1.3.1 Psychological Theories of Depression

Psychological theories of depression have been developed, usually within one of the three main schools of psychological thought: psychoanalytic; behavioural; and cognitive. The main focus of the current thesis is on cognitive functioning in depression. Therefore, psychoanalytic and behavioural approaches will be outlined briefly in this section, while cognitive theories will be considered in detail in section 2.1.3.3 below.

2.1.3.1.1 Psychoanalytic theories of depression

In his classic paper "Mourning and Melancholia", Freud (1917/1965) compared depression with grief. He suggested the symptoms of depression and mourning differ only in the absence of lowered self-esteem in mourning. They are both reactions to the loss of a loved object. However, Freud argued, while mourning clearly follows a death, melancholia may follow losses where the true nature or meaning of the loss are unavailable even to the depressed individual. Freud suggested unconscious rather than conscious processes underlie melancholia, and argued that it is the combination of lowered self-esteem, expressed through self-reproaches, with the apparent obscurity of the source of the depression which make the condition inexplicable to others. Freud suggested depression arises when the individual forms an attachment to another person, but then receives a rejection, real or imagined, which shakes this attachment. Instead of withdrawing the attachment and forming a new one, the depressed individual identifies the lost attachment with a part of their own self. Therefore all the negative symptoms of depression, including suicidal impulses, are directed against the internalised object (Freud, 1917/1965). Thus, Freud (1917/1965) conceptualised depression as an unconscious process of 'giving up' the internalised attachment akin to the process in mourning of 'giving up' the person who is mourned .

Following on from Freud's theory of depression (1917/1965), there have been further contributions from writers within the psychoanalytic school (for a collection of the most important papers in this field see Gaylin, 1968). Many of these have focused on the role of the early mother-child relationship in the later development of depression (e.g. Klein, 1934/1968). The influential work of John Bowlby on attachment (e.g. 1978; 1981) draws on many of the ideas expressed in earlier psychoanalytic writings. He has argued that depression in adulthood is related to the failure of the child to form a stable, secure attachment with the parents, or the experience of an actual loss of a parent. These experiences are associated with feelings of helplessness, and colour the individual's internal representations of self and others.

Thus, psychoanalytic theorists draw on the normal process of mourning as an analogy for depression, with the themes of loss, and 'letting go' of the lost object used to explain the process and symptoms of depression. Psychoanalytic theorists have also considered the influence of the early parent-child relationship on the later onset of depression.

2.1.3.1.2 Behavioural theories of depression

Behavioural approaches have tried to understand depression in terms of specific learning histories and the environmental responses or stimuli which act to maintain maladaptive behaviours. The social reinforcement theory put forward by Lewinsohn and his colleagues (e.g. Lewinsohn, 1974; 1975; Lewinsohn, Youngren & Grosscup, 1979) and a complementary theory put forward by Coyne (1976) will be considered in this section, while for broader reviews of behavioural theories of depression see Gilbert (1992, pp.413-433) and Gotlib and Hammen (1992, pp.71-74).

Lewinsohn (1974; 1975) relates depression to a low rate of response-contingent positive reinforcement (RCPR). He notes that it is the lack of contingency on the behaviour rather than the lack of positive reinforcement per se which is the crucial factor. Thus, giving positive reinforcement that is non-contingent does not reduce depression. The total amount of RCPR received by an individual is presumed to be a function of three sets of variables: a) the number of events that are potentially reinforcing for the individual; b) the number of potentially reinforcing events that can be provided by the environment; c) the extent to which the individual possesses the skills and emits the behaviours that will elicit reinforcement from the environment. In particular, deficits in social skills, defined as the emission of behaviours which are positively reinforced by others, are seen as important in the development and maintenance of depression.

Coyne (1976) also emphasised the importance of interpersonal interactions in understanding depression. He contended that changes in the individual's social structure, such as the loss of a significant relationship or a change in employment status, may precipitate depressive symptomatology, which he interprets as the individual signalling their need for reassurance regarding their social position. However, he argues, while depressive symptoms may initially elicit positive responses from others, these may become more negative and rejecting as time goes on and the depressed individual continues to make demands on those around. This has the effect of increasing the distress and insecurity of the depressed individual, and increasing their need for reassurance. Coyne (1976) suggests this situation may escalate until either the hospitalisation of the depressed individual, or the complete withdrawal of those in social contact with the individual, breaks the stalemate.

This brief description of two behavioural theories of depression illustrates the contrast between psychoanalytic theories, with their emphasis on early experiences of attachment and loss, and behavioural theories, which focus on the influence of the reinforcers in the individual's current environment. However, it should be noted that while these theories may differ in emphasis, there are common themes, with both acknowledging the importance of interpersonal interaction in depression While the psychoanalytic theories focus on the intrapersonal conflicts that arise as a result of dysfunctional interpersonal interactions, behavioural theories are concerned with the impact of social reinforcement on the behaviour of the depressed individual.

2.1.3.2 Other Theories of Depression

Psychoanalytic and behavioural theories of depression postulate a primary role for psychological processes in the aetiology of depression, but other theories have placed a different emphasis. As noted above, both psychoanalytic and behavioural theories acknowledge the importance of interpersonal interactions in the development and maintenance of depression, but some theorists have gone further, giving the social environment a primary role. In contrast, other theories of depression have sought physical causes within the individual to explain the disorder. Both social and physical theories of depression are considered in brief in this section.

2.1.3.2.1 Social theories of depression

The important work of G.W. Brown and his colleagues places primary emphasis on the social circumstances of the individual. This theory is outlined here, while Gilbert (1992, pp.435-457) provides a review of other social theories of depression. Brown and his colleagues (e.g. Brown, Adler & Bifulco, 1988; Brown, Andrews, Harris, Adler & Bridge, 1986; Brown & Harris, 1978) were interested in the role of life events, and their interaction with the personal resources of the individual (e.g. social support). For example, Brown and Harris (1978) carried out a large-scale study of psychiatric disorder in women. The authors viewed depression as a largely social phenomenon, and postulated a model which explains the aetiology of depression in terms of the presence or absence of three factors: provoking agents; vulnerability factors; and symptom-formation factors. Provoking agents are essentially either severe life-events associated with loss or disappointment which have long-term consequences for the individual, or major ongoing difficulties. Brown and Harris (1978) argued provoking agents are rarely sufficient to

bring about the onset of a depressive episode, although they do determine *when* an episode will occur. The influence of provoking agents is thought to be mediated by vulnerability factors (Brown & Harris, 1978). These are aspects of the individual's social situation, such as the presence or absence of a confiding relationship, that can either buffer the individual from the impact of the provoking agent, or exacerbate its effect. Brown and Harris (1978) also examined factors which influenced depressive symptoms after the onset of a depressive episode. With regard to severity, they reported severe events occurring *after* onset, past bereavements, and a previous episode of depression were all associated with increased severity of symptoms. Thus, Brown and Harris (1978) assume complex interactions between aspects of the individual's social environment can explain the onset and nature of a depressive episode.

2.1.3.2.2 Biological theories of depression

Following the development of successful drug treatments for depression, research into physical changes in depression initially focused on biochemical systems (for reviews see Delgado, Price, Heninger & Charney, 1992; Paykel, 1989b; Tucker & Liotti, 1989). Much of this work may have implications for understanding cognitive function in depression, but at present the overall picture is still very unclear and a full review is beyond the scope of this thesis. The current review will be confined to a brief outline of the work done using modern imaging techniques, and the implications of this work for understanding cognitive function in depression.

The advent of modern neuro-imaging techniques provided a new approach to the investigation of brain structure and function in depression. Initially, it was anticipated that examining the brain structure of depressed patients using computed tomography (CT) and magnetic resonance imaging (MRI) would reveal significant structural abnormalities. Early studies did find evidence of structural brain abnormalities in depression in the form of enlarged ventricles (e.g. Standish-Barry, Bouras, Bridges & Bartlett, 1982). However, this abnormality is rather nonspecific in terms of its implications for brain function, and its significance has remained unclear. Furthermore, as noted by George, Ketter and Post (1993), while differences between depressed and nondepressed individuals may be discernable when comparing group data, differences are not readily detectable when individual scans are inspected. Thus, CT and MRI cannot be used to make a positive diagnosis of depression in individual patients.

Investigations of depression using CT and MRI have been less informative than anticipated, but these techniques have played an important role in increasing our understanding of depression. Important structural clues about the neuroanatomy of depression have come from CT and MRI studies of patients with depression secondary to other brain disorders, such as strokes and tumours. These have generally been consistent in finding left anterior lesions are associated with an increased likelihood of developing a secondary depression (e.g. Robinson, Kubos, Starr, Rao & Price, 1984; for a review see Cummings, 1993).

While CT and MRI permit detailed examination of the structure of the brain, other imaging techniques allow visualisation and measurement of brain function. Two techniques of this kind are used most often: single photon emission computed tomography (SPECT) and positron emission tomography (PET). These can be used to measure regional cerebral blood flow and cerebral metabolic rate, both of which are indicators of brain activity. Studies carried out using these techniques with depressed subjects (Ss) have found evidence of abnormal function in several brain regions, with the frontal lobes of the brain implicated most consistently (e.g. Bench, Friston, Brown, Scott, Frackowiak & Dolan, 1992; Sackeim, Prohovnik, Moeller, Brown, Apter et al., 1990).

This outline, while brief, nevertheless illustrates the importance of biological approaches to depression. Modern techniques allow sophisticated analysis of brain structure and function, and have made an important contribution to our current understanding of depressive disorders. The evidence suggests that while depression may not be clearly associated with detectable structural brain damage, it is associated with abnormal brain function, with the frontal lobes of the brain identified most consistently. Further support for the role of the frontal lobes in depression comes from the evidence that damage to the left frontal lobe is associated with an increased likelihood of developing a secondary depression. The next step is to relate abnormal brain function in depression to the symptoms of depression, such as cognitive impairment (e.g. Dolan, Bench, Brown, Scott & Frackowiak, 1994). In the current thesis, the function of the frontal lobes of the brain will be considered in detail in section 2.2.3.

2.1.3.3 Cognitive Theories of Depression

In section 1.2 it was noted that cognitive theories of depression may offer important insights into reasoning processes in depression, and these are considered in this section. Two early models of cognitive function in depression will be considered in detail: Beck's cognitive model (e.g. Beck, 1967); and the reformulated learned helplessness model (e.g. Abramson, Seligman & Teasdale, 1978). Both models incorporate hypotheses regarding the nature of depressive thought, and are therefore of direct relevance to the current thesis. Following consideration of these two theories, subsequent cognitive theories of depression will be reviewed. A large number of studies have been carried out to test the predictions of the cognitive theories of depression reviewed in this thesis. Due to considerations of space, it will not be possible to review this literature in its entirety, but work relevant to the question of reasoning in depression will be assessed.

At this point, it is important to address a methodological issue which is pertinent to the work carried out to assess cognitive theories of depression. Two populations have been utilised to study cognitive processes in depression: (a) clinical populations of individuals who have a diagnosis of depression; (b) nonclinical populations of individuals. When using the latter group, the experimenter may either induce a manifest emotional state using standard mood induction procedures (e.g. Velten, 1968), or evaluate current mood state and allocate Ss on this basis. The implications of these different approaches will be considered in section 2.3.1 with regard to measuring cognitive function in depression in relation to neutral stimuli. For the following review of depressive processing of emotional stimuli, relevant studies using all these methodological approaches will be considered.

2.1.3.3.1 Beck's cognitive model of depression

The work of A.T. Beck (e.g. Beck, 1967; Beck, 1976; Beck, Rush, Shaw & Emery, 1979; Kovacs & Beck, 1978) is a major influence in the field of depression, and has attracted a great deal of interest and research. Although there has been some variation in the details of Beck's theory over time, the basic principles have remained generally intact, and these are outlined below.

The main premise of Beck's cognitive model is that cognitive processes translate external events into meaningful internal representations, and that it is these, rather than the events themselves, which underlie pathology. Negative automatic thoughts play a key role in

Beck's theory. These are defined as follows (Beck, Rush, Shaw & Emery, 1979): 1) they are automatic in that they habitually appear without 'reasoning' thought processes taking place; 2) they are involuntary, and the individual may have great difficulty in inhibiting them; 3) they are irrational and dysfunctional; 4) they are accepted as the truth by the individual and it is only on reflection that their irrational nature can be perceived.

The cognitive theory of depression consists of three concepts which explain how automatic thoughts are formed, and their role in depression. These are schemata, the cognitive triad, and cognitive errors. Beck (1967) used the term 'schemata' to refer to stable cognitive structures which are the basic components of cognitive organisation. Schemata actively interpret incoming information to make sense of and encode it. Dysfunctional schemata are thought to develop as a result of early life experiences, and take the form of excessively rigid and inappropriate beliefs about the self and the world (Beck, Rush Shaw & Emery, 1979). Beck (1967) postulated dysfunctional schemata may lie dormant in people vulnerable to depression, but once triggered they may override more appropriate schemata, and, as their activity level increases, may be activated by increasingly inappropriate stimuli.

When depressogenic schemata are activated, negative thoughts characterised by a negative view of the self, the world, and the future are produced: the cognitive triad. Beck, Rush, Shaw and Emery (1979) suggest the depressed person perceives himself as "defective, inadequate, diseased or deprived", and therefore worthless. He interprets his experiences as examples of his inability to achieve anything or perform in an adequate manner, and sees his environment as difficult and full of problems. With regard to the future, the depressed individual believes it will be unpleasant and full of failure.

It is postulated that the contents of the depressogenic schemata are translated into the negative thoughts seen in depression by 'thinking errors'. These are systematic errors in thinking which help to maintain negative beliefs in the face of contradictory evidence. Beck, Rush, Shaw and Emery (1979) described this pattern of thinking as 'primitive' in that it is associated with relatively simplistic interpretations of events, such as interpreting situations as unvarying and irreversible. This was contrasted with more mature thought patterns that recognise, for example, that situations may change over time, and that actions can be taken to effect the change. Beck identified the following specific thinking errors

associated with depression:

1. Arbitrary inference - drawing a negative conclusion in the absence of supporting data.
2. Selective abstraction - focusing on a detail out of context, often at the expense of more salient information.
3. Overgeneralisation - drawing a conclusion that relates to a wide variety of things on the basis of a single event.
4. Magnification and minimisation - making errors in evaluating the importance and implications of events.
5. Personalisation - relating external (often negative) events to the self when there is little reason for doing so.
6. Absolutistic, dichotomous thinking - thinking in polar opposites (black and white). Something is all good, or totally bad and a disaster.

Both Beck and other theorists have added to this list over time. For example, 'catastrophising' is the belief that one should always expect the worst since this is the most likely outcome. However, the basic principle of the concept of thinking errors has remained relatively unchanged. Thus, Beck's theory of depression (e.g. 1967; 1976; Beck, Rush, Shaw & Emery, 1979) clearly postulates that depression is associated with faulty reasoning processes which may play a role in causing and/or maintaining depression.

2.1.3.3.2 Evaluation of Beck's cognitive model of depression

Beck's cognitive model of depression has been the subject of a large volume of research. This has focused on several hypotheses associated with the model: whether dysfunctional schemata can be identified in individuals who are vulnerable to depression at times when they are not depressed; the efficacy of cognitive therapy in relation to other treatments; and whether thinking errors can be identified and measured in depressed individuals. Constraints of space mean it is not possible to evaluate all this research. Therefore this review will be limited to work relevant to the issue of reasoning processes, or 'thinking errors', in depression, and this is considered below.

Much of the work concerned with studying depressive 'thinking errors' has focused on developing measures intended to assess them. The Cognitive Bias Questionnaire (CBQ; Hammen & Krantz, 1976; Krantz & Hammen, 1979) consists of six vignettes of problematic situations involving interpersonal or achievement themes. For each one, Ss are asked to select one of four possible responses as being the closest to the feelings of

the main character in the vignette. The response options reflect two dichotomous and crossed dimensions in terms of logical inference based on the information provided in the story: depressed versus nondepressed; and distorted versus non-distorted. Several studies have been able to distinguish depressed and nondepressed Ss on the basis of the depressed-distorted responses in both clinical (Krantz & Gallagher-Thompson, 1990; Krantz & Hammen, 1979) and nonclinical samples (Blaney, Behar & Head, 1980; Hammen, 1978; Hammen and Krantz, 1976; Krantz & Hammen, 1979).

Lefebvre (1981) described the Cognitive Error Questionnaire (CEQ). This consists of 24 vignettes followed by a statement relating to the vignette and designed to represent one of the cognitive errors posited by Beck (see section 2.1.3.3.1). Ss were asked to rate how similar the statement was to the thought they would have themselves in response to the situation described in the vignette. Lefebvre (1981) compared clinically depressed patients with chronic pain patients who were assigned to depressed or nondepressed groups on the basis of their score on the BDI, and a normal control group. He reported that depressed Ss with and without chronic pain showed similar levels of cognitive distortion, and significantly more than nondepressed Ss.

The Cognitive Response Test (CRT; Watkins & Rush, 1983) is a less structured measure of cognitive distortion in depression. The items are presented in an open-ended, sentence completion format, and responses are scored on the basis of standardised rules as rational, irrational-depressed, irrational-other, or unscorable. Watkins and Rush (1983) found the CRT distinguished depressed patients from three nondepressed control groups (nondepressed psychiatric patients; nondepressed medical patients; normal controls), with depressed Ss making a significantly greater number of irrational-depressed responses, and significantly fewer rational responses. The groups did not differ on the other two response types. Wilkinson and Blackburn (1981) also found currently depressed patients made more irrational-depressed responses on the CRT than recovered depressed patients, patients who had recovered from psychiatric disorders other than depression, and normal controls. The recovered depressed Ss did not differ significantly from the other two control groups.

Fennell and Campbell (1984) developed the Cognitions Questionnaire (CQ) which was intended to provide a more comprehensive measure of depressive cognition. The test consists of eight scenarios: two positive, two neutral and four negative; each followed by

five questions designed to assess five response dimensions: emotional impact, attribution of causality, generalisation across time (implications for the future), generalisation across situations (how far events were seen as typical), and perceived uncontrollability (what could be done about the situation). Fennell and Campbell (1984) reported on the performance of currently depressed and previously depressed Ss drawn from patient and community populations, depressed patients with a diagnosis other than depression, nondepressed psychiatric patients, and nondepressed community-living Ss. The CQ was found to discriminate between currently depressed and nondepressed Ss, with depressed Ss showing greater distortion. Generalisation was the most powerful discriminator between the groups. Distortion did not characterise all psychiatric patients, but was specific to those with depressed mood. Previously depressed Ss did not differ significantly from nondepressed Ss, except on generalisation from negative scenarios.

In summary, several measures designed to measure thinking errors in depression have been found to discriminate between depressed and nondepressed Ss. However, these findings are not entirely consistent with Beck's notion of a range of different thinking errors (see section 2.1.3.3.1). It was reported that during the development of three of the measures described above (CBQ; CEQ; and CRT), independent judges had difficulty in reliably separating one type of thinking error from another (Krantz & Hammen, 1979; Lefebvre, 1981; Watkins & Rush, 1983). As a result, two of the measures (CBQ; Krantz & Hammen, 1979; CRT; Watkins & Rush, 1983) used a single category of 'distorted' responses, while Lefebvre (1981) reported the initial seven cognitive errors had to be condensed into four, and the final analysis found little difference in Ss' endorsement of the different error types. Watkins and Rush (1983) commented that the common denominator in responses which judges identified as 'distorted' appeared to be the process of exaggeration in which Ss drew conclusions unwarranted by the premises or facts presented. Perhaps as a result of these problems, interest in measuring cognitive thinking errors in this way seems to have declined during the last ten years, although some of these measures have been adapted to investigate the effects of information valence on depressive cognition (e.g. Krantz & Gallagher-Thompson, 1990) .

One of the most serious criticisms of Beck's (e.g. 1967) cognitive theory of depression is that it underestimates the contribution of the environment to the experience of depression (Coyne & Gotlib, 1983). Thus, Brown and Harris (1978), whose social theory of

depression was outlined in section 2.1.3.2.1, suggest it is unnecessary to postulate distorted thinking in depressed individuals, since their depression could often be attributed to an accurate assessment of their current environment (Brown & Harris, 1978, p.83). Studies have provided evidence that supports this criticism. These have demonstrated that the 'distorted' thinking shown by depressed Ss when compared with nondepressed Ss may reflect real differences in their circumstances or experiences. For example, Dunning and Story (1991) required depressed and nondepressed students to predict the likelihood of a list of personal events which varied in how desirable and controllable they were. The participants were asked at the end of the semester whether the events had occurred. They found that while depressed Ss were more pessimistic in their predictions, being significantly more likely to predict the occurrence of aversive events than controls, this pessimism was justified because depressed Ss were significantly more likely to experience these negative events than controls. Using a different approach, Dykman, Horowitz, Abramson and Usher (1991) found depressed Ss' negative self-ratings of their social competence during a group discussion were actually consistent with their evaluation by other members of the discussion group and by independent observers, again suggesting a basis in reality for their 'distorted' interpretations.

In the light of this kind of evidence, Beck (e.g. Haaga & Beck, 1995) has revised his notion that depressed individuals show 'distorted' thinking in the sense that they reach conclusions that are inconsistent with objective reality. This has been replaced with the concept of 'biased' thinking which is defined as a tendency to make judgements in a consistent manner across different situations, and which would manifest itself in depression as a tendency to draw negative conclusions about oneself (Haaga & Beck, 1995). This means that while the failures and losses described by depressed individuals may be perceived accurately, the meanings and generalisations attached to them may still show a negative bias. For example, following a job redundancy, an objectively negative event, a depressed person may draw the conclusion that he is a worthless failure who will never succeed in obtaining another job, while someone who is not experiencing depression may feel sad, but may attribute the loss to the economic climate and feel hopeful of future employment. The issue of causal attributions in depression is dealt with in more detail in the next section in the context of another important cognitive model of depression, the reformulated learned helplessness model (Abramson, Seligman & Teasdale, 1978).

2.1.3.3.3 The Reformulated Learned Helplessness Model of depression

Like Beck's cognitive model (see section 2.1.3.3.1), the reformulated learned helplessness model of depression (Abramson, Seligman & Teasdale, 1978) is also concerned with the way in which depression affects thinking and reasoning. The first formulation of this model (e.g. Seligman, 1975) was based on the finding that animals exposed to repeated and uncontrollable aversive stimuli later showed a reduction in avoidance behaviour in situations where escape was possible. This was labelled 'learned helplessness'. Seligman (1975) suggested depression and learned helplessness might share the same aetiology, that is, the experience of an uncontrollable event leading to the expectation that future events will also be outside control. Such a belief is likely to result in a decreased rate of responding, leading to the passivity of depression and depressed affect.

Seligman's (1975) learned helplessness model provoked debate and research, but was also subject to heavy criticism (e.g. Costello, 1978). Then Abramson, Seligman and Teasdale (1978) published a critique and presented a reformulation which introduced the concept of attributions as a mediating factor between the experience and interpretation of uncontrollable aversive events. They posited that an individual who experiences an uncontrollable event asks *why* he is helpless in this situation. The causal attribution made determines the generality and chronicity of helplessness deficits and any effects on self-esteem. Abramson, Seligman and Teasdale (1978) postulated three important attributional dimensions: 1) the internal-external dimension refers to whether the cause for the situation is perceived as being within the individual or in the external environment; 2) the stable-unstable dimension refers to whether it is perceived as long-standing or short-lived; 3) the global-specific dimension refers to whether the cause is perceived as likely to influence a wide range of situations, or only a few specific situations. When an individual asks *why* a situation has arisen, these three dimensions can be joined in any combination, giving a total of eight types of attribution. Each will have very different consequences for the individual's expectations for the future.

The reformulated learned helplessness model (Abramson, Seligman & Teasdale, 1978) proposed negative events attributed to internal (personal), stable (unchanging) and global (wide ranging) attributions would be the most distressing (e.g. "I failed the exam because I am stupid."). Depression is assumed to result if an aversive state of affairs is believed to be likely (or if a highly desired state of affairs is believed to be unlikely), and the

individual perceives the outcome is uncontrollable, and makes a global, stable and internal attribution for this. Later conceptualisations of this hypothesis have emphasised that it should be considered as a diathesis-stress model of depression (e.g Metalsky, Halberstadt & Abramson, 1987). That is, the tendency to attribute negative events to internal, stable, global causes is a diathesis for depressive reactions accompanied by lowered self esteem, while negative life events are a stress for depressive reactions.

In summary, the reformulated learned helplessness model is concerned with the role of attributional processes in the onset and maintenance of depression. In section 1.2 reasoning was defined as the ability to draw inferences. In order to identify accurately the cause of an outcome, the individual will need to apply their knowledge and experience to the situation by means of drawing relevant inferences. For example, if the individual knows from past experience that people who are "not very clever" fail exams, then on failing an exam the individual may infer "I am not very clever". Alternatively, if the individual knows from past experience that "clever" people may sometimes fail exams because they are unlucky, then on failing an exam the individual may infer "I was unlucky". Therefore the attributional processes central to the reformulated learned helplessness model are essentially reasoning processes, and a review of the evidence relating to this model is likely to inform the question of whether depression is associated with reasoning deficits.

2.1.3.3.4 Evaluation of the Reformulated Learned Helplessness Model

Much of the work carried out to evaluate the reformulated learned helplessness model has focused on testing the hypothesis that attributional style is a predictor of later depressed mood (see section 2.1.3.3.3). However, for the purposes of the current thesis the evidence will be reviewed simply in relation to whether depression is associated with a specific attributional style, such as the tendency to make internal, stable and global attributions for uncontrollable aversive events postulated by Abramson, Seligman & Teasdale (1978).

Several measures have been developed to assess the nature of attributional processes in depression. Seligman, Abramson, Semmel and von Baeyer (1979) developed the Attributional Style Questionnaire (ASQ; also see Peterson, Semmel, von Baeyer, Abramson, Metalsky & Seligman, 1982). Ss are presented with twelve hypothetical situations, six describing good outcomes and six describing bad outcomes, and each

involving themes of either achievement or affiliation. They are asked to identify the major cause of the event, and to make attributional judgements on the three relevant dimensions (stable-unstable, internal-external, global-specific). Peterson and Seligman (1984) cited three studies (Eaves & Rush, 1984; Persons & Rao, 1981; Raps, Peterson, Reinhard, Abramson & Seligman, 1982) that found the predicted pattern of responding on the ASQ in clinically depressed Ss. However, Hargreaves (1985) compared clinically depressed Ss and normal controls on the ASQ, and found a significant difference only on stable-unstable attributions for positive events.

Three studies have compared depressed and nondepressed students selected on the basis of their BDI scores on the ASQ. Seligman et al. (1979) found depressed students made attributions for negative outcomes that were significantly more internal, stable and global than those of nondepressed students. Blaney et al. (1980) found a highly significant difference in the predicted direction between depressed and nondepressed students in terms of the globality of their attributions for negative events, with a significant difference for attributions of stability, and no group difference for attributions of internality. Ingram, Kendall, Smith, Donnell and Ronan (1987) found depressed students made significantly more internal and stable attributions for negative events than both anxious and normal controls; there was no significant difference for globality scores. With regard to positive events, the depressed Ss made significantly less internal, stable, and global attributions than anxious and normal control groups. Another study looked at correlations between student scores on the ASQ and the BDI. Brewin and Furnham (1986, Experiment 1) found a significant correlation between student's BDI score and both internal and global attributions for negative outcomes on the ASQ, but not for stability attributions. When a subset of the same sample were asked to make attributions for actual negative life events (Brewin & Furnham, 1986, Experiment 2), the correlation with BDI score reached significance for internality, but not stability or globality.

Studies have also investigated Ss' attributions for success and failure on laboratory tasks where, unbeknown to Ss, the outcome is experimenter-controlled. The findings from these studies have, at best, found only limited support for the reformulated learned helplessness model. Consistent with the model, these studies have generally found evidence to suggest depressed students make more internal attributions for failure than nondepressed students (e.g. Kuiper, 1978; Rizley, 1978, Experiment 1; Zemore & Johansen, 1980; Zuroff, 1981).

However, two studies of clinical patients failed to find the predicted group differences (Abramson, Garber, Edwards & Seligman, 1978; Gotlib & Olson 1983); and Zuroff (1981) reported that while depressed students might make more internal attributions for failure than nondepressed students, their absolute preference was for external over internal attributions. The evidence that depression is associated with more stable attributions for failure is also mixed. Zemore and Johansen (1980) reported some evidence of a significant correlation between BDI score in students and stable attributions for failure. Zuroff (1981) found while depressed students showed an absolute preference for stable attributions for failure, they did not differ significantly from nondepressed students on this measure. Other studies have found no evidence that either depressed students (Kuiper, 1978; Rizley, 1978 Experiment 1) or depressed patients (Gotlib & Olson, 1983) make more stable attributions for failure than nondepressed controls.

With regard to attributions following success, the findings have been even more inconsistent. Depressed students have been found to make both more (Zuroff, 1981) and less internal attributions (Rizley, 1978 Experiment 1) than nondepressed controls, while Kuiper (1978) found depressed and control groups did not differ. All three studies found depressed students and controls did not differ in the degree to which they made stable attributions for success (Kuiper, 1978; Rizley, 1978 Experiment 1; Zuroff, 1981). Gotlib and Olson (1983) found depressed patients did not differ from controls in their attributions following success on the dimensions of either internality or stability.

As noted above, the central question of this thesis is whether depressed Ss have reasoning deficits. To this end, the evidence relating to the nature of attributional processes in depression has been considered simply in terms of whether depressed and nondepressed Ss show different attributional processes. The evidence gives only weak support to the notion that depression is associated with internal, global and stable attributions for negative events as postulated by the reformulated learned helplessness theory (Abramson, Seligman & Teasdale, 1978), with highly inconsistent findings. However, the evidence does indicate differences between depressed and nondepressed Ss in their attributional processes. Of the studies reviewed above, only Gotlib and Olson (1983) found no significant group differences in attributional processes. This suggests depressed Ss may differ from nondepressed Ss in their attributional processes, but that the reformulated learned helplessness model does not give the best account of these differences. In the

next section, the hypothesis that the attributions of nondepressed Ss may be subject to bias and that depressed Ss may be more accurate in their attributions than nondepressed Ss, often labelled the depressive realism hypothesis, will be considered.

2.1.3.3.5 Depressive realism

Alloy and Abramson (1979) reported that depressed Ss might be more accurate in their attributions than nondepressed Ss. Ss were required to judge the degree of contingency between their responses and the outcomes that followed. While depressed students judged accurately the degree of control their responses exerted over outcomes, nondepressed students showed a pattern of responding that has been described as an "illusion of control" (Langer, 1975). The nondepressed Ss overestimated their control over objectively uncontrollable outcomes when these were frequent or associated with success (e.g. winning money), and underestimated their degree of control over undesirable outcomes. This finding has been replicated using the same paradigm (e.g. Martin, Abramson & Alloy, 1984). In addition, Alloy, Abramson and Viscusi (1981) reported that inducing an elated mood in depressed students resulted in an illusion of control effect, while inducing a depressed mood in nondepressed Ss produced greater accuracy.

Theorists have tried to account for the findings described by Alloy and Abramson (1979). In a review of the area, Ackermann and DeRubeis (1991) noted that some theorists have argued that self-deception and bias contribute to a nondepressed person's sense of well-being, and are therefore adaptive in the long-term. In support of this, Alloy and Clements (1992) found students who showed a greater illusion of control on the Alloy & Abramson (1979) noncontingent-win task were less likely to show negative mood reactions after failure on a laboratory task, or to show discouragement or depression after real-life negative events. Furthermore, having reviewed the literature on the relationship between mental health and positive self-evaluations, Taylor and Brown (1988) concluded that overly positive self-evaluations, exaggerated perceptions of control or mastery, and unrealistic optimism are characteristic of normal human thought. Moreover, these illusions appear to promote other criteria of mental health, such as the ability to care about others, and the ability to feel happy or contented.

In their review, Ackermann and DeRubeis (1991) noted that some authors have gone further than simply arguing there may be advantages to positive illusions, and have

suggested depressed persons may be more accurate in their perceptions and judgements than their nondepressed counterparts. This stronger version of the depressive realism hypothesis has important implications since it contradicts the widely held view that psychopathology in general, and depression in particular, is associated with an impaired ability to test reality. Any evidence that suggests depressed Ss are more realistic than their nondepressed counterparts requires careful consideration.

As noted above, the basic effect reported by Alloy and Abramson (1979) has been replicated using similar paradigms (e.g. Martin et al., 1984). However, Ackermann and DeRubeis (1991) note that many experiments cited in support of the depressive realism hypothesis do not represent a valid test of the hypothesis because they do not permit a comparison of Ss' judgements with an objective reality. For example, the study reported by Lewinsohn, Mischel, Chaplin and Barton (1980), which is often cited as evidence of depressive realism, relies on a comparison between self-ratings of social skills during an experimental 'group interaction' and ratings made by observers. While the poorer ratings of depressed relative to nondepressed Ss may suggest that the negative self-ratings of depressed Ss have some real basis (see section 2.1.3.3.2), it is difficult to argue that one group is closer to reality than the other. The original Alloy and Abramson (1979) paradigm does permit a comparison with objective reality.

Ackermann and DeRubeis (1991) argue if consideration is limited to those studies which permit an objective assessment of accuracy, the evidence suggests depressed students do tend to be more accurate in their judgement of contingency and in their self-other judgements than nondepressed students, who tend to show evidence of a self-serving bias. However, there are several lines of evidence that contradict the depressive realism hypothesis. Thus, other paradigms have found nondepressed Ss to be more accurate than both depressed students and patients, such as in their recall of evaluative information (e.g. Kuiper, 1978). Furthermore, when Dunning and Story (1991) addressed the question of depressive realism in a more naturalistic setting using objective reality as a standard comparison they found depressed Ss were significantly less accurate and more over-confident in their predictions than nondepressed Ss, although nondepressed Ss also displayed an unrealistic confidence in the accuracy of their predictions. Finally, Benassi and Mahler (1985), using the Alloy and Abramson (1979) paradigm, found the presence of an observer reversed the basic depressive realism effect with the depressed students

showing an illusion of control in the observer condition. Overall, the evidence suggests that depressed Ss can only reliably be found to be more realistic than nondepressed Ss when the original Alloy and Abramson (1979) paradigm is employed. Whether depressed or nondepressed Ss are found to be more realistic is highly dependent on the methodology of the particular study, and future work should be engaged in mapping the effects of experimental manipulations.

While the question of depressive realism remains to be resolved, there are interesting accounts of the existing evidence. Several authors (e.g. Haaga & Beck, 1995) have noted the majority of evidence for depressive realism comes from investigations using college students rather than clinically depressed patients. Haaga and Beck (1995) note the Alloy and Abramson (1979) task has yet to be tested with clinically depressed patients. They speculate that nondepressed Ss might show a slight positive bias, that is reduced in mild depression to something close to reality, but may swing to a negative bias in more severely depressed Ss. More studies with clinical patients are needed to test this.

Following their review, Ackermann and DeRubeis (1991) hypothesised there might be a combination of factors underlying the pattern of findings, with the self-serving bias observed in nondepressed Ss resulting from motivational factors (e.g. desire to maintain self-esteem), while the negative bias found in depressed Ss' recall of evaluative information could be due to cognitive factors (e.g. faulty information processing). In support of this they note nondepressed Ss are generally found to show a self-serving bias when required to interpret information (e.g. how much control they have), but they may be able to recall facts, such as evaluative information, accurately. By this logic, the performance of depressed Ss could be explained as a breakdown in motivation to maintain self-esteem coupled with a cognitive deficit in their ability to recall evaluative information accurately. These hypotheses are post-hoc in nature and largely speculative, and therefore the authors need to make specific, testable predictions.

Finally, Dykman and his colleagues (e.g. Dykman, Abramson & Albright, 1991; Dykman, Abramson, Alloy & Hartlage, 1989) argue that neither depressed or nondepressed Ss characteristically process information in a particular way. Instead, they argue, both engage in "schematic processing", that is, they interpret information in a way that matches their existing schemata. Thus, either group may exhibit positive or negative biases, or may

appear unbiased, depending on the information-to-schema match of the particular situation. Group differences would only be predicted in situations where depressed and nondepressed Ss had different pre-existing task-relevant schemata. Several studies have been carried out to test this model (e.g. Dykman, Abramson & Albright, 1991; Dykman et al., 1989), and the findings thus far have been consistent with its predictions.

In summary, the question of depressive realism remains to be resolved, but it has raised interesting questions about the nature of cognition in depressed and nondepressed states. In particular, it raises the possibility that depressive thinking might be subject to the same biasing and distorting processes that characterise human cognition in general. Nevertheless, the weight of evidence suggests there are differences between depressed and nondepressed Ss in their cognitive functioning. Since the early models of Beck (e.g. 1967) and Abramson, Seligman and Teasdale (1978), a great deal of work has been carried out to investigate the role of cognition in depression, and a number of competing theories now exist. In the following section the development of cognitive theories of depression since 1980 will be reviewed.

2.1.3.3.6 Cognitive theories of depression since 1980

While the theories of Beck (e.g. 1967) and Abramson, Seligman and Teasdale (1978) focused on thinking in depression, later theories have considered the impact of depression on all aspects of cognition, including attention and memory. The application of concepts from the field of cognitive psychology has resulted in increasingly sophisticated models of depressive cognition. The following review provides a brief outline of the developments most relevant to the current thesis.

2.1.3.3.6.1 Bower's Associative Network Theory

In 1981, Bower put forward his influential associative network theory of mood and memory. Network theories had previously been developed by theorists (e.g. Anderson & Bower, 1973) to explain long-term memory (LTM) function. The basic premise of network approaches is that information in LTM is stored as nodes in a network, with related nodes sharing associative connections. Accessing information involves activating the appropriate node in LTM. Such activation will spread, through the associative connections, to partially activate (or 'prime') related information, which will thereafter become disproportionately available to the cognitive system. Bower (1981; 1987)

proposed that within this cognitive framework each emotion has a specific memory node which acts as a focusing point for all associated aspects of that emotion, including emotion-related memories and cognitions. Although some connections to the emotion node are believed to be innate, cognitive linkages are largely learned, being established through contiguity during life events. When a particular node is activated, the emotion is experienced and activation is channelled through its connections to evoke the emotion's other manifestations (subjective feelings, physiological response etc.). Activation of an emotion node also spreads activation throughout connected memory structures.

Bower's (1981; 1987) theory formed part of a growing body of work on the effects of mood on the processing of emotional material. Initially, interest was focused on memory, with two phenomena of particular interest: mood-state dependent retrieval and mood-congruent learning. Mood-state dependent retrieval is a process whereby memory for material is enhanced if the mood at encoding and retrieval are the same; the affective valence of the material is irrelevant. Mood congruency refers to enhanced encoding and/or retrieval of affectively valanced material while the S is in the corresponding mood state; concordance between mood at exposure and mood at recall is not required or relevant. Bower's (1981) theory incorporated both phenomena, and made strong predictions about their role in emotional disorders such as depression and anxiety. A large volume of work has been carried out to investigate the two phenomena in depression and anxiety and the area has been reviewed on a regular basis (e.g. Blaney, 1986; Bower, 1987; Dalgleish & Watts; 1990; Ellis & Ashbrook, 1989; MacLeod, 1990; Mathews & MacLeod, 1994; Williams, Watts, MacLeod & Mathews, 1988). A full review of this work will not be included in the current thesis, instead a brief review will be given, and the reader is referred to the reviews cited above for a fuller description.

With regard to depression, reviewers have generally noted that state dependent recall has proved difficult to demonstrate (e.g. Blaney, 1986; Bower, 1987; Dalgleish & Watts, 1990), a fact that highlights a major weakness in Bower's (1981; 1987) associative network theory. Mood-congruent recall in depression has been reported, although its exact nature is open to debate. Initially studies of mood-congruent recall tended to be concerned with recall of real events from Ss' lives. For example, Teasdale and Fogarty (1979) used a mood induction technique (Velten, 1968) to compare Ss in 'happy' and 'sad' conditions. They found a bias in the latency to retrieve positive or negative memories,

with slowed recall of positive material in the depressed group being the predominant feature. This effect appears to be robust, having been replicated consistently (e.g. Teasdale & Taylor, 1981; Teasdale, Taylor & Fogarty, 1980).

The study of recall from autobiographical memory has several associated difficulties: mood at encoding is not known, so the influence of state dependent learning cannot be estimated; individuals differ in the number of positive and negative experiences encountered, so Ss may recall more negative memories because they have had more negative experiences; it is difficult to differentiate between a recall bias (greater recall of negative memories) and a response bias (equal recall of positive and negative memories, but a greater tendency to report the negative memories). Researchers have tried to overcome these difficulties using a range of experimental techniques. Thus, supplying Ss with affectively-valanced material for recall overcomes the problem of Ss differing in available autobiographical memories. Most often the material has consisted of lists of negative, positive or neutral words (e.g. Bradley & Mathews, 1983; Dunbar & Lishman, 1984; Teasdale & Russell, 1983), and the majority of studies have found some evidence of lower recall of positive material and/or higher recall of negative material among depressed individuals compared with controls. Alternatively, the use of mood induction allows control of Ss' mood at encoding and retrieval.

While studies have generally been consistent in finding evidence of a mood congruency effect in depressive recall, the evidence is far from conclusive with regard to the effects of depression on attention to affectively valanced material (e.g. Dalgleish & Watts, 1990; MacLeod, 1990). This has been in direct contrast to the findings relating to anxiety, with strong evidence of an attentional bias in anxiety, but inconclusive evidence regarding memory bias (e.g. Dalgleish & Watts, 1990; MacLeod, 1990).

2.1.3.3.6.2 An information-processing model: Williams et al., 1988

Oatley and Johnson-Laird (1987) put forward a general theory of emotion which posited that basic emotions have evolved to serve important biological and social functions, and to determine priorities when conflicts arise in ongoing plans and goals. Within this model, sadness/depression is thought to result from the failure of a major plan or the loss of a goal. Emotions impose a relatively stereotyped mode of operation on the cognitive system, consistent with the evolutionary function of that emotion. Their theory predicts

that the mode of processing, as well as the type of information being processed, will vary across emotional disorders. This is consistent with the differences described above between depression and anxiety. However, MacLeod (1990) notes this framework is not articulated to a level that permits detailed specification of the various cognitive modes, something addressed by the model proposed by Williams et al. (1988).

The model put forward by Williams et al. (1988) is based on Mandler's distinction between activation (or priming) and elaboration (e.g. Graf & Mandler, 1984). Activation can take place at a 'preattentive' stage of processing. Williams et al. (1988) use the term 'preattentive' to refer to processes that operate automatically, do not depend on awareness of the stimuli, and influence the allocation of subsequent processing resources. Mogg, Bradley and Williams (1995) advocate the use of 'preconscious' rather than 'preattentive' since it is not certain that such processes are independent of attentional selectivity. The preconscious stage allows attentional capture of information and is reflected in implicit memory tests. The process of elaboration involves the formation and strengthening of associative links between the representation currently being processed and other existing representations in memory. This is a strategic operation that will occur over an extended period of time. Williams et al. (1988) distinguished between depression and anxiety, arguing that biases operate at different stages of processing in the two disorders. They postulated anxiety may be associated with a preconscious bias that favours threat stimuli and automatically directs the focus of attention toward the location of such stimuli in the environment. In contrast, depression is thought to involve the biased use of mnemonic cuing at the elaboration stage (Williams et al., 1988), resulting in greater ease of recall or recognition of mood-congruent material.

The model proposed by Williams et al. (1988) was developed to account for the pattern of evidence outlined above which seems to suggest depression is associated with biases at retrieval but not encoding, and the converse for anxiety. With regard to the aims of the current thesis, the most pertinent question is whether depression is associated with preconscious biases in addition to the well-established mood-congruent memory biases found on explicit memory tasks. This question presents certain methodological difficulties in terms of establishing whether any observed biases are due to conscious elaborative processes or automatic preconscious processes, and it is only relatively recently that authors have attempted to address it. Two areas have been addressed: 1) whether

depressed Ss show attentional biases that can be attributed to preconscious processes; 2) whether depressed Ss show mood-congruent memory bias on implicit memory tests.

With regard to the question of attentional bias in depression, perhaps the most stringent test devised thus far has been achieved by presenting test material in both subthreshold and suprathreshold conditions. Several studies have been carried out using this approach (Mathews, Ridgeway & Williamson, 1996; Mogg et al., 1995; Mogg, Bradley, Williams & Mathews, 1993), with conflicting results. Mogg et al. (1993) reported that, consistent with the predictions of Williams et al. (1988), clinically anxious but not clinically depressed Ss showed attentional biases for negative words in both subthreshold and suprathreshold conditions. Mathews et al. (1996) and Mogg et al. (1995) both compared clinically depressed, clinically anxious and control Ss on a probe detection task of the type developed by MacLeod, Mathews and Tata (1986). In this task two words are presented simultaneously, and on some trials one is replaced by a dot. The latency to detect the probe dot gives a measure of which word was attended to by the S. Both Mathews et al. (1996) and Mogg et al. (1995) reported depressed Ss did show an attentional bias to suprathreshold negative words; while a depressive preconscious bias was unproven.

Studies have also investigated processing biases in implicit versus explicit memory tasks. Implicit memory is measured by performance change as a result of prior exposure, or priming, regardless of awareness. A typical task requires Ss to carry out word-stem completion (e.g. don___) with the first idea that comes to mind; performance is 'primed' by previous study of relevant items (e.g. presentation of the word 'donkey' increases the likelihood of stem completion with this word). Denny and Hunt (1992) and Watkins, Mathews, Williamson and Fuller (1992) found no evidence of significant mood-congruent implicit memory in clinical depression, although Roediger and McDermott (1992) noted both studies did find non-significant trends toward depression-congruent priming effects.

While implicit memory tasks by definition do not involve intentional retrieval instructions, if Ss have been conscious of the priming material the possibility exists that they have used explicit memory processes to carry out the task. This difficulty has been addressed by Brendan Bradley and his colleagues (Bradley, Mogg & Williams, 1994; Bradley, Mogg & Millar, 1996; Bradley, Mogg & Williams, 1995) in a series of studies comparing the priming effects of material presented in sub- and suprathreshold conditions. Bradley et

al. (1995) and Bradley et al. (1996, Experiment 2) both reported clinically depressed Ss showed evidence of implicit memory bias for depression-relevant information in both subthreshold and suprathreshold conditions. Bradley et al. (1994) and Bradley et al. (1996, Experiment 1) reported dysphoric students showed the depression-congruent priming effect in the subthreshold, but not the suprathreshold condition. Thus, all three studies found evidence of depressive memory biases that are not reliant on explicit memory processes.

In summary, these studies have found a pattern of evidence that is inconsistent with the predictions of the model proposed by Williams et al. (1988). Williams et al. (1988) predicted depression should be associated with a different pattern of biases to those shown by anxious Ss, with biases in memory rather than attention, and in effortful rather than automatic processing. The hypothesis that depression and anxiety will be associated with different patterns of bias has found consistent support. However, there is now evidence that depression may be associated with attentional biases, at least when the material is presented suprathreshold, and with biases in automatic memory processes. Clearly more work will be needed before this question is fully resolved.

2.1.3.3.6.3 A problem-solving formulation of depression

A. M. Nezu and colleagues (e.g. Nezu, 1987; Nezu, Nezu & Perri, 1989) have proposed a problem-solving formulation of depression which draws on many of the ideas incorporated in the early cognitive models of depression reviewed above (e.g. Abramson, Seligman & Teasdale, 1978; Beck, 1967). It is proposed that the onset of depression can occur when an individual experiences a problem. Failure to resolve the problem effectively may create a host of negative consequences, resulting in decreased personal and social reinforcement, while effective resolution will serve to decrease the likelihood of a depressive episode. The basic premise is that depression can result from deficiencies or reduced effectiveness in any or all of the five major components of problem-solving (problem orientation, problem definition and formulation, generation of alternatives, decision-making, and solution implementation and verification).

Depressive problem-solving difficulties have often been investigated using hypothetical social situations which require Ss to generate means of achieving an end goal, such as the Means-Ends Problem-Solving test devised by Platt and Spivack (MEPS; 1972). The

MEPS is a paper-and-pencil test presenting Ss with a series of problem situations, usually interpersonal, giving both an initial situation and a goal to be achieved. Scoring is based on factors such as the number and kinds of means generated to achieve the goal, the degree of elaboration of detail, and the relevance of the solutions to the problem situation. More recently, the Social Problem-Solving Inventory (SPSI; D'Zurilla & Nezu, 1990) has been developed. This assesses problem orientation and problem-solving skills, and has seven sub-scales. The SPSI items are self-statements depicting either positive or negative cognitive, affective or behavioural responses to real-life problem-solving situations. D'Zurilla and Nezu (1990) reported it has adequate reliability and validity.

Shaw and Dobson (1981) cited unpublished work by Giles which found clinically depressed inpatients to be significantly poorer than normal controls in generating solutions on the MEPs test. Marx, Williams and Claridge (1992, Experiment 1) compared depressed, anxious and normal control Ss and found depressed Ss were impaired relative to the other two groups on the MEPS, although the anxious group also showed some impairment. There was no significant difference between the groups in Ss' ratings of the effectiveness of their own solutions, but when they were judged by two independent raters, those of the depressed Ss were rated as less effective than those of the other two groups. Marx et al. (1992, Study 2) also reported on a modified version of the MEPS using Ss' own problems. Ss were asked to describe the actual strategy used, and their ideal strategy in retrospect. Both depressed and anxious Ss had actually used less effective strategies to solve their problems, but only the depressed Ss were unable to devise ideal strategies that were more effective. Goddard, Dritschel and Burton (1996) found depressed Ss were impaired relative to a hospital control group (patients receiving treatment for relatively mild, physical complaints) on the MEPS both in terms of the number of solutions generated and the effectiveness of the solutions. Thus, the few studies carried out with clinically depressed Ss have been consistent in finding evidence of poorer social problem-solving in depressed Ss.

Studies carried out with nonclinical Ss have produced mixed findings. Gotlib and Asarnow (1979) compared depressed and nondepressed students selected on the basis of the BDI on the MEPS test, and found significantly worse performance in the depressed group. Zemore and Dell (1983) reported a significant correlation between performance on the MEPS and score on the BDI, and also with a self-report rating of depression-

proneness which remained when current mood-state was partialled out. Lakey (1988) administered measures of mood and problem-solving ability to a group of students on two occasions ten weeks apart. While there was no correlation between measures of mood and problem-solving on the first occasion, problem-solving ability at this time proved to be an independent predictor of dysphoric mood on the second occasion. Conflicting results have been reported by other studies. Shaw and Dobson (1981) cited unpublished work by Krumm and by Giles, Dobson and Shaw which failed to find significant differences on the MEPS test between depressed and nondepressed students, although no details were provided regarding selection of Ss. Doerfler, Mullins, Griffen, Siegel and Richards (1984) compared depressed and nondepressed adults and children classified by self-report measures, including the BDI or an equivalent form for children, and found no significant depressive deficit in MEPS test performance. Cane and Gotlib (1985) compared depressed and nondepressed students assessed on a self-report measure of depression, and reported no significant group differences in dealing with hypothetical social problem situations. Blankstein, Flett and Johnston (1992) compared depressed and nondepressed students on an adapted college version of the MEPS and found that while depressed Ss had lower expectations of their performance and lower self-evaluations after carrying out the task, they did not in fact differ from nondepressed Ss in their performance.

In summary, the four experiments carried out with clinical samples (Giles, cited by Shaw & Dobson, 1981; Goddard et al., 1996; Marx et al., 1992, Experiments 1 & 2) found consistent evidence of a depressive deficit in performance on measures of social problem-solving, while of the eight studies carried out with nonclinical samples, only three found evidence of a depressive deficit (Gotlib & Asarnow, 1979; Lakey, 1988; Zemore & Dell, 1983). This pattern of findings could be due to a relationship between severity of depression and performance on social problem-solving tasks. This would explain the inconsistent findings of studies carried out with nonclinical Ss who may be only mildly depressed. Another possibility is that the use of subjective judgements in the scoring procedure for the MEPS and some of the other measures may have contributed to the inconsistency of the findings. Finally, it should be noted that while these measures may have the advantage of being designed to assess problem-solving in real-life situations, they do not elucidate the possible causes of any deficits or how they might relate to other cognitive processes, such as attention and memory.

2.1.3.3.6.4 The Interacting Cognitive Subsystems approach

Another important cognitive theory of depression has been developed by J.D. Teasdale (1983; 1988), later in conjunction with P.J. Barnard (e.g Barnard & Teasdale, 1991; Teasdale & Barnard, 1993). Teasdale (1983; 1988) proposed the Differential Activation hypothesis which suggests that, in vulnerable people, temporary mood disturbance can activate patterns of cognitive activity that turn the mood disturbance into severe and persistent depression. Studies carried out to test the predictions of this hypothesis have generally found support for it (e.g. Teasdale & Dent, 1987).

Later, Teasdale and Barnard (e.g Barnard & Teasdale, 1991; Teasdale & Barnard, 1993) proposed the interacting cognitive subsystems (ICS) approach to depression. This retains aspects of the Differential Activation hypothesis, with a broader framework. The ICS approach postulates interconnected modules, each coding for a different type of information. The most important, with regard to depression, are the propositional and implicational modules. Propositional codes represent specific meanings or knowledge of the type conveyed by a sentence ("cold" cognition), while knowledge held in implicational code represents schematic mental models having personal implications ("hot" cognition). Propositional schemas integrate information from various knowledge sources, while implicational models represent a more generic, holistic, level of meaning integrating information from all other mental codes, including body-state information.

With regard to depression, the ICS approach (e.g Barnard & Teasdale, 1991; Teasdale & Barnard, 1993) proposes emotion-related schematic models encode the prototypical features extracted from previous situations eliciting a particular emotion. When these patterns of implicational code are processed, emotion is produced. So a depressed emotional state might result from the synthesis of schematic models encoding themes such as 'globally negative view of the self' that have been extracted as prototypical of previous depressing situations. Furthermore, elements of the implicational code derived from sensory information can have a direct effect on these higher level meanings, and so on the production of emotion. Therefore, bodily feedback of a bowed posture and frowning expression may enhance the effects of propositional loss-related meanings, and increase the likelihood of an implicational level schematic model of a depressing kind (e.g. 'globally negative view of the self') being created. ICS differs from associative networks of the type postulated by Bower (1981; 1987) in that ICS require emotional effects to

operate at the implicational schema level, rather than at the level of propositional or lexical items such as single words (Teasdale & Barnard, 1993). Emotional effects are expected only when information is encoded in relation to its implications for oneself, and not when responses are based on propositional knowledge alone.

2.1.3.3.7 Implications of Cognitive Models of depression

The purpose of the current thesis is to investigate reasoning processes in depression. Early cognitive theories of depression (e.g. Abramson, Seligman & Teasdale., 1978; Beck, 1967) were concerned with thinking processes in depression, and are therefore highly relevant to the purposes of the current thesis. Beck (e.g. 1967; 1976) postulated depression is associated with 'thinking errors' which are essentially failures to draw the correct conclusions. This can be rephrased as an hypothesis that depression is associated with reasoning deficits. The reformulated learned helplessness model (Abramson, Seligman & Teasdale, 1978) postulated depression is associated with a tendency to make internal, stable and global attributions for negative events. Again, this can be rephrased as an hypothesis that depression is associated with deficits in reasoning processes. However, the current review of the evidence relating to Beck's cognitive model (e.g. Beck, 1967; 1976) and the reformulated learned helplessness model (Abramson, Seligman & Teasdale, 1978) has raised questions about whether depressive thinking processes are actually impaired. Instead, the evidence seemed to suggest that thinking in both depressed and nondepressed states may be influenced by the contents of LTM via schematic processing. This can lead to bias in both depressed and nondepressed thinking, with depression more likely to be associated with a negative bias. This suggests a qualitative difference between depressed and nondepressed Ss in their reasoning ability rather than a true impairment in depressive reasoning.

With regard to subsequent cognitive theories of depression, these have focused on attention and memory processes rather than considering depressive thinking directly. There is evidence depression may be associated with biases in memory and possibly attention when processing emotional material. The problem-solving formulation of depression proposed by Nezu (e.g. Nezu, 1987; Nezu et al., 1989) suggests depression is associated with deficiencies in each of the stages of problem-solving. Problem-solving is likely to involve reasoning processes. However, there has been no attempt to explore the mechanisms by which depressive impairments in problem-solving might arise.

The evidence reviewed thus far suggests depression may be associated with qualitative differences in the processing of emotional material, but it is not sufficient to answer the question of whether depression might be associated with a reasoning deficit. However, the review has raised interesting questions that clearly need to be addressed. The hypothesis that nondepressed Ss may show bias in their reasoning suggests a need to review the nature of human reasoning. Furthermore, the evidence that reasoning in both depressed and nondepressed Ss may be influenced by schematic processes in LTM, and the evidence that depressed Ss show bias in attentional and memory processes highlights the need to explore the relationship between reasoning and cognitive processes such as attention and memory. Finally, most of the evidence considered thus far has related to depressive processing of emotionally salient material. The question of whether depression is associated with reasoning deficits may be best answered by considering performance on neutral tasks which should be less subject to any biases. As a first step towards addressing these issues, the next section will present models of reasoning taken from the fields of cognitive psychology and neuropsychology, and these will be considered in relation to the evidence reviewed thus far on reasoning processes in depression.

2.2 REASONING

As noted in section 1.2, while reasoning has long been a subject of concern to philosophers, it has only existed as a topic within psychology during the past thirty years. It has been studied within the fields of cognitive psychology and neuropsychology; both approaches will be considered below. Reasoning refers to the ability to draw inferences. Inferences are the means by which humans apply their knowledge and experience to specific situations. For example, if it is known it snows only when it is cold, then on seeing snow the inference "it must be cold today" can be drawn. Within the field of psychology, the study of reasoning has frequently overlapped with the study of problem-solving and intelligence, and the relationship between these three areas will be considered before moving on to a more detailed consideration of reasoning.

2.1.1 THE RELATIONSHIP BETWEEN REASONING, PROBLEM-SOLVING AND INTELLIGENCE

Problem-solving was relatively neglected until the influential work of Newell and Simon (1972). They introduced the concept of a problem-solving *state*. Subsequent problem-solving research has usually involved the study of Ss' performance on experimental

problems that have a clear starting state and a clearly defined goal state, and which require Ss to move through a series of steps or intervening states. Means-ends analysis is the most common type of problem-solving method (e.g. Anderson, 1993) whereby the problem-solver identifies the biggest difference between the current state and the goal state, and tries to reduce that difference.

The study of reasoning is linked to problem-solving for at least two reasons. First, at its most general, problem-solving refers to the activities that take place when an individual is trying to achieve a goal. In this sense, measures of cognitive function, including reasoning tasks, are problems to be solved by the S, and models of problem-solving can be applied to task performance. Secondly, reasoning is often used to solve a problem, so models of reasoning may increase understanding of Ss' problem-solving performance.

The study of intelligence evolved as part of the psychometric approach to psychology. 'Psychometric' means measurement of the mind. This approach grew from a desire to select people on the basis of their abilities for jobs or educational opportunities, and with regard to intelligence it resulted in the development of intelligence or "IQ" tests. One of the major controversies has been whether intelligence has a central component (*g*, for general intelligence) which underlies individual performance on a range of tasks, or whether there are separate abilities to solve different types of problem e.g. spatial versus verbal abilities. The technique commonly used to research this question is a complex development of correlational techniques called factor analysis. If a general ability underlies solution of a range of problems then there should be a high positive correlation between performance on different tasks, while a low correlation would be consistent with separate abilities. However, the results of factor analysis have proved inconclusive in deciding between the two alternatives. Proponents of *g* have often divided it into two types: fluid versus crystallised intelligence. Fluid intelligence is thought to be an enduring capacity which underlies people's ability to deal with novel situations, while crystallised intelligence refers to acquired knowledge.

More recently, cognitive approaches to intelligence have been developed. Perhaps most notable has been the work of Sternberg and his colleagues (e.g. Sternberg, 1988). This has focused on the cognitive processes or 'components' that might underpin intelligent behaviour. More detail of Sternberg's componential approach is given in section 2.2.2.3.3

below. Sternberg (1988) concluded some information-processing components seem to underlie performance on a range of tasks, and this might explain *g*.

Reasoning is generally accepted as an important constituent in the study of intelligence, with reasoning tasks featuring in most intelligence tests, and performance on these tasks showing high correlations with *g* (Marshalek, Lohman & Snow, 1983). Many experimental reasoning tasks are utilised as measures of fluid intelligence. Sternberg (1988) considers reasoning ability to be just one of a number of abilities that might underlie intelligence. However, it is possible that the relationship between reasoning and intelligence is mediated by a shared dependence on a more general ability, such as working memory (WM; see section 2.2.6 below). In conclusion, both psychometric and information-processing approaches to intelligence have identified a strong link between reasoning and intelligence, although both approaches conceptualise intelligence as subsuming reasoning along with a number of other abilities.

2.2.2 THE COGNITIVE PSYCHOLOGY APPROACH TO REASONING

Historically, reasoning has been divided into two separate fields of study, namely deductive and inductive reasoning, and this division has been maintained within cognitive psychology. In order to be consistent with the existing literature, deductive and inductive reasoning will first be described separately before moving on to consider models of reasoning which are applicable to both.

2.2.2.1 Deductive Reasoning

Deductive reasoning refers to the process of drawing an inference which is latent, but implicit, in the information given, that is, where the conclusion necessarily follows from the premises or assumptions. It does not result in an increase in semantic information, but only 'draws out' the information available in the premises. For example, "All swans are white, therefore this swan is white" constitutes a deduction. The study of deductive reasoning has its roots in philosophy and the study of logic, and originated in the writings of Aristotle. More recently, logic has been used to provide models of how people would reason if they were able to do so without making errors, and these have been used by psychologists for comparison with Ss' actual performance. Most commonly, Ss are presented with a series of problems which can be solved using logic, and their performance is then compared with the logically correct pattern of responding.

Syllogistic reasoning is the most extensively studied type of deductive reasoning. Syllogisms typically contain two statements (premises) that are assumed to be true, such as "All dogs are mammals" and "All corgis are dogs". One premise relates the subject (corgis) to the middle term (dogs), and the other premise relates the middle term to the predicate (mammals). The task is to relate the subject and predicate terms to each other to decide what conclusion can be reached, if any, using the possible terms 'all', 'some', 'none', and 'some not'. In this case, a valid conclusion can be reached: "All corgis are mammals". The conclusion is made on the basis of the logic of the premises only, and not on the basis of semantic knowledge about the concepts involved. However, studies using syllogisms, and other deductive reasoning tasks, have consistently reported two findings which require explanation: 1) people make many logical errors on such tasks (see Johnson-Laird & Byrne, 1991, Table 6.1); 2) responses may be influenced by the semantic content or context of the problem (content and context effects), or biased by their beliefs about the world (belief bias effects), despite their logical irrelevance (e.g. Evans, Barston & Pollard, 1983). So, for example, given two sets of premises which have the same logical structure, but differ in their semantic content:

Set 1: All of the Frenchmen are wine drinkers.
Some of the wine drinkers are gourmets.

Set 2: All of the Frenchmen are wine drinkers.
Some of the wine drinkers are Italians.

Ss would be more likely to draw the incorrect conclusion "Some of the Frenchmen are gourmets" for Set 1 because of its consistency with their prior knowledge about the world, but to correctly conclude "No valid conclusion" for Set 2. Any successful model of deductive reasoning must be able to account for these findings.

2.2.2.2 Inductive Reasoning

The concept of inductive reasoning is more complex than deductive reasoning, and this is reflected by the lack of a generally agreed definition of induction. For the purposes of this thesis, the definition suggested by Johnson-Laird (1993: p.60) will be adopted. He notes that in contrast to deductive inferences, an inductive inference is one which adds new information, going beyond the information given in the premises. The drawback is

that the conclusion may be false even if the premises are true, while a deductive inference will only be false if the premises are false. In fact, the conclusion of an inductive inference should be regarded as a hypothesis, which must be tested to see if it is correct. An example of an inductive inference is "All the swans I have ever seen are white, therefore all swans are white".

Because inductive inferences create new information, they are crucial to human thinking, and fundamental to learning, problem-solving, and concept formation. They are involved in detecting co-variation between elements, drawing analogies, generating hypotheses to describe or explain phenomena, scientific reasoning, understanding causality, and many other related processes. It is not possible in the confines of this review to detail all the work which has been carried out on induction, and therefore a brief review of perhaps the most fundamental inductive process, concept formation, will be given as an illustration.

Our knowledge of the world is composed largely of categories and concepts, and the relationships between them. 'Categories' are mental representations of a class of instances, such as dogs, while 'concept' is a broader term including mental representations that do not so obviously refer to classes, such as 'my pet fish' or 'love'. Early theories of how concepts are formed assumed each concept has a common, defining element learned by experiencing the noun and the object in different pairings, allowing abstraction of the common element (Hull, 1920). Bruner and his colleagues (Bruner, Goodnow & Austin, 1956) carried out the seminal work in this field. They reported a series of studies in which Ss were shown an array of stimuli which varied on dimensions such as colour and shape, and were asked to discover a concept, such as 'red and square', by selecting cards from the array and receiving feedback as to whether they were examples of the concept. Ss were found to use systematic strategies in their efforts to discover an unknown concept, and it was postulated that concepts are formed via an active, strategic, hypothesis-testing process. Several hypothesis-testing models were developed as a result of this work (e.g. Levine, 1966). More recently, these hypothesis-testing models have been supplanted, but Schustack (1988, p. 109) notes that although this type of model may not be representative of how people form concepts in the real world, they are very relevant to issues of hypothesis-generation, -testing and -revision. Later theories of concept formation have focused on the role of prototypes (e.g. Rosch, 1973).

It is clear inductive reasoning is a very important, but also a very complex field of study. Like deductive reasoning, there is evidence that humans make errors on inductive reasoning tasks and are prone to biases, with the most influential work in this area carried out by Amos Tversky and Daniel Kahneman (e.g. Tversky & Kahneman, 1974; 1983; see section 2.2.2.3.5). This again highlights the need for models of reasoning to take into account the role of content effects and belief bias effects.

2.2.2.3 Models of Reasoning

Having defined deductive and inductive reasoning, and given examples of the kind of studies which psychologists have used to study them, it is now important to consider some of the models postulated to explain the way people reason. These models have generally been developed within the field of either deductive or inductive reasoning. However, as Rips (1990) notes, while it is useful to divide the area into more manageable segments, it is unlikely the human brain has two separate systems for dealing with deductive and inductive inferences, and therefore the models will be considered together.

2.2.2.3.1 Formal rule theories

Formal rule theories were the earliest of the contemporary models of reasoning. They have their roots in the study of deductive reasoning within the field of logic. Formal rule theorists (e.g. Braine, 1978; Rips, 1989) assume humans possess an inherent mental logic comprising a set of general purpose, abstract inference rules or schemata which are applied in all contexts. According to this type of theory, reasoning is accomplished by translating problems into some form of abstract code so the rules of logic can be applied. For example, all the formal rule theories assume people possess the rule known as Modus Ponens by logicians. This states "If p then q", so if told:

> If the letter is an A then the number is a 3
> The letter is an A

Modus Podens is applied directly to give the answer "The number is a 3". In contrast, it is assumed people do not have the rule known as Modus Tollens, which states "If p then q, not q, then not p", since they are invariably poorer at solving problems which require this type of inference to be drawn. There are a number of criticisms of formal rule theory (see Garnham & Oakhill, 1994, pp.77-79), not least is its incompatibility with

the content, context and belief bias effects described in section 2.2.2.1 above. Thus, if humans reason by applying a set of logical rules, then there should be no mechanism by which the content of the problem or the beliefs of the S could influence the final response.

2.2.2.3.2 Mental models theory

The mental models approach was first postulated by Johnson-Laird (1983) to explain deductive reasoning and was later extended to cover inductive reasoning (Johnson-Laird, 1993). Like formal rule theories, it has its roots in the philosophical study of logic. It differs from formal rule approaches in that reasoning is conceptualised as semantic rather than syntactic in nature, and dependent not upon inference rules, but upon systematic procedures for constructing and evaluating mental models. The mental models approach proposes that reasoning proceeds through three main stages (e.g. Johnson-Laird & Bara, 1984): 1) Ss formulate a mental model to represent a possible state of the world consistent with the information supplied by the premises; 2) they formulate a putative conclusion by generating a description of the model that is semantically informative (not a repetition of a premise or a statement less informative than a premise); 3) the putative conclusion may be tested by trying to construct alternative models in which the premises of the argument are true but the conclusion is false. If no such counter-example is found, then the conclusion is inferred to be valid. Reasoning performance is thought to be constrained by available processing resources, so errors are predicted on problems which require more models to be constructed, and thus make greater processing demands. Errors may also arise if Ss fail to construct all the necessary models. The mental models approach is able to account for apparent biases and content effects in human reasoning because prior knowledge adds information to the models constructed. However, Evans, Newstead and Byrne (1993) note it does assume implicit understanding of some logical rules.

2.2.2.3.3. The componential approach

The componential approach is most closely associated with R. Sternberg (e.g. Sternberg, 1982; 1988). It has been applied to a range of cognitive processes, including reasoning. A component is defined as an "elementary information process that operates on internal representations of objects or symbols" (Sternberg, 1982, p.414). The number of components posited by Sternberg varies between accounts. Sternberg (1988) identified three types: performance components, metacomponents, and knowledge-acquisition components. Performance components are lower order processes involved in task

execution; metacomponents are executive processes used in the planning and monitoring of a task; knowledge-acquisition components are used in learning and storing new information. Individual differences are explained in terms of deficits relating to component use, while high correlations between performance on a range of tasks are explained in terms of shared performance components. Sternberg (1986; 1988) argued that three types of knowledge-acquisition components are crucial to reasoning: selective encoding, by which relevant information is distinguished from irrelevant information; selective comparison, by which relevant stored knowledge is retrieved from memory; selective combination, by which the selectively encoded and selectively compared information is selectively combined in WM. Sternberg (1986) proposed that both inductive and deductive reasoning are likely to involve all three. He suggests the difficulty of inductive problems derives from the selective encoding and selective comparison processes, both involve sorting relevant from irrelevant information when the limits on what is relevant may be unconstrained by the problem. In contrast, the difficulty of deductive problems derives from the selective combination process because of the need to identify the logically correct combination(s) from a number of possibilities.

Sternberg (1986) argues that in addition to the three processes (selective encoding, comparison, and combination), reasoning problems require the use of inferential rules, such as heuristics, mental guidelines, algorithms, and so on, and that "mediators" will influence the way in which the three processes (selective encoding, comparison and combination) can be applied to the inferential rules. Mediators are defined as "any intervening variable that increases or decreases the availability or accessibility of the inferential rules for use in a particular problem" (Sternberg, 1986, p. 290). Sternberg (1986) provides a list of mediators. For example, "prior knowledge" refers to the fact that if an inferential rule is simply unknown to an individual, then he will be unable to complete the task (unless it is possible to infer the rule). Thus, this approach does include a mechanism that can account for content and belief bias effects.

2.2.2.3.4 Domain-sensitive rules or schemata

This type of theory assumes reasoning is achieved by domain-sensitive rules, or by schemata, elicited by the context, which contain rules that can be applied to reasoning in a particular domain. The theory of pragmatic reasoning schemata (Cheng & Holyoak, 1985) is perhaps the best known theory of this type. It is specifically concerned with the

explanation of content effects in reasoning, and has been restricted mainly to the explanation of findings with one particular task (the four card selection task; Wason, 1966). It is suggested people learn to reason in certain contexts and formulate schemata to abstract this knowledge. These schemata consist of a set of generalised, context-sensitive rules defined in terms of classes of goals and relationships to these goals. It is argued that tasks that may be difficult to solve when presented in abstract form, may become easier when they are set within a context that elicits an appropriate schema. This type of theory has the advantage of explaining the effects of content and context on reasoning. Its weakness is that it cannot explain how people reason on abstract or artificial tasks, although it is possible that people have a general-purpose method of reasoning based on mental models or some other mechanism for use with unfamiliar problems, but tend to use schemata in domains where they have relevant experience.

2.2.2.3.5 Heuristics and biases

The terms 'heuristic' and 'bias' are frequently used interchangeably, but Evans et al. (1993) distinguish between them. They define 'heuristic' as a theoretical construct to describe reasoning processes which may lead to a quick solution, but are liable to error. They define 'bias' as empirical observations that systematically attend to logically irrelevant information, or neglect to attend to relevant information. Heuristics may be useful to reduce the difficulties of making predictions from complex data, but their overuse may lead to a variety of inferential errors. The work of Tversky and Kahneman (e.g. Tversky & Kahneman, 1974) is most strongly associated with the concept of heuristics. They identify three heuristics used to assess probabilities and predict values. For example, the availability heuristic refers to situations in which people assess the frequency or probability of an event by the ease with which instances can be brought to mind. The classic example of bias resulting from this type of strategy is demonstrated by asking Ss to judge whether words that start with the letter 'r' are more or less common than those that have 'r' as their third letter. Because it is easier to call to mind words that begin with 'r', Ss frequently conclude incorrectly that these are more common.

Pollard (1982) considered how the availability heuristic might play a role in the content effects and belief biases commonly seen in performance on deductive reasoning tasks (see section 2.2.2.1). He argued that the availability of salient problem features, or associated information retrieved from memory, influences the response to deductive reasoning tasks,

and that a number of known error tendencies lend themselves to an explanation in terms of availability. For example, a number of studies (e.g. Evans et al., 1983; see section 2.2.2.1) have found Ss are more likely to accept an invalid syllogism as valid when the content makes it 'true' in terms of their pre-existing knowledge of real-life. Pollard (1982) notes the truth status of a syllogism can only be judged on the basis of what the S can retrieve from their experience, or, in Tversky and Kahneman's (1974) terms, on the basis of the availability of relevant information. Pollard (1982) argues the availability of the conclusion may directly mediate evaluation of both truth status and validity.

Evans (1984) noted the importance of identifying what information Ss take into consideration when carrying out a reasoning task. He distinguished between pre-attentive 'heuristic' processes involved in the selection of 'relevant' and 'irrelevant' information, and 'analytic' processes that operate on the selected items to generate inferences or judgements. This theory emphasises relevance rather than availability. Factors that contribute to heuristic selection include perceptual salience, linguistic suppositions and semantic associations. Biases occur because relevant features of the task are not selected or irrelevant features are selected during the heuristic stage.

Nisbett and Ross (1980) conducted an extensive review of the literature relating to inferences in the area of social judgments. Their review suggests humans make widespread use of a whole range of heuristics and knowledge structures, such as schemata, when making social judgements. Nisbett and Ross (1980) concluded most judgmental errors among lay people were due to the *over-use* of these inferential heuristics and knowledge structures. They noted the use of heuristics and knowledge structures was optimal in many situations, such as when the judgement was trivial ("Shall I have vanilla or strawberry ice-cream?") or when a more appropriate strategy was unavailable. However, they concluded humans also used heuristics and knowledge structures in preference to more appropriate strategies. For example, when deciding which car to buy, an individual may give undue weight to anecdotal evidence about one particular type of car and pay less attention to performance statistics. Even when more appropriate strategies are used for a particular judgemental task, the undue influence of the simpler, more intuitive strategies may persist.

2.2.2.3.6 Summary of cognitive psychology models of reasoning

It is clear from this brief review of reasoning models that while this area has advanced very quickly over the past thirty years, it is still some distance from achieving a unified approach. This can be explained in part by the way the study of reasoning has been conducted. In particular, the detailed study of a small number of tasks has resulted in models that account for performance on one task, but are difficult to generalise. For example, there are more than five models of syllogistic reasoning (for review see Garnham & Oakhill, 1994, pp.100-116). More recently, theorists have recognised the need for theories with a wider scope (e.g. Evans, 1991; Johnson-Laird, 1993; Rips, 1990; Sternberg, 1986), and have considered the possibility of integrating some of these approaches. For example, Evans et al. (1993) note the model proposed by Evans (1984), which suggests reasoning involves both heuristic and analytic processes, does not specify the mechanism which carries out analytic processing. Therefore it is not incompatible with other theories, such as mental models theory (Johnson-Laird, 1983), that are concerned only with analytic processes.

2.2.3 NEUROPSYCHOLOGICAL MODELS OF REASONING/PROBLEM-SOLVING

The evolution of neuropsychological approaches to reasoning and problem-solving has followed a rather different path to the cognitive psychology approach. While the cognitive approach began with formal models of reasoning from the field of logic which were then tested and modified in the light of human performance, the neuropsychological approach grew from work carried out to delineate the deficits associated with damage to the frontal lobes of the brain. The frontal lobes have been associated with a range of cognitive abilities (see e.g. Levin, Eisenberg & Benton, 1991), but the current review will focus on the role of the frontal lobes in reasoning and problem-solving. Studies of frontal lobe performance on measures of reasoning and problem-solving are reviewed below.

2.2.3.1 Performance of Frontal Lobe Patients on Reasoning and Problem Solving Tasks

Frontal lobe damage has been found to be associated with deficits on rule-finding tasks, with the Wisconsin Card Sorting Test (WCST, Grant & Berg, 1948) being the most frequently reported. In this test Ss are given a series of cards and asked to sort them into one of four piles. The experimenter gives the S feedback about the correctness of each

sort, and the rule the experimenter is following changes in a pre-determined fashion, although the S is not aware of this. In one study, Nelson (1976) found frontal lobe patients tended to identify fewer rules, and make more errors on a modified version of the task. In particular they were characterised by perseverative errors, that is, they carried on sorting to a previously correct rule regardless of negative feedback from the experimenter. Cicerone, Lazar and Shapiro (1983) compared anterior and posterior lesion groups on a concept discrimination learning task. The task was a modification of a paradigm developed by Levine (e.g. Levine, 1966) which is described in more detail in chapter IV. In essence, Ss had to identify which rule was in operation by testing hypotheses and receiving feedback from the experimenter. Frontal lobe patients achieved fewer correct solutions and tested fewer appropriate hypotheses, and again there was a tendency for frontal Ss to make perseverative errors.

Frontal lobe deficits have also been found on tasks that require planning and goal-setting. Vilkki and Holst (1989) found frontal lobe patients were impaired in their learning of spatial sequences when the sequence length was subject-determined, but were no different to patients with posterior lobe damage when the sequence length was set by the experimenter. Similarly, Petrides and Milner (1982) found frontal lobe patients to be impaired on 'self-ordered pointing tasks' in which Ss were presented with a series of arrays of pictures, words, or abstract designs. Each array in a particular series contained the same items in different spatial locations. The Ss' task was to point to one item on each array in the series until all of the items had been touched once only. This finding has been replicated by Wiegersma, van der Scheer and Hijman (1990) and by Owen, Downes, Sahakian, Polkey & Robbins (1990) using modified versions of the task.

Shallice (1982) investigated planning and goal-setting in frontal and posterior patients and normal controls using the Tower of London task. In this task, three beads are arranged in various starting positions on sticks of unequal length, and Ss are required to move them to a goal position in the minimum number of moves, with certain constraints on the types of possible move. Patients with left anterior lesions solved significantly fewer problems, and this was explained as a deficit in the planning component of the task, although Shallice (1988, p.347) reported a study carried out by Shallice, Warrington, Watson and Lewis which failed to replicate this finding. Owen, Downes, Sahakian, Polkey & Robbins (1990) used a modified and computerised version of the Tower of London task, and found

frontal lobe patients to be impaired relative to a group of normal controls. The frontal lobe patients took significantly more moves to solve the problems and a yoked motor control condition revealed movement times were significantly increased in this group. Analysis of the results showed initial planning time was unimpaired in the frontal group, but thinking time subsequent to the first move was significantly prolonged. The authors interpreted this as evidence that the frontal lobe patients made their first move impulsively, before they had generated an adequate solution to the problem.

Shallice and Burgess (1991) described the performance of three patients with frontal lobe damage on two tests that require planning and monitoring of behaviour. The Six Element Test (SE) consists of six open-ended tasks that Ss must carry out in a fixed amount of time in such a way as to maximise their score without breaking certain rules. The Multiple Errands Test (ME) is similar in principle to the SE task in that Ss must carry out a number of different tasks with minimal environmental cues to guide performance. Ss are taken to a shopping precinct and given a list of simple tasks to carry out, such as buying certain items, while following certain rules. On both these tasks the three frontal Ss were found to be both quantitatively and qualitatively impaired relative to normal controls. On the SE task frontal Ss attempted fewer sub-tasks than the controls, while on the ME task they were more likely than controls to use an inefficient strategy, to break the rules of the task, and to fail to achieve a sub-task.

General problem-solving skills were assessed in patients with anterior versus posterior lesions by Shallice and Evans (1978) using a cognitive estimates task that requires Ss to make best-guess estimates regarding facts such as "The length of the average man's spine." This requires Ss to generate possible solutions and evaluate their likelihood, and thereby provides a measure of problem-solving ability. Patients with anterior lesions produced responses significantly more outside the normal range of responses than posterior Ss. Frontal patients have also shown deficits on verbal fluency tasks (e.g. Perret, 1974), where Ss are asked to generate as many words as possible according to a rule given by the experimenter, such as "words beginning with the letter 'F'". Again, these deficits are thought to result from a failure to generate and test appropriate solutions. Burgess and Shallice (1996) examined the performance of patients with anterior and posterior lesions on the Hayling test. This is a sentence completion task with two conditions. In the first condition Ss are required to give a word that completes the sentence appropriately, while

in the second condition they are required to produce a word unrelated to the sentence. In comparison with patients with lesions elsewhere, patients with frontal lobe involvement showed longer response latencies in the first condition and produced more words related to the sentence in the second condition. Furthermore, in the second condition patients with frontal lobe lesions were less likely to produce words that suggested the use of a strategy during response generation.

In summary, frontal lobe patients have been found to show deficits on a range of reasoning and problem-solving tasks, including measures of rule-finding, hypothesis-testing, planning, goal-setting and monitoring of behaviour. In general, clinical descriptions of the behaviour of frontal lobe patients have noted perseveration, distractibility, failure to inhibit inappropriate responses, and failure to initiate actions (e.g. Milner, 1964). This pattern of responding is known as frontal lobe or dysexecutive syndrome, and deficits on reasoning and problem-solving tasks form just a part of this. The diversity of deficits associated with damage to the frontal lobes has resulted in a profusion of theoretical accounts of frontal lobe function. A brief and selective review of the most important and relevant of these is given below.

2.2.3.2 Theories of Frontal Lobe Deficits

Luria (e.g. 1973) provided one of the most influential of the earlier descriptions of frontal lobe function, and many of his theoretical ideas have been incorporated in later theories. Luria (1973) postulated the frontal lobes play an essential role in modulating the activity of 'lower' brain systems when the individual is carrying out complex tasks. He believed the frontal lobes have a crucial role in directing non-routine mental activity requiring the initiation and maintenance of a plan. He suggested the frontal lobes are necessary when forming a plan of activity, executing a complex "program of activity", organising a strategy, and evaluating the action taken. His description of the frontal lobe syndrome included a reduced ability to direct attention appropriately; an increase in "stereotyped" behaviour and habitual routines; and a reduced ability to plan, organise and reflect on a course of action. Luria (1973) believed that studying the performance of frontal lobe patients on tasks requiring complex intellectual activity, such as problem-solving tasks, would be most revealing of the deficits associated with damage to the frontal lobes.

Milner (1982) refined Luria's (e.g. 1973) notion of a modulatory function for the frontal

lobes. She hypothesised that frontal lobe lesions in man are associated with two dissociable types of deficit: one in the ability to modulate self-generated plans, the other in the ability to monitor external events. This argument was based on the results of a series of studies comparing the performance of patients with unilateral frontal lobe lesions (i.e. lesions of either the right or left frontal lobe) on a range of tasks. On self-ordered pointing tasks of the type described above (see section 2.2.3.1), which are thought to measure planning ability, there was evidence that the left frontal lobe played a major role. In contrast, Ss with right frontal deficits were more impaired than left frontal Ss on a task of recency discrimination (deciding which of two stimuli had been presented more recently), suggesting a failure to monitor external events.

One of the most prominent theories of recent years is was posited by Norman and Shallice (1986), and has been described in detail in a number of other publications (e.g. Shallice, 1982, 1988). This theory elaborates on some of the ideas put forward by Luria (e.g. 1973) described above. The Norman and Shallice (e.g. Shallice, 1982; 1988) model assumes both cognition and action depend upon the 'running' of highly specialised routine schemata, each controlling a specific over-learned action or skill. Schemata may be activated in various ways, such as 'triggers' from perception, or the output of other schemata, and, since they are activated in ways that are independent of each other, several schemata may run at once. The novel aspect of this model is the inclusion of two processes in the selection of schemata. Contention scheduling is held to involve routine selection between the schemata, while the Supervisory Attentional System (SAS) is required in non-routine or novel situations. The model predicts a specific deficit of the SAS should not affect the performance of routine tasks even if they require considerable processing resources, but would result in difficulty in coping with novelty, in planning, or in overcoming a strong habitual response. The predicted impairments fit the classical view of frontal lobe dysfunction. If the SAS is inoperative, then behaviour will be controlled by the contention scheduling process, and this will lead to behaviour characterised by perseveration, inability to initiate activities, distractibility, failures of planning, and inability to inhibit habitual patterns of behaviour. These are all explained as a failure of the SAS to initiate appropriate activities, or interrupt inappropriate ones.

Although the Norman and Shallice (e.g. Shallice, 1982, 1988) model of frontal lobe function described above goes some way to capturing the 'flavour' of frontal lobe

dysfunction, it also has at least one severe limitation. McCarthy and Warrington (1990, p.363) note the single supervisory system proposed by the model is not compatible with evidence of dissociations between different frontal lobe functions (e.g. Milner, 1982; see above). This point has been noted in later work by Shallice and Burgess (1991; 1993) who found evidence of fractionation whilst investigating possible measures of the SAS. As a result, Shallice and Burgess (1991; 1993) have modified the concept of a single resource SAS to encompass the possibility of fractionation. It is now postulated to consist of an unspecified number of sub-components (Shallice & Burgess, 1991; 1993).

Other approaches to frontal lobe function have placed different emphases on the various aspects of behaviour associated with frontal lobe syndrome. For example, Fuster (1989) emphasised the role of the frontal lobes in the temporal organisation of behaviour. He suggested there are at least three cognitive functions that can be identified as specific to the frontal lobes: 1) WM, which permits referral to relevant preceding events; 2) anticipatory set, which uses past experience to anticipate and prepare for future events; and 3) an interference control that inhibits the disruption of goal-directed behaviour by behaviour that is incompatible with it. Fuster argues the unifying purpose of these three components is the temporal organisation of behaviour and the structuring of goal-directed behaviour. The frontal lobes serve to bind together temporally distal events and behaviours for the purpose of goal attainment. Fuster emphasises it is the discontinuity between environmental events, behavioural responses, and their goals which make the role of the frontal lobes essential. In a different approach, Dempster (1991) emphasised the inhibitory function of the frontal lobes He argued the suppression of irrelevant stimuli or associations may play an important, and often unrecognised, role in determining "intelligent" behaviour. Many tasks require the suppression of task-irrelevant information for effective performance, and the suppression of task-irrelevant thoughts and perceptions may be an important factor in the acquisition of higher-level knowledge structures such as strategies. Dempster suggests individual and group differences in the capacity for inhibition are manifestations of frontal lobe function.

Kimberg and Farah (1993) have criticised existing models of frontal lobe function, including those of Luria (e.g. 1973), Milner (1982), Norman and Shallice (e.g. Shallice, 1982; 1988; Shallice & Burgess, 1991), Fuster (1989) and Dempster (1991) reviewed above, on the basis that no single model proposed thus far is able to account for the full

range of deficits associated with frontal lobe damage. Instead, each model may provide a good fit to deficits shown on one or two tasks, but be less applicable to deficits on other tasks. Kimberg and Farah (1993) propose a model of their own which, they argue, is capable of explaining a wider range of frontal lobe deficits than earlier models.

Unlike some of the earlier models (e.g. Luria 1973; Shallice, 1982; 1988), Kimberg and Farah (1993) do not posit a central executive or SAS function for the frontal lobes. Instead, the central tenet of their theory is that the frontal lobes are involved in maintaining associations among the elements of WM. These elements include representations of goals, stimuli in the environment, and stored declarative knowledge. It is argued that these representations are intact in frontal patients, but the associations between them are impaired. Frontal lobe damage would therefore be associated with a lowered sensitivity to the mutual relevance of goals and stimulus attributes and to the relations between facts and their contexts. In situations where several sources of information could potentially influence behaviour, connections between internal representations may be critical in determining which is successful. For example, responding to a stimulus in the environment that is relevant to a particular goal. If these connections are weakened by frontal lobe damage, other sources become more important in determining behaviour. Thus, behaviour may be influenced by stimuli irrelevant to current goals. Normally these stimuli would be ignored because of their lack of association in WM with any goals.

The model proposed by Kimberg and Farah (1993) has some advantages over earlier models of frontal lobe function in that it avoids the conceptual difficulties, outlined above, associated with the notion of a single executive or SAS. Furthermore, Kimberg and Farah (1993) have successfully modelled their approach and simulated performance on a range of tasks sometimes impaired by frontal lobe damage. These simulations produced patterns of performance qualitatively similar to those often seen in frontal lobe patients. However, there is still much work that needs to be done in extending this model to other tasks, in exploring dissociations between tasks, and in specifying the way WM functions might be organised within the frontal lobes.

In summary, the neuropsychological approach to understanding reasoning and problem-solving is closely related to theories of frontal lobe function. The frontal lobes are still

relatively poorly understood in neuropsychology, chiefly because the effect of frontal lobe damage has been observed on a diverse range of tasks. This has resulted in a proliferation of theories, each pertinent to understanding specific aspects of frontal lobe function, but without the capacity to explain the full range of frontal lobe deficits. Nevertheless, the field has made progress in the last thirty years, and it is interesting to note there are increasing links with theories of reasoning from the field of cognitive psychology discussed above (see section 2.2.2). In particular, the central role of some form of WM is emerging from both cognitive and neuropsychological approaches. With regard to this, current models of WM will be explored in more detail in section 2.2.5.2, but first the implications of these approaches to reasoning and problem-solving will be considered in relation to the models of thinking in depression outlined in section 2.1.3.3.

2.2.4 MODELS OF REASONING IN RELATION TO DEPRESSION

Having reviewed current models of reasoning in the fields of cognitive psychology and neuropsychology, the next step is to consider how these can increase understanding of reasoning in depression. From the evidence relating to cognitive theories of depression (see section 2.1.3.3) it was concluded that a simple distinction between distorted, biased depressive thinking and logical, accurate 'normal' thinking was not supported. While there was evidence that depressed and nondepressed Ss differed in their 'thinking', with studies consistently finding significant group differences, neither was found to be consistently more accurate than the other, with findings highly susceptible to experimental manipulation. Similar conclusions have been reached by other reviews (e.g. Haaga and Beck, 1995; Power & Champion, 1986; Watts, 1992). The cognitive psychology approach to reasoning is highly informative with regard to understanding these findings. This work suggests humans do not reason in a logically perfect way, with strong evidence that errors, biases, and failure to attend to all the relevant information may be the norm.

Hayes and Hesketh (1989) postulated that the use of 'normal' heuristics may lead to many of the inferential errors made by depressed patients. For example, the availability heuristic (see section 2.2.2.3.5) might lead them to overestimate the frequency or probability of a negative event. Thus, a 'normal' heuristic, but with a negative bias, may underlie depressive 'thinking errors'. Therefore depression might be associated with qualitative differences in reasoning rather than an actual impairment, contradicting both Beck's (e.g. 1967; 1976) original assertion that depression is associated with impaired

thinking, and the subjective reports of depressed patients that their thinking is impaired. However, the research considered thus far has been carried out with tasks containing emotionally-salient or personally-relevant material which is likely to elicit thinking biases. Therefore it is important to explore any effects of depression on reasoning when the task material is neutral. Until this is done it is not possible to reject the hypothesis that depression is associated with impairments in reasoning ability.

The neuropsychological literature provides information about the brain areas involved in reasoning, and the nature of reasoning deficits following damage to these. This is of interest because of recent advances in the neuropsychology of depression. Work using advanced neuro-imaging techniques suggests the frontal lobes may play a central and important role in depression (e.g. Bench et al. 1992; Cummings, 1993). Frontal lobe dysfunction is associated with quantitative impairments on reasoning tasks using neutral material. Thus, if depression is associated with frontal lobe dysfunction, this may lead to a prediction of quantitative depressive reasoning deficits on neutral tasks.

The review of cognitive psychology and neuropsychology approaches to reasoning revealed a number of competing models. However, there is consensus on the factors that limit reasoning performance: attention, WM, and retrieval of knowledge from long-term memory (LTM). When reasoning tasks exceed available processing resources, Ss are more likely to make errors or to use heuristic strategies. If depression is associated with reduced processing resources then this could result in reasoning deficits. Depressed patients frequently complain of difficulties with attention and memory in everyday life. With regard to emotionally-salient or personally-relevant material the evidence suggests depression may be associated with biases in memory, and possibly attentional, processes. Thus, depressed Ss may allocate processing resources to negative material, leaving fewer resources available for performing neutral tasks. This is one possible mechanism by which quantitative depressive reasoning deficits on neutral reasoning tasks might arise.

Laboratory reasoning tasks are usually abstract and novel. They are deliberately chosen to limit the effects on performance of past experience and existing knowledge and therefore usually place greater demands on attention and WM than on LTM. Therefore, in the next section current models of attention and WM will be outlined, and their implications for reasoning will be considered in section 2.2.6.

2.2.5 MODELS OF ATTENTION AND WORKING MEMORY

2.2.5.1 Attention

2.2.5.1.1 The concept of attention

Early theorists assumed that attention was a single function. More recently, theorists have come to accept that attention is a multidimensional rather than a unitary construct, involving both cognitive and psychophysiological variables (Barkley, 1996). Perhaps the most common distinction within the field of attention is between selective and sustained attention. That is, the ability to selectively attend to some inputs rather than others versus the ability to sustain attention over a period of time. However, orienting, vigilance, focusing, arousal, divided attention, inhibition, and shifting have all been studied within the field of attention. Psychophysiological approaches to attention have included assessment of variables such as heart rate, skin conductance, pupillary dilation, measures of brain electrical activity derived from the scalp electroencephalogram (EEG), and event-related brain potentials (ERPs). The latter are measures of small but reliable signals produced by the brain in response to environmental events.

2.2.5.1.2 Automatic versus effortful processing

Models of attention usually incorporate the concept of limited resources, and a mechanism to select what is processed. These two notions are intrinsically linked, since if resources were not limited, then there would be no need to attend selectively (Schneider & Shiffrin, 1977). It is because there is an overload of information in many situations that a sub-set of information must be selected for attention.

Theorists have postulated that some cognitive processes make only minimal demands on attentional resources (Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). These are known as automatic processes. At the other end of the continuum are processes demanding of attentional resources. These have been labelled "conscious" (Posner & Snyder, 1975), "controlled" (Schneider & Shiffrin, 1977), or "effortful" (Hasher & Zacks, 1979). Automatic and effortful processes are thought to differ in a number of important functional characteristics, such as their correlation with awareness and intention, their susceptibility to interference, and the effects of stimulus load (Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). There is some disagreement as to the exact nature of these differences.

Hasher and Zacks (1979) outlined the functional characteristics of both automatic and effortful processes, and claimed their definition incorporates that of Posner and Snyder (1975), and most of the points made by Shiffrin and Schneider (1977). Hasher and Zacks (1979) suggest automatic processes operate continually to encode certain attributes of whatever information is the focus of attention. They do not require either awareness or intention, and therefore make minimal demands on attentional resources, although the products of automatic processes may enter consciousness. When a process is automatised, it cannot be improved upon by practice or feedback about performance, and it cannot be inhibited. It is possible to pay attention to information that would otherwise be encoded automatically, but this serves only to reduce available resources without improving performance. Because automatic processes make minimal demands on attentional resources, they allow the organism to operate even when high demands are made upon capacity, as in moments of stress. Some automatic processes are thought to result from innate factors, such as encoding the frequency, location and timing of events, while others are the result of extensive practice, such as certain aspects of driving in an experienced driver. Those that result from practice are thought to show more between-individual variation and to be more susceptible to disruption than those that are genetically prepared. Hasher and Zacks (1979) suggest these "learned" automatic processes may fall on a continuum between automatic and effortful processes.

Hasher and Zacks (1979) contrasted automatic processes with effortful processes which require attentional resources, and so limit ability to carry out other effortful processes simultaneously. The efficiency of effortful operations increases with practice, and their use is voluntary, often occurring only with specific instructions. The individual is almost always aware of the activity of effortful processing mechanisms, and a wide range of individual differences may be seen. Hasher and Zacks (1979) see their definition of effortful processes as analogous to Posner and Snyder's (1975) "strategies", and Shiffrin and Schneider's (1977) description of "accessible controlled processes".

Another important, and related, distinction is that between implicit and explicit processes (for reviews see e.g. Berry & Dienes, 1993; Richardson-Klavehn & Bjork, 1988; Roediger, 1990). Implicit memory refers to demonstrations that Ss' performance is influenced by prior experience with particular stimuli, although they may not be able to recall it deliberately. This contrasts with explicit memory, which refers to the deliberate,

conscious recollection of previously studied material. Relating implicit and explicit processing to the automatic-effortful distinction described above, Parkin and Russo (1990) suggest explicit measures such as free recall reflect effortful processes and implicit measures of performance may reflect the operation of automatic processes.

2.2.5.1.3 Models of attention

As noted in section 2.2.5.1.2, models of attention usually incorporate the concept of limited resources and a mechanism to select what is processed. Early models of attention (e.g. Broadbent, 1958) postulated a series of stages between input and response, with selective attention acting at a particular stage, depending on the model, thereby creating a bottleneck. Kahneman (1973) proposed a highly influential alternative. He suggested attentional limits were not due to a 'bottleneck' at a specific stage of processing, but instead that there are general resources that can be allocated flexibly to different stages of processing or to different tasks. Within this model, limits to attentional processing are explained in terms of finite resources. Kahneman (1973) did not see capacity as permanently fixed, but as varying between individuals, and also within individuals as a result of changes in factors such as mood and arousal.

Kahneman (1973) proposed cognitive processes differ in the amount of attentional resources they require, with processes such as sensory analysis requiring none or very little, and rehearsal making heavy demands, and that tasks differ in their demands, with easy tasks making fewer demands than difficult tasks. Kahneman (1973) argued performance of a task should be positively related to the resources available to it, and that when the supply of attention did not meet the demands, performance should falter, or fail entirely. Thus, an activity could fail because of insufficient resources, or because resources were engaged by other activities. Hasher and Zacks (1979) note that variation in attentional capacity should have major effects on the efficiency with which effortful processes occur.

The concept of task difficulty is important and needs to be considered. As noted above, it is assumed that difficult tasks require more attentional resources than easy tasks, but this argument can become circular if tasks that seem to make heavy demands on resources are then labelled as difficult. Wickens (1989) suggests the concept of difficulty can be defined in terms of task characteristics and degree of automaticity. For example,

increasing the complexity of mapping from stimuli to responses should increase difficulty, while increasing the automaticity of the task should decrease difficulty. However, it is not always possible to determine these a priori, and at present task difficulty is usually determined by empirical observation of performance levels. Some authors (e.g. Logie, Gilhooly & Wynn, 1994) have argued this reduces its usefulness as a concept.

In a development of Kahenman's (1973) theory, Norman and Bobrow (1975) noted task performance depends on both the quality of the data and upon the processing resources used. They distinguished between resource-limited and data-limited processes: when an increase in the amount of processing resources results in improved performance, then performance on the task is labelled as resource-limited; and when performance is independent of processing resources then it is labelled as data-limited. In general, most tasks will be resource-limited up to the point where all the processing that can be done has been done, and data-limited from then on, and so almost all processes will have regions that are resource-limited and regions that are data-limited. Norman and Bobrow (1975) suggested that failure to recognise this distinction may explain the apparent discrepancies in research findings. If one study finds two tasks interfere with each other, and a second study finds no interference, then the difference may be because one studied the tasks at levels where they were resource-limited, whereas the other did not.

2.2.5.1.4 Central versus multiple resources

One of the fundamental debates within the study of attention concerns the division of attention among concurrent mental activities. Both structural (e.g. Broadbent, 1958) and capacity (e.g. Kahneman, 1973) models predict that concurrent activities are likely to be mutually interfering, but they ascribe the interference to different causes. In a structural model, interference occurs when the same mechanism is required to carry out two incompatible operations at the same time. In a capacity model, interference occurs when the demands of two activities exceed available resources. Thus a structural model implies that interference between tasks is specific, and depends on the degree to which the tasks call for the same mechanisms, whilst in a capacity model interference is nonspecific, and depends only on the demands of the tasks. Studies show both types of interference occur, indicating that deployment of attention is more flexible than expected under the assumption of a structural bottleneck, but more constrained than expected under the assumption of free allocation of resources (Kahneman, 1973).

Other authors (e.g. Navon & Gopher, 1979; Wickens, 1984) have argued neither structural or capacity models are able to account for all the findings. Instead, it has been postulated (e.g. Navon & Gopher, 1979; Schneider & Detweiller, 1987; Wickens, 1980; 1984; 1989) that multiple resources may exist, each with its own capacity which may be shared by concurrent processes, and that different tasks may require different resources in various compositions. Navon and Gopher (1979) noted this approach has elements of both structural and capacity models. It is structural because it identifies the limit on performance with the availability of any one of several processing mechanisms and ascribes task interference to the overlap in engaged mechanisms, but it is a capacity approach in that it posits the sharing of resources by several processes. Wickens (1989) argued the level of dual-task performance is a joint function of the amount of resources invested, as determined by task difficulty, and the degree of resource competition between the tasks within a pair. If the difficulty of one task becomes too great, then parallel processing may be abandoned in favour of serial task performance.

The challenge for multiple-resource theories is to define the functional composition of these separate resources. Wickens (1980) argued resources may be defined by a three-dimensional metric consisting of stages of processing (e.g. encoding versus responding), hemisphere of processing (e.g. verbal versus spatial), and modalities of processing (e.g. manual versus vocal responding). Schneider and Detweiller (1987) assume resources vary in terms of the type of material stored, the time required to store material, proactive interference effects, retrieval time, trace decay, and the robustness of the storage. The WM model of A. Baddeley and G. Hitch (Baddeley, 1986; Baddeley & Hitch, 1974) is also based on the concept of multiple resources, and this will be presented in more detail in section 2.2.5.2.2.

Although resource theory has been extremely popular over the past twenty-five years, it is not without its critics. Navon (1984) launched a strong attack on resource theory, arguing it is circular and self-reinforcing, and that effects usually explained within the resource theory framework can be explained quite adequately without calling on this concept. He suggests resource theory may actually act as a useful metaphor for the way cognitive processes operate, rather than reflect an actual mechanism.

The debates relating to attention have yet to be resolved, but a common theme is the

increasing overlap between concepts of attention and concepts of WM. In the next section, the topic of WM will be considered, and this will be followed by a discussion of the implications of issues relating to attention and WM for understanding reasoning processes.

2.2.5.2 Working Memory

Models of WM have their roots in the early models of short-term memory, and therefore these will be considered before moving onto models of WM.

2.2.5.2.1 Early models of short-term memory

Early models of information-processing assumed at least two memory systems: short-term memory (STM) and LTM. Models proliferated in the 1960s and early 1970s, typified by the models of Atkinson and Shiffrin (1968) and Waugh and Norman (1965). Miller (1956) had earlier postulated STM might be limited to the "magic number" of seven plus or minus two items, and so STM was conceptualised as a brief storage system of limited capacity. Waugh and Norman (1965) described STM capacity as a hypothetical buffer with a fixed number of slots, all incoming information was either rehearsed or forgotten. If rehearsed, information transferred to secondary memory or LTM from which it decayed more slowly. The Atkinson and Shiffrin (1968) model assumed three components. The first was a sensory register capable of holding information from the different sense modalities. Information from the sensory register was thought to feed into the second component, STM, which was postulated to be crucial to a range of tasks, and to be a necessary intermediate stage in transferring information to the third component - LTM.

2.2.5.2.2 The concept of working memory

Baddeley and Hitch (1974) proposed a more complex STM system than the earlier models described in section 2.2.5.2.1 above, and this has proved to be extremely influential. The term WM developed as a way to refer to a more active part of the human processing system (e.g. Atkinson & Shiffrin, 1968). WM is assumed to have processing as well as storage functions: it serves as the site for executing processes and for storing the products of these processes (Baddeley, 1986; Baddeley & Hitch, 1974). While the concept of STM as a WM system was not original to Baddeley and Hitch (1974), they did introduce the notion of WM as several separable subsystems. The Baddeley and Hitch (1974) model therefore represents a multiple-resource model (see section 2.2.5.1.4).

The WM model of Baddeley and Hitch (Baddeley & Hitch, 1974; Baddeley, 1986) consists of a central executive (CE) and two slave systems. More is known about the latter than the former, probably because they are relatively simple, and they are thought to function as short-term stores of verbal and visuo-spatial material. The phonological loop (PL), also known as the articulatory loop, is made up of a phonological input store and an articulatory rehearsal process, while the visuospatial sketch pad (VSSP) is specialised to maintain and manipulate visuo-spatial images. Both are basically input stores, but are also active in the sense that memory traces are thought to be regenerated by a process outside the store itself. Finally, the CE is described as a limited-capacity processor which is assumed to co-ordinate information from the PL and VSSP. It serves the role of allocating attentional resources and of selecting and operating central control processes and strategies. It plays a crucial role in mediating complex cognitive processes such as reasoning and problem-solving (Baddeley, 1986). Baddeley (e.g. 1990, p.132) has suggested the CE actually functions more like an attentional system than a memory store. Therefore, WM is essentially the limited capacity system discussed in section 2.2.5.1.3 above. Furthermore, Baddeley (e.g. 1986; 1990) has argued the SAS postulated by Norman and Shallice (e.g. Shallice, 1982; 1988) and described in section 2.2.3.2 gives a good account of the functioning of the CE.

Having considered theories and issues pertaining to attention and WM, it is now possible to consider their role in reasoning. This will be addressed in the next section.

2.2.6 THE RELATIONSHIP BETWEEN ATTENTION, WORKING MEMORY AND REASONING

As noted in section 2.2.4, while theorists may disagree on the extent of human competence in reasoning, there is agreement that performance is constrained by underlying factors such as attention, WM, and LTM. The role of LTM in reasoning is considered briefly below, while attention and WM are considered in more detail.

All of the reasoning models outlined in section 2.2.2.3 describe processes that rely on access to information in a long-term store. For example, formal rule theories (see section 2.2.2.3.1) assume people reason by applying existing logical rules, presumably stored in LTM, to new situations. Similarly, the componential approach put forward by Sternberg (e.g. 1986; see section 2.2.2.3.3) assumes the pre-existence of inferential rules that can be

applied to a current problem. The mental models theory put forward by Johnson-Laird (1983; see section 2.2.2.3.2) postulates that people use prior knowledge to "flesh out" the mental models they construct of a particular situation. Theories that postulate domain-sensitive rules or schemata (e.g. Cheng & Holyoak, 1985; see section 2.2.2.3.4) assume people reason using schemata learned in previous situations and stored in LTM. Theories that posit a role for heuristics or biases in reasoning (e.g. Evans, 1984; Sternberg, 1986; see section 2.2.2.3.5) also clearly assume a role for LTM in reasoning. However, as noted above, reasoning tasks chosen for study by psychologists, including those outlined in chapters IV-VIII of the current thesis, are selected to minimise the role of LTM processes. This is achieved by choosing abstract, novel tasks. The aim is to reduce the impact of Ss' prior experience, since this represents an uncontrollable variable. Therefore, although the role of LTM in reasoning is acknowledged, the focus of the current review is on attentional and WM processes in relation to reasoning.

As noted in section 2.2.5.1.1, attention is assumed to be a multi-dimensional rather than a unitary concept. Furthermore, it is now recognised that attentional and WM processes are closely interlinked, with WM fulfilling the role of the limited capacity resource posited by many theories of attention (e.g. Kahneman, 1973; see section 2.2.5.1.3). All the processes associated with the concept of attention (see section 2.2.5.1.1) are likely to play a role in reasoning. In particular, as noted by both Evans (1984) and Sternberg (1986), the selection of information considered to be relevant to the reasoning task is crucial. Evans (1984; see section 2.2.2.3.5) suggests people select information by means of heuristics such as perceptual salience or semantic associations. This can result in the selection of irrelevant information, and the omission of relevant information. Sternberg (1986) emphasised the importance of selective encoding of relevant versus irrelevant information, and suggested this might be particularly important in inductive reasoning since tasks often place minimal constraint on what constitutes relevant information.

The notion of a limited-capacity processing resource such as WM (e.g. Baddeley & Hitch, 1974) is central to various theoretical explanations of reasoning. Reasoning is generally assumed to be an effortful rather than an automatic process, and therefore performance should be influenced by the availability of processing resources. Models of reasoning have attempted to identify the mechanism by which WM capacity might limit reasoning performance. Explanations have been put forward in terms of the number of mental

models that have to be held simultaneously in WM (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991), the number of inferential steps required (Braine, 1978; Rips, 1989), and the need to selectively encode, selectively compare and selectively combine information from potentially unconstrained inputs (Sternberg, 1986) (see section 2.2.2.3). In each case, it is assumed errors may arise if WM capacity is exceeded. Evidence to support such assertions has been derived from manipulation of the number of models or inferential steps required to solve a problem (Johnson-Laird & Bara, 1984; Rips, 1989) and correlational analyses relating individual WM capacity to performance on reasoning tasks (Oakhill & Johnson-Laird unpublished, cited by Johnson-Laird, 1983).

As noted in section 2.2.5.2.2, the Baddeley and Hitch (1974) model of WM assumes three components: the CE and two slave systems (the PL and the VSSP). Several studies have used dual-task methodology to examine the role of these in reasoning performance. Hitch and Baddeley (1976) used very simple verbal reasoning problems consisting of sentence verification. They found six-item concurrent loads slowed verification speeds considerably, suggesting CE involvement, while a concurrent PL task had a minor impact. Gilhooly, Logie, Wetherick and Wynn (1993) studied performance on syllogistic-reasoning tasks, and found performance was disrupted by a concurrent CE task (random number generation), but not by tasks tapping the PL and VSSP. Furthermore, the concurrent syllogism task interfered with random number generation, and to some extent with the PL task, but not with the VSSP measure. The authors concluded the CE played a major role in syllogistic task performance, the PL had a lesser role, and the VSSP did not appear to be involved. Toms, Morris and Ward (1993) reported a similar pattern of findings on a set of conditional reasoning problems of the type "If p then q". The VSSP has been implicated in simple spatial reasoning (Farmer, Berman & Fletcher, 1986) and in drawing inferences from descriptions of spatial layout (Oakhill & Johnson-Laird, 1984).

As described in section 2.2.5.2.2, the PL and VSSP are involved in the storage and manipulation of verbal and visuospatial material respectively under the control of the CE. The role of the PL and VSSP in reasoning is likely to reflect these functions, and this is supported by evidence outlined above, with the VSSP implicated in tasks involving spatial material (Farmer et al., 1986; Oakhill & Johnson-Laird, 1984), and the PL in tasks requiring the storage of verbal material (Gilhooly et al., 1993). The evidence also suggests the CE plays a significant role in reasoning performance (Gilhooly et al., 1993;

Hitch & Baddeley, 1976). The precise role of the CE in reasoning is more difficult to delineate due to its complexity as a concept and has been widely debated. Engle, Cantor and Carullo (1992) suggest four possibilities. Individuals may vary in the efficiency with which they carry out processing operations (e.g. Baddeley, 1986); in the skill of task-specific operations (e.g. Daneman & Carpenter, 1980); in the intelligent use of strategies; or in storage capacity general to a variety of tasks (e.g. Turner & Engle, 1989).

Most of the functions attributed to the CE are likely to be important to reasoning performance. The CE is assumed to co-ordinate information from the PL and VSSP, both of which have been found to be involved in reasoning (see above). It has been postulated the CE is responsible for allocating attentional resources which, as has been discussed above, are likely to be a crucial to reasoning. The CE has also been postulated to be the seat of the executive processes outlined in section 2.2.3. These include the selection of strategies, self-regulation, planning and monitoring, and are most likely to be called upon when faced with a novel situation. By definition, reasoning involves the combination of existing elements to generate new information, and this suggests executive processes are likely to be important. However, it may be possible for the process to take place in a relatively automatic fashion, without conscious effort. For example, if someone is asked to define a particular word, if the word is known the response may be simple fact retrieval, but if the word is not known, the answer may require greater reasoning involvement, such as the generation of strategies (e.g. "thinking of words that sound similar"). Sternberg (1986) recognised the importance of novelty in his definition of reasoning. He argued that when a task requires either selective encoding, selective comparison or selective combination then it can be identified as a reasoning task, but only if the process is carried out in a controlled rather than an automatic fashion. Thus tasks may vary in their reasoning demands, depending on how many of the three defining processes are present, and individuals may vary in the degree to which they carry out a task in a controlled or automatic fashion. Therefore, a 'prototypical' reasoning task would be one that required all three processes, and was executed in a highly controlled fashion.

Two studies (Kyllonen & Christal, 1990; Tirre & Pena, 1993) have used factor analysis to explore the relationship between WM and reasoning ability. Kyllonen and Christal (1990) reported consistently high estimates of the correlation between WM capacity and reasoning ability factors, and concluded reasoning is little more than WM capacity. Tirre

and Pena (1993) found evidence that reasoning is a function of specific processing components as well as general WM capacity, and concluded Kyllonen and Christal (1990) had somewhat overestimated the role of WM in reasoning.

In his work on cognitive deficits in normal ageing, T. Salthouse (e.g. Salthouse, 1992; 1993) has explored the relationship between WM and reasoning on the basis that reduced WM resources in older adults might underlie their performance deficits on reasoning tasks. Salthouse (1993) reported that statistical control of an index of WM reduced the age-related variance in performance on Raven's Progressive Matrices Test, a measure of abstracting ability, by approximately 70%. Salthouse and his colleagues have reported similar findings on an integrative reasoning task (Salthouse, 1992; Salthouse, Mitchell, Skovronek & Babcock, 1989) described in more detail in chapter V.

It seems clear that attentional, WM and LTM processes are all involved in reasoning. In particular, the role of WM, as conceptualised by Baddeley (e.g. 1986; 1990), appears to be inextricably linked with reasoning, so that Kyllonen and Christal (1990) concluded reasoning is little more than WM capacity. However, this conclusion depends on the particular conceptualisation of WM adopted, and on the particular model of reasoning. In the past attentional, LTM and executive functions were conceptualised as separate from WM, suggesting reasoning ability was dependent on several separate systems. However, in recent years theorists from a range of disciplines in the fields of attention, WM and executive function have recognised the close links between the three (for review see Lyon & Krasnegor, 1996), and this is reflected in current models (e.g. Pennington, Bennetto, McAleer & Roberts, 1996) which assume WM has a central role, including both attentional and executive functions. Therefore, the close relationship between reasoning and WM suggested above is unlikely to be disputed by current theorists.

With regard to the relationship between reasoning and LTM, all the models considered in the current review clearly implicate LTM in reasoning. In the past, LTM has generally been conceptualised as a separate system from STM or WM (see section 2.2.5.2.1). More recently theorists have begun to conceptualise WM and LTM as different aspects of the same system. For example, Engle et al. (1992) postulate the contents of WM consists of information in LTM above a critical threshold of activation, and that WM resources are determined by limitations on the total amount of LTM activation, which varies between

individuals. If WM is conceptualised in this way the links between reasoning and WM are strengthened rather than weakened by the assumption of LTM involvement in reasoning performance.

In conclusion, in terms of current models of WM, which see it as closely linked with attentional, executive function and LTM processes, reasoning may well be little more than WM capacity. However, it should be noted reasoning is only one of many functions carried out by WM, and therefore the relationship between reasoning and WM is not reciprocal (reasoning may be little more than WM capacity, but WM capacity is much more than reasoning). Also, as conceptualisations of WM become more complex, it is increasingly difficult to identify single measures that capture the full 'flavour' of WM capacity. It may be that tasks designed to measure particular aspects of WM, such as measures of simultaneous storage and processing (e.g. Daneman & Carpenter, 1980), will continue to be useful in the future.

In summary, section 2.1 contained a review of 'thinking' in depression that ended with the conclusion that, with regard to emotionally-salient material at least, there was no evidence that depressed and nondepressed people differ in their capacity for logical thought, with both subject to the use of heuristics and biases depending on the particular circumstance. This contradicts Beck's original assertion that depression is associated with a thinking impairment, and the subjective reports of depressed patients that their thinking is impaired.

Section 2.2 began with a review of what has been learnt about reasoning within the fields of cognitive psychology and neuropsychology. The evidence from the field of cognitive psychology confirmed that in general people do make many errors on reasoning tasks, and this may be attributed to the use of heuristics, and the influence of biases. There was consensus that errors arise when the demands of the reasoning task exceed available WM resources. The evidence from the field of neuropsychology identified the frontal lobes as the area of the brain most closely linked to performance on reasoning tasks and converged with the evidence from the cognitive psychology models of reasoning in identifying the importance of WM in reasoning performance.

Current models of attentional and WM processes were reviewed in section 2.2.5, and their relationship to reasoning considered in section 2.2.6. It was concluded that in terms of

current models of WM, which see it as closely linked with attentional, executive function and LTM processes, reasoning may well be little more than WM capacity. Recent models of cognitive function in depression reviewed in section 2.1.3.3 suggest depressed Ss show a strong tendency to allocate processing resources to emotionally salient or personally relevant material. This could act to reduce available WM resources, and thereby result in impaired functioning on tasks involving neutral material. In order to address this, the depressive performance on tasks containing neutral material will be reviewed next.

2.3 COGNITIVE FUNCTION IN DEPRESSION

2.3.1 METHODOLOGICAL ISSUES

Studies carried out to investigate the relationship between depression and cognitive function have generally used one of four experimental designs. Each of these designs is associated with methodological problems or difficulties of interpretation, and these are considered in turn.

The comparison of a group of clinically depressed patients with a control group is the most common design. The main methodological problem is the selection of an appropriate control group. 'Normal' controls differ from depressed Ss in both their level of depression and their nonpatient status. An alternative is to use a group of psychiatric patients with diagnoses other than depression. This has the advantage of controlling for patient status, but the disadvantage that other psychiatric disorders may be associated with cognitive deficits, so that if no group difference is found then it may be because both groups are impaired. Finally, some studies have used medical patients as controls. This again has the advantage of controlling for patient status, but medical conditions associated with possible brain damage need to be excluded. Furthermore, it is sometimes difficult to separate out the effects of the medical condition and any mood disorder on measures of depression (Snaith, 1987).

It is essential to ensure the control group(s) do not differ from the experimental group on variables, other than depression, likely to influence task performance. Depressed and control Ss are commonly matched on age, sex, and educational level. Studies may also try to match Ss on a measure of intelligence, although because of the close link between intelligence and reasoning outlined in section 2.2.1, a measure of crystallised intelligence, such as vocabulary or general knowledge, may be most appropriate. However, the most

common and problematic difference between groups is the medication taken by the depressed Ss (and any psychiatric or medical control Ss). Anti-depressant medication has effects on cognitive function independent of the effects of depression (Deptula & Pomara, 1990), and these depend on the particular drug. If depressed Ss in an experimental sample are taking a variety of drugs then it may not be possible to partial out the effects of medication. This is a serious problem and one that has often been ignored.

The second study design utilises depressed patients as their own controls, comparing their performance during the depressive episode with performance on recovery. While this removes the problem of selecting an appropriate control group, it does have other flaws. First, since Ss are always tested first when they are depressed, there is the possibility that any improvement in performance upon recovery may be due to practice effects or to statistical regression toward the mean. The second problem is that depressed Ss may fail to make a full recovery, and therefore at re-test may present as a heterogeneous group ranging from fully recovered to profoundly depressed. The third problem is that Ss are likely to be on medication, necessitating the separation of the effects of medication from the effects of depression on cognitive function.

The third study design is the use of Ss, usually students, who have not sought treatment for depression, but who are selected on the basis of their score on a measure of depression such as the BDI. It has been recommended (Kendall, Hollon, Beck, Hammen & Ingram, 1987) that these Ss are labelled as dysphoric rather than depressed in recognition of the fact they have not been diagnosed as depressed (see section 3.2.1.1 for a more detailed discussion of this issue). The disadvantage of this method is that Ss have not been formally diagnosed as depressed, so they may have other conditions that elevate their score on the BDI. Furthermore, these Ss often have relatively mild symptomatology, and therefore may not exhibit deficits in cognitive function. The advantage of this type of methodology is that Ss are unlikely to be taking anti-depressant medication, and therefore the problem of confounding the effects of depression and medication are eliminated. Another advantage is that selecting the sample from a population of undergraduate students means the depressed and nondepressed Ss are likely to be similar on variables such as age, general intelligence and non-patient status. Studies using this methodology will be included in the current review, but will be separated from clinical samples, except where the overall number of studies in an area is particularly low.

A fourth design uses mood induction techniques to compare depressed, neutral and euphoric moods. Such studies are able to demonstrate clearly the causal direction of any effect of mood on performance. Furthermore, confounding variables, such as medication, that are problematic when patient samples are used, are eliminated. This approach has been used extensively in the investigation of depressive processing of emotional material as reviewed in section 2.1.3.3.6 above. However, questions have been raised regarding the validity of this approach. For example, Perrig and Perrig (1988) investigated the mood-congruity effect, which refers to facilitated processing of information when the affective valence of this information is congruent with Ss' mood (see section 2.1.3.3.6.1). They were able to replicate effects usually attributed to mood induction by simply asking Ss to behave as if they were depressed or happy. They suggested Ss used the mood instruction as a context cue constructed during learning and used again at the time of retrieval. Perrig and Perrig (1988) argued that mood may be a sufficient but not a necessary condition to produce the mood-congruity effect of selective learning. This highlights a lack of clarity regarding the relationship between induced and naturally occurring depressed mood.

One argument in favour of mood induction to study depression is that findings relating to induced mood processing of emotional material often parallel those using naturally occurring mood (e.g. Ellis & Ashbrook, 1989). However, with regard to processing of neutral material, two studies (Hertel & Rude, 1991a, Experiment 2; Kwiatkowski & Parkinson, 1994) have carried out direct, between-subject comparisons of induced and naturally occurring *dysphoric* mood, and found differences in the pattern of performance. The theoretical significance of the actual pattern of findings was unclear, but the important point is that induction procedures may provide poor models of natural moods, at least with regard to neutral tasks. There are no direct comparisons of induced mood and naturally occurring *depression*. However, since the relationship between induced and naturally occurring mood is, as yet, poorly understood, studies using mood induction methodology will not be included in the current review of depressive performance on neutral cognitive tasks.

2.3.2 SUBJECTIVE REPORTS OF COGNITIVE DEFICITS

Complaints of difficulty with memory and concentration by depressed patients are so common they are usually included as a diagnostic indicator in current classificatory

systems (e.g. DSM-III-R, 1987, APA; see Table 2.1). Studies have examined the nature of subjective complaints made by depressed Ss and their relationship to performance on objective measures of cognitive function.

2.3.2.1 Clinical Samples

Squire and Zouzounis (1988) found depressed and amnesic patients could be distinguished on a self-report measure of cognitive performance in everyday life. They found amnesic patients reported more problems overall, and the pattern of responding was different, with amnesic patients reporting difficulties with learning and recall of new material, while depressed Ss reported impaired attention and concentration. Watts and Sharrock (1985) used a structured interview that confirmed depressed patients frequently report concentration problems affecting their ability to perform a range of everyday tasks. Complaints of concentration problems whilst reading or watching television correlated significantly with the number of concentration lapses reported by Ss during ten minutes of silent reading, and with cued recall of a story. Unfortunately, neither of these two studies included a normal control group, which reduces their usefulness.

In a review of other studies comparing subjective complaints with actual memory performance, Watts (1993) concluded depression may be associated with memory complaints rather than deficits in objective memory performance. However, he noted many of these studies were carried out with elderly Ss, and the relationship between subjective and objective aspects of memory impairment may differ between young and old depressed Ss.

In contrast to the hypothesis that depression is associated with memory complaints rather than objective memory performance, Dalla Barba, Parlato, Iavarone and Boller (1995) found evidence that depressed Ss *underestimate* their memory difficulties. The authors labelled this "anosognosia", suggesting depressed Ss are unaware of their memory deficit. Dalla Barba et al. (1995) found that on a self-rating scale of memory function depressed Ss did not differ significantly from control Ss when rating their own memory ability, but depressed Ss did show significant impairment relative to controls on two objective memory measures. The authors calculated an Anosognosia Index by subtracting Ss' objective memory scores from their subjective memory rating, and found the depressed Ss had a significantly higher score than controls.

2.3.3 PSYCHOMOTOR SPEED

2.3.3.1 Clinical Samples

Studies of speech rate have found depressed patients speak slowly (Hinchcliffe, Lancashire & Roberts, 1971; Pope, Blass, Siegman & Raher, 1970). Later studies have used more accurate recording techniques, and these have revealed increased pause time (Szabadi, Bradshaw & Besson, 1976; Greden & Carroll, 1980) which normalises on recovery (Greden, Albala, Smokler, Gardner & Carroll, 1981).

The most common measure of psychomotor speed used with depressed samples is the Digit Symbol sub-test of the WAIS (Wechsler, 1955) or WAIS-R (Wechsler, 1981). Slow rates of performance on this task are associated with depression (Austin, Ross, Murray, O'Carroll, Ebmeier & Goodwin, 1992; Friedman, 1964; Hart, & Kwentus, 1987; Ilsley, Moffoot & O'Carroll, 1995; Rogers, Lees, Smith, Trimble & Stern, 1987; Sackeim, Freeman, McElhiney, Coleman, Prudic & Devanand, 1992), with some exceptions (e.g. Beatty, Wonderlich, Staton & Ternes, 1990). Depressive psychomotor retardation has also been found using measures other than Digit Symbol (e.g. Byrne, 1976a; Gunther & Kryspin-Exner, 1991; Friedman, 1964; Hall & Stride, 1954; Hemsi, Whitehead & Post, 1968; Martin & Rees, 1966; Weckowicz, Tam, Bay, Collier & Beelen, 1981; Weckowicz, Tam, Mason & Bay, 1978). However, Beck, Feshbach and Legg (1962) reported scores on Digit Symbol were not significantly related to measures of depressive symptomatology (psychiatrists' ratings of depth of depression; and scores on the BDI), but were related to severity of psychiatric illness regardless of nosological group.

The Sternberg short-term memory scanning procedure has been used to measure cognitive and psychomotor speed in depression. In this paradigm, several stages of information processing can be assessed independently by analysing the relation between task factors and their effect on reaction time. Thus, a linear relation holds between the time taken to compare a series of items with similar items in memory, and the number of items held in memory (as memory load). In representing this relationship, the intercept is thought to measure the rate of perception and output factors, while the slope of the reaction-time/set-size function is a measure of the memory scanning process. Most studies (Glass, Uhlenhuth, Hartel, Matuzas & Fischman, 1981; Hart & Kwentus, 1987; Koh & Wolpert, 1983) have found no association between depression and slow scanning performance. However, Brand and Jolles (1987) reported slower scanning in depressed Ss. The authors

attributed this to their use of newer versions of the task which were shorter. In contrast, there is agreement that depression is associated with impaired performance on the non-scanning aspects of the comparison task, that is, encoding the information, making a decision, and responding. Glass et al. (1981) concluded depressed Ss appeared to maintain accuracy at the expense of slowing during this part of the task.

Cornell, Suarez and Berent (1984) also attempted to differentiate between cognitive and motor components of behaviour that might underlie psychomotor slowing. They compared psychomotor and cognitive speed using three tasks: 1) a simple reaction time task; 2) a version involving an additional motor component; 3) a version involving an additional cognitive component. Comparing melancholic (see section 2.1.1.2) and non-melancholic depressed patients with normal controls, there was evidence of a motor component to slowing in both depressed groups, while only the melancholic Ss showed evidence of cognitive slowing. However, the melancholic Ss were rated as significantly more depressed on the Hamilton Depression Rating Scale than the non-melancholic patients, and the difference might therefore reflect symptom severity rather than sub-type.

In summary, the evidence supports the hypothesis that clinical depression is associated with psychomotor slowing, with the majority of studies finding evidence that depressed Ss both speak and move more slowly than nondepressed Ss. This finding has been confirmed using a range of measures. However, the functional basis of psychomotor slowing is more difficult to deduce from the available evidence. Cornell et al. (1984) found evidence of motor slowing, while cognitive slowing was confined to melancholic (or possibly more severely depressed) Ss. In contrast, studies using the Sternberg short-term memory scanning procedure have suggested depression is associated with slowing in encoding information, making a decision, and responding. Finally, it should be noted that the findings of Beck et al. (1962) suggest slowing may not be specific to depression, but may instead be an indicator of the severity of psychiatric disturbance.

2.3.3.2 Nonclinical Samples

Few studies have considered psychomotor speed in nonclinical samples. Berndt and Berndt (1980) reported dysphoric students were slower than controls on a digit symbol test. Clearly, more studies are needed before any conclusions can be reached regarding psychomotor slowing in nonclinical samples.

2.3.4 MEMORY AND LEARNING IN DEPRESSION

2.3.4.1 Learning

2.3.4.1.1 Clinical samples

Many studies have reported that depressed patients show impairment on measures of learning. Depressive deficits have been reported on word-learning tasks (Coughlan & Hollows, 1984; Friedman, 1964; Henry, Weingartner & Murphy, 1973; Sternberg & Jarvik, 1976), with the exception of Rohling and Scogin (1993). Depressed patients have also shown deficits on the Wechsler Memory Scale (WMS; Wechsler, 1945) (Breslow, Kocsis & Belkin, 1980; Danion, Willard-Schroeder, Zimmerman, Grange, Schlienger & Singer, 1991; Stromgren, 1977; Williams, Little, Scates & Blockman, 1987), and on recall of nonsense syllables (Cohen, Weingartner, Smallberg, Pickar & Murphy, 1982), and tests of visual retention (Elliott, Sahakian, McKay, Herrod, Robbins & Paykel, 1996; Friedman, 1964; Robertson & Taylor, 1985).

Several studies have compared learning of verbal and nonverbal material, with conflicting findings. Peselow, Corwin, Fieve, Rotroson and Cooper (1991) reported depressive deficits for recall of both words and pictures, while Steif, Sackeim, Portnoy, Decina and Malitz (1986) found depressive impairment in recognition of both words and faces. Richards and Ruff (1989), using a battery of verbal and nonverbal tasks, found depressed Ss were differentiated from controls by some of the nonverbal tasks, but by none of the verbal tasks. A similar pattern was reported by Boone, Lesser, Miller, Wohl, Berman et al. (1995), with depressed and control Ss differing significantly on a composite score of visual memory, but not on a composite score of verbal memory. Calev, Korin, Shapira, Kugelmass and Lerer (1986) used verbal and nonverbal tasks matched for difficulty, and found depressed Ss equally impaired on both relative to controls. Deptula, Manevitz and Yozawitz (1991) found evidence of differential depressive impairment on both free recall and recognition of nonverbal relative to verbal serial learning tasks. Although the tasks in the latter study were not matched in the same way as those of Calev et al. (1986), the authors reported the control Ss performed at a similar level on both. They suggested the inconsistency in the findings might result from their use of a serial learning task compared with the single trial task used by Calev et al. (1986), since the differential deficit was revealed only on later trials. An alternative explanation is that the tasks used by Deptula et al. (1991) were not sufficiently sensitive to detect differences in the control group.

2.3.4.2 Free Recall Versus Recognition Memory

2.3.4.2.1 Clinical samples

Several studies have included a within-subject comparison of performance on tasks measuring different types of retrieval process, with conflicting results. Some studies reported depressed Ss showed impairment on measures of free recall but not recognition memory when compared with normal controls (Beatty et al., 1990; Brand, Jolles & Gispen-de Wied, 1992, Experiment 1; Ilsley et al., 1995; Watts & Sharrock, 1987; Williams et al., 1987), suggesting free recall is a more sensitive measure of depressive deficits. Other studies found depressed Ss to be impaired on both recognition and free recall tasks (Austin et al., 1992; Backman & Forsell, 1994; Brown, Scott, Bench & Dolan, 1994; Silberman, Weingartner, Laraia, Byrnes & Post, 1983; Wolfe, Granholm, Butters, Saunders & Janowsky, 1987). One study failed to find depressive deficits on either recognition or recall tasks (Davis & Unruh, 1980); however, there is evidence the Ss were less severely depressed than Ss in other studies, and the control group consisted of nondepressed psychiatric outpatients rather than normal controls.

Inconsistency in the findings is likely to reflect differences between the studies in terms of sample characteristics, and may also reflect variations between the tasks on dimensions other than recognition versus free recall, such as differences in difficulty. This is supported by the study carried out by Brand et al. (1992, Experiment 2) who manipulated the difficulty of the Rey Auditory Verbal Learning Test by varying the number of learning trials (1, 3, or 5) before testing both free recall and recognition. They found depressed Ss showed impaired recall in all three conditions, but deficits in recognition memory emerged only in the most difficult (1-trial) condition. Calev and Erwin (1985) attempted to overcome this problem by using recognition and free recall tasks matched for difficulty. They found while depressed Ss were impaired on both tasks relative to normal controls, the degree of impairment was differentially greater on the free recall task. The findings suggest depression is usually associated with deficits on tests of free recall, and less consistently with deficits on recognition memory tasks.

There are at least two possible explanations for the apparent differential depressive deficit on free recall versus recognition memory tasks. First, theories of recognition memory suggest it may rely on two separable processes: at retrieval Ss may base their response either on conscious recollection of the item, or on feelings of familiarity that may occur

in the absence of conscious recollection (for review see Gardiner & Java, 1993). The former is associated with effortful processes, while the latter is thought to result from automatic processes (see section 2.2.5.1.2). This means if conscious recollection fails during recognition memory, Ss may still be able to achieve the correct response by relying on the 'familiarity' response. If automatic responses are spared in depression, then this could explain why free recall is more sensitive to depressive deficits than recognition memory. This explanation is supported by a study carried out by Hertel & Milan (1994) using the process dissociation paradigm developed by Jacoby (1991). In standard recognition test paradigms, correct responding may be based on either familiarity or recollection, or a combination of both, and it is not possible to separate the two. Jacoby and his colleagues (e.g. Jacoby, 1991) developed a way of dissociating the two processes by putting them in opposition to each other during recognition by means of different task instructions. Hertel and Milan (1994) found dysphoric students were impaired on recollection but not familiarity compared with normal controls when these two processes were measured in opposition, but there was no group difference when both familiarity and recollection were working in the same direction, as they do in a standard recognition task.

A second, related possibility is that free recall may benefit more than recognition memory from the use of strategies, such as attempts to make semantic links between words or the use of imagery techniques, during either encoding or recall. Hertel and her colleagues (e.g. Hertel, 1994; Hertel & Hardin, 1990; Hertel & Rude, 1991a,b) have postulated depression may be associated with reduced initiative, leading to failure to use strategies spontaneously or engage in elaborative thinking (see section 2.3.9.2). If true, then it could have a differential impact on free recall relative to recognition memory performance in depressed Ss.

2.3.4.3 Response Bias

It has been suggested that depressive deficits on memory tasks may be due to a conservative response style. That is, depressed Ss may have the information available to them but, because of an overly cautious response strategy, may be unwilling to report the information. A conservative response bias of this type would be expected to result in errors of omission rather than commission on free recall tasks, and reduced rates of hits and false alarms on recognition memory tasks.

2.3.4.3.1 Clinical samples

On tests of free recall, two studies (Henry et al., 1973; Whitehead, 1973) have shown depression to be associated with errors of omission. However, four studies (Brand et al., 1992; Manschreck, Maher, Rosenthal & Berner, 1991; Dalla Barba et al, 1995; Peselow et al., 1991) have reported depressed patients do not differ from normal controls in the number of intrusion errors made.

On tasks of recognition memory, reduced rates of both hits and false alarms would indicate a conservative response style, or, if signal detection analysis is used, depression would be expected to lead to an effect on Beta (an index of response bias) rather than d' (an index of sensitivity). Both Miller and Lewis (1977) and Dunbar and Lishman (1984) found depression was associated with lower levels of both hits and false alarms. Furthermore, when signal detection analyses were carried out, depressed Ss were found to differ from controls on Beta but not d', indicating depressed Ss showed a more cautious response criteria rather than a true memory deficit. The study reported by Larner (1977) has sometimes been quoted in support of a conservative response style in depression. He found depressed Ss did not differ from physically ill controls on d', and both groups showed an equally cautious response style. However, some of the 'physically ill' group had disorders associated with cognitive deficits (e.g. two cardiovascular attacks, one parkinsonism), and therefore did not constitute an adequate control group.

Other studies contradict the conservative response bias hypothesis. Silberman, Weingartner, Laraia, Byrnes and Post (1983) and Calev and Erwin (1985) both reported that depressed Ss achieved significantly fewer hits than controls, but did not differ on false alarms. Several studies (Channon, Baker & Robertson, 1993a; Brown et al, 1994; Deijen, Orlebeke & Rijsdijk, 1993; Watts, Morris & MacLeod, 1987) have carried out signal detection analyses and found depressed Ss to be impaired relative to controls on d', but not to differ significantly on Beta. Two studies (Backman & Forsell, 1994; Corwin, Peselow, Feenan, Rotrosen & Fieve, 1990) reported impairment on d', but also found depressed Ss to be significantly more conservative than controls on Beta. Another two studies found evidence of a more liberal response bias in depressed Ss (Deptula et al., 1991; Wolfe et al., 1987), although the study carried out by Wolfe et al. (1987) combined the results for hits and misses, making interpretation difficult. Overall the evidence suggests depressed Ss may be more likely than controls to show a conservative response

criteria, but they also show a reduction in memory sensitivity which cannot be explained in terms of a response bias.

2.3.4.3.2 Nonclinical samples

Two studies (Henriques, Glowacki & Davidison, 1994; Hertel & Milan, 1994), carried out with nonclinical samples, may shed light on the findings in relation to recognition memory test performance. Henriques et al. (1994) studied response bias in three pay-off conditions: neutral, punishment and reward. In both the reward and punishment conditions, earnings were maximised by a liberal response criterion. Dysphoric students were not impaired on d' in any of the experimental conditions; nor on Beta in the neutral or punishment conditions. However, dysphoric Ss were significantly more conservative than controls in the reward condition, and there was a nonsignificant tendency for them to show a more liberal bias in the punishment condition. This suggests dysphoric Ss were more sensitive to punishment than reward, with the converse true for the controls. Although in this study payment was used to create the different conditions, the findings may explain the conflicting findings described above. It is possible that procedural variations between the studies resulted in some appearing more threatening or encouraging than others, thereby potentially influencing Ss' response criteria in different ways.

The second study was carried out by Hertel and Milan (1994) using a process dissociation paradigm to test recognition memory. The study is described in detail in section 2.3.4.2.1 above. In essence, Hertel and Milan (1994) found that when the processes of remembering by recollection versus familiarity were put into opposition during retrieval on a recognition task, dysphoric Ss were impaired on recollection but not familiarity compared with normal controls. There was no group difference when both familiarity and recollection were working in the same direction (as they do in a standard recognition task). This finding suggests variations between studies in the degree to which both depressed and control Ss rely on the two processes (recollection and familiarity) during task performance could be a factor in the variable findings outlined above.

2.3.5 EFFORTFUL ENCODING

As described in section 2.2.5.1.2, it is assumed tasks vary in their demands on cognitive resources, and that performance reflects the resources invested in the task. Based on this premise, Tyler, Hertel, McCallum and Ellis (1979) explored the effect of different

encoding processes on performance. They concluded increased effort at encoding leads to higher recall. Studies have compared the performance of depressed and nondepressed Ss on tasks which vary effort during encoding, and also during retrieval.

2.3.5.1 Clinical Samples

Structuring material during encoding or retrieval is an effortful strategy likely to improve recall. Studies have found control Ss are more likely than depressed Ss to recall words in clustered form even when they are presented in unclustered form (Backman & Forsell, 1994; Koh, Kayton & Berry, 1973, Experiment 1; Russell & Beekhuis, 1976; Calev & Erwin, 1985; Weingartner, Cohen, Murphy, Martello & Gerdt, 1981, Experiment 3), with two exceptions (Calev et al., 1986; Silberman, Weingartner, Targum & Byrnes, 1985). Watts and Cooper (1989) assessed recall of a prose passage and included an internal comparison of story-units that differed in how central they were to the gist. Depressed Ss failed to show the usual superior recall of the central aspects, suggesting a failure to use structure to organise the stories.

Other studies have considered the relationship between depression and performance on tasks where the degree of organisation or structure has been manipulated by the experimenter, and these have produced conflicting results. Two studies used approximation to text as a means of manipulating task structure (Levy & Maxwell, 1968; Manschreck et al., 1991). Ss are presented with a series of word-lists that at one extreme form a normal sentence, and at the other consist of random words. An example of intermediate approximation to text would be, "They saw the play Saturday and sat down beside him." Both studies compared depressed, schizophrenic and control Ss on free recall. Levy and Maxwell (1968) found all three groups had equivalent recall of less structured material, but depressed and schizophrenic Ss showed less benefit from increasing approximation to text. However, since the word-lists were presented to all Ss in order of increasing level of contextual constraint, this finding may be explained by fatigue effects. In contrast, Manschreck et al. (1991) found no significant differences between depressed and control subjects on any measure, in fact the depressed group performed marginally better, and this could not be explained in terms of group differences in age, sex or education.

Weingartner et al. (1981, Experiment 3) found depressed Ss were impaired relative to

controls in their recall of a word-list when there was no structure to the list, or when the structure was not immediately apparent, but achieved similar levels of recall when the word-list was highly structured. Tancer, Brown, Evans, Ekstrom, Haggerty et al. (1990) compared depressed patients with a group of psychiatric patients of mixed diagnosis and found the depressed Ss to be significantly impaired in their recall of a list of random words, but no different when the list consisted of categorised words clustered together. However, since Ss always received the clustered list first this finding, like that of Levy and Maxwell (1968), may be due to the effects of fatigue.

Watts, Dalgleish, Bourke and Healy (1990) studied both approximation to text and clustering by semantic categories, and found while depressed Ss showed an overall decrement in performance, this did not interact with structure. Exploratory post-hoc analyses suggested the groups differed more on medium than low levels of structure. Watts et al. (1990) suggested materials of medium structure might be more sensitive to memory deficits in depression than very unstructured materials because the latter do not repay efforts to restructure them. Support for this hypothesis was found by Channon et al. (1993a). Depressed and nondepressed Ss were compared on their retrieval of high, medium and low structured material, using word-lists consisting of uncategorised words and categorised words presented in randomised and clustered order. The depressed Ss were found to be impaired only on the medium level of structure (randomised categories list). Backman and Forsell (1994) compared performance on word-lists equivalent to the low and medium structured lists described by Channon et al. (1993a). They found depressed Ss were equally impaired on both lists, showing no benefit from the structure available in the medium list, while the controls showed significant improvement on the medium relative to the low structured condition. The authors do not report whether the group difference was significantly greater in the medium structure condition, so it is not clear whether this was actually more sensitive to depression. Finally, Brown et al. (1994) compared depressed and normal control Ss on three word-lists of similar structure to those of Channon et al. (1993a) and found depressed Ss to be impaired overall compared with controls. However, both groups showed the same pattern of performance, with recall highest on the medium structured list and lowest on the list of random words.

It has been suggested depressed Ss may impose organisation during encoding, but may not take advantage of this at retrieval. Russell and Beekhuis (1976) found depressed,

schizophrenic and normal control Ss did not differ in their initial categorisation of a list of nouns, but the recall and clustering of both patient groups was inferior to that of the controls at retrieval. Weingartner et al. (1981, Experiment 2) reported depressed patients imposed more organisation than controls when instructed to sort random words into categories, but then recalled fewer, although they were equivalent to controls in their recall of a list of related words.

Weingartner and his colleagues also found evidence of depressive deficits in recalling semantically processed material Weingartner et al. (1981, Experiment 1) reported when Ss were instructed to produce either a semantically related or an acoustically related response to each of a series of stimulus words, depressed Ss were impaired relative to controls in their recall of semantically but not acoustically processed words. Roy-Byrne, Weingartner, Bierer, Thompson and Post (1986, Procedure 1) found depressed Ss to be impaired relative to nondepressed Ss in their ability to recall noun-pairs after making comparative judgements about them. Silberman, Weingartner, Laraia, Byrnes and Post (1983) instructed Ss to rate the emotional impact of a list of words counterbalanced with respect to high versus low emotionality and concreteness. While there was no group difference in word-ratings, depressed Ss were impaired in their retrieval of the material on both recognition and free recall tasks. Group differences were greater for free recall of high concrete-emotional words, and for recognition of low concrete-emotional words. Thus, the depressed Ss benefitted from salient stimulus qualities less than normals in free recall, and more than normals in recognition. The authors concluded low impact material (low concrete emotional) was not stored with strong enough traces to be retrieved, even in the recognition condition, while high impact (high concrete-emotional) was processed deeply enough for recognition, but not free recall.

This review indicates a proliferation of studies, resulting in complex and often contradictory findings. There is evidence depressed Ss differ from nondepressed Ss in performance on this type of task, with most studies finding significant group differences. The exact nature of these differences is less clear-cut. The hypothesis that depressed Ss make less use of encoding and retrieval strategies such as structuring, organisation and clustering of material is supported, although whether these deficits occur at encoding, retrieval or both remains to be fully investigated. The relationship between depressive deficits and degree of structuring has proved difficult to identify, with authors arguing

materials with high (Levy & Maxwell, 1968), medium (Channon et al., 1993a; Watts et al., 1990) and low (Weingartner et al., 1981, Experiment 3) structure are most sensitive to depressive deficits. It is possible that for each study, the material with the highest pay-off in terms of improved recall resulting from effort expended in encoding or retrieval strategies will be most sensitive to depressive deficits. This needs further investigation.

2.3.5.2 Nonclinical Samples

Few studies have investigated the use of encoding and retrieval strategies in nonclinical samples. The findings have been mixed, with two studies (Hasher & Zacks, 1979, Experiment 4; Potts, Camp & Coyne, 1989) finding evidence of group differences, and two studies (Hertel & Rude, 1991a, Experiment 3; Kwiatkowski & Parkinson, 1994) finding no group differences. Hasher and Zacks (1979, Experiment 4) compared dysphoric and nondysphoric students, selected on the basis of the BDI, on their ability to recall a list of words. Ss were presented with a recognition test consisting of the studied words, semantic and acoustic associates of the studied words, and unrelated words. Dysphoric Ss chose fewer incorrect semantic associates than the controls, suggesting less semantic processing of the studied words. Potts et al. (1989) compared dysphoric and control students on their recall of target words from simple and elaborated sentences. They found control Ss recalled more elaborate sentences than simple, while dysphoric Ss recalled equal numbers of both, but there was no overall difference in level of recall. Hertel and Rude (1991a, Experiment 3) presented dysphoric and nondysphoric students with a series of incomplete sentences, and asked them to judge whether a word presented simultaneously would fit sensibly into the sentence. Half were judged to be easy and half difficult. It was predicted dysphoric Ss should be differentially impaired in their recall of the more difficult items, but in fact no group differences were found on any of the task measures. Kwiatkowski and Parkinson (1994) found no differences between dysphoric and control students in their recall of target words from simple and elaborated sentences.

2.3.6 AUTOMATIC PROCESSING

2.3.6.1 Frequency Encoding and Spatial Location Encoding

2.3.6.1.1 Clinical and nonclinical samples

There is evidence that event frequency and information about the spatial location of an object are both encoded automatically (Hasher & Zacks, 1979). Several studies have reported depressed Ss are equivalent to normal controls in their ability to judge frequency.

For example, Golinkoff and Sweeney (1989) showed Ss a list of 30 words that each appeared from one to seven times. Immediately after presentation, Ss were asked to make relative frequency judgements for 15 word-pairs taken from the list. Depressed and personality-disordered patients performed as well as normal control Ss on this task. Similar findings have been reported by Roy-Byrne et al. (1986, Procedure 2), and by Hasher and Zacks (1979, Experiment 3) who compared depressed and nondepressed community adults on the basis of a median split of their scores on the BDI. Some of these studies found the same depressed Ss to be impaired on an effortful task (Golinkoff & Sweeney, 1989; Roy-Byrne et al., 1986) which increases the strength of these findings. Rohling and Scogin (1993) compared young and old groups of depressed patients, mixed psychiatric controls and normal controls on measures of memory for frequency and spatial location, and found no significant effects of mood. On two effortful memory tasks there were significant effects of age but not mood, weakening this finding.

2.3.6.2 Priming

A second type of task used to measure automatic processing is the priming task described in section 2.1.3.3.6.2. This paradigm permits investigation of the effects of prior experience in circumstances where Ss are not required to refer to the prior experience, and may be unable to recall it. A typical task requires Ss to carry out word-stem completion (e.g. don___) with the first idea that comes to mind; performance is facilitated by previous study of relevant items (e.g. presentation of the word 'donkey' increases the likelihood of stem completion with this word).

2.3.6.2.1 Clinical samples

Danion et al. (1991) reported depressed patients performed a word-stem completion task at normal levels, but were impaired on a measure of explicit memory. Similar findings were reported by Denny and Hunt, (1992) using a word-fragment completion test. Watkins et al. (1992) and Danion, Kauffmann-Muller, Grange, Zimmermann and Greth (1995) found depressed Ss unimpaired on both implicit and explicit measures, although Watkins et al. (1992) did report a trend toward depressive impairment on the explicit task.

Bazin, Perruchet, de Bonis and Feline (1994) noted the tasks used by Danion et al. (1991) and Denny and Hunt (1992) confound the distinction between implicit and explicit memory tasks with the provision of retrieval cues, since providing a word-fragment

essentially turns the implicit test into a cued-recall measure. Bazin et al. (1994) carried out a comparison of implicit and explicit memory tasks that differed only in their instructions, and found a dissociation, with depressed Ss impaired relative to controls only on the explicit task. Furthermore, when Ss were retested 4 weeks later, when the depressed group were clinically improved, the groups did not differ on either measure. Elliott and Greene (1992) reported depressed patients were impaired on both implicit and explicit memory measures, but Bazin et al. (1994) noted their implicit memory task procedure increased the likelihood Ss would use explicit retrieval processes.

Hertel (1994) used a word identification task in which depressed and nondepressed Ss tried to identify previously rated and unrated words that were presented briefly and masked. Hertel reported a depressive deficit in word identification when the rating task required Ss to refer to physical characteristics of the word, but not when the rating task required Ss to make a semantic judgement. Hertel interpreted these findings as suggesting depressed Ss were less likely to read the target words spontaneously (when rating physical characteristics), but when required to read the words, implicit memory was intact. The same depressed Ss were impaired relative to controls on free recall of the same words.

With regard to the studies reviewed above, the possibility cannot be excluded that explicit recall of material may have contributed to performance on the implicit memory task. Bradley et al. (1995) carried out a stringent test of implicit memory processes by introducing a subthreshold priming condition to a lexical decision task which compared neutral and emotional words. If material is presented subthreshold then any priming effect must be due to automatic processes. They found control Ss showed significantly greater priming effects for neutral words than depressed Ss in both the suprathreshold and subthreshold conditions, with the converse pattern for depression-relevant words.

2.3.6.2.2 Nonclinical samples

Several studies using nonclinical samples have found evidence of intact implicit memory with deficits in explicit memory performance in dysphoric Ss. This pattern of findings was reported by Ruiz-Caballero and Gonzalez (1994, Experiment 1) using a word-stem completion task and Hertel and Hardin (1990, Experiment 3) using a homophone spelling task. Both studies compared dysphoric and nondysphoric students selected on the basis of the BDI. Watkins, Vache, Verney, Muller and Mathews (1996) compared students

diagnosed as depressed with nondepressed controls on a task designed to investigate conceptual rather than perceptual repetition priming. Following the presentation of the target words, Ss were required to free-associate to a series of words that were semantically related to the targets. Implicit memory was evidenced by the production of the target words during free association. Again, the dysphoric Ss were not impaired on the implict memory task.

Of the clinical studies reviewed in section 2.3.6.2.1, only Bradley et al. (1995) fully excluded the possibility that explicit recall of material contributed to performance on the implicit memory tasks. Two studies using nonclinical samples have attempted to address this issue. Bradley et al. (1994) carried out a lexical decision task similar to that of Bradley et al. (1995) (see section 2.3.6.2.1). While there was no difference between dysphoric and nondysphoric students on a suprathreshold priming task, there was a difference when the primes were presented subthreshold. The control group showed significant effects of subthreshold priming on neutral and anxious words, but not on depressed or positive words, while the dysphoric Ss showed a converse pattern. Hertel and Milan (1994), as described in section 2.3.4.2.1, used a process dissociation paradigm to separate automatic and effortful processes in recognition memory, and found dysphoric students to be unimpaired on recognition based on the automatic process of familiarity, but to be impaired on recognition based on effortful recollection of items.

In summary, the weight of evidence suggests that provided depressed Ss do process the target words (Hertel, 1994), the priming effect is intact in depressed Ss. However, the results of the studies carried out by Elliott and Greene (1992) and Bradley et al. (1994; 1995) bring a cautionary note to this conclusion. Bazin et al. (1994) noted some problems with the methodology of Elliott and Greene's (1992) study, but it is more difficult to ignore the findings of the two studies carried out by Bradley et al. (1994; 1995). The use of subthreshold presentation of target words is a powerful approach in terms of excluding the possibility of explicit processing of material. It is clearly important to replicate the procedures used studies that found no evidence of a depressive deficit in implicit memory using subthreshold presentation of the target words. Until this work is done, it is not possible to reach an unequivocal conclusion about priming in depression.

2.3.7 ATTENTION AND WORKING MEMORY

2.3.7.1 Sustained Attention

2.3.7.1.1 Clinical samples

As noted in section 2.2.5.1.1, a common distinction with regard to attention is between the ability to sustain attention over a period of time versus the ability to selectively attend to some inputs rather than others. Vigilance refers to the ability to sustain attention over a period of time whilst trying to detect particular events, for example, monitoring a radar screen. Two studies have tested vigilance in depression. Byrne (1976b) demonstrated deficits in depressed patients on a vigilance task. She reported different patterns of deficits for psychotic and neurotic depressed groups. Byrne had hypothesised that psychotic depressed Ss would be under-aroused while neurotic depressed Ss would be over-aroused. Consistent with these predictions, the psychotic depressed Ss had a low level of correct detections and showed marked deterioration over time. In contrast, the neurotic depressed Ss were less impaired in their vigilance performance, and their errors were mostly false positives. However, caution needs to be exercised because the results have not been replicated (Byrne, 1977). Frith, Stevens, Johnstone, Deakin, Lawler & Crow (1983) reported depressed Ss seemed impaired on a vigilance task, which required Ss to respond to a signal over a 15 minute period, when compared with task norms. Therefore, both studies found evidence that depression may be associated with an impairment in this aspect of attention, although more work is clearly needed in this area.

2.3.7.2 Selective Attention

2.3.7.2.1 Clinical samples

While a range of paradigms are available to measure selective attention, only a few have been investigated in relation to depression. The paradigms that have received most attention are those relating to 'filtering' relevant from irrelevant information (Frame & Oltmanns, 1982; Hemsley & Zawada, 1976; Knott, Lapierre, Griffiths, de Lugt & Bakish, 1991; Pogue-Geile & Oltmanns, 1980). Knott et al. (1991) used a dichotic listening task that required Ss to respond to a target tone when it was received by the designated ear. In addition, a physiological measure of attention, event-related potentials (ERPs), was also recorded. ERPs are measures of small but reliable signals produced by the brain in response to environmental events. Depressed patients were able to attend selectively and detect target signals as efficiently as normal Ss, and they had appropriate ERPs.

Pogue-Geile and Oltmanns (1980) compared depressed, manic, schizophrenic and normal controls on their ability to shadow short stories presented to a single ear in a dichotic listening paradigm, with a competing message presented to the other ear on some trials. Recall of the stories was also tested. The shadowing performance of all three patient groups was equivalent to the controls', and was not affected by the presence of a competing message. Only the schizophrenic group differed from the controls in ability to recall the material.

Frame and Oltmanns (1982) presented depressed, schizophrenic and control Ss with lists of items to be recalled in two experimental conditions. In the neutral condition all the items were read by a female voice, but in the 'filter' condition items alternated between female and male voices, and Ss were instructed beforehand to recall only those items read by the female voice. Both patient groups were tested on two occasions: soon after hospitalisation, and again when they were much improved and/or ready for discharge. The controls were not re-tested. The depressed Ss did not differ from the controls on any aspect of their task performance, while the schizophrenic Ss showed some impairment in their overall level of recall. The depressed Ss showed no significant change in performance from the first to the second occasion of testing, while the schizophrenic Ss did show improved recall performance. Neither patient group was found to be more distractible than the controls on either occasion.

In the study carried out by Hemsley and Zawada (1976) all the lists had the same format as the 'filter' condition of Frame and Oltmann's (1982) study, with the items read alternately by male and female voices. Ss had to recall one set of items (those read by either the male or female voice),. and were informed which items to recall either before or after the items were presented. The performance of the control group was improved by telling Ss *before* rather than *after*, but the two patient groups did not show this effect. Both patient groups had impaired recall relative to the control group in both conditions.

There is clearly a discrepancy in the findings in this area, with the studies carried out by Knott et al. (1991), Pogue-Geile and Oltmanns (1980), and Frame and Oltmanns (1982) finding no evidence of depressive impairment on tasks which, in the latter two studies, were sensitive to deficits in schizophrenic Ss. Hemsley and Zawada (1976) found depressive deficits on every aspect, while the other three studies found no evidence of

depressive deficits, even on measures of recall usually sensitive to depression (see section 2.3.4.2), suggesting the discrepancy may be due to a factor associated with S selection rather than the differential sensitivity of the tasks used. All four studies were carried out with clinically depressed patients. Knott et. al. (1991), Pogue-Geile and Frame (1980), and Frame and Oltmanns (1982) all used standardised diagnostic criteria to select Ss, while Hemsley and Zawada (1976) recruited Ss on the basis of diagnoses made by clinicians. This does not suggest Ss in the Hemsley and Zawada (1976) study were likely to be more severely depressed than Ss in the other three studies. Furthermore, with the exception of Knott et al. (1991), all the studies controlled for age and intelligence. Therefore, on the basis of the information available, it is difficult to explain the discrepancy noted above. On balance, the failure of Knott et. al. (1991), Frame and Oltmanns (1982) and Pogue-Geile and Oltmanns (1980) to find any evidence of depressive deficits in 'filtering' information must be taken to outweigh the evidence presented by Hemsley and Zawada (1976) of a depressive deficit, although the paradigm reported in the latter study is clearly worthy of further exploration.

Two recent studies (Lemelin, Baruch, Vincent, Laplante, Everett & Vincent, 1996; Trichard, Martinot, Alagille, Masure, Hardy, Ginestet & Feline, 1995) have examined the performance of clinically depressed Ss on the Stroop Colour-Word Test (SCWT; Stroop, 1935) measure of selective attention. The basic SCWT paradigm requires the S to name the colour of a word, which is actually the name of an incongruent colour e.g. the word RED printed in green ink. In order to give the correct response the S must selectively attend to the colour of the word, and inhibit the lexical response. Lemelin et al. (1996) also compared Ss on a modified SCWT task designed to be less demanding. Both studies (Lemelin et al., 1996; Trichard et al., 1995) reported depressed Ss showed significantly greater interference on the standard SCWT task, and Lemelin et al. (1996) also reported a depressive deficit on the modified task. Trichard et al. (1995) found depressed Ss remained impaired when they were reassessed on recovery. These two studies therefore provide strong evidence of a distractor inhibition disturbance in clinical depression

2.3.7.3 Secondary Tasks

As noted in section 2.2.5.1.1, models of attention usually incorporate the concept of limited resources, although there is wide debate about the exact nature of the resources and the nature of any limiting mechanism. The secondary task or dual task paradigm has

been used to investigate this. Essentially, this approach examines the effect on performance of a primary task of performing a second task simultaneously.

2.3.7.3.1 Clinical samples

The use of secondary tasks has been reported by several studies using clinically depressed Ss. Foulds (1952) showed a simple secondary task (repeating digits after the experimenter at approximately every 2 seconds) *increased* the speed of maze performance in depressives, anxiety states and obsessionals, but not in hysterics or psychopaths. This has been labelled the 'distraction effect'. In a further experiment, Foulds (1952) reported that following treatment with electroconvulsive therapy (ECT) the effect was reduced in depressed patients, but since there was no control group, practice effects cannot be excluded. Shapiro, Campbell, Harris and Dewsbery (1958) attempted to replicate Fould's (1952) findings as part of a larger study. Ss were depressed patients who had been prescribed a course of ECT treatment. All Ss were tested on two occasions, and half underwent ECT in the intervening time. The 'distraction effect' observed by Foulds was replicated in both groups with faster performance when performing the secondary counting task. However, the treatment group failed to show a reduction in the 'distraction effect' after ECT as predicted by Foulds (1952).

Blackburn (1975) reported a more complex pattern of findings. She compared current and recovered bipolar depressed, unipolar depressed and bipolar manic patients, but had no normal control group. Ss carried out a maze task in three conditions: without a secondary task; counting upwards at a rate of one digit every 2 seconds; and while a pre-recorded news item was played. This latter condition is rather difficult to interpret since Ss were given no specific instructions, thereby allowing the possibility that some may have chosen to ignore it whilst others made it the focus of their attention. Therefore, the results of this condition will not be considered further. The counting condition was associated with an increase in speed in all three of the 'current' groups, but there were no significant differences either between the different diagnostic groups, or between 'current' and 'recovered' groups of the same diagnosis. Thus, the results are consistent with a 'distraction effect'. However, it should be noted that the 'current' Ss were unmedicated, while the 'recovered' Ss were receiving maintenance doses of medication, making interpretation of these results more difficult.

When Foulds (1952) first reported the 'distraction effect' he attributed it to the disruption of, or distraction from, depressive preoccupations, thereby freeing processing capacity for the primary task. Support for this position is provided by Fennell, Teasdale, Jones and Damle (1987). They reported that asking clinically depressed patients to concentrate on a series of pictures significantly reduced the frequency of depressing thoughts compared with asking Ss to simply look at a white light for the same length of time, suggesting that a requirement to carry out a capacity-demanding task may reduce the frequency of negative thoughts. However, as noted by Williams et al. (1988, pp.36-37) in their review, this does not explain why the secondary task does not disrupt performance on the primary task to the same degree as the depressive ruminations it replaces. Williams et al. (1988) offer an alternative explanation. They note there is evidence that while a secondary task may increase the speed of performance on the primary task, accuracy may be reduced. This would suggest a change in performance strategy rather than a simple improvement in performance. Williams et al. (1988) suggest depressives may normally adopt a more conservative strategy than controls, maintaining accuracy at the expense of speed. This would give them more scope to increase the speed of performance, with only marginal effects on accuracy, while the controls might be more likely to adopt a speed close to their maximum before the addition of a secondary task.

Krames and McDonald (1985) reported an interesting study using a secondary task paradigm with depressed Ss. They used a group of patients who had all been diagnosed as depressed in the past, and divided them into currently depressed and nondepressed groups on the basis of the BDI. Ss were tested on their ability to recall word-lists under conditions of varying simultaneous memory load (recalling lists of digits). The depressed Ss did not differ from controls in word-list recall with a high simultaneous memory load, but were impaired at easier levels of secondary task performance. The authors attributed the results to a 'distraction effect', arguing that as the demands of the secondary task increase, it is the depressive thoughts that are displaced. This study is not open to the criticisms levelled by Williams et al. (1988) regarding a possible speed/accuracy trade-off. The depressed Ss made more errors overall on the secondary task, but this did not interact with difficulty, and so does not suggest a change in strategy as postulated by Williams et al. (1988). The results of this study taken alongside those of Foulds (1952), Shapiro et al. (1958) and Blackburn (1975) provides reasonably strong evidence that the cognitive performance of normal Ss given a secondary task is like that of depressed Ss without one.

2.3.7.4 Working Memory - The Phonological Loop and Visuospatial Sketchpad

With the exception of Channon, Baker and Robertson (1993b), the two slave systems of the WM model described in section 2.2.5.2.2 have received little direct attention within the depression literature. However, a number of studies have used tasks that are generally accepted as measures of PL and VSSP function and these are reviewed below.

2.3.7.4.1 Clinical samples

Channon et al. (1993b) administered tasks taken directly from the cognitive psychology WM literature and designed to investigate the length and phonological store of the PL. Depressed patients did not differ from normal controls on these measures. Forward digit span from the WAIS or WAIS-R (Wechsler, 1955; 1981) is generally accepted as a measure of the PL, and a number of studies have assessed depressive performance on this task. The majority have not found depressive impairment on this measure (Austin et al, 1992; Backman & Forsell, 1994; Channon et al., 1993b; Ilsley et al., 1995; Kopelman, 1986; Peselow et al., 1991; Richards & Ruff, 1989; Warren & Groome, 1984; Whitehead, 1973; Williams et al., 1987), while a few have found depressive impairment (Breslow et al., 1980; Gruzelier, Seymour, Wilson, Jolley & Hirsch, 1988).

With regard to the VSSP, Channon et al. (1993b) reported depressed Ss showed normal recall of forward spatial sequences using the Block Span Task described by Milner (1971), as did Beats, Sahakian and Levy (1996). Richards and Ruff (1989) also reported depressed Ss were unimpaired on the standard version of this task, but found depressed Ss to have a significantly shorter span than controls in a delayed recall condition. However, Elliott et al. (1996) found depressed Ss to be impaired on this task.

2.3.7.4.2 Nonclinical samples

Gass and Russell (1986) compared dysphoric and nondysphoric medical patients selected on the basis of the Minnesota Multiphasic Personality Inventory (MMPI), and with or without an organic brain disorder, on the digit span test. They found organic brain disorder, but not depression, impaired performance. Colby and Gotlib (1988) presented Ss with lists of 6, 8, 10 or 12 digits, and tested recall either immediately or after a delay. They found dysphoric students were impaired relative to nondysphoric students only on the 8 and 10 length strings. The authors attributed this to ceiling and floor effects on the 6 and 12 length strings.

Taking into consideration studies of both clinical and nonclinical Ss, the evidence suggests the function of the two WM slave systems is usually unimpaired by depression. However, deficits have been reported by some studies, and it is important to consider why this discrepancy has arisen. Two studies (Colby & Gotlib, 1988; Richards & Ruff, 1989) found depressive deficits when the standard task procedures were altered, making the task more difficult, and potentially making demands on other memory systems. Breslow et al. (1980), Elliott et al. (1996) and Gruzelier et al. (1988) reported depressive deficits using the standard task procedure, and differences in S selection may offer an explanation for the discrepancy. Both Breslow et al. (1980) and Gruzelier et al. (1988) used hospitalised patients, who were tested soon after admission, suggesting Ss were severely and acutely depressed at the time of testing. Elliott et al. (1996) tested a mixed group of inpatients and outpatients. Of the studies reporting no depressive deficits, Warren and Groome (1984) and Whitehead (1973) studied newly hospitalised patients, Channon et al. (1993) and Williams et al. (1987) tested mixed samples of inpatients and outpatients, Backman and Forsell (1994), Peselow et al. (1991) and Richards and Ruff (1989) studied outpatients. Therefore, with the exception of Warren and Groome (1984) and Whitehead (1973), the Ss in the remaining studies are likely to have been less severely depressed than those of Breslow et al. (1980) and Gruzelier et al. (1988), but not necessarily less severely depressed than those of Elliott et al. (1996). It is possible that depressive deficits are related to symptom severity, but more work to study this issue directly is needed.

2.3.7.5 Working Memory - The Central Executive

Measures of CE function come from one of two backgrounds: clinical tasks known to be sensitive to frontal lobe damage; and measures of CE developed by cognitive psychologists which require Ss to store and process information simultaneously, such as the WM span task (Daneman & Carpenter, 1980). Few studies have used tasks from the cognitive psychology literature to investigate CE function in depression, while rather more work has been done with tasks sensitive to frontal lobe deficits.

2.3.7.5.1 Clinical samples

Gunther and Kryspin-Exner (1991) compared depressed, dysthymic and control Ss on a random number generation task taken from the cognitive psychology literature. This task is thought to draw upon the resources of the CE component of WM (Baddeley, 1986). They reported the two patient groups had a higher perseveration tendency than controls.

With regard to tasks sensitive to frontal lobe deficits, Channon et al. (1993b) administered a battery of tasks, including backward digit and spatial span, a paced and unpaced serial addition test (PASAT), a trail-making test, and a letter cancellation task. They reported the depressed Ss were unimpaired on the majority of tasks, with evidence of depressive deficits only on backward digit span and the PASAT. Other studies have measured backward digit span in depressed Ss, with most finding them unimpaired relative to normal controls (Austin et al., 1992; Backman & Forsell, 1994; Ilsley et al., 1995; Peselow et al., 1991; Richards & Ruff, 1989), with one exception (Breslow et al., 1980).

Rush, Weissenburger, Vinson and Giles (1983) administered a battery of tasks likely to involve CE function, and found depressed Ss were not impaired relative to test norms. Austin et al. (1992) reported endogenous, but not nonendogenous, depressed Ss were significantly slower than controls on a trail-making test. Beats et al. (1996) found some evidence of depressive deficits on a self-ordered pointing task of the type described in section 2.2.3.1. The authors concluded the pattern of depressive deficit differed from that typically shown by Ss with frontal lobe damage. Elliott et al. (1996) reported depressive deficits on a spatial self-ordered searching task that required Ss to search through an array of coloured boxes for 'tokens'. The depressed Ss were significantly less likely to use an efficient strategy to guide their searches. Boone et al. (1995) reported depressed Ss were impaired relative to controls on a summary measure of executive function which comprised four tests sensitive to frontal lobe function: Stroop; Auditory Consonant Trigrams; verbal fluency (see below); and WCST (see section 2.3.8.2.2.1). Similarly, Dalla Barba et al. (1995) found depressed Ss to be impaired relative to controls on measures of frontal function: Graphic Sequences; Cognitive Estimates (see section 2.2.3.1); verbal fluency (see below); and a modified card sorting test (see section 2.3.8.2.2.1).

Verbal fluency is a sensitive measure of frontal lobe deficits (see section 2.2.3.1). It is measured by asking Ss to list as many words as possible which begin with a certain letter. Three studies have found depressed Ss to be unimpaired on this measure when compared with normal controls (Austin et al., 1992; Johnson & Crockett, 1982; Wolfe et al., 1987). Two of these included other psychiatric patient groups who showed impairment (schizophrenics, Johnson & Crockett, 1982; bipolar depressed, Wolfe et al., 1987). Calev, Nigal and Chazan (1989) compared depressed, manic, stable bipolar, and normal controls

using either a letter or a semantic category as prompts. They found depressed Ss were impaired relative to the other groups only on the semantic task, which they suggest is more difficult than the letter version of the task. Ilsley et al. (1995) found depressed Ss to be unimpaired on both letter and semantic category versions, and Beatty et al. (1990) found no depressive impairment on letter, semantic category and design versions, but depressed Ss were impaired on a famous people version. Five studies (Beats et al., 1996; Brown et al., 1994; Elliott et al., 1996; Robertson & Taylor, 1985; Trichard et al., 1995) found depressed Ss to be impaired on both letter and semantic versions. Two studies (Boone et al., 1995; see above; Dalla Barba et al. 1995; see above) found impairment on the letter version, but did not include a semantic version, while Peselow et al. (1991) found impairment on a semantic version, but did not measure letter fluency. Peselow et al. (1991) also reported depressed Ss who responded to treatment showed significantly more improvement than nonresponders on this task when retested, and Trichard et al. (1995) found deficits had disappeared when Ss were retested after successful treatment. The results of these studies suggest depression is sometimes associated with impairment on letter versions, with impairment found more consistently on semantic category retrieval, which may be more difficult. However, any deficits seem to remit with recovery.

In summary, there have been relatively few studies of attention and WM in depression. Studies of attention have consistently showed depressive deficits in vigilance, while studies of selective attention have been less robust. With regard to WM, the strongest evidence relates to CE function as measured by tasks sensitive to frontal lobe deficits, although the findings have not been consistent. The effects of distraction on depression suggest the cognitive performance of normal Ss given a secondary task may resemble that of depressed Ss without one. Overall, there is evidence that depression may be associated with deficits on attentional and WM processes likely to underlie performance on reasoning tasks. In the next section the evidence relating to the performance of depressed Ss on tasks designed to measure reasoning processes will be reviewed.

2.3.8 REASONING/PROBLEM-SOLVING

2.3.8.1 Deductive Reasoning

Few studies have examined depressive performance on tasks traditionally associated with deductive reasoning. Channon and Baker (1993) compared dysphoric and nondysphoric students selected on the basis of the BDI on a syllogistic reasoning task of the type

described in section 2.2.2.1. The findings revealed a significant difference between the groups in ability to solve the problems correctly, and an examination of the types of errors made showed the dysphoric Ss made significantly more errors which involved a failure to integrate information from the two premises to solve the problems.

2.3.8.2 Inductive Reasoning

As described in section 2.2.2.2, inductive inferences are an integral part of human thinking, and are associated with a range of processes often making it difficult to categorise what is being measured by a task. Therefore, this section is divided using the headings commonly found in reviews of inductive reasoning, but recognising some tasks may overlap these sections.

2.3.8.2.1 Rule-learning

2.3.8.2.1.1 Clinical samples

Gruzelier et al. (1988) employed conditional associate learning tasks (Petrides, 1982) which require inductive inferences for acquisition. They compared affective disorder Ss, schizophrenics and controls on spatial and nonspatial tasks. Affective disorder Ss performed the nonspatial task better than the spatial task, but were significantly poorer than the controls; however, the authors failed to report separate analyses for manic and depressive Ss.

Abas, Sahakian and Levy (1990) compared a group of older depressed patients with patients with early dementia of the Alzheimer-type (DAT) and normal controls on a spatial conditional associate learning task which varied the number of associations to be learned. The data revealed the control group reached criteria at each level of difficulty, while the depressed group performed at the same level as the control Ss when only one or two associations had to be learned, but started to show deficits when this increased to three. When eight associations had to be learned the depressed group performed at the same level as the DAT patient group. The depressed Ss were retested on recovery, when they showed improved performance, although they were still impaired in their learning of eight associations.

2.3.8.2.2 Hypothesis-testing

2.3.8.2.2.1 Clinical samples

As noted in section 2.3.7.5.1, Boone et al. (1995) compared depressed and control Ss on the WCST (see section 2.2.3.1). The WCST consists of a series of cards which Ss are asked to sort into one of four piles. The experimenter gives the S feedback about the correctness of each sort, and the rule which the experimenter is following changes in a pre-determined fashion, although the S is not aware of this. Boone et al. (1995) found depressed Ss were impaired relative to controls in successfully identifying the rules, but did not differ from controls on number of perseverative errors (continuing to test a rule following negative feedback from the experimenter). Beats, Sahakian and Levy (1996) used a set-shifting task they described as being similar to the WCST. They found clinically depressed Ss were impaired in identifying the rule in operation and shifting from one rule to another in response to feedback. The depressed Ss were retested after showing clinical improvement, and showed substantial but incomplete recovery of performance. Elliott et al., (1996) reported on the same set-shifting task and found no significant differences between depressed and nondepressed Ss. Dalla Barba et al. (1995) used a modified version of the WCST in which Ss were explicitly informed of a change in the rule. They reported depressed Ss achieved significantly fewer categories than controls, and made significantly more errors.

Silberman, Weingartner and Post (1983) used a concept discrimination learning task (Levine, 1966) and found clinically depressed Ss to be impaired relative to controls in using feedback to eliminate possible solutions and to identify the correct ones. King and Phillips (1985) used Levine-type tasks to compare affective disorder Ss with process and reactive schizophrenic groups. They reported impaired performance in all groups, including poor focusing and perseveration with negative feedback, with poorest performance in the reactive schizophrenics. However, the fact that there was no control group, and the inclusion of both depressed and manic patients in the affective group, means the findings are difficult to interpret.

2.3.8.2.2.2 Nonclinical samples

Smith, Tracy and Murray (1993) carried out experiments with students (Experiment 1) and patients of mixed diagnoses (Experiment 2) allocated to groups on the basis of their BDI scores. In both experiments they found evidence that dysphoric Ss were impaired on a

concept discrimination learning task, but performed similarly to controls on a 'holistic categorisation' task which the authors argued did not require a hypothesis-testing approach. Direct comparison of the two tasks is, however, hampered by the finding that the two tasks differed in level of difficulty. Dobson and Dobson (1981) studied the strategies used in a concept attainment task of the type described by Bruner et al. (1956), and found dysphoric students selected on the basis of the BDI performed more poorly than a control group in identifying more difficult solutions, and had a greater tendency to seek redundant confirmatory information.

2.3.8.2.3 Abstracting ability

2.3.8.2.3.1 Clinical samples

Andreasen (1976) assessed depressed patients on three measures of abstracting ability: the Goldstein-Scheerer Object Sorting Test (OS; Goldstein & Scheerer, 1941); the Ravens Progressive Matrices (RPM; Raven, 1965); and the Shipley-Hartford Scale (S-H; Shipley, 1940). Patients were tested upon admission and discharge and there was no control group. The OS test (Goldstein & Scheerer, 1941) requires Ss to sort a variety of objects into categories, and was designed to measure ability to abstract. The RPM requires Ss to abstract the relationship between items in an incomplete matrix so as to correctly identify the item that will complete it. The S-H Scale (Shipley, 1940) consists of incomplete verbal and numerical sequences. Ss are asked to detect the relationship between the items in order to complete the series. It also has a separate vocabulary test. Andreasen (1976) reported that on all three tests there was a nonsignificant tendency for improved performance at discharge. However, the failure to include a control group means this could be attributable to the effects of practice.

Salzman, Goldstein, Atkins and Babigian (1966) compared groups of neurotic and psychotic depressed Ss with three other groups of psychiatric patients on the S-H Scale (see above) and on the Gorham Proverbs test (Gorham, 1956) which asks Ss to interpret proverbs. There was no significant difference between the groups when vocabulary level was taken into account. Again, the failure to include a control group makes it difficult to interpret the findings of this study. Braff and Beck (1974) compared depressed, schizophrenic and control Ss on the S-H Scale, and also compared the depressed and schizophrenic groups on the Gorham Proverbs test. They found depressed Ss were impaired on the abstracting, but not the vocabulary measure of the S-H Scale compared

with the controls, but the schizophrenic group showed a greater degree of impairment than the depressed Ss, and were impaired on both measures. The schizophrenic group also performed more poorly than the depressed Ss on the Proverbs test.

Shipley, Kupfer, Spiker, Shaw, Coble et al. (1981) assessed a relatively large sample of depressed patients on the S-H Scale, and compared their performance with the task norms. Ss should exhibit the same level of performance on both the vocabulary and abstracting sub-tests, but they found the depressed Ss showed relatively poor performance on the abstracting measure, and this deficit remained when Ss were re-tested on recovery.

The Category Test from the Halstead-Reitan Battery (Halstead, 1947) requires Ss to identify visuospatial relationships such as position or orientation in a series of geometric figures. Savard, Rey and Post (1980) studied the performance of unipolar and bipolar inpatients on this test, and found the depressed Ss made more errors overall than controls, but this was attributable to deficits in the bipolar rather than the unipolar Ss. Donnelly, Waldman, Murphy, Wyatt and Goodwin (1980) also compared unipolar, bipolar and control Ss on the same test. They reported significant differences between bipolar depressed patients and normal controls which were not attributable to factors such as age and IQ. Unipolar depressed patients did not differ significantly from the bipolars during the acute stage, but the comparison with normal controls did not quite reach significance. The depressed Ss showed improvement after remission, but details were not reported.

2.3.8.2.3.2 Nonclinical samples

Sprock, Braff, Saccuzzo and Atkinson (1983) allocated a group of patients being treated for pain to dysphoric and nondysphoric groups on the basis of their BDI scores, and also selected a group of normal controls. They found the dysphoric pain Ss had significantly poorer performance on a forced choice version of the Gorham Proverb test than normal or nondysphoric pain controls. They tended to give less abstract definitions, and these differences were not attributable to age, education or vocabulary level. The dysphoric Ss also scored significantly lower on the Similarities sub-test of the WAIS, another measure of abstracting ability.

2.3.8.2.4 Problem-solving

As noted in section 2.1.1, there are strong links between problem-solving and reasoning.

Carrying out a reasoning task can be interpreted as a problem-solving activity, while problem-solving frequently requires reasoning. The following tasks are classified as problem-solving tasks in that they require several steps to solution, and they require Ss to generate a strategy. Task solution is likely to require reasoning, but it is less certain what type of inferences Ss might draw and when, since the tasks are less constrained than the reasoning tasks described above.

2.3.8.2.4.1 Clinical samples

Watts MacLeod and Morris (1988) used the Tower of London task (Shallice, 1982) to examine problem-solving ability in depressed and nondepressed Ss. In this task, three beads are arranged in various starting positions on sticks of unequal length, and Ss are required to move them to a goal position in the minimum number of moves, with certain constraints on the types of possible move. The groups did not differ significantly in the number of problems solved, but time taken both to plan and execute the moves was significantly longer for the depressed Ss. Beats et al. (1996) compared depressed and normal controls on a computerised version of this task. The depressed Ss took significantly more moves overall to solve the problems than controls, although the groups did not differ significantly in the number of problems solved in the minimum number of moves. The depressed Ss were significantly slower than controls for both initial and subsequent movement times. Beats et al. (1996) calculated a measure of thinking times from the movement time data, and found the groups did not differ in initial thinking time, but depressed Ss did spend significantly longer than the controls thinking about the task after making the first move. There was evidence of improvement in performance when depressed Ss were re-tested on recovery, but they remained significantly slower than the controls. These two studies (Beats et al., 1996; Watts et al., 1988) are similar in finding the accuracy of depressive performance tended to be less impaired than latency. Elliott et al. (1996) reported a different pattern, with depressed Ss showing global impairment in performance accuracy, but no difference in movement time. When thinking times were calculated, depressed Ss were found to spend more time than controls planning the simpler problems, but less time planning the more difficult problems. The depressed Ss spent significantly longer than the controls thinking about the task after making the first move.

2.3.8.2.4.2 Nonclinical samples

Hiroto and Seligman (1975) developed an anagram task in which a series of twenty

anagrams of equal length are arranged in a standard letter sequence. Whilst they can be solved individually, recognition of the standard sequence considerably improves performance, assessed in terms of response latency, number of unsolved anagrams, and number of trials to learn the pattern. This task was originally developed to test the learned helplessness model of depression described in section 2.1.3.3.3, and a series of studies employed 'helplessness' manipulations in nondepressed Ss by exposing them to uncontrollable aversive events, and assessing subsequent task performance. Some of these studies have also included a straightforward comparison of dysphoric and nondysphoric Ss on the task. Of these, several have found evidence of deficits in task performance in dysphoric compared with control Ss (Klein, Fencil-Morse & Seligman, 1976; Miller & Seligman, 1975; Willis & Blaney, 1978, Experiment 3), while others have not (Gotlib & Asarnow, 1979; Sacco & Hokanson, 1978). All five of these studies used similar methods of S recruitment, and Ss showed similar scores on the BDI. The tasks used in each study were also very similar, although the studies did vary experimental conditions such as the addition of a 'helplessness' condition. The variability of the findings therefore suggests any depressive deficit on this task is not robust, at least when nonclinical Ss are used.

Hertel and Knoedler (1996, Experiments 1, 2A & 2B) compared dysphoric and nondysphoric students selected on the basis of the BDI on their ability to solve problems by analogy. In Experiment 1, Ss first solved a series of training problems that were analogous to subsequent target problems. In the hint condition, the relationship between the target problem and the preceding training problem was made explicit in the task instructions. In the no-hint condition no reference was made to this relationship. The nondysphoric Ss solved more target problems in the no-hint relative to the hint condition, suggesting the hints inhibited performance, perhaps by focusing efforts on remembering irrelevant details of the training problems. In contrast, the dysphoric Ss solved a similar number of target problems in the two conditions. Of more interest, while there was no difference between the groups in the no-hint condition, the dysphoric Ss solved more target problems than the nondysphoric Ss in the hint condition. Hertel and Knoedler (1996) suggest post-hoc that the dysphoric Ss failed to concentrate on trying to recall the training problems when prompted to by the hints, and therefore did not experience the disadvantage shown by the nondysphoric Ss. In Experiment 2B, dysphoric and nondysphoric Ss were compared on the same set of problems with all Ss receiving hints, but during the training phase Ss were either instructed to attempt solution of the training

problems (as in Experiement 1) or they were instructed to try to remember the training problems. The results showed nondysphoric Ss solved more problems in the memory-oriented versus the problem-oriented training, while dysphoric Ss showed no difference between the two conditions. When comparing the groups, dysphoric Ss solved more problems than nondysphoric Ss in the problem-oriented condition, replicating the findings of Experiment 1, but the groups did not differ in the memory-oriented condition. The authors suggest the memory-oriented training provided a better basis, for the nondysphoric Ss at least, for transfer of the solution, reducing the interfering effects of remembering irrelevant details, while the results suggest dysphoric Ss did not attempt to make use of the strategies built into the experimental design, and therefore showed an advantage in performance when these strategies proved to be unhelpful.

In summary, there have been relatively few studies of reasoning and problem-solving on neutral tasks, and many of those carried out are difficult to interpret due to problems of design. For example, some studies (e.g. King & Phillips, 1985; Salzman et al., 1966) have compared depressed patients with groups of other psychiatric diagnoses, but have not included a normal control group, while others have included both manic and depressed Ss in the same group (e.g. Gruzelier et al., 1988). Overall, the weight of evidence suggests a deficit on reasoning and problem-solving tasks associated with depression, but the studies reviewed above provide very little information about the qualitative nature of any deficits, and have made little attempt to explore the processes which might underlie depressive performance on these tasks.

2.3.9 MODELS OF COGNITIVE FUNCTION IN DEPRESSION

Overall, the evidence suggests depression is often, but not always, associated with deficits on a range of tasks. The variable findings are likely to reflect problems in study design, many of which were outlined in the section on methodological issues (see section 2.3.1). More recent studies have tended to be better designed than earlier work and the situation is likely to be clarified in the future.

A variety of hypotheses and models have been put forward to explain the mechanisms which might underlie any cognitive impairments associated with depression. In recent years the debate has centred on the question of whether depression is associated with an actual reduction in processing resources or whether deficits can be explained by a lack of

motivation or "cognitive initiative" (Hertel & Hardin, 1990; Hertel & Rude, 1991b) or a response bias (Johnson & Magaro, 1987).

2.3.9.1 Reduced Capacity

The notion of reduced capacity in depression was first proposed by Hasher and Zacks (1979). The main assumption of resource allocation theory is that encoding operations vary in the demands they make on limited attentional resources, and that individuals vary in the amount of attentional resources they have available (see section 2.2.5.1.). Hasher and Zacks (1979) suggested the pattern of cognitive deficits associated with depression was compatible with reduced attentional resources, that is, deficits are more likely on tasks which make high demands on attentional resources. While Hasher and Zacks (1979) did not specify the means by which resources might be reduced in depression, others have suggested biological mechanisms might be implicated. For example, Roy-Byrne et al. (1986) suggested depression may selectively impair effortful processes as opposed to automatic processes by interfering with the effects of the neurotransmitter dopamine.

Ellis and Ashbrook (1988) proposed a model which built on the ideas posited by Hasher and Zacks (1979). Like Hasher and Zacks, Ellis and Ashbrook (1988) assumed tasks vary in the demands they place on attentional resources, and depression might be a factor in regulating the amount of available resources. They introduced the idea that increased processing of irrelevant aspects of the task, or processing of non-task material, such as material relating to personal concerns, might be one mechanism by which depression reduces available resources. Thus, Ellis and Ashbrook suggest resources might be diverted as well reduced in depression, but the overall effect will be one of reduced availability of attentional resources. This model has strong links with the cognitive models of depression reviewed in section 2.1.3.3.6 (e.g. Williams et al., 1988) which posited depression is associated with biases in memory, and possibly attention, for negative material which could divert depressive allocation of resources. Hasher and Zacks (1988) put forward a model which assumes narrowing of attentional focus to be the underlying causal mechanism. They suggested in both normal ageing and depression, normal inhibitory mechanisms may become less efficient. This would permit irrelevant information to enter WM and receive sustained activation, leading to reduced resources.

The hypothesis that depression is associated with reduced and/or diverted cognitive

resources predicts performance should suffer with increasing task demands on attentional capacity. If processing resources are reduced or partially occupied, fewer resources are available for task performance. If the task is resource-limited (see section 2.2.5.1.3), impairment in performance may occur, and the extent of this should be related to the amount of resources needed for its performance, so it is predicted depressed Ss should not show deficits on tasks which make no, or only minimal demands on processing resources. Hypotheses of both reduced and diverted resources make identical predictions in most instances, although Hartlage, Alloy, Vazquez and Dykman (1993) have suggested some ways in which they could best be distinguished in future research.

The evidence reviewed in section 2.3 suggests depressed Ss are differentially impaired on effortful relative to automatic tasks consistent with the prediction of a reduced resources model, although the inconsistency of the findings in relation to implicit memory tasks (see section 2.3.6.2) suggest this conclusion cannot yet be fully accepted. Studies which have varied task demands have frequently reported depressed Ss show relatively more impairment as demands increase (see section 2.3). However, Robbins, Joyce and Sahakian (1992) postulate studies may have confounded task demands with task difficulty, so that, for example, differential performance on automatic and effortful tasks might simply be a function of task difficulty. They suggest difficult tasks might be more discriminating in separating patients and controls, but the concept of resource allocation may not have any additional explanatory power.

The theories of reduced and/or diverted resources in depression (Ellis & Ashbrook, 1988; Hasher & Zacks, 1979) outlined above seem to assume a single pool of attentional resources. However, as noted in section 2.2.5.1.4, theorists (e.g. Navon & Gopher, 1979; Wickens, 1984) have argued multiple resources may exist, each with its own capacity. It is therefore important to consider the implications of this type of attentional system for current theories of reduced/diverted resources in depression. The assumption of a single pool of resources predicts the effects of a reduction/diversion of resources will be determined simply on the basis of the total demands made on resources. In contrast, the predictions of a multiple resource model of attention would be far more complex, since any deficits would depend on which of the resources were reduced/diverted in depression and which were required by the current task, as well as whether the available resources were exceeded. In section 2.3.7, the evidence of depressive task performance was

assessed in relation to one model of multiple resources: the WM model posited by Baddeley and Hitch (1974). While there was some evidence of depressive impairment on measures of the CE component, the two slave systems, the PL and the VSSP, appeared relatively intact. Other models of multiple resources are not yet sufficiently well-defined to permit this type of detailed assessment, although this is likely to change in the future.

2.3.9.2 Response Style

Other models have challenged the notion depression is associated with reduced availability of resources, and have suggested instead their style of performance differs from non-depressed Ss in ways that lead to the appearance of reduced resources. Johnson and Magaro (1987) proposed depressed Ss recall as much as controls, but are less willing to report the information. The evidence for the hypothesis that depressed Ss are characterised by a conservative response bias was examined in section 2.3.4.3, and on the whole it seemed to suggest depressed Ss do have a tendency to respond in a conservative manner, but nevertheless they also exhibit true deficits on some tasks.

One of the earliest hypotheses regarding depressive deficits was the proposal that depressed individuals are simply not motivated to do well when given tasks to perform, or that, even though motivated, depressed individuals may be unable to sustain any prolonged motivation. This is consistent with the reduction in motivation identified as a salient feature of depression by several theorists (e.g. Abramson, Seligman & Teasdale, 1978; Beck, 1967). In his early review of the field of cognitive function in depression Miller (1975) identified reduced motivation as one possible explanation of the pattern of findings, and this was supported by McAllister (1981) in his review. Subsequently, the simple motivation explanation became less popular, and little research effort has been expended to test the specific predictions of this position. One exception was reported by Richards and Ruff (1989) who compared depressed patients and normal controls on a range of cognitive tasks in motivation and non-motivation conditions. They found a manipulation designed to increase motivation did not differentially improve performance in the depressed Ss, even though the manipulation enhanced speed of responding to the same degree as in normal Ss. The authors concluded depressive deficits were not attributable to motivational factors. The findings of studies described in section 2.3.7.3 that the addition of a secondary task does not necessarily impair performance in depressed Ss, and may even improve it, is consistent with the hypothesis of reduced motivation.

Recently, Elliott et al. (1996) have proposed that depressive performance on measures of cognitive function may be influenced by a highly specific form of motivational deficit involving the response of patients to perceived failure. They reported that on two tasks, negative feedback given immediately following an experimental test trial, increased the likelihood that depressed patients would fail the next item. Elliott et al (1996) speculate that the depressed Ss' response to the negative feedback interfered with their performance on the next test item. However, this explanation cannot account completely for depressive deficits since depressed Ss were significantly impaired relative to controls after controlling for the effects of failure.

In a more sophisticated version of the motivational hypothesis, Hertel and her colleagues (e.g. Hertel & Hardin, 1990; Hertel & Rude, 1991b) have challenged the processing resources account of depression. They argue depressed Ss are characterised by reduced initiative, leading to failure to use strategies spontaneously or engage in elaborative thinking, but they are capable of this when directed. They emphasise depression impairs initiative to use strategies, but not the ability to do so. On many tasks, the initiative hypothesis makes predictions which are similar to the processing resources model. For example, reduced initiative would also predict depressed Ss should be unimpaired on automatic tasks which do not require the use of a strategy.

In order to test their model, Hertel and her colleagues have conducted a series of experiments in which they manipulated whether the appropriate strategy was implicit in the task, or had to be generated spontaneously. They found any deficits shown by depressed Ss relative to controls tended to show improvement with specific direction as to the use of an appropriate strategy. They therefore suggested tasks sensitive to depressive deficits would be those which permit but do not specify the spontaneous use of strategies, rather than those which direct or bypass the use of strategies. Hertel and Knoedler (1996) demonstrated this lack of initiative can be beneficial when attempts to use a strategy interfere with task performance (see section 2.3.8.2.4.2). Hertel and her colleagues contend that an account in terms of reduced or unavailable processing resources alone is inadequate, since performance of their tasks under directed strategy conditions was in some cases likely to be at least as effortful as undirected conditions. However, their findings were not clear-cut since directed strategy manipulations did not always prove to be completely effective in removing performance deficits. Thus, Hertel and Rude

(1991b) found depressive Ss remained slower than controls on a secondary probe latency task, although primary task performance improved to normal levels. Hertel and Rude (1991b) acknowledged that impairment in initiating strategies may reflect either motivational deficiencies, or a true cognitive deficit in planning and generating appropriate performance strategies (which may in turn require processing resources).

In summary, there are a number of competing models of depressive deficits, and on the basis of current evidence it is difficult to conclude which model is best supported. This situation has in part arisen from the fact that predictions made by the models are indistinguishable for many of the tasks which have been utilised, while only a small number of studies have attempted to test hypotheses which would separate the models. Furthermore, since the models are not incompatible with each other, it is possible the current situation has arisen because more than one of the models is correct. For example, Watts (1993) concluded both reduced processing resources and under-deployment of remaining resources may characterise depressed Ss.

The models of cognitive function in depression outlined above have given some attention to the possible biological mechanisms which might underlie any deficits in cognitive function associated with depression (e.g. Roy-Byrne et al., 1986). However, the recent work using neuro-imaging techniques to investigate physical changes in depression which was reviewed in section 2.1.3.2.2 is likely to inform theories of cognitive function to a far greater degree in the future. Furthermore, this work may potentially provide a link between the competing theories. For example, it is possible both reduced capacity and a particular response style might result from impaired brain function. The studies reviewed in section 2.1.3.2.2 suggested impairments in frontal lobe function have most consistently been associated with depression. Frontal lobe function was reviewed in section 2.2.3. The evidence links the frontal lobes with executive function, postulating a role in both the allocation of attentional capacity and in the implementation of strategies. While hypotheses about possible links between depression and brain function must remain tentative at present, this is likely to be a fruitful direction for future research, and needs to be taken into consideration.

2.4 SUMMARY OF LITERATURE REVIEW

The material reviewed above suggests evidence from a range of theoretical perspectives would lead to a prediction of reasoning deficits associated with depression. These can be summarised as follows:

1. Depressed patients frequently complain of difficulties with concentration, memory and "thinking".
2. Early cognitive models suggested depression may be associated with distorted (e.g. Beck, 1967) or biased (Beck, 1967; Seligman, Abramson & Teasdale, 1978) thinking which may act to maintain the depressed state.
3. More recently, cognitive models (e.g. Williams et al., 1988) have suggested depression may be associated with biased memory, and possibly attentional, processing. This diversion of cognitive resources toward emotionally-salient material could potentially reduce the cognitive resources available for performing reasoning tasks.
4. The frontal lobes of the brain are thought to play an important role in the performance of reasoning tasks, and there is increasing evidence depression may be associated with frontal lobe dysfunction.
5. There is evidence depression is associated with deficits in a range of cognitive functions, such as memory, learning and attention. This has been attributed to reduced or diverted WM capacity or reduced initiative.
6. There is evidence reasoning ability is closely linked to WM, and that in normal circumstances people may use heuristics to reduce demands on WM, while their pattern of errors may reflect the demands placed on WM.
7. A small number of studies have examined depressive performance on reasoning tasks, and found evidence of deficits.

Thus, there is strong evidence to support the hypothesis depression is associated with reasoning deficits, and this may be mediated by reduced or diverted WM resources and/or by reduced initiative to generate appropriate strategies. However, as yet this hypothesis has not been tested directly.

CHAPTER III

METHODOLOGY

3.1 AIMS OF THE CURRENT STUDY

As noted in section 2.4, there is evidence from a range of theoretical perspectives to support the hypothesis that depression is associated with reasoning deficits, and that this may be mediated by reduced or diverted WM resources and/or by reduced initiative to generate appropriate strategies. The aim of the current study was threefold: 1) to establish whether depression is associated with deficits on reasoning tasks; 2) to explore the possible contribution of reduced or diverted processing resources in depression by varying the demands made on WM by particular reasoning tasks; 3) to evaluate the use of appropriate strategies by depressed and nondepressed Ss by examination of their pattern of performance. Reasoning tasks were selected from the fields of cognitive psychology and neuropsychology. The five experiments are described in chapters IV to VIII. Before introducing the first of these experiments, there are several methodological issues that are germane to the entire study, and these are considered below.

3.2 METHODOLOGICAL ISSUES

3.2.1 EXPERIMENTAL DESIGN

As described in section 2.3.1, there are different approaches to studying cognitive function in depression, with advantages and disadvantages to each. At the outset, there was the option of carrying out the current study with either clinically depressed patients or students selected on the basis of the BDI, with the latter being available in greater numbers than the former. In section 2.3.1, the disadvantages of using clinically depressed patients were highlighted, including the influence of both medication and patient-status on cognitive performance, and the difficulty of selecting adequate control groups. Issues relating to the use of nonclinical samples were described in section 2.3.1, and are considered in more detail below.

3.2.1.1 Nonclinical Samples

As noted in section 2.3.1, the use of nonclinical samples circumvents several of the problems associated with studying patient samples. Usually nonclinical Ss have not sought treatment for their depression, and therefore are not receiving treatment that might affect their cognitive performance. Nonclinical samples are often recruited from student populations and therefore tend to be relatively homogeneous in terms of age, intelligence

and social background meaning the control group will be more closely matched to the experimental group on variables that might influence cognitive functioning. Nevertheless, there are difficulties associated with this approach relating to both the nature of nonclinical samples, and to the method of recruitment.

In two influential papers, Coyne and Gotlib (1983) and Gotlib (1984) put forward methodological criticisms of the use of nonclinical samples. On the basis of the existing literature they concluded the depression experienced by students is qualitatively different from that manifested by psychiatric patients, in particular student depression seemed to be associated with fewer somatic sympoms. However, Vredenburg, Flett and Krames (1993) noted there were serious flaws in the design of the studies on which Coyne and Gotlib (1983) based their conclusions. Furthermore, Hill, Kemp-Wheeler and Jones (1986) factor-analysed the BDI responses of psychiatric patients and students and found comparable factors for both groups, including a 'somatic disturbance' factor.

Gotlib (1984) noted students selected on the basis of high scores on a measure of depression also tend to score highly on measures of constructs other than depression, such as anxiety, and therefore may be suffering from general psychological distress rather than true depression. Vredenburg et al. (1993) challenged this conclusion on several grounds. First they note that many measures have an overlap of items which confound the findings. This is especially true of anxiety measures, such as the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch & Lushene, 1970) which has many items clearly related to depression rather than anxiety (see e.g. Endler, Cox, Parker & Bagby, 1991). Vredenburg et al. (1993) also note depression may be associated with other constructs, such as anxiety, in clinical populations as well (see section 2.1.1.1).

Coyne and Gotlib (1983) suggested depressed mood in student samples is relatively transient. Hatzenbuehler, Parpel and Matthews (1983) reported evidence, for the BDI at least, that this apparent transience may be due to an effect of retesting. They found Ss scoring in the depressed range tended to obtain a lower score on a second administration of the BDI, even if this was carried out on the same day, but this was removed if two different measures of depression were used. Vredenburg et al. (1993) postulated this effect might be the result of a fundamental law of statistics referred to as 'regression to the mean'. This law states in repeated testing, extreme scores will tend to be pulled

toward the mean as testing is repeated. Therefore, when Ss are selected on the basis of high or low scores on a test, a later reassessment is likely to yield mean scores that are closer to the mean of the original full sample. However, this explanation would also predict the mean of the control group should rise on the second occasion, a prediction not borne out by the findings (e.g. Hatzenbuelhler et al. 1983).

It is possible to introduce selection procedures that reduce the likelihood of selecting Ss with transient moods. One approach is to measure depressed mood on two occasions, that is, at the initial screening, and again when the experimental measures are administered. Ss are included only if they score within the criterion ranges on both occasions (e.g. Kendall et al., 1987). This approach increases the sample homogeneity by eliminating those with transient or unstable depressive symptoms (Deardorff & Funabiki, 1985).

One way of evaluating the use of nonclinical samples is to compare the results of studies carried out with nonclinical and clinically depressed Ss. In the current review (see section 2.3), the pattern of findings was similar, although the clinical studies tended to reveal greater impairment, presumably reflecting the greater severity of the symptomatology in those samples. Vredenburg et al. (1993) compared clinical and nonclinical findings in other fields of study, and concluded there were no major differences between the two.

On the basis that students were more readily available, and that this type of methodology has been widely and successfully used, nonclinical samples were recruited in the current study. The next issue was how to select Ss.

3.2.1.2 The Beck Depression Inventory

The most common measure used to select nonclinical samples is the Beck Depression Inventory. This was introduced in 1961 (Beck et al., 1961), and the revised version was developed in 1971 and copyrighted in 1978 (Beck, Rush, Shaw & Emery, 1979) (see Appendix 1). It is a 21-item self-rating measure originally designed to assess the severity of depressive symptomatology in persons already diagnosed as depressed, and therefore not designed to diagnose depression. However, it has been used widely as a screening instrument for detecting the presence of depressive symptomatology in other populations, such as students. When used in this way it is important to be clear about its limitations as a measure, and to emphasise that its design means it does not give a diagnosis of

depression. It has been recommended that samples selected solely on the basis of BDI scores should be labelled dysphoric rather than depressed (Kendall et al., 1987).

Recruiting Ss on the basis of their BDI score is quick, and avoids observer bias. It is also relatively unobtrusive, and may reduce feelings of discomfort associated with revealing personal details when compared with a long and detailed psychiatric interview. In studies comparing scores on the BDI with diagnoses made on the basis of structured interviews, it has been found to be relatively sensitive to depression, but only moderately specific, resulting in a proportion of false positives (e.g. Oliver & Simmons, 1984). However, as noted by Kendall et al. (1987), in terms of the base rate of different psychopathologies, students selected on the basis of depressive symptomatology are statistically more likely to be depressed than, for example, suffering from schizophrenia. The psychometric properties of the BDI have been reviewed on a number of occasions, and it has been reported as having good reliability and high validity (e.g. Beck, Steer, & Garbin, 1988).

3.2.1.3 Other Clinical Measures

As noted in section 2.3.1, when comparing groups of depressed and nondepressed Ss on measures of cognitive function it is essential to match the groups on variables, other than depression, that might influence cogntive function, such as age, sex and intellectual functioning. For the latter, some studies have matched Ss on educational level, but for student samples, this would not allow adequate differentiation between Ss. It was therefore necessary to select a measure of intelligence. However, as noted in section 2.2.1, there is a strong relationship between intelligence and reasoning, and many tests of intelligence include reasoning tasks, so it was necessary to ensure the measure was not confounded with the experimental tasks. In their review, Hartlage et al. (1993) noted the Vocabulary, Picture Completion and Information sub-scales of the WAIS (Wechsler, 1955) all primarily involve retrieval of information encoded before the depressive episode. Since this is a relatively automated process, they predicted depressive performance on these tasks should be preserved. They argued further that the Vocabulary subtest is the most automated of the three since it essentially entails the automatic activation of the meanings of words, and they concluded the evidence supported their prediction that depressed Ss would not show impairment on this task. Therefore the Vocabulary subtest from the WAIS-R (Wechsler, 1981) was chosen as a measure of intellectual ability. It also has the best individual correlation with WAIS-R Verbal IQ.

3.2.2 STATISTICAL APPROACH AND ANALYSIS

Each of the experiments to be described in the current thesis has one independent variable (IV): the allocation of Ss to dysphoric and nondysphoric groups. The two groups were then compared on at least one dependent variable (DV). The DVs were one or more measures of performance on a reasoning task. Where only one I V was measured, the comparison could be carried out using univariate statistics (e.g. t-test). However, the complexity of many of the experimental designs meant the most important and revealing comparisons involved two or more I Vs, thereby necessitating the use of multivariate procedures (e.g. repeated measures analysis of variance; ANOVA). For example, many of the experimental designs employed in the present series of experiments involved the comparison of two groups on a task carried out in two experimental conditions. Furthermore, within each condition there might be two or more meaures, either one variable measured at different stages of the task (e.g. number of correct responses after the first, second and third feedback), or measures taken at different levels of the same variable (e.g. number of correct responses on problems with one, two or three premises).

When selecting a test to analyse a data-set, cetain issues need to be addressed. The aim of analysis is usually to get as close as possible to the 'truth' as revealed by the data. This is formalised by setting up an experimental hypothesis, which predicts a difference in the DVs as a result of the experimental manipulation of the IV(s). The null hypothesis is also set up, which predicts there will be no effect of the experimental manipulation. When selecting a statistical test, two possible kinds of error are of concern, known as Type I and Type II errors. A Type I error occurs when the experimental hypothesis is accepted and the null hypothesis is rejected, even though the null hypothesis is true. On the other hand, a Type II error involves not rejecting a null hypothesis that is in fact false. The selection of a statistical test is influenced in part by the need to achieve a balance between the possibility of making one of these two errors. That is, the aim is to identify an experimental effect, but only if it really exists.

Statistical tests are often divided into two types: parametric and nonparametric. Parametric tests require the data to meet more stringent assumptions than do nonparametric tests. When carrying out data analysis, parametric tests are often preferred to nonparametric tests because of their greater power. The 'power' of a test refers to the probability of correctly rejecting a false null hypothesis and accepting the experimental

hypothesis. In general, when the assumptions of a parametric test are met, the nonparametric test requires more observations than the parametric test for the same level of power. Thus, for a given set of data, the parametric test is more likely to lead to rejection of a false null hypothesis than is the corresponding nonparametric test. For this reason, in the current study, parametric tests were adopted whenever appropriate.

As mentioned above, parametric tests require the data to meet certain assumptions. The parametric tests used in the current study, analysis of variance (ANOVA) and the t-test, make two basic assumptions. The first assumption is that the populations from which the samples were taken are normally distributed. The second assumption is that the samples are drawn from populations of equal variances. This is known as the *homogeneity of variance* assumption. (The variance is a measure of the dispersion of a distribution, and is calculated by summing the squared deviations of each observation from the mean of the distribution; for more detail see e.g. Howell, 1987, p. 39.) Although these two assumptions are theoretical requirements of parametric tests, in practice the tests are robust to violations of these assumptions, provided they are not too extreme. The use of samples that are relatively large and equal in size offers protection against the effects of any violations. For example, with regard to homogeneity of variance, the general conclusion is that provided sample sizes are equal, violation of the assumption of homogeneity produces very small effects (Howell, 1987, p. 179). In the current study, the use of student samples had the advantage that Ss were available in relatively large numbers, and it was therefore possible to achieve adequate sample sizes with equal numbers in each group. It was therefore assumed that provided the data passed the checks described below, it was safe to proceed with parametric analysis on the basis that any violation of these two assumptions was unlikely to lead to serious problems.

While noting parametric tests are generally robust to violations of their underlying assumptions, Tabachnick and Fidell (1983; 1989) have made recommendations about issues that need to be considered before proceeding with multivariate parametric tests. Many of these are also relevant to other parametric tests. Therefore, before carrying out data analysis, the following issues were considered, and remedial steps taken where necessary. If it was found that a particular data-set was not appropriate for parametric analysis, then an alternative was sought. This is also outlined below.

Tabachnick and Fidell (1983; 1989) identified unequal sample size and missing data as potential problems when attempting multivariate analyses. Fortunately, in the current study sample sizes were equal in each experiment, and there were no missing data.

As noted above, parametric tests make the assumption that the populations from which the samples were drawn are normally distributed. In multivariate tests, this becomes an assumption of multivariate normality, implying that the sampling distributions of the mean of the various DVs in each cell and all linear combinations of them are normally distributed. The sampling distribution of the mean is the distribution of values that would be obtained for that statistic if an infinite number of samples were drawn from the population in question and the mean was calculated for each sample. All the important information about the sampling distribution of the mean is summed up by the Central Limit Theorem. In its simplest form, this states the sampling distribution of the mean approaches normal as N, the sample size, increases. With univariate F and large samples, the central limit theorem suggests the sampling distribution of the mean approaches normality even when the raw data do not. Tabachnick and Fidell (1989) note univariate F is robust to modest violations of normality as long as the violations are not due to outliers (see below).

Tabachnick and Fidell (1983; 1989) note multivariate analyses assume linear relationships among all pairs of DVs, and deviations from linearity reduce the power of the test. The assumption of linearity is that the relationship between two variables, between one variable and a combination of others, or between combinations of variables from each of two sets can be described using a straight line. Normal distribution of each DV increases the chances of a linear relationship. The only way to establish a linear relationship exists is to plot each pair of DVs on a bivariate scatterplot, and then make a subjective judgement about their relationship. With a large number of variables this is both time-consuming and likely to be inexact. Therefore in the current study, each variable was inspected for its normality by screening for outliers, kurtosis and skewness (see below), and either taking steps to achieve a normal distribution or by using alternative nonparametric analyses. Homoscedasticity is the assumption that the variability in scores on one variable is roughly the same at all values of the other variable. This assumption is met when both variables have a normal distribution, and therefore the steps taken to ensure linearity (screening for outliers, skewness and kurtosis) should also ensure the data

meet the assumption of homoscedasticity, or that any data-set failing this assumption will be identified, and appropriate steps taken (see below).

On the basis of the recommendations put forward by Tabachnick and Fidell (1983), the following procedure was adopted in the current study for each data-set. The data were first inspected for unequal sample sizes and missing data; neither were found on any occasion in the current study. Next, each variable was inspected for normality of distribution. The following procedures were carried out to identify and deal with outliers, skewness, and kurtosis, all of which can lead to a failure of normality in a variable, and have deleterious effects on the robustness of parametric tests.

T-test and ANOVA are sensitive to outliers. Outliers are cases with such extreme values on a variable that they unduly influence statistics. They can lead to both Type I and Type II errors, with no clue as to which has occurred and they lead to results that do not generalise because of being overly determined by the outlier(s). In the current study, outliers were detected by converting each variable to standardised scores, and identifying any cases that had standardised scores in excess of +/-3.00. The influence of any outliers was reduced by either transforming the data as described below, or by replacing the raw score of the outlier with the score-plus-1 of the next most extreme case in the distribution, as recommended by Tabachnick and Fidell (1983). This process was carried out separately for each group.

Skewness has to do with the symmetry of a distribution in that a skewed variable is one whose mean is not in the centre of the distribution. The skewness of each variable in the current study was examined using the equation recommended by Tabacknick and Fidell (1983) to compare it with the standard error for skewness and test whether it differed significantly from a normal distribution (zero). The standard error for skewness was calculated using the equation:

$$s_s = \text{sqrt. } 6/N$$

where N is the number of cases. The probability of obtaining a skewness value of the size given if the data came from a normal distribution was then calculated using the *z* distribution, where:

$$z = S - 0/s_s$$

Where S is the value reported for skewness.

At the 1% level, a z value in excess of +/-2.58 would lead to rejection of the assumption of normality.

$$S = +/-2.58 \times s_s$$

Kurtosis has to do with the peakedness of a distribution: a distribution is either too peaked (with too few cases in the tails) or too flat (with too many cases in the tails). The kurtosis of each variable was examined using the equation recommended by Tabacknick and Fidell (1983) to compare it with the standard error for kurtosis and test whether it differed significantly from a normal distribution (zero). The standard error for kurtosis was calculated using the equation:

$$s_k = \text{sqrt. } 24/N$$

where N is the number of cases. The probability of obtaining a kurtosis value of the size given if the data came from a normal distribution was then calculated using the z distribution, where:

$$z = K - 0/s_k$$

Where K is the value reported for kurtosis. At the 1% level, a z value in excess of +/-2.58 would lead to rejection of the assumption of normality.

$$K = +/-2.58 \times s_k$$

When a variable was found to have a skewness or kurtosis value in excess of the calculated acceptable level then steps were taken to reduce them. As described above, all the variables were subjected to a test for the presence of outliers, and treating them as described above frequently reduced skewness and kurtosis to an acceptable level. If skewness or kurtosis remained then an appropriate transformation of the data was performed. Different transformations can be carried out, and they vary in their strength and effect. In each case, the transformation resulting in skewness and kurtosis values closest to zero was selected. Tabachnick and Fidell (1983) discuss the transformations most likely to correct positive and negative skewness of different degrees. The most common transformation for positive skewness is either a square root or logarithmic transformation, depending on the severity of the skewness, although stronger transformations are possible. For negative skewness, Tabacknick and Fidell (1983) recommend a "reflex" strategy. This involves subtracting each sample value from the largest score+1 in the distribution, thus converting a variable with negative skewness to one with positive skewness, and the application of an appropriate transformation for positive skewness.

Violation of the assumptions underlying parametric tests can result in misleading conclusions with regard to the significance of the results. Therefore, when transformation of the data failed to reduce skewness and/or kurtosis to an acceptable level, an appropriate nonparametric comparison was carried out. However, there are data sets for which nonparametric analyses are not yet available, for example, when a comparison of two independent groups on two or more dependent variables is required. In these cases, repeated measures ANOVA was employed, and to check this approach was not leading to false conclusions about the data, the analysis was repeated as far as possible with nonparametric tests.

In line with much of the data collected in the course of psychological research, some of the variables in the current study were discrete rather than continuous. In particular, a number of variables were proportional in nature, such as 'number of errors out of ten.' Data of this type may not fit a normal distribution, and may be closer to a binomial distribution. The optimal method for dealing with data of this kind would therefore be statistical tests designed for a binomial distribution. However, there was no binomial test available to perform multivariate analyses. The alternative solution, which is commonly applied, is to transform the data. The arcsine transformation has been recommeded for proportions (Winer, 1971), and this was adopted in the current study. Variables were tested for skewness as described above. Where skewness was above the 1% level, then the data were transformed using the formula recommended by Winer (1971). The raw scores were first converted to proportions, and these were transformed using the equation:

$$2 \arcsin \sqrt{N}.$$

If this transformation failed to reduce skewness to an acceptable level then nonparametric analyses were the most appropriate solution, except in situations where a multivariate analysis was required, as described above.

When ANOVA is carried out with both between- and within-subject variables, then there is an additional assumption which must be met known as *sphericity*. For a full discussion of this concept see Winer (1971) or Greenhouse and Geisser (1959). When the sphericity assumption was violated then degrees of freedom were adjusted using the Greenhouse-Geisser adjustment (Greenhouse & Geisser, 1959).

The level of significance was set at 0.05 throughout. When post-hoc tests were carried out to explore interactions, the level of significance was adjusted by dividing 0.05 by the number of post-hoc tests. All analyses were carried out using the Statistical Package for the Social Sciences (SPSS).

CHAPTER IV

EXPERIMENT 1

4.1 INTRODUCTION

The material reviewed in chapter II identified a body of evidence supporting the prediction of reasoning deficits in depression and these were summarised in section 2.4. The review of cognitive function in depression contained in section 2.3 highlighted that there is a relatively small body of empirical work on reasoning processes in depression. Of the studies carried out, the most informative have used measures of rule-learning where Ss are required to formulate a hypothesis and test it over a series of trials to determine which guiding rule or concept is in operation (see section 2.3.8.2.1). Tasks of this type have also been found to be sensitive to frontal lobe deficits (see section 2.2.3.1), and as noted in section 2.1.3.2.2, there is increasing evidence of frontal lobe dysfunction in depression. The studies of rule-learning in depressed Ss are considered in more detail below.

Dobson and Dobson (1981) used a concept attainment task of the type described by Bruner et al. (1956; see section 2.2.2.2) to investigate the strategies used in learning three new logical rules of varying difficulty, and then identifying which rule was in operation. Using the BDI to classify Ss, they found dysphoric students performed more poorly than a nondysphoric group on both learning and identifying the two more difficult rules, while there was no significant difference between the groups on the simpler rule. On the more difficult rules, the performance of the dysphoric Ss was characterised by a greater tendency to collect redundant confirmatory information. Dobson and Dobson (1981) considered that this pattern of responding could be explained by either a conservative responding style, or by a failure to remember information that had been obtained on earlier trials, but the design of their study did not permit them to assess directly the contribution of memory to performance.

Silberman, Weingartner and Post (1983) used a concept discrimination learning task (see section 2.2.2.2) first described by Levine (1966) to examine the strategies employed by depressed Ss in testing and discarding hypotheses. This paradigm consists of a series of cards which each contain two stimuli (e.g. letters). The stimuli vary on four dimensions, so that, for example, one is white, the other black; one is big, the other small; one is on the left, one on the right; one is letter A, the other is letter B. The S is informed that one of these eight possibilities is the correct solution to the problem. On each trial the S

points to the stimulus that he thinks is correct, and on some trials receives feedback from the experimenter. The remaining nonfeedback trials provide information as to the hypotheses being tested by the S. Levine (1966) provided a detailed analysis of the processes involved in the successful solution of this type of discrimination learning problem, and the evidence generally supports his model (e.g. Levine, 1966; Eimas, 1970). Levine (1966) suggested that when presented with the task stimuli the first step is to code and store the dimensions believed to be correct (e.g. black, X, big, on the right). Having pointed to this stimulus, positive feedback should result in the continued storage of the coded possibilities, while negative feedback necessitates recoding the remaining logically correct set of dimensions (e.g. white, Y, small, on the left). Successful performance on the succeeding feedback trials requires both the retention of the coded or recoded list of possible solutions, and the intersection of information from all the feedback trials.

Silberman et al. (1983) compared clinically depressed and normal control Ss on a series of discrimination learning problems of the type described above. Ss also carried out some problems in a report condition where they were asked to list the potential solutions after each feedback trial. The depressed Ss were found to be significantly impaired relative to controls in using feedback to eliminate incorrect solutions and identify the correct ones. The performance of the depressed Ss was characterised by two types of error: an inability to narrow down the set of possible solutions, and perseveration on hypotheses that should have been disconfirmed. The report condition revealed accurate initial performance in the depressed Ss, but progressive difficulties in eliminating hypotheses correctly. This pattern of findings is consistent with a depressive deficit in the logical processes of recoding and intersection. However, it is also consistent with poorer retention of information across successive trials.

The findings of the studies carried out by Dobson and Dobson (1981) and Silberman et al. (1983) are suggestive of deficits associated with depression in carrying out the logical processes necessary for hypothesis testing and elimination. The tasks used in both studies require the processing and manipulation of information, thought to be carried out by the CE, and also the temporary storage of relevant information. Impairments in reasoning might be attributable either to deficient CE processing, or to reduced storage capacity. Using Levine's discrimination learning task, Eimas (1970) found facilitation of performance with normal children aged 7-9 years when feedback was left on display to

reduce the need to store relevant information, and he concluded that their deficits were largely the result of limited storage capacity.

The current experiment used Levine-type concept discrimination learning problems similar to those used by Silberman et al. (1983) to investigate further the nature of any reasoning deficits associated with depression. There were two main aims: first, to investigate whether the depressive deficits reported by Silberman et al. (1983) could be replicated using a dysphoric student sample rather than clinically depressed patients; and second, to vary the storage demands of the task, to see whether the performance of depressed Ss could be facilitated to normal levels. Two conditions were used which varied memory load by comparing aural feedback (standard condition) with visual feedback (memory-aid condition), in which visual feedback remained on display throughout each problem.

4.2 EXPERIMENTAL HYPOTHESES

If depression is associated with deficits on reasoning tasks then dysphoric students should show impairments relative to nondysphoric students on Levine-type discrimination learning problems.

If dysphoric Ss are impaired in their capacity to store information in WM then an experimental manipulation which reduces the storage load should differentially facilitate the performance of the dysphoric group.

If dysphoric Ss are impaired in their ability to process and manipulate information in WM then an experimental manipulation which reduces the storage load should not facilitate their performance to normal levels.

4.3 METHOD

4.3.1 EXPERIMENTAL MEASURES

4.3.1.1 Experimental Stimuli

The concept discrimination learning (Discrimination Learning) problems were of the type described by Levine (1966). Each problem consisted of a series of trials, requiring a choice between two stimuli on each trial (see Figure 4.1). The two stimuli varied on up to four dimensions (big-small, left-right, A-B, black-white), such that one of the two stimuli in each pair would be large, the other small, and so on. The two stimuli were

drawn on 7 x 12cm. cards, 3.5cm. apart. The large stimulus on each card was 2cm. high, and the small stimulus was 1cm. high. In view of the difficulty of the four-dimensional version of the task, a pre-training procedure of one- and two-dimensional problems was used based on that described by King and Phillips (1985). The experimental task therefore consisted of eight one- and two-dimensional pre-training problems, and eight four-dimensional experimental problems. Each of the one- and two-dimensional problems required a different set of stimuli. The four-dimensional problems were given using four sets which differed only in the letters used (A-B, X-T, O-F, S-H) (see Figure 4.1).

4.3.1.2 One and Two-Dimensional Problems

The first four problems administered varied only in one dimension (either size, position, letter or colour). For example, if a problem varied on the size dimension, then the two stimuli would be identical except that one would be big and the other small. With the first card on view, Ss were given the following instructions:

"On this card you will notice that there are two letters, and that one letter is big and one letter is small. The correct solution to this problem is either 'big' or 'small' and your task is to find out which it is. I want you to point to the one you think is correct, and I will tell you whether you are correct or wrong. Either the big letter will be correct all the time or the small letter will be correct all the time. Point to the one you think is correct. Your first choice will be a guess."

Each problem consisted of ten trials, with feedback given by the experimenter on each trial. There was a criterion of two consecutive correct responses or completion of the series. Ss were asked to verbalise the solution, and the same problem was presented again immediately if this proved to be wrong. A different dimension was used for each of the four one-dimensional problems and the solutions were as follows: small, white, X, right.

The next four problems varied on two dimensions, for example, colour and position, such that each stimulus would be either black or white, and either on the left or the right. Again, only one of the four possible solutions (e.g. either black, white, left or right) was correct. The instructions to the Ss were the same as those for the one-dimensional problems, except that there were now four possible solutions instead of two, and these were listed for the S.

The stimuli for the two-dimensional problems were designed so that after the first feedback, two of the hypotheses could be eliminated (e.g. black, right), and after the second feedback a further possibility was eliminated (e.g. white), leaving only one logically correct solution (left). Ss were not informed of this, or given any further help with performing the task. The first two problems followed the same procedure as the one-dimensional problems. On the third problem, non-feedback trials were introduced. In addition to the usual instructions, Ss were told:

"On the previous problems I told you whether you were correct or wrong after each card. This time I will not always tell you whether you're correct or wrong, after some cards I'll say nothing. Don't let that disturb you, and try to be correct on all the cards."

Feedback was given on alternate trials until Ss responded correctly on five consecutive feedback trials or completed the series of twenty trials. The fourth problem consisted of forty trials, with feedback every fifth trial, and was otherwise identical to the previous problems. The correct solutions in the two-dimensional problems were selected to complement those of the one-dimensional problems (i.e. big, T, left, black), so that having completed the pre-training problems, each of the eight possible solutions had been experienced as correct.

4.3.1.3 Four-Dimensional Problems

The four-dimensional problems varied along all four dimensions simultaneously, giving eight possible solutions, and the task was to determine which was the correct one. Ss were shown the first card, and given the following instructions:

"On this card notice that there are two different letters, one of the letters is big and one's small, one's black and one's white, and one is on the left and one's on the right. Therefore, the solution to this problem could be either A, B, big, small, black, white, left or right, and your task is to find out which one is the solution. For each card I want you to point to the one you think is correct. For some cards I'll tell you whether you're correct or wrong, and for other cards I'll say nothing. Try to be correct on all the cards. Point to the one you think is correct."

Each problem consisted of fifteen cards (see Figure 4.1), presented one at a time, and each card was removed after Ss had made a response. Feedback as to whether the chosen stimulus was correct or wrong was given every fifth trial (i.e. on the first, sixth and eleventh trials), according to a pre-determined sequence in order to control exposure to positive and negative feedback. There are eight possible sequences of correct (C) or wrong (W) feedback for each series of three feedback trials: CCC; WWW; CCW; WWC; CWW; WCC; CWC; WCW; the latter four were selected. For the four problems within each condition, the feedback sequences were presented in a different random order for each S, matched across groups.

For every possible combination of feedback and response pattern it is possible to arrive at a single correct solution; the particular solution to each problem depends upon both Ss' choices and upon the pre-determined feedback sequence. There are always four logically correct hypotheses after the first feedback, two after the second, and one after the third (the correct solution). The remaining trials were non-feedback trials which were included to permit examination of the Ss' strategies. At the end of the fifteen trials, Ss were asked to verbalise the solution. Each problem was presented only once, and Ss were not informed whether their solutions were correct.

The stimuli were constructed as described by Levine (1966). The two stimuli on each card varied on four dimensions (large-small, left-right, A-B, black-white), and they were internally orthogonal, that is, every dimension appeared exactly twice with every other dimension. Two sets of stimuli were used, with four cards in each set, and the two sets differed only in the reversal of the left-right position of the stimuli on each card. One set was used for the non-feedback trials, and the stimuli were presented in a different fixed random order for each set of four non-feedback trials. Three of the four cards from the other set were presented on the feedback trials. The use of two sets ensured that a card presented on a non-feedback trial had not previously received feedback.

Optimal performance dictates that on feedback trials the possible solutions should be retained or discarded appropriately, and an example of correct performance in solving a four-dimensional problem is shown in Figure 4.1. As a consequence of the internal orthogonality of the stimuli, there are four logically correct hypotheses after the first feedback trial, two after the second feedback trial, and one logically correct solution after

the third. This is true regardless of the Ss' response pattern or the feedback sequence. The pattern of responses on each set of four non-feedback trials revealed whether the S was testing one hypothesis consistently, or responding in a more random manner. On any set of four non-feedback trials there are sixteen possible response patterns. Eight of these reflect the consistent selection of one hypothesis, with a different pattern for each. The remaining eight patterns are indeterminate and suggestive of random or inconsistent responding. Thus for each problem, the trials were as follows:

Trial 1	First feedback trial
Trials 2-5	First non-feedback set
Trial 6	Second feedback trial
Trials 7-10	Second non-feedback set
Trial 11	Third feedback trial
Trial 12-15	Third non-feedback set

4.3.1.4 Experimental Manipulations

In addition to the standard procedure, an additional manipulation was introduced to examine the effects of reducing the memory storage load. The memory-aid condition was similar to that of Eimas (1970), in that visual feedback was given which removed the necessity to remember the information from the three feedback trials. The cards were once again presented serially, but unlike the standard condition, cards used on the three feedback trials remained on display throughout the problem; the other cards were removed as in the standard condition once Ss had made a response. On each feedback trial the word 'correct' or 'wrong' was placed next to the stimulus which had been selected. This provided a visual record of the feedback given by the experimenter, indicating which of the two stimuli had been selected by the S. Therefore, by the end of each problem in the memory-aid condition, all three feedback cards and their respective visual feedback were displayed for the S, thereby removing the need to remember the response made or the feedback received (see Figure 4.1).

Eight four-dimensional problems in all were used, four in each of the two conditions: standard and memory-aid. The conditions were presented in a fixed order, with the standard condition always given first since it was possible that the memory-aid condition might alter Ss' strategies and thus affect performance on subsequent trials.

4.3.2 CLINICAL MEASURES

The revised BDI (Beck, Rush, Shaw & Emery, 1979) was used as a measure of depression (see Appendix 1). A criterion of 5 or below was used to select nondysphoric control Ss. To select the dysphoric group, a criterion of 11 or above was used, as recommended for British Ss by Metcalfe and Goldman (1965).

As described in section 3.2.1.3, the Vocabulary sub-test of the WAIS-R (Wechsler, 1981) was chosen as a measure of intelligence. Age-scaled scores were calculated using the WAIS-R manual (Wechsler, 1981), and Ss scoring below the average range (less than 8) were not included in the study.

A brief history was taken in which Ss were asked about any sensory impairments, physical illness or injury, dyslexia and previous psychiatric illness.

4.3.3 SUBJECT SELECTION

4.3.3.1 Selection Criteria

Ss were initially selected to participate in the experiment on the basis of their score on a first administration of the BDI (BDI1) (see section 4.3.2 for criteria). Those who were selected carried out the experiment and completed the BDI again (BDI2). Ss were excluded from the study if their score was no longer within the criterion range on BDI2 (see section 4.3.2). In addition, Ss had to score in the average range or above on the Vocabulary test (see section 4.3.2.), and they were excluded if they had any significant sensory impairment, physical illness, diagnosis of dyslexia, or any history of psychiatric illness other than depression in the dysphoric group.

4.3.3.2 Selection Procedure

Ss were recruited from a pool of first- and second year university undergraduate volunteers. Initially, students were asked to fill in the BDI as part of a lecture (BDI1). They were also given an information sheet to fill in which asked for basic biographical

information, and a means of contacting them on an individual basis, either a telephone number or a contact address (see Appendix 2). To preserve confidentiality, students were not required to put their name on the BDI. The experimenter was able to identify individuals by means of a code number which matched the BDI with the information sheet. Prior to filling out the questionnaires, students were informed by the experimenter that the questionnaires were part of a research study on mood and cognitive function. They were asked to complete the BDI as honestly as they could, and warned they would be unlikely to find responses that described their feelings exactly, and should therefore select the response they felt was closest. They were told the information sheet would be used to contact a sub-set of the group in the near future, and that those approached would be asked to take part in an experiment on an individual basis. It was emphasised that the research would require Ss with a range of scores on the BDI. This was intended to reduce the likelihood of students falsifying their scores in an attempt to influence their likely inclusion in the study. It also meant that being selected for participation in the study did not give information regarding the likely score of the student. If students had been informed that the purpose was to find high- and low-scoring individuals, then selection would have implied membership of one of these groups. The students were then given the opportunity to ask questions, and these were answered by the experimenter. The BDI and information sheets were collected by the experimenter as soon as they were completed so as to minimise the opportunity for group discussion of the responses given.

The BDI was scored by the experimenter in the standard way, by summing the ratings given to each of the 21 items (Beck, Rush, Shaw & Emery, 1979). If a questionnaire had not been filled in properly then it was excluded from further consideration; this was true of only a small minority. Individuals who scored within the two criterion ranges (see section 4.3.2) were contacted by the experimenter and invited to take part in an experiment on "memory and reasoning".

4.3.3.3 Subjects

Approximately 200 volunteers initially completed BDI1. Of the 46 who scored 11 or above, 11 did not want to participate, and 35 carried out the experiment. Eight of these were excluded since they did not score in the dysphoric range on BDI2, and two others were excluded since they did not meet the selection criteria. Ninety Ss initially scored five or below, and individuals were selected from this pool until a sample size which

matched that of the dysphoric Ss was collected. Eight Ss did not want to take part, five scored above the criterion on BDI2, and two did not meet the selection criteria. The final sample consisted of twenty-five Ss in each group.

T-tests confirmed that the dysphoric group scored significantly higher than the control group on both BDI1 and BDI2. The groups did not differ significantly in vocabulary or age (see Table 4.1).

Table 4.1 Means (and standard deviations) for age, vocabulary and BDI scores, and information on sex in the two groups

	Dysphoric	**Controls**
Sex	7m, 18f	6m, 19f
Age	21.6 (3.6)	20.6 (1.5)
Vocab	12.5 (1.6)	13.0 (2.0)
BDI 1	18.8 (5.2)	1.90 (1.6)
BDI 2	17.7 (3.9)	1.6 (1.8)

4.3.4 PROCEDURE

All Ss who agreed to take part carried out the experimental task. They were given an individual appointment at their convenience. On arrival, Ss were welcomed by the experimenter and shown to the test room. They had already received some information about the study during recruitment (see section 4.3.3.2), and they were given an

information sheet to read (see Appendix 3), as required by the Ethics Committee. They then filled out a consent form (see Appendix 4), carried out the experimental and selection measures, and completed the BDI again (BDI2). In order to encourage honesty in responding, they were informed that they would not be expected to discuss their BDI responses, and the experimenter busied herself with other tasks rather than watching them. They were paid a small sum for participation. Finally, any further questions were answered by the experimenter. The most common question was the reason for selection, and this was answered (as described in section 4.3.3.2) by saying that people with a range of scores were selected. Ss often asked to be told more about the nature of the research, and this was done in general terms, without referring to specific predictions. Each S was asked not to discuss the research with other students, and in particular not to pass on any insights they might have gained in relation to the nature of the tasks and how to do them.

4.4 RESULTS

4.4.1 ONE- AND TWO-DIMENSIONAL PROBLEMS

The means and standard deviations for the performance measures on the one- and two-dimensional problems are shown in Table 4.2.

On the four one-dimensional training problems, all Ss in each group performed at ceiling level, achieving the criterion of two consecutive correct responses on each problem, and verbalising the correct solutions.

On the two-dimensional training problems, most Ss again performed at ceiling level. All the control Ss and all but three of the dysphoric Ss achieved the learning criterion on each of the four problems. All Ss, except one dysphoric S, were able to verbalise the correct solutions. The problem was repeated for this single dysphoric S, and the correct solution was achieved on the second administration.

Table 4.2 Means (and standard deviations) for performance measures on the one- and two-dimensional problems for both groups

	Dysphoric	Controls
1-dimensional problems		
Reached learning criterion /4	4.00	4.00
	(0.00)	(0.00)
Correct solutions /4	4.00	4.00
	(0.00)	(0.00)
2-dimensional problems		
Reached learning criterion /4	3.84	4.00
	(0.37)	(0.00)
Correct solutions /4	3.96	4.00
	(0.20)	(0.00)

4.4.2 FOUR-DIMENSIONAL PROBLEMS

The means and standard deviations for the performance measures in the standard and memory-aid conditions of the four-dimensional problems are shown in Table 4.3.

4.4.2.1 Correct Responding on Non-Feedback Trials

As described in section 4.3.1.3, on each set of non-feedback trials there are sixteen possible response patterns. Half correspond to the selection of a single hypothesis (one pattern for each hypothesis), while the remaining eight are indeterminate and suggest that no single hypothesis is being tested. The initial stage in scoring each problem was carried out by matching the response pattern on each set of non-feedback trials with the key in Appendix 5. This revealed which hypothesis, if any, was being tested. The next stage was to determine whether the hypotheses tested were logically correct or not. This was determined for each individual problem by combining the pre-determined feedback sequence with the S's responses on the feedback trials. The key in Appendix 6 was used

to deduce which hypotheses were still potentially correct and which were logically incorrect after each feedback trial. For every combination of feedback and S responses there are four logically correct hypotheses after feedback 1, two after feedback 2, and one logically correct solution after feedback 3. In summary, each response pattern on a non-feedback trial could be classified as either indeterminate or consistent with a hypothesis, and each hypothesis could be classified as either logically correct or incorrect.

The groups were first compared on the number of logically correct hypotheses tested on the non-feedback trials. Thus, a hypothesis was scored as correct if it was one of the four logically correct hypotheses after feedback 1, one of the two logically correct hypotheses after feedback 2, or the only remaining hypothesis after feedback 3. Since there were three sets of non-feedback trials, and the set of potentially correct hypotheses became smaller at each stage, any effect of stage of testing (after 1st, 2nd or 3rd feedback trial) was examined to see whether the groups differed in the efficiency with which they progressively ruled out incorrect hypotheses. Figure 4.2 shows the results for the two groups in the two conditions. Repeated measures ANOVA revealed that the group x stage of testing x condition interaction was not significant ($F=0.24$, $df=2,96$, $p=0.78$); nor were the two way interactions of group with stage of testing ($F=0.09$, $df=2,96$, $p=0.92$) or condition ($F=0.00$, $df=1,48$, $p=0.96$) significant. There was a significant overall effect of group ($F=10.90$, $df=1,48$, $p=0.002$), and Figure 4.2 shows that the dysphoric Ss performed below the level of the controls.

There was a significant condition by stage interaction ($F=5.73$, $df=2,96$, $p=0.004$), and this was explored by comparing the number of correct hypotheses tested after each feedback in the two conditions collapsed across groups, using a post-hoc significance level of 0.016. T-tests showed no significant difference between the conditions after feedback 1 ($t=1.74$, $df=49$, $p=0.09$), but Ss achieved significantly more correct hypotheses in the memory-aid condition than in the standard condition after both feedback 2 ($t=3.54$, $df=49$, $p=0.001$) and feedback 3 ($t=5.60$, $df=49$, $p<0.001$). The proportion of correct hypotheses declined with successive feedback trials in the standard condition, whilst in the memory-aid condition performance improved on successive trials. Since all Ss carried out the two experimental conditions in fixed order (standard followed by memory-aid), this finding might reflect either the effects of condition or a practice effect in the memory-aid condition.

There was a significant overall effect of condition ($F=21.63$, $df=1,48$, $p<0.001$), but no significant effect of stage of testing ($F=0.26$, $df=2,96$, $p=0.77$).

4.4.2.2 Incorrect Responding on Non-Feedback Trials

The analyses described in section 4.4.2.1 established that the dysphoric group made significantly fewer logically correct responses on non-feedback trials. This could indicate that dysphoric Ss made more indeterminate responses than controls, or more logically incorrect responses. Both possibilities were explored by comparing the rate of each type of response in the two groups. Separate analyses were carried out since the two response types were not independent of each other.

Figure 4.3 shows the mean number of logically incorrect responses made by dysphoric and nondysphoric Ss in the standard and memory-aid conditions. The data were found to be positively skewed and to have positive kurtosis beyond acceptable limits. An arcsine transformation failed to reduce either to an acceptable level. The analysis of the data required a test that was capable of comparing the two groups across the two experimental conditions, and therefore a repeated measures ANOVA was carried out, and compared with nonparametric analysis of the same data. ANOVA revealed no significant group x condition interaction ($F=0.15$, $df=1,48$, $p=0.70$), but there was a significant effect of group on the number of logically incorrect responses made ($F=5.86$, $df=1,48$, $p=0.02$). Figure 4.3 shows the dysphoric group made more logically incorrect responses than the control group. There was a significant effect of condition ($F=32.04$, $df=1,48$, $p<0.001$), and both groups made fewer logically incorrect responses in the memory-aid relative to the standard condition, again reflecting either facilitation or order effects.

Nonparametric analyses were also carried out. Mann-Whitney U-tests were used to compare the number of logically incorrect responses given by the two groups in the two conditions. There was a significant difference between the groups in both the standard ($U-208.0$, $p=0.04$) and memory-aid ($U=183.0$, $p=0.009$) conditions which confirmed the significant group difference found using the ANOVA. A Wilcoxon test was used to compare the number of logically incorrect responses given in the standard and memory-aid conditions collapsed across the two groups. This revealed a significant difference between the two conditions ($Z=-4.53$, $p<0.0001$), and confirmed that Ss in both groups made fewer logically incorrect responses in the memory-aid than in the standard condition.

The logically incorrect responses made by Ss on the non-feedback trials were examined to see whether the performance of the dysphoric Ss was characterised by a particular type of incorrect response. Instances where a S tested a hypothesis, received negative feedback, and then continued to test the same hypothesis were identified as reflecting a perseverative style of responding. Across the two conditions, there was a total of eight occasions when this could possibly occur for each S. This pattern was found on 4.5% of occasions in the dysphoric group, and on 2.5% of occasions in the control group. It was therefore rare in both groups, and unlikely to account for the group differences. The data were also inspected for instances when a S tested a hypothesis, received positive feedback, and then shifted to a different, incorrect hypothesis. Again, there was a total of eight instances for each S when this pattern might be observed. The pattern was found on 2.5% of occasions in the dysphoric group, while the controls did not show this pattern on any occasion. Again, although the dysphoric group made more of this type of error than the controls, it was rare, and could not account for the overall group differences.

4.4.2.3 Indeterminate Responses on Non-Feedback Trials

The evidence described in sectons 4.4.2.1 and 4.4.2.2 suggests that dysphoric Ss test fewer correct hypotheses and test more incorrect hypotheses on non-feedback trials than control Ss. The groups were also compared on the number of indeterminate responses made and the results are shown in Figure 4.4. As noted in section 4.3.1.3, indeterminate responses were those that were not consistent with a single hypothesis. Assuming that by chance, consistent response patterns would be generated on 50% (6/12) of trials, it is possible to compare the actual number of hypothesis patterns with a random responding rate. Levine (1966) reported that a group of normal undergraduate students produced indeterminate response patterns on 8% of trials during a standard administration of the task. In the current study, the figures were 26% and 25% for the dysphoric group, and 17% and 12% for the control group in the standard and memory-aid conditions respectively. These figures are clearly lower than would be expected if Ss were responding randomly. Repeated measures ANOVA showed that the proportions for the dysphoric Ss were significantly lower than those of the control group ($F=6.92$, $df=1,48$, $p=0.01$). There was no significant group x condition interaction ($F=0.59$, $df=1,48$, $p=0.45$); nor was there a significant effect of condition ($F=1.04$, $df=1,48$, $p=0.31$).

4.4.2.4 Correct Solutions

The correct solution to each problem was worked out using the procedure described in section 4.4.2.1. The mean number of correct solutions for each group in the two conditions is shown in Figure 4.5. The data were found to be positively skewed, and to have positive kurtosis beyond the accepted limit. An arcsine transformation failed to correct these. Since a test was needed to compare the number of correct solutions achieved by the two groups in the two conditions, repeated measures ANOVA was carried out in addition to nonparametric analyses. ANOVA revealed a significant group difference ($F=7.46$, $df=1,48$, $p=0.009$), and this did not interact significantly with condition ($F=0.67$, $df=1,48$, $p=0.42$). Figure 4.5 shows the dysphoric Ss reported fewer correct solutions than controls in both conditions. There was a significant main effect of condition ($F=37.50$, $df=1,48$, $p=<0.001$), and the Figure 4.5 shows the Ss achieved fewer correct solutions in the standard relative to the memory condition, again suggesting either facilitation or order effects.

Nonparametric analyses were also performed. Mann-Whitney U-tests were used to compare the number of correct solutions achieved by the two groups in the two conditions. There was no significant difference between the groups in the number of solutions achieved in the standard condition ($U=248.0$, $p=0.20$), but the dysphoric group achieved significantly fewer correction solutions relative to the control group in the memory condition ($U=194.0$, $p=0.008$). This was not consistent with the results of the ANOVA which found no significant group x condition interaction. A Wilcoxon test was used to compare the number of correct solutions achieved in the standard and memory-aid conditions collapsed across the two groups. This revealed a significant difference between the two conditions ($Z=-4.47$, $p<0.0001$) confirming that Ss in both groups achieved more correct solutions in the memory-aid than in the standard condition.

4.4.2.5 Correct Responding on Non-Feedback Trials After Positive and Negative Feedback

As outlined in section 4.1, different types of processing are necessitated by positive and negative feedback trials (Levine, 1966). Although both types of feedback should result in the elimination of hypotheses which are no longer logically consistent with the feedback, positive feedback allows the retention of the current hypothesis, while negative feedback requires the selection of a new one. In each condition (standard versus memory-

aid) there were three feedback trials for each of the four problems giving a total of twelve, half of which were negative. Figure 4.6 shows the mean number of logically correct responses on non-feedback trials immediately following positive and negtative feedback for each group in the two conditions. Repeated measures ANOVA with two within-subjects factors (condition: standard vs. memory-aid; type of feedback) was carried out. The three-way group x condition x type of feedback interaction was not significant ($F=0.07$, $df=1,48$, $p=0.80$); nor were the two way interactions of group x type of feedback ($F=0.01$, $df=1,48$, $p=0.94$) or group x condition ($F=0.14$, $df=1,48$, $p=0.71$) significant. There was a significant effect of group ($F=11.58$, $df=1,48$, $p=0.001$). The dysphoric group made fewer correct responses following both positive and negative feedback.

There was a significant condition x type of feedback interaction ($F=7.98$, $df=1,48$, $p=0.007$). Post-hoc t-tests with an adopted significance level of 0.025 showed that the number of correct responses after positive and negative feedback collapsed across groups was significantly different in both the standard ($t=5.98$, $df=49$, $p<0.001$) and memory-aid ($t=3.03$, $df=49$, $p=0.004$) conditions, with Ss testing more logically correct hypotheses after positive feedback, but this was less pronounced in the memory-aid condition, suggesting either facilitation or order effects.

There were also significant overall effects of condition ($F=28.80$, $df=1,48$, $p<0.001$) and type of feedback ($F=34.59$, $df=1,48$, $p<0.001$).

4.4.2.6 Consistency of Non-Feedback Trials With Earlier Feedback

As described in section 4.4.2.1, it was possible to determine whether each hypothesis tested on a non-feedback trial was logically correct or incorrect when the information from all the previous feedback trials was integrated. However, if Ss failed to integrate the information from the feedback trials then they might take into account just one of the feedback trials. Eimas (1970) described a measure which gives information about which of the feedback trials Ss were actually taking into account when responding on non-feedback trials. This was done by judging each hypothesis in the light of each of the preceding feedback trials taken independently. If a S was unable to retain information across trials, then a hypothesis which should have been discarded on the basis of earlier feedback might be tested because it is consistent with the most recent feedback information (0 intervening feedback trials). Alternatively, if a S failed to take into

account the information provided on successive feedback trials, then a hypothesis which is consistent with initial feedback trials (2 intervening feedback trials), but which is inconsistent with more recent feedback information might be tested. Each hypothesis tested on a non-feedback trial was judged for its consistency with feedback trials with zero, one and two intervening feedback trials, of which there were 12, 8 and 4 instances respectively in each experimental condition. The mean scores of the two groups in the two conditions are shown in Figure 4.7.

Repeated measures ANOVA with one between-subjects factor (group) and two within-subject factors (condition; 0, 1, or 2 intervening feedback trials) was carried out. The three-way interaction was not significant ($F=1.44$, $df=1.64,96$, $p=0.24$). There were no significant two-way interactions of group with condition ($F=0.00$, $df=1,48$, $p=0.97$) or group with number of intervening feedback trials ($F=0.04$, $df=1.54,96$, $p=0.92$). There was a significant overall effect of group ($F=11.42$, $df=1,48$, $p=0.001$). Figure 4.7 shows the hypotheses tested by dysphoric Ss were less likely to be consistent with feedback, regardless of the number of intervening feedback trials, than those tested by the controls.

There was no significant interaction of condition with number of intervening feedback trials ($F=2.21$, $df=1.64$, $p=0.13$). There was a significant overall effect of condition ($F=18.84$, $df=1,48$, $p<0.001$), and inspection of the means revealed that both groups tested hypotheses that were consistent with feedback more often in the memory-aid than in the standard condition. There was no significant effect of the number of intervening feedback trials ($F=0.68$, $df=1.54,96$, $p=0.47$).

4.4.2.7 Correlation of Performance with Severity of Depressive Symptomatology

Pearson correlation coefficients were calculated between each of the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2). This was carried out within each group because the selection criteria for the two groups (dysphoric >10; nondysphoric <6) ensured that, taken together, the scores on the BDI could not have a normal distribution. There were no significant results with either the BDI1 or the BDI2. Table 4.4 shows the results for the BDI2. This is the measure which was taken during the actual test session, and might therefore be expected to show a higher correlation with task performance than the BDI1 which was completed as part of the initial screening process (see section 4.3.3.2).

Table 4.3 Means (and standard deviations) for performance measures in the standard and memory-aid conditions on the four-dimensional problems

	Dysphoric		Controls	
	Standard	Memory-aid	Standard	Memory-aid
Responding on non-feedback trials:				
Correct after feedback 1 /4	1.96 (1.02)	2.16 (1.14)	2.56 (1.19)	2.92 (1.00)
Correct after feedback 2 /4	1.80 (1.22)	2.52 (1.08)	2.44 (1.33)	3.04 (1.10)
Correct after feedback 3 /4	1.68 (0.85)	2.64 (0.95)	2.36 (1.29)	3.32 (0.80)
Total incorrect /12	3.60 (1.96)	1.72 (1.46)	2.64 (2.34)	1.04 (1.84)
Arcsine	1.10 (0.43)	0.69 (0.38)	0.88 (0.51)	0.41 (0.49)
Total indeterminate /12	3.08 (2.29)	3.00 (2.04)	2.00 (2.33)	1.44 (1.64)
Correct solutions /4	2.20 (0.76)	3.20 (0.87)	2.52 (1.33)	3.76 (0.52)
Arcsine	1.69 (0.45)	2.40 (0.66)	1.94 (0.93)	2.86 (0.45)
After +ve feedback /6	3.28 (1.49)	3.92 (1.32)	4.28 (1.93)	5.00 (1.08)
After -ve feedback /6	2.16 (1.21)	3.40 (1.15)	3.08 (1.53)	4.52 (1.45)
Hs on non-feedback trials consistent with feedback:				
0 intervening %	62.67 (17.2)	68.00 (17.13)	72.00 (20.82)	84.00 (16.30)
1 intervening %	59.00 (18.23)	73.50 (18.86)	74.00 (23.36)	85.50 (15.17)
2 intervening %	59.00 (22.68)	76.00 (18.37)	73.00 (27.88)	87.00 (14.65)

Table 4.4 Correlation of BDI2[+] with performance measures in the standard and memory-aid conditions on the four-dimensional problems

	Dysphoric		Controls	
	Standard	Memory-aid	Standard	Memory-aid
Responding on non-feedback trials:				
Correct after fedback 1 /4	-.05	-.01	-.06	.13
Correct after feedback 2 /4	.16	-.07	-.08	-.05
Correct after feedback 3 /4	.04	.13	.01	.19
Total incorrect /12	.03	-.13	.27	.09
Arcsine	.06	-.21	.27	.01
Total indeterminate /12	.18	-.06	.20	.09
After +ve feedback /6	.02	-.18	.04	.10
After -ve feedback /6	.14	.22	-.16	-.09
Correct solutions /4	-.23	.04	-.08	-.23
Arcsine	-.22	.06	-.08	-.18
Hs on non-feedback trials consistent with feedback:				
0 intervening %	.09	.03	.04	-.10
1 intervening %	.10	.02	.25	-.07
2 intervening %	.07	-.05	.10	.03

*p<.01; **p<.001, two-tail.

[+]BDI2 = Beck Depression Inventory given on the second occasion (see section 4.3.3.1).

4.5 SUMMARY OF RESULTS

1. Both groups were at ceiling on the one-dimensional problems.

2. The control Ss were at ceiling and the dysphoric Ss were mildly impaired on the two-dimensional problems.

3. The dysphoric Ss made significantly fewer logically correct responses on the non-feedback trials than the controls, and this did not interact with the stage of testing or experimental condition.

4. The dysphoric Ss made significantly more logically incorrect responses on the non-feedback trials than the controls, and this did not interact with condition.

5. Dysphoric Ss made significantly more indeterminate responses than controls on the non-feedback trials, and this did not interact with condition.

6. There was a discrepancy between the parametric and nonparametric analyses of the data. Analysis of the data using parametic tests revealed that the dysphoric Ss achieved significantly fewer correct solutions relative to the control group, and this did not interact with condition. The nonparametric analysis found that dyphoric Ss achieved significantly fewer correct solutions than controls in the memory-aid but not the standard condition.

7. The dysphoric Ss tested significantly fewer logically correct hypotheses than controls after both positive and negative feedback, and this did not interact with condition.

8. The dysphoric Ss tested significantly fewer hypotheses on non-feedback trials that were consistent with earlier feedback trials relative to the control group, and this did not interact with condition.

9. There were no significant correlations between the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2).

4.6 DISCUSSION

The aim of the present experiment was to investigate the nature and extent of any reasoning deficits in a group of dysphoric Ss on a Discrimination Learning task, and to elucidate the role of storage in any deficits. It was predicted that if dysphoric Ss are impaired in their capacity to store information in WM, then an experimental manipulation which reduces the storage demands of the task should differentially facilitate the performance of the dysphoric group.

The findings showed the dysphoric Ss to be impaired relative to the controls on the four-dimensional problems, with intact performance on the simpler one-dimensional problems, and mild impairment on the two-dimensional task. This finding supports the hypothesis that depression is associated with deficits on reasoning tasks. Both groups performed better in the memory-aid condition relative to the standard condition, although this could be due to either facilitation or the effects of practice, or both (see below). The crucial finding is the lack of interaction of group with condition, showing that aiding memory failed to facilitate performance differentially for the dysphoric group.

The deficits shown by the dysphoric Ss in this experiment appeared to be relatively robust, since despite the memory-aid they still had difficulty in testing hypotheses and reaching correct solutions. This was consistent with the findings of Silberman et al. (1983) using clinically depressed Ss. The results of the one- and two-dimensional problems in the pretraining phase of the experiment showed that the dysphoric Ss had minimal impairment on these, although they were based on the same logical principle as the four-dimensional problems. This finding suggests that the increasing complexity of the task may be related to its sensitivity to depressive deficits, which were only reliably found on the four-dimensional problems. The more complex problems differed from the simpler problems in both the storage and processing demands which they made. The more complex problems required Ss to retain more information during the course of the problem, since there were more potential solutions; and the degree of processing complexity was also greater, since Ss had to manipulate and integrate more information.

The role of memory storage in carrying out the Discrimination Learning task was investigated in the present study by introducing the memory-aid condition. There is evidence that the performance of both groups was better in the memory-aid condition than

in the standard condition. The proportion of correct hypotheses tested after each feedback grew smaller on successive trials in the standard condition, while in the memory-aid condition the reverse pattern was observed. Both groups also gave more correct solutions, and made more correct responses after feedback, in the memory-aid relative to the standard condition. This may reflect a facilitation effect of providing a visual record of feedback in the memory-aid condition, although it is not possible to exclude the effects of practice as a result of the fixed order of presentation, since the memory-aid condition always followed the standard condition. This procedure was adopted because of the possibility that the memory-aid condition might alter Ss' strategies and thus affect performance on subsequent trials. Regardless of which explanation is correct, the failure to facilitate depressive performance to normal levels suggests the deficits associated with depression are unlikely to be explained solely in terms of an underlying storage deficit.

Since the deficits shown by the dysphoric Ss did not seem to be explicable in terms of the storage demands of the task, the findings were examined further. There are a number of specific performance deficits which are consistent with particular patterns of findings, and these were examined. The first step was to consider the pattern of hypothesis testing on the non-feedback trials. Dysphoric Ss were found to test fewer correct hypotheses overall. This could be the result of testing incorrect hypotheses or of making indeterminate responses. These two possibilities have different implications for understanding the nature of the performance deficit, since testing incorrect hypotheses suggests errors in eliminating hypotheses correctly, while indeterminate responses are suggestive of random or inconsistent responding. In fact, the dysphoric Ss were found to make significantly more of both types of error than controls.

The fact that dysphoric Ss tested more incorrect hypotheses on non-feedback trials suggests they were responding to feedback in a different manner to the controls. This might reflect a failure to integrate feedback from successive feedback trials. This should result in successful performance after feedback 1, at which point no integration is required, but impairment after later feedback trials. However, there was no significant interaction between group and stage of testing in the number of correct hypotheses tested, suggesting that the deficits shown by the dysphoric Ss were present throughout the problem, and could not be accounted for as a failure to integrate information on later trials. Furthermore, if the depressive deficits had reflected a failure to integrate

information, then their hypotheses should have shown greater consistency with initial relative to more recent feedback trials, but again, the findings did not support this. The performance of the dysphoric Ss therefore seems to be consistent with a more general difficulty in responding appropriately to feedback. In addition, the finding that the dysphoric Ss were less likely than controls to test a single hypothesis on the non-feedback trials raises the possibility that they were not carrying out the task in the same way as the controls, and may sometimes have guessed or used simple heuristics rather than an appropriate hypothesis-testing strategy.

The possibility that the groups differed in their ability to respond appropriately after positive and negative feedback was also considered. As described in section 4.4.2.5, the processing required after a negative feedback trial is more complex than after a positive feedback trial since it necessitates the selection of a new hypothesis, while positive feedback confirms the existing one. Dysphoric Ss were found to test significantly fewer correct hypotheses after both positive and negative feedback, but there was no significant interaction between group and type of feedback. The data were also inspected for evidence of perseverative responding after negative feedback, which has been found to be characteristic of patients with frontal lobe deficits (see section 2.2.3.1), but this was found to be extremely rare in both groups, and did not account for the group differences.

The dysphoric Ss were clearly impaired relative to the control group on every task measure, but the pattern of findings is not consistent with any of the specific sources of performance deficit considered above. This suggests the performance of the dysphoric Ss either reflected some specific performance deficit which is not apparent in the information available, or that it reflected some more general failure to carry out the task appropriately. The latter is supported by the finding that the dysphoric Ss were more likely than controls to make indeterminate responses on the non-feedback trials, suggesting that they may have been using simple, heuristic strategies to perform the task.

The discrepancy[1] between the findings of the parametric and nonparametric analyses of the number of correct solutions achieved by the groups in the two conditions needs to be considered. The two analyses were carried out because the data showed unacceptable skewness and kurtosis even after appropriate transformation, but there was no nonparametric test suitable for testing the crucial interaction between group and condition. ANOVA found no significant group by condition interaction, while there was a significant group difference in the memory-aid condition, with the control group achieving more correct solutions than the dysphoric group, but no group difference in the standard condition using Mann-Whitney U tests. Neither analysis found support for the hypothesis that dysphoric Ss would be differentially facilitated by the memory-aid manipulation. Therefore, the discrepancy does not cause a problem in deciding whether to accept or reject the null hypothesis.

The failure to find any significant correlations between performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2) also needs to be considered. This suggests either that there was no relationship between severity of depressive symptomatology (as measured by the BDI) and task performance, or that a relationship did exist, but that the measures used made it difficult to detect. Considering the first alternative, much of the work on cognitive function in depression reviewed in section 2.3 suggested that deficits in cognitive function are related to the severity of depressive symptoms. In the current experiment there was strong evidence that the dysphoric group

---

[1] The discrepancy in the results of the parametric and nonparametric analyses of the number of correct solutions achieved by the groups in the two conditions needs to be considered. The two analyses were carried out because the data showed unacceptable skewness and kurtosis even after appropriate transformation, but there was no nonparametric test suitable for testing the crucial interaction between group and condition. ANOVA found no significant group by condition interaction. However, Mann-Whitney U tests found a significant group difference in the memory-aid condition, with the control group achieving more correct solutions than the dysphoric group, but no significant group difference in the standard condition.

There are three possible explanations for the discrepancy between the parametric and nonparametric analyses, and these are considered in turn. The first possibility is that the skewness in the data reduced the accuracy of the ANOVA and that its findings are not reliable. However, Mardia (1971) cited by Tabachnick and Fidell (1983, p.232) reported that multivariate analyses are robust to a moderate violation of normality if the violation results from skewness rather than outliers, as is the case in the present experiment. Multivariate analyses are also more likely to be robust to these violations if the sample size is not small and the groups are of equal size (Tabachnick & Fidell, 1983); again the current experiment meets these criteria. It can therefore be argued that the discrepancy is unlikely to be the result of unreliability of the MANOVA. A second possibility is that since the nonparametric test is less powerful that the MANOVA, it has a greater chance of resulting in a Type 2 rather than a Type 1 error. The nonparametric test may not be powerful enough to detect a group difference in the standard condition. The third possibility is that when repeated measures ANOVA is carried out with more than one dependent variable, there are a number of simultaneous comparisons of the data, and the findings on one may obscure the findings on another. In this case, the highly significant overall effect of group may have reduced the chance of detecting a significant interaction. It is not possible to distinguish between these last two alternatives, and this raises the question of whether the use of the two parallel analyses is an appropriate strategy. The alternative is to use only one analysis, but the difficulty is choosing one when neither is entirely appropriate. Using both analyses does at least provide a check on the accuracy of the analysis, and a way of highlighting any problems.

was impaired relative to the controls on all the task measures, and therefore a relationship with the severity of depressive symptoms was predicted. With regard to the possibility that the measures may have obscured a relationship, it should be noted that correlations are more difficult to detect if the range of scores is restricted in some way. This was a particular problem for the control group, where there was a possible range of only 0-5 on the BDI, making it unlikely that any meaningful relationship with depressive symptomatology could be identified. The dysphoric group had a greater range of scores on the BDI (11-26 on BDI2), but there was a restricted range on many of the task measures. For example, the number of correct solutions had a range of 0-4 in each condition. Therefore the measures were not well-suited to identifying correlations, although it is possible that a relationship between depressive symptomatology and task performance simply did not exist.

In summary, this experiment found that dysphoric Ss showed deficits on a Discrimination Learning task, thereby supporting the hypothesis that depression is associated with deficits on reasoning tasks. The performance of the dysphoric Ss was not differentially facilitated by reducing the storage demands of the task, suggesting that depressive deficits cannot be explained simply in terms of reduced storage capacity. By simple process of elimination, this gives indirect support to the hypothesis that dysphoric Ss may be impaired in their ability to process and manipulate, rather than simply store, information. Further work is needed to explore this in a more direct fashion.

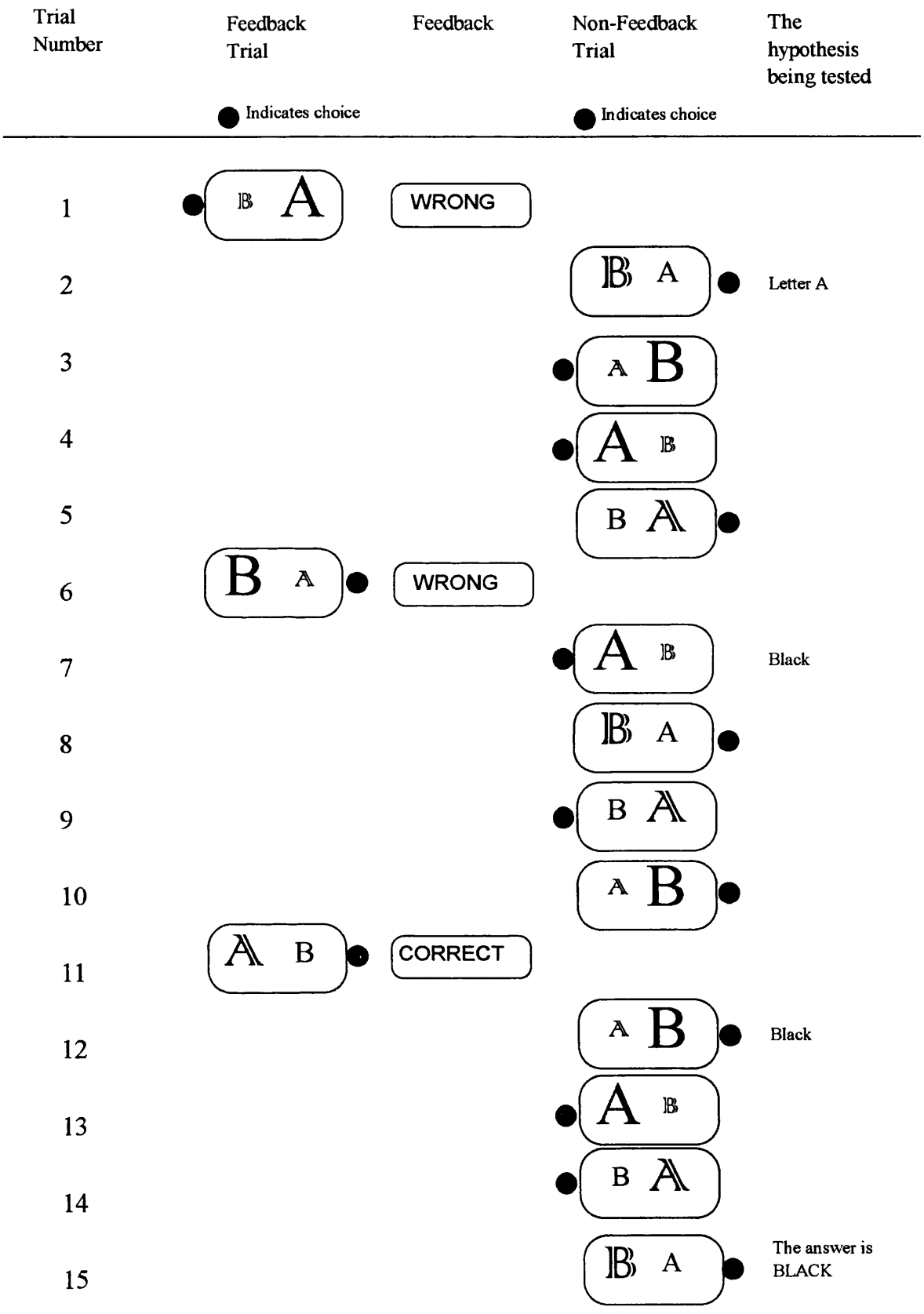

Figure 4.1 An example of a four dimensional problem showing optimal performance in the memory-aid condition.

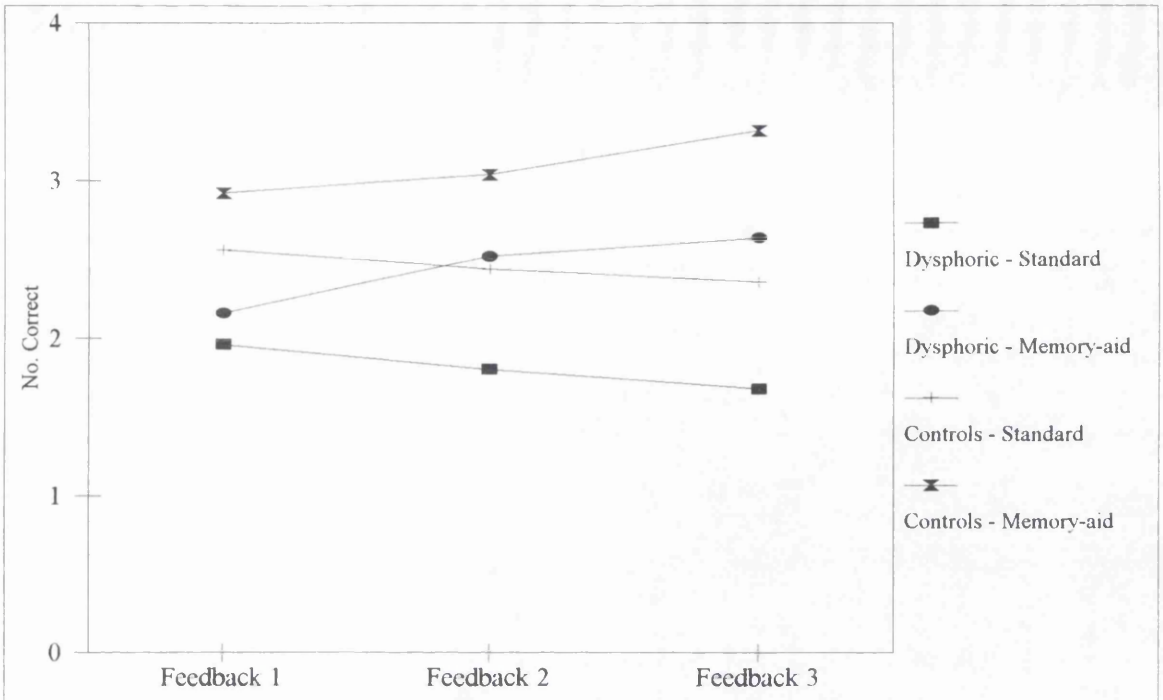

Figure 4.2 Responding on non-feedback trials - number correct /4 after feedbacks 1-3 in the standard and memory-aid conditions of the Discrimination Learning task

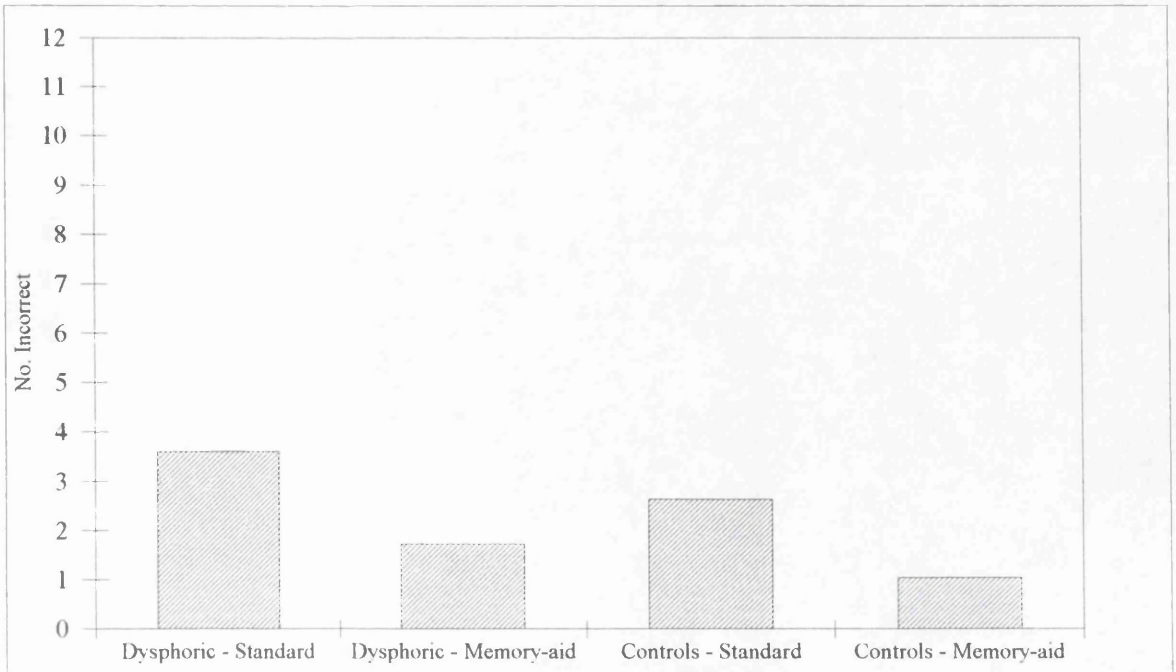

Figure 4.3 Total number of logically incorrect responses /12 in the standard and memory-aid conditions of the Discrimination Learning task

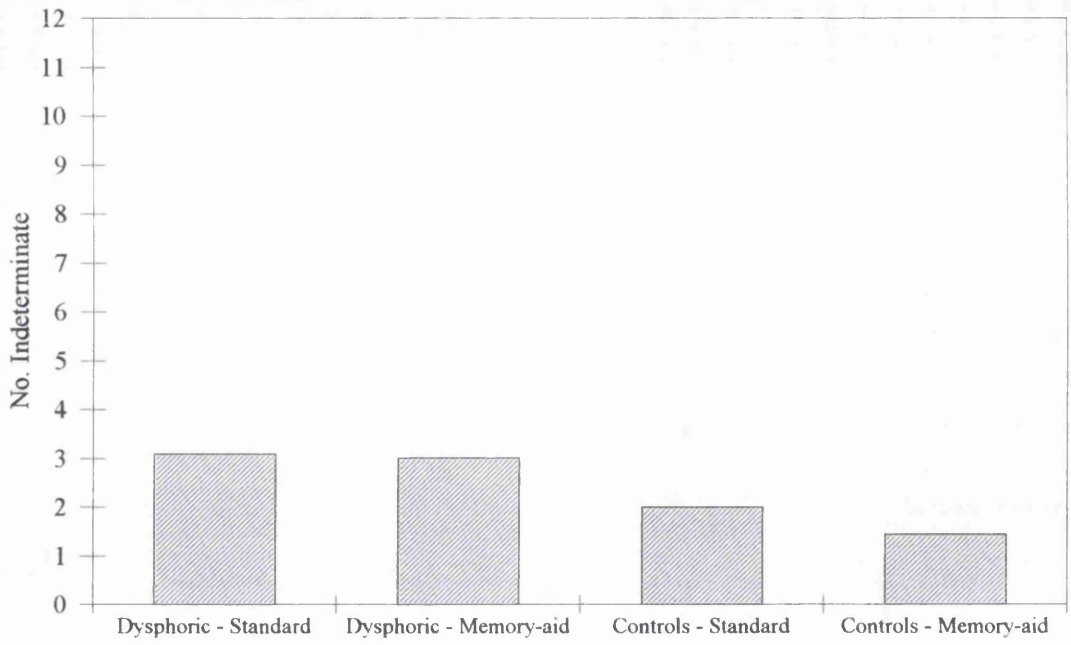

Figure 4.4 Total number of indeterminate responses /12 in the standard and memory-aid conditions of the Discrimination Learning task

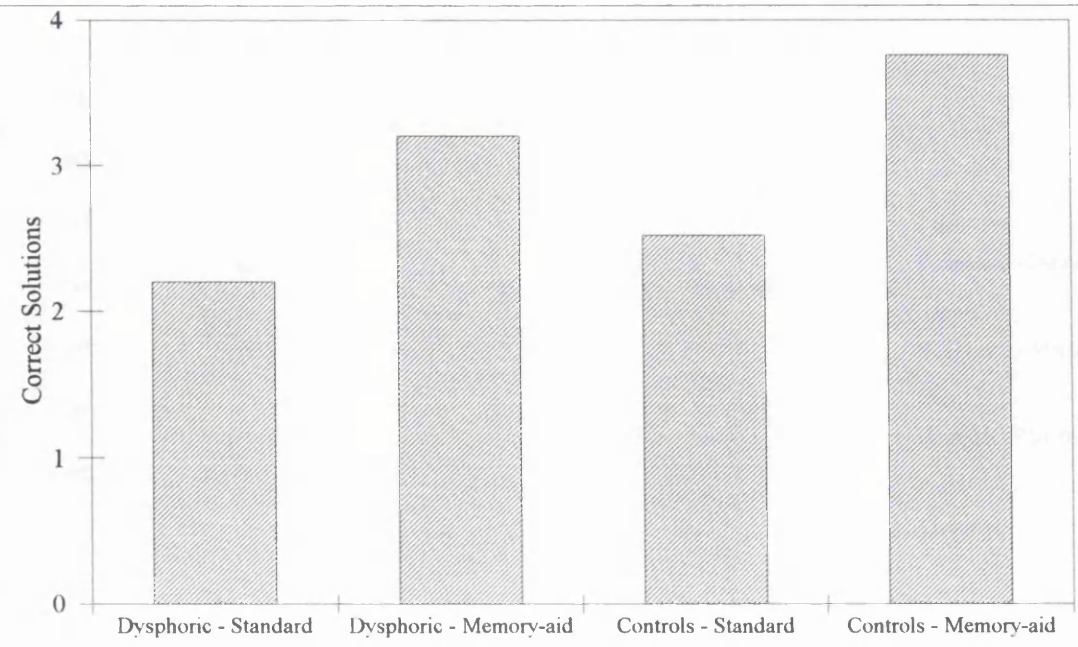

Figure 4.5 Total number of correct solutions /4 in the standard and memory-aid conditions of the Discrimination Learning task

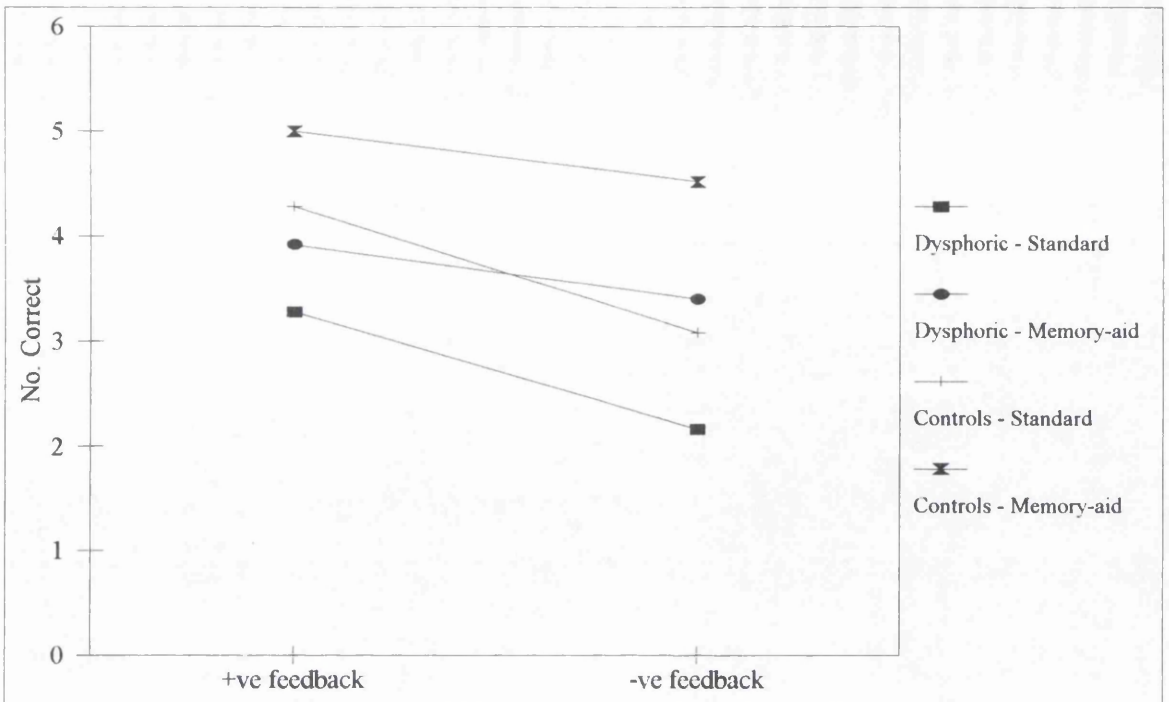

Figure 4.6 Logically correct responding /6 on non-feedback trials after positive and negative feedback in the standard and memory-aid conditions of the Discrimination Learning task

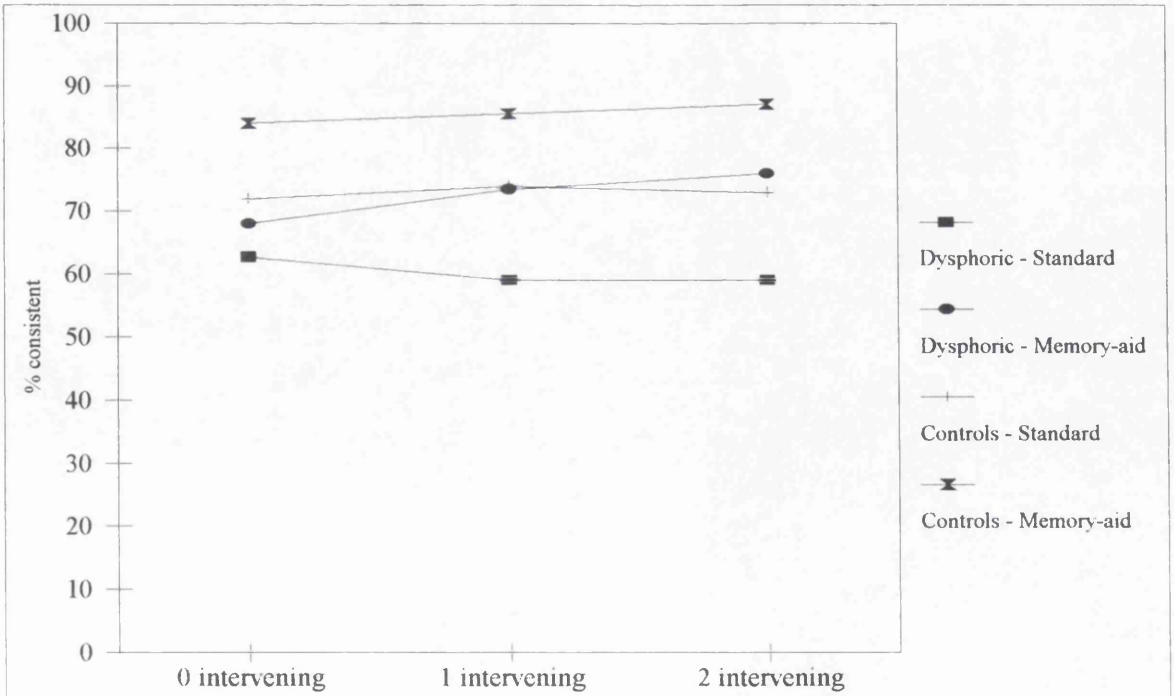

Figure 4.7 Percentage of hypotheses on non-feedback trials consistent with prior feedback in the standard and memory-aid conditions of the Discrimination Learning task

CHAPTER V

EXPERIMENT 2

5.1 INTRODUCTION

The experiment reported in chapter IV compared dysphoric and control Ss on Discrimination Learning problems which required Ss to eliminate hypotheses in response to feedback in order to identify the correct solution. Solution of the problems required both storage and processing, since Ss needed to store relevant information, and carry out manipulation and integration of the information in order to achieve the correct solution. Therefore, any performance deficits might reflect either reduced storage capacity, or reduced processing capacity, or both. The contribution of depressive storage deficits to any impairment in performance was investigated by the introduction of a memory-aid condition which reduced the storage demands of the task, and it was predicted that this manipulation would differentially facilitate the performance of the dysphoric Ss.

Overall, the findings of the experiment reported in chapter IV indicated that dysphoric Ss were impaired on the more complex Discrimination Learning problems. These made greater demands on both storage and processing capacity than the simpler problems. Reducing the storage demands of the more complex problems did not differentially facilitate dysphoric Ss. It was therefore considered important to explore further the effects of varying storage and processing demands on depressive performance. A task which would permit storage and processing demands to be varied independently was sought.

Salthouse and his colleagues (e.g. Salthouse, 1992; Salthouse, Legg, Palmon & Mitchell, 1990; Salthouse, Mitchell, Skovronek & Babcock, 1989) have reported on an integrative verbal reasoning task which, the authors argue, allows storage and processing demands to be varied independently. Ss are presented with a series of premises which each describe the relationship between two variables (e.g. A and B do the same). These are followed by a question which asks what would happen to one variable if another variable was changed in a specified way (e.g. If A increases, then what happens to B?). On some trials the information necessary to answer the question is contained in just one of the premises (one-relevant), while on others, information has to be integrated from two or more separate premises (all-relevant). Salthouse et al. (1990) suggested that optimal performance on this task involves encoding and retaining each successive premise, encoding the question, searching and retrieving relevant information from the stored

premises, integrating the information across premises when necessary, and evaluating the information in order to reach a decision. The task is therefore similar to the Discrimination Learning task described in chapter IV, and seems likely to be sensitive to depressive deficits.

Salthouse et al. (1990) argued that a comparison of performance on one-relevant trials with performance on all-relevant trials should allow a distinction to be made between storage and processing of the premises. The rationale for this argument can be summarised as follows. All trials in which only one premise is relevant should involve the same decision processes, because they are based on the same information (i.e. the relationship between two variables as presented in one premise), and differ only in the context in which the information is presented. In trials where additional irrelevant premises are presented, it may take more time to access the relevant premise, and this may reduce the probability that it can be successfully retrieved, but provided the relevant premise is available in memory, exactly the same decision process is required regardless of the number of other premises presented in that trial. Salthouse et al. (1990) concluded that therefore any variations in decision time or decision accuracy with one-relevant trials as a function of the number of premises presented could be attributed to storage processes (e.g. time to search and retrieve or failure to retain), rather than to processing limitations associated with ability to integrate the information.

Salthouse et al. (1990) went on to argue that, in contrast, all-relevant problems, where the final question refers to variables presented in different premises, necessitate integration of information, in addition to storage and retrieval of the relevant information. Therefore, on all-relevant problems, any decline in accuracy associated with an increasing number of premises could result either from a failure to retrieve the critical information from memory, or because of a failure to integrate the information from the multiple premises. Similarly, an increase in decision time with additional premises could be because there is more information to search, or because additional time is required to reorder and integrate the information (Salthouse et al., 1990).

On the basis of the above arguments, Salthouse et al (1990) suggested that the relative contributions of storage and processing or integration factors to performance in this task should be distinguishable by contrasting the effects of the number of presented premises

in trials with only one relevant premise, with trials in which all premises are relevant. All of the effects with one-relevant trials are attributed to storage factors, and thus any greater effects with all-relevant trials can be assumed to be due to the requirement to integrate information (Salthouse et al., 1990). It should be noted that this argument rests on the assumption that Ss recognise that integration of premises is redundant on one-relevant problems. Failure to recognise this could lead to fruitless attempts to integrate premises, thereby increasing the processing demands of the task. This possibility was not acknowledged by Salthouse and his colleagues.

Salthouse and his colleagues (e.g. Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990) have used the verbal integrative reasoning task to test the hypothesis that ageing is associated with a reduction in cognitive resources. These studies revealed that storage of information was an important factor in successful task performance for both younger and older Ss, with all three studies finding that decision accuracy decreased as the number of premises increased. Of particular interest was the fact that decision accuracy decreased on one-relevant trials when additional irrelevant premises were presented. This was attributed by the authors to reduced availability of information in memory. Salthouse et al. (1990) reported that decision time overall was found to increase consistently as the number of premises increased, indicating that retrieval of the relevant information may take longer when there are more premises held in memory. A consistent finding of these studies (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990) was that older adults were differentially impaired in their decision accuracy relative to younger Ss as the storage demands of the task increased (i.e. as the number of premises increased).

The number of *relevant* premises was not found to affect decision accuracy for either young or old Ss (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990), although decision time did increase significantly when more premises had to be taken into account (Salthouse et al., 1990). Salthouse et al. (1990) concluded that the integration of information from multiple premises did not cause any further loss of information, but required additional time for processing. There was no significant interaction between age and number of relevant premises. Salthouse et al. (1990) concluded that ageing-related processes influence the likelihood that information will be available, but not the success with which it can be integrated given that it is available. Some caution is necessary in accepting this conclusion since the pattern of results is also consistent with a failure of

the experimental manipulation, a possibility not acknowledged by Salthouse and his colleagues. As noted above, it is possible that fruitless attempts to integrate premises on one-relevant trials could result in heavy processing demands on these problems.

In summary, Salthouse and his colleagues (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990) claim to have developed a task that permits independent variation of storage and processing demands. However, the rationale for the task relies on certain assumptions about how the task is performed, assumptions that may or may not be justified. The pattern of findings reported by Salthouse and his colleagues do not permit any firm conclusions to be made about the capability of the task to measure storage and processing independently. However, for the purposes of the current experiment, due to the lack of available tasks capable of separating storage and processing effects, it was decided to proceed with the verbal integrative reasoning task as described above, but to remain cautious in any interpretations. Furthermore, it was decided to include a memory-aid version of the task, equivalent to the memory-aid condition of the Discrimination Learning task, to give additional information about the role of storage in task performance.

The current experiment was designed to investigate the role of storage and processing in depression using an integrative verbal reasoning (Integrative Reasoning) task similar to that described by Salthouse and his colleagues (e.g. Salthouse et al., 1989). On the basis that dysphoric Ss were found to be generally impaired on the Discrimination Learning task described in chapter IV, it was predicted that they would show deficits on the current task. In order to investigate the role of storage and processing, the number of premises, and the number of relevant premises were varied independently in the way described above. In addition, the problems were presented in two conditions which were analogous to the conditions described in the Discrimination Learning task, and designed to vary further the storage demands of the task. In one condition the premises were presented serially (standard condition), while in the other the premises were presented in parallel (memory-aid condition), and remained on display throughout the problem.

5.2 EXPERIMENTAL HYPOTHESES

If dysphoric Ss are impaired in their ability to carry out reasoning tasks then they should show impairment relative to controls in their performance on the Integrative Reasoning task.

If dysphoric Ss are impaired in their ability to store information in WM then:
i) an experimental manipulation which increases the storage load by increasing the number of premises should differentially impair the performance of the dysphoric group
ii) an experimental manipulation which reduces the storage load should differentially facilitate the performance of the dysphoric group.

If dysphoric Ss are impaired in their ability to process information in WM then:
i) an experimental manipulation which increases the processing load by increasing the number of premises relevant to the solution should differentially impair the performance of the dysphoric group;

5.3 METHOD

5.3.1 EXPERIMENTAL MEASURES

5.3.1.1 Experimental Stimuli

Ss completed a set of forty Integrative Reasoning problems of the type described by Salthouse et al. (1990). Each problem consisted of one to three premises followed by a question about the information contained in the premises. Each premise described the relationship between two variables represented by letters of the alphabet. The relationship was always either "does the opposite" or "does the same". For example:

Example 1

K and L do the opposite.

Or:

K and L do the same.

The question always asked what would happen to one letter if a specified change occurred to another. The change was always "increases" or "decreases", and the correct answer to each problem was also either "increases" or "decreases". So for example 1, the question would be either:

If K increases, what happens to L?

Or:

If K decreases, what happens to L?

Thus:

K and L do the opposite.

If K increases what happens to L?

Answer: L decreases.

Between one and three premises were presented for each problem. The problems varied in how many premises were actually relevant to the solution, but Ss were not informed of this. In some problems the relevant information was contained in just one premise (see Example 1), while in others all the premises were relevant to the answer, and the premises had to be integrated, as follows:

Example 2

L and M do the same.

K and L do the opposite.

If K decreases, what happens to M?

Answer: M increases.

With regard to the number and relevance of the premises, there were five problem types: 1-1, 2-1, 2-2, 3-1, and 3-3, where the first number refers to the number of premises presented, and the second number refers to the number of relevant premises. Thus, Example 1 would be a 1-1 problem, while Example 2 would be a 2-2 problem. For the 2-1 and 3-1 problems, only one premise was relevant to the solution, while the remaining premises contained redundant information. The amount of information intervening between the relevant premise and the question was controlled, with the relevant premise presented first in half the problems and last in the remaining problems. For the 2-2 and 3-3 problems all premises were relevant, and it was important to control the amount of storage and processing of the premises carried out prior to presentation of the question.

Example 3

A and B do the same. (i)

B and C do the opposite. (ii)

C and D do the same. (iii)

If A increases, what happens to D?

Answer: D decreases.

In Example 3, Ss could carry out integration of the information after each successive trial, thereby reducing the storage burden. To discourage this, 2-2 problems were presented as in Example 2 so that integration could not be carried out until after presentation of the second premise. The 3-3 problems were presented in one of two orders either i,iii,ii or iii,i,ii where the numbers refer to the premises shown in Example 3. This meant that, again, full integration could not proceed until presentation of all the premises.

Each S carried out eight of each problem type, and these were divided into two equivalent sets with four of each type in each set, matched for the variables described above. Within each set, the problems were presented in a fixed-random order such that no more that three successive problems had the same answer (i.e. 'increases' or 'decreases') In order to reduce interference between the problems, no two successive problems used the same letters of the alphabet. Both sets were constructed so that each premise and each question were typed on a separate card measuring 7 x 12cm. Appendix 7 contains the full set of material.

5.3.1.2 Experimental Manipulations

There were two experimental conditions: standard and memory-aid. In the standard condition, each premise was presented one at a time for four seconds, followed by the question. In the memory-aid condition, all the premises and the question were presented simultaneously. The two sets of stimuli were used an equal number of times within each condition. The order of the conditions was balanced, and matched across groups.

5.3.1.3 Practice Trials

Each experimental condition was introduced by a series of practice trials, with one practice trial of each problem type. The exact instructions varied depending on the order in which the S carried out the two experimental conditions.

5.3.1.3.1 Standard condition

The first practice problem was of the 1-1 type. For Ss who carried out the standard condition first, the following instructions were given:

"In this test you'll be presented with a number of statements, which will be followed by a question about them. Each statement describes the relationship between two letters such as: 'K and L do the opposite'. You'll then be presented with a question, such as: 'If K increases, what happens to L?' The correct answer to the problem will always be either 'increases' or 'decreases'."

For those Ss who had already carried out the memory-aid condition, the initial instructions were as follows:

"You should do the following problems in the same way, except that this time you'll see one statement at a time and then the question."

The remaining instructions and procedure were the same for Ss in both orders.

"Please read each statement carefully and then answer the question. We'll try some practice trials first. Remember, the correct answer to the problem will be either 'increases' or 'decreases'."

The first premise was presented for four seconds and was then replaced with the question. The question was left on view until the S responded, and the time from the presentation of the question until a response was given was recorded by the experimenter. If an incorrect response was given then the experimenter said: "That's the wrong answer, please look at the cards and try again." The problem was repeated as many times as necessary until the S gave the correct answer.

The next problem was at the 2-1 level (2 premises, only 1 relevant to the answer). The instructions were the same, except the S was informed there would be more than one statement.

The next practice problem was at the 2-2 level. Again, the instructions were the same, except Ss were told: "This time I want you to work as quickly as you can without making mistakes." The problem was presented in the same way as described above, but if the S gave an incorrect response, the problem was repeated once only. Ss then carried out practice problems at the 3-1 and 3-3 levels using the same procedure. Prior to the 3-1 problem, Ss were informed that there would be three premises.

5.3.1.3.2 Memory-aid condition

The practice trials in the memory-aid condition were carried out using the same procedure, except that for each problem all the premises and the question were presented simultaneously. This was achieved by holding a sheet of card in front of the test stimuli while the experimenter arranged them on the desk, and then removing the card to reveal the complete problem to the S. The experimenter timed from presentation of the stimuli until a response was given.

For Ss who carried out the memory-aid condition first, the instructions were the same as those given to Ss carrying out the standard condition first. For Ss who had already received the standard condition, the initial instructions were:

"You should do the following problems in the same way, except that this time you will see all the statements and the question at once."

5.3.1.4 Experimental Problems

In each condition, the experimental problems were presented immediately after the relevant practice trials.

5.3.1.4.1 Standard condition

Having completed the five standard practice problems, the experimenter gave the following instructions:

"I'm now going to show you the rest of the problems and you should do these in the same way. Remember to read each statement carefully and then answer the question at the end as quickly as you can without making mistakes. The correct answer will always be either 'increases' or 'decreases'."

The twenty standard problems were administered in the same way as the practice trials except that each problem was administered only once, and no feedback about performance was given.

5.3.1.4.2 Memory-aid condition

Having completed the five memory-aid practice problems, the experimenter gave the same instructions as for the standard condition. Again, the twenty memory-aid problems were administered in the same way as the practice trials, except that each problem was administered only once, and no feedback was given.

5.3.2 CLINICAL MEASURES

These were the same as those described in section 4.3.2.

5.3.3 SUBJECT SELECTION

5.3.3.1 Selection Criteria and Procedure

The selection criteria and procedure were the same as those described in section 4.3.3.1 and 4.3.3.2.

5.3.3.2 Subjects

Approximately 190 volunteers initially completed the BDI1. Of the 38 who scored 11 or above, nine did not want to take part and 29 carried out the experiment. Five of these were later excluded since they did not score in the dysphoric range on administration of the BDI2. Eighty-seven Ss initially scored five or below, and individuals were selected from this pool until a sample size which matched that of the dysphoric Ss was collected. Ten Ss did not want to take part, three scored above the criterion on the BDI2, and one did not meet the selection criteria. The final sample consisted of 24 Ss in each group .

T-tests revealed that the dysphoric group scored significantly higher than the control group on both the first and second administration of the BDI. The groups did not differ significantly in vocabulary or age (see Table 5.1).

Table 5.1 Means (and standard deviations) for age, vocabulary and BDI scores, and information on sex in the two groups

	Dysphoric	Controls
Sex	4m, 20f	6m, 18f
Age	22.4	22.6
	(4.4)	(4.4)
Vocab	12.6	12.6
	(1.5)	(2.2)
BDI 1	18.7	2.8
	(7.4)	(1.6)
BDI 2	17.3	1.9
	(6.4)	(1.7)

5.3.4 PROCEDURE

The procedure was the same as that described in section 4.3.4.

5.3.5 DATA SCORING

Within experimental conditions, the number of correct solutions for each problem type was scored out of a maximum of four correct, and the mean time taken was recorded.

5.4 RESULTS

5.4.1 NUMBER OF PREMISES

The means and standard deviations for the two groups on the performance measures for number of premises on the Integrative Reasoning problems are shown in Table 5.2.

5.4.1.1 Decision Accuracy

Salthouse et al. (1990) reported that increasing the number of premises reduced decision accuracy. In the current study, Ss carried out problems with either 1, 2 or 3 premises in two experimental conditions: standard and memory-aid. For the purpose of these analyses, the 2-1 and 2-2 problems were considered together as 2-premise problems, and the 3-1 and 3-3 problems were considered as 3-premise problems. The results are shown in Figure 5.1. Several of the variables were found to have unacceptable skewness and kurtosis, and an arcsine transformation failed to correct this. An analysis that would allow comparison of the three types of problem in the two conditions was needed. Since no such nonparametric analysis was available, repeated measures ANOVA was carried out alongside individual nonparametric analyses.

ANOVA with two within-subject factors (condition; number of premises: 1,2 or 3) found that the three-way group x condition x number of premises interaction was not statistically significant ($F=0.75$, $df=2,92$, $p=0.48$); nor were the two-way interactions of group x condition ($F=2.10$, $df=1,46$, $p=0.15$) or group x number of premises ($F=0.49$, $df=2,92$, $p=0.61$) significant. There was no overall significant effect of group ($F=0.04$, $df=1,46$, $p=0.85$) (see Figure 5.1).

There was a significant condition x number of premises interaction ($F=25.97$, $df=2,92$, $p<0.001$). Paired t-tests were carried out to explore this, with a post-hoc significance level of 0.016 (0.05/no. of tests). There was no significant difference between the two conditions in the number of correct solutions achieved for the 1-premise problems ($t=-1.66$, $df=47$, $p=0.103$), but there were significant differences between the conditions for both the 2-premise ($t=-4.89$, $df=47$, $p<0.001$) and 3-premise ($t=-8.30$, $df=47$, $p<0.001$) problems. The means showed that Ss in both conditions achieved fewer correct solutions as the number of premises increased, but this was far less pronounced in the memory-aid condition (see Figure 5.1). There were also significant overall effects of condition ($F=75.52$, $df=1,46$, $p<0.001$) and number of premises ($F=79.12$, $df=2,92$, $p<0.001$).

Nonparametric analyses were carried out. Number of correct responses on the 1-, 2- and 3-premise problems in the two conditions was compared using Mann-Whitney U tests. There were no significant group differences in the number of correct responses on any of the problem types in either condition. In the standard condition there was no significant group difference on the 1-premise ($U=276.0$, $p=0.64$), 2-premise ($U=231.0$, $p=0.23$) or 3-premise ($U=281.0$, $p=0.88$) problems. Again, in the memory-aid condition, there was no significant group difference on the 1-premise ($U=276.0$, $p=0.32$), 2-premise ($U=242.5$, $p=0.29$) or 3-premise ($U=275.0$, $p=0.77$) problems.

When the different problem types (1-, 2- or 3-premises) were compared within each of the two experimental conditions using Wilcoxon tests, there were significant differences between the 1-premise and 2-premise problems in the standard condition ($Z=-4.05$, $p=0.0001$), and also between the 1-premise and 3-premise problems ($Z=-5.58$, $p<0.0001$), and the 2-premise and 3-premise problems ($Z=-4.91$, $p<0.0001$). In each case, inspection of the means revealed that Ss achieved more correct solutions when there were fewer premises. In the memory-aid condition there was a significant difference between the 1-premise and 2-premise problems ($Z=-3.12$, $p=0.0018$), and between 1-premise and 3-premise problems ($Z=-6.03$, $p<0.0001$), and also between 2-premise and 3-premise problems ($Z=-6.03$, $p<0.0001$). Again, in each case, there were more correct solutions when the problems contained fewer premises.

Finally, Wilcoxon tests were used to compare the different problem types across the two conditions. For the 1-premise problems, there was no significant difference between the standard and memory-aid conditions ($Z=-1.47$, $p=0.14$). There were significant effects of condition on the 2-premise ($Z=-3.76$, $p=0.0002$) and 3-premise problems ($Z=-5.29$, $p<0.0001$). Inspection of the means revealed that Ss achieved more correct solutions in the memory-aid relative to the standard condition.

The nonparametric analysis was therefore consistent with the parametric analysis in finding no significant effect of group on the number of correct solutions. The findings were also consistent in finding that Ss achieved more correct solutions when the problem contained fewer premises, and in the memory-aid relative to the standard condition on problems with more premises.

5.4.1.2 Response Time

Salthouse et al. (1990) also found that response time increased with increasing numbers of premises. In the current study response time was not directly comparable in the two conditions, since it was measured from presentation of the question until a response was given in the standard condition, and from presentation of the entire problem until a response was given in the memory-aid condition. Therefore, any overall effects of condition on decision time cannot be interpreted. However, they were both included in a single analysis in order to reveal any interactions with condition.

Several variables showed unacceptable skewness and kurtosis. A log 10 transformation reduced these to acceptable levels, so parametric analysis of the data was appropriate. Repeated measures ANOVA with two within-subject factors (condition; number of premises) found the three-way group x condition x number of premises interaction was not statistically significant ($F=1.11$, $df=2,92$, $p=0.33$); nor was the two-way interaction of group x condition ($F=0.16$, $df=1,46$, $p=0.70$) significant. The group x number of premises interaction approached significance ($F=2.78$, $df=2,92$, $p=0.07$). Exploratory t-tests were carried out with a post-hoc significance level of 0.016. The groups were compared on 1-, 2- and 3-premise problems collapsed across condition. Positive skewness was reduced to an acceptable level with a log 10 transformation. The groups did not differ significantly on 1-premise ($t=0.26$, $df=46$, $p=0.79$), 2-premise ($t=-0.78$, $df=46$, $p=0.44$), or 3-premise problems ($t=-1.48$, $df=46$, $p=0.15$). The means revealed a trend for dysphoric Ss to have faster response times than the controls as the number of premises increased, as shown by Figure 5.2. There was no overall significant effect of group ($F=0.69$, $df=1,46$, $p=0.41$).

There was a significant interaction between condition and number of premises ($F=48.12$, $df=2,92$, $p<0.001$). Post-hoc t-tests carried out with a post-hoc significance level of 0.016 revealed that while there was no significant difference in response time for the 2-premise problems in the two conditions ($t=-1.69$, $df=47$, $p=0.10$), there were significant differences for the 1-premise problems ($t=4.66$, $df=47$, $p<0.001$), and for the 3-premise problems ($t=-5.60$, $df=47$, $p<0.001$). An inspection of the means revealed a crossover, with Ss faster on the 1-premise problems in the memory-aid condition, but slower on the 3-premise problems in the memory-aid condition relative to the standard condition, as illustrated by Figure 5.2. There was a significant overall effect of number of premises ($F=297.76$, $df=2,92$, $p<0.001$), but no significant effect of condition ($F=2.94$, $df=1,46$, $p=0.93$).

Table 5.2 **Means (and standard deviations) for number of premises in the standard and memory-aid conditions on the Integrative Reasoning problems**

	Dysphoric		Controls	
	Standard	Memory-aid	Standard	Memory-aid
% Correct solutions				
1-premise problems	96.88	98.96	97.92	100.00
	(8.45)	(5.10)	(7.06)	(0.00)
Arcsine	2.96	3.04	3.00	3.08
	(0.33)	(0.20)	(0.28)	(0.00)
2-premise problems	87.50	93.75	84.38	95.31
	(14.28)	(7.37)	(11.80)	(8.89)
Arcsine	2.55	2.77	2.42	2.86
	(0.48)	(0.36)	(0.43)	(0.37)
3-premise problems	68.23	93.23	66.15	92.71
	(20.85)	(9.01)	(15.85)	(12.18)
Arcsine	2.07	2.75	1.92	2.77
	(0.52)	(0.39)	(0.39)	(0.45)
Mean response time secs.				
1-premise problems	2.98	2.56	3.10	2.39
	(0.54)	(0.98)	(1.18)	(0.88)
Log 10	0.47	0.38	0.47	0.35
	(0.08)	(0.15)	(0.14)	(0.16)
2-premise problems	5.67	5.97	6.17	6.69
	(3.00)	(2.82)	(3.44)	(2.82)
Log 10	0.71	0.74	0.73	0.79
	(0.19)	(0.18)	(0.22)	(0.18)
3-premise problems	6.00	9.27	7.74	10.13
	(2.95)	(4.76)	(4.73)	(4.24)
Log 10	0.74	0.92	0.83	0.97
	(0.19)	(0.20)	(0.21)	(0.17)

5.4.1.3 Correlation of Performance with Severity of Depressive Symptomatology

Pearson correlation coefficients were calculated between each of the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2). This was carried out within each group because the selection criteria for the two groups (dysphoric >10; nondysphoric <6) ensured that, taken together, the scores on the BDI could not have a normal distribution. There were no significant correlations with either the BDI1 or the BDI2. Table 5.3 shows the results for the BDI2. This is the measure which was taken during the actual test session, and might therefore be expected to show a higher correlation with task performance than the BDI1 which was completed as part of the initial screening process (see section 4.3.3.2).

Table 5.3 Correlation of BDI2[+] with number of premises in the standard and memory-aid conditions on the Integrative Reasoning problems

	Dysphoric		Controls	
	Standard	Memory-aid	Standard	Memory-aid
% Correct solutions				
1-premise problems	.14	.21	.07	/
Arcsine	.14	.21	.07	.00
2-premise problems	-.15	-.11	.03	.18
Arcsine	-.25	-.14	.11	.13
3-premise problems	.10	-.34	-.13	.06
Arcsine	.06	-.26	-.02	.09
Mean response time secs.				
1-premise problems	-.09	-.09	.12	.25
Log 10	-.10	-.12	.15	.26
2-premise problems	-.19	-.23	.13	.21
Log 10	-.21	-.18	.18	.24
3-premise problems	-.32	-.03	.16	.26
Log 10	-.35	-.01	.20	.29

*p<.01; **p<.001, two-tail. "/" indicates that a coefficient could not be computed.
[+]BDI2 = Beck Depression Inventory given on the second occasion (see section 4.3.3.1).

5.4.2 NUMBER OF RELEVANT PREMISES

The means and standard deviations for the two groups on the performance measures for number of relevant premises on the Integrative Reasoning problems are shown in Table 5.4.

5.4.2.1 Decision Accuracy

To explore the effect of number of relevant premises, performance one-relevant problems (2-1 and 3-1) was compared with problems where all premises were relevant (2-2 and 3-3). The 1-1 problems were not included in this analysis since they do not fall clearly into either category. The results are shown in Figure 5.3. Several of the variables had skewness and kurtosis that were unacceptable, and an arcsine transformation failed to reduce these to acceptable levels. Since no appropriate nonparametric test was available, repeated measures ANOVA was carried out in addition to nonparametric tests. ANOVA with two within-subject factors (condition; number of relevant premises) found that the three-way interaction was not statistically significant ($F=0.24$, $df=1,46$, $p=0.63$); nor were the two-way interactions of group x condition ($F=1.05$, $df=1,46$, $p=0.31$), or group x number of relevant premises ($F=1.65$, $df=1,46$, $p=0.21$) significant. There was no significant main effect of group ($F=0.17$, $df=1,46$, $p=0.68$).

There was no significant interaction between condition and number of relevant premises ($F=0.22$, $df=1,46$, $p=0.64$), but there was a significant main effect of number of relevant premises ($F=38.71$, $df=1,46$, $p<0.001$), and the means revealed that in both conditions Ss got more one-relevant than all-relevant problems correct (see Figure 5.3). There was a significant effect of condition ($F=98.54$, $df=1,46$, $p<0.001$), with both groups getting more problems correct in the memory-aid relative to the standard condition (see Figure 5.3).

Nonparametric tests were also carried out. Mann-Whitney U-tests were used to test group differences on each problem type in the two conditions. The groups were not significantly different in the number of correct answers achieved on one-relevant problems in either the standard ($U=218.0$, $p=0.13$) or the memory-aid ($U=276.0$, $p=0.56$) conditions; nor were there significant group differences on all-relevant problems in the standard ($U=281.5$, $p=0.89$) or memory-aid ($U=247.0$, $p=0.36$) conditions. The effect of the number of relevant premises was assessed using Wilcoxon tests to compare performance across groups on one-relevant versus all-relevant problems. The difference was found to be

significant in both the standard ($Z=-4.15$, $p<0.0001$) and memory-aid ($Z=-3.89$, $p=0.0001$) conditions, with better performance on one-relevant problems. Finally, the effect of condition was assessed using Wilcoxon tests to compare performance across groups in the standard versus memory-aid conditions. The difference was found to be significant for both the one-relevant ($Z=-4.77$, $p<0.0001$) and all-relevant ($Z=-4.48$, $p<0.0001$) problems, with more correct solutions in the memory-aid condition in both cases. This pattern of findings was consistent with those of the parametric analysis reported above in finding no group differences, but that Ss achieved more correct solutions on one-relevant than all-relevant problems, and in the memory-aid relative to the standard condition.

5.4.2.2 Response Time

The effect of number of relevant premises on decision time was explored by comparing decision time for one-relevant (2-1 and 3-1) and all-relevant (2-2 and 3-3) problems, and the results are shown in Figure 5.4. Some of the variables showed unacceptable skewness and kurtosis, but a log 10 transformation reduced this to acceptable levels, and parametric tests were therefore appropriate. As described in section 5.4.2.1, an overall effect of condition on decision time would be difficult to interpret because of the different measures used in the two experimental conditions, but both conditions were included in the analysis so as reveal any interactions with condition. Repeated measures ANOVA with two within-subject factors (condition; number of relevant premises) showed that the three-way group x condition x number of relevant premises interaction was not significant ($F=0.04$, $df=1,46$, $p=0.83$); nor were the two-way interactions of group with condition ($F=0.14$, $df=1,46$, $p=0.71$) or group with number of relevant premises ($F=0.70$, $df=1,46$, $p=0.40$) significant. There was no significant main effect of group ($F=1.60$, $df=1,46$, $p=0.21$).

There was a significant condition x number of relevant premises interaction ($F=39.85$, $df=1,46$, $p<0.001$), and paired t-tests were carried out with a post-hoc significance level of 0.0125. There were significant differences in response time between one- and all-relevant problems in both the standard ($t=7.53$, $df=47$, $p<0.001$) and memory-aid ($t=16.48$, $df=47$, $p<0.001$) conditions. The means revealed that these differences were in different directions. On one-relevant problems, Ss had faster response times in the memory-aid condition relative to the standard condition, whereas for the all-relevant problems the reverse was true (see Figure 5.4). There were significant effects of condition ($F=10.39$, $df=1,46$, $p=0.002$), and number of relevant premises ($F=210.73$, $df=1,46$, $p<0.001$).

Table 5.4 Means (and standard deviations) for number of relevant premises in the standard and memory-aid conditions on the Integrative Reasoning problems

	Dysphoric		Controls	
	Standard	Memory-aid	Standard	Memory-aid
% Correct solutions				
One-relevant premise	88.54	99.48	82.81	98.96
	(12.18)	(2.55)	(13.70)	(3.53)
Arcsine	2.58	3.05	2.37	3.02
	(0.46)	(0.14)	(0.44)	(0.19)
All-relevant premises	67.19	87.50	67.71	89.06
	(20.13)	(14.28)	(22.70)	(18.55)
Arcsine	1.99	2.58	2.02	2.37
	(0.54)	(0.51)	(0.61)	(0.44)
Mean response time secs.				
One-relevant premise	4.28	4.26	4.86	4.53
	(1.36)	(1.71)	(2.10)	(1.60)
Log 10	0.61	0.60	0.65	0.63
	(0.12)	(0.16)	(0.16)	(0.16)
All-relevant premises	7.40	10.98	9.04	12.29
	(4.67)	(5.92)	(5.62)	(5.56)
Log 10	0.80	0.99	0.88	1.05
	(0.24)	(0.20)	(0.25)	(0.19)

5.4.2.3 Correlation of Performance with Severity of Depressive Symptomatology

Pearson correlation coefficients were calculated between each of the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2). This was carried out within each group because the selection criteria for the two groups (dysphoric >10; nondysphoric <6) ensured that, taken together, the scores on the BDI could not have a normal distribution. There were no significant correlations with either BDI1 or BDI2. Table 5.5 shows the results for BDI2. This is the measure which was taken during the actual test session, and might therefore be expected to show a higher correlation with task performance than BDI1 which was completed as part of the initial screening process (see section 4.3.3.2).

Table 5.5 Correlation of BDI2^{+} with number of relevant premises in the standard and memory-aid conditions on the Integrative Reasoning problems

	Dysphoric		Controls	
	Standard	Memory-aid	Standard	Memory-aid
% Correct solutions				
One-relevant premise	.16	.01	.24	-.40
Arcsine	-.22	.01	.25	-.40
All-relevant premises	.09	.28	-.21	.20
Arcsine	.09	-.18	-.20	.20
Mean response time secs.				
One-relevant premise	-.34	.10	.06	.40
Log 10	-.33	-.05	.08	.46
All-relevant premises	-.22	-.10	.19	.19
Log 10	-.28	-.07	.22	.20

5.5 SUMMARY OF RESULTS

1. There was no significant effect of group or interaction with group on any of the measures.

2. Ss in both groups achieved fewer correct solutions on problems with more premises, but this effect was reduced in the memory-aid relative to the standard condition.

3. Ss in both groups took longer to respond to problems with more premises. In the memory-aid condition, Ss were significantly faster in responding to 1-1 problems, but significantly slower on 3-3 problems, relative to the standard condition.

4. Ss in both groups achieved fewer correct solutions on all-relevant relative to one-relevant problems, and in the standard compared with the memory-aid condition.

5. Ss in both groups were faster on one-relevant premises and slower on the all-relevant problems in the memory-aid condition compared with the standard condition.

6. There were no significant correlations between the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2).

5.6 DISCUSSION

This experiment was designed to examine the effects of storage and processing demands on an Integrative Reasoning task which varied the number of premises, the number of relevant premises, and the need to store information. There were no significant differences between the dysphoric and control Ss on any task measures, and so the general prediction of dysphoric impairment was not supported. The failure to find any significant group differences on a reasoning task similar to other tasks, such as the Discrimination Learning task reported in chapter IV, found to be sensitive to depressive deficits was the most interesting aspect of this experiment, and one which needs to be examined in more detail. Several specific predictions were made regarding the performance of dysphoric Ss on the Integrative Reasoning task, and these will be considered in turn.

It was predicted that dysphoric Ss would be differentially impaired by the increased storage demands associated with increasing the number of premises. The findings showed that Ss in both groups generally achieved fewer correct solutions and had longer response times as the number of premises increased. This was consistent with the findings reported by Salthouse (1992) and Salthouse et al. (1989; 1990), suggesting that this task manipulation was successful. There were no significant group differences or interactions with group for decision accuracy or response time associated with number of premises. However, for response time, the interaction between group and number of premises did approach significance, and the means indicated that dysphoric Ss tended to be faster than controls as the number of premises increased. This was an unexpected finding, and will be given further consideration later in this discussion. However, there was no evidence that the dysphoric Ss were differentially impaired by increasing storage demands.

Storage demands were also manipulated in the current experiment by the introduction of a memory-aid condition, analogous to the memory-aid condition for the Discrimination Learning problems in chapter IV, in which the premises and the question were presented simultaneously, and left on view until the S made a response. This removed the necessity for Ss to store the premises prior to giving a response. It was predicted that if dysphoric Ss have reduced storage capacity then they should be differentially facilitated by the memory-aid condition. Ss in both groups achieved significantly more correct solutions on 2- and 3-premise problems in the memory-aid relative to the standard condition, while there was no difference between the conditions on the 1-premise problems. This suggests that reducing the need to store the premises in the memory-aid condition facilitated performance, and attenuated the effects of increasing the number of premises for both groups. There were also significant interactions between condition and number of premises and number of relevant premises in response time, and these will be considered below. However, again there were no significant interactions between group and condition, indicating that dysphoric Ss were not differentially facilitated by the reduced storage demands of the memory-aid condition. This is consistent with the finding that dysphoric Ss were not differentially facilitated by a similar memory-aid manipulation in the Discrimination Learning task described in chapter IV.

It was predicted that dysphoric Ss would be differentially impaired by the increased processing demands when all rather than only one premise was relevant to the answer. Ss in both groups achieved fewer correct solutions and had longer response times on all-relevant relative to one-relevant problems. This differs from the findings reported by Salthouse et al. (1990). They reported that Ss' decision accuracy did not differ significantly on one-relevant compared with all-relevant problems, although their response times were significantly longer for the all-relevant problems. The reason for this discrepancy between the current findings and those of Salthouse et al. (1990) will be considered in more detail below. Again, there were no significant group differences or interactions with group for decision accuracy or response time associated with the number of relevant premises. This suggests the requirement to integrate premises increased task difficulty for Ss in both groups, but dysphoric Ss did not appear to be differentially impaired by the increased demands on processing capacity. There was no interaction of condition with number of relevant premises for decision accuracy, suggesting that Ss' ability to integrate the premises successfully on all-relevant problems was not facilitated

by the memory-aid condition. This is consistent with the fact that the memory-aid condition reduced storage, but not processing, demands.

There were significant interactions for response times between condition and number of premises, and between condition and number of relevant premises, and inspection revealed both were cross-over interactions. Ss were faster on one-relevant problems, but slower on all-relevant problems in the memory-aid relative to the standard condition. Similarly, Ss were faster on the 1-premise problems, but slower on the 3-premise problems in the memory-aid relative to the standard condition. The fact that Ss were faster on 1-premise and one-relevant problems in the memory-aid relative to the standard condition is consistent with the predicted facilitation of performance in the memory-aid condition. However, the fact that Ss were slower on the more demanding 3-premise and all-relevant problems in the memory-aid condition is unexpected, and requires explanation.

The fact that Ss were faster on the more demanding problems in the standard condition suggests that they may have adopted a less demanding strategy for these problems. In both the standard and memory-aid conditions, the solution to each problem was always either 'increases' or 'decreases', and so there was a 50% chance of achieving the correct solution by use of a simple guessing strategy, and therefore guessing could potentially be a relatively successful strategy. An inspection of the mean scores reveals that decision accuracy on the 3-premise problems and the all-relevant problems in the standard condition was above 50%, but much lower than any of the other problem types. It is therefore possible that some Ss adopted a guessing strategy when the task demands exceeded capacity in the standard condition.

The failure to find any significant correlations between performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2) also needs to be addressed. In section 4.6 it was noted that this could suggest either that there was no relationship between severity of depressive symptomatology (as measured by the BDI) and task performance, or that a relationship did exist, but that the measures used made it difficult to detect. Considering the first alternative, in the current experiment there was no evidence that the dysphoric group was impaired relative to the controls on any measure of performance on the Integrative Reasoning task, and therefore the lack of a relationship between task performance and severity of depressive symptoms is perhaps unsurprising.

Having outlined the main findings of this experiment, there are two important points that have been identified as needing further consideration: 1) the failure to find any significant group differences or interactions with group on the Integrative Reasoning task; 2) the discrepancy with the findings of Salthouse and his colleagues in finding a significant reduction in decision accuracy associated with increasing the number of *relevant* premises in the current experiment.

With regard to the failure to find any evidence of depressive deficits on the Integrative Reasoning task, there are a number of possible explanations to be considered. One possibility is that the dysphoric students were not sufficiently depressed to show impairment on the experimental task. This explanation seems unlikely since the dysphoric Ss in this experiment had similar mean BDI scores to those of the Ss in the experiment reported in chapter IV, and also to Ss in a number of other studies where significant group differences have been reported (e.g. Channon & Baker, 1994; see section 2.3.8.1).

Another possibility is that the Integrative Reasoning task is not sensitive to the effects of dysphoric mood on cognitive function. This might result if the task was either too demanding, resulting in floor effects, or not demanding enough, resulting in ceiling effects. An inspection of the results indicates that the latter may have been the case, with Ss in both groups close to ceiling on the easier problems (those with fewer premises, fewer relevant premises, and those in the memory-aid condition). Decisions about task design were made on the basis of the findings reported by Salthouse and his colleagues (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990). The current task was very similar to that reported by Salthouse et al. (1990, Study 2). A comparison of the performance of the control group (non-dysphoric undergraduates in their 20s) in the current experiment with the control group (undergraduates in their 20s) in the Salthouse experiment (Salthouse et al., 1990, Study 2) reveals comparable decision accuracy and response time in relation to number of premises. Salthouse et al. (1990, Study 2) reported significant group differences, suggesting that having the control group perform at this level should not preclude the possibility of finding group differences. Thus, it must be concluded that while the Integrative Reasoning task might be sensitive to age-related deficits, it was not sensitive to dysphoric deficits in the current experiment.

A third possibility is that the nature of the task in some way disguised actual differences between the groups in their task performance, and the pattern of findings does offer some tentative support for this hypothesis. It was predicted that dysphoric Ss would be differentially impaired by increases in either storage or processing demands, or both. Both storage and processing demands would be greatest on the 3-premise and all-relevant problems in the standard condition, and therefore group differences would be most strongly predicted on these. However, as discussed above, there is some evidence to suggest that Ss may have adopted a guessing strategy on these problems. Since there was a 50% chance of guessing the correct solution, this was a potentially successful strategy, and could obscure group differences on these problems. There is some evidence to support this hypothesis in the finding of a near-significant interaction between group and condition for response times when the number of premises was considered. There was some evidence that dysphoric Ss were faster than controls as the number of premises increased, suggesting either that they had greater available capacity than the controls, or a greater tendency to guess as the problems exceeded capacity. Since the dysphoric Ss did not show greater decision accuracy than the controls, a finding that would suggest greater capacity, the possibility that both groups used a guessing strategy, but that dysphoric Ss were more likely to guess seems the most probable explanation.

Having considered the possible reasons for the failure to find the expected depressive deficit on the Integrative Reasoning task, the second question which needs to be addressed is why the current experiment differed from those reported by Salthouse and his colleagues (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990) in finding a significant interaction between decision accuracy and number of relevant premises. Both the current and earlier Salthouse studies reported a significant effect of number of relevant premises on response time, with Ss slower to respond as the number of relevant premises increased. This rules out a simple explanation in terms of different speed-accuracy trade-offs. Therefore, the first step is to establish whether the discrepancy reflects an actual performance difference, or whether it could be attributable to an artifact of data analysis.

In order to address the question of whether the discrepancy reflects an actual performance difference, a direct comparison was made of the relevant data in the current experiment with that reported by Salthouse et al. (1990, Study 2) and Salthouse (1992). These two experiments were selected as being the most similar to the current experiment. In the

current experiment, Figure 5.3 shows a clear separation between one-relevant and all-relevant problems, with poorer performance on the latter. The equivalent figures reported by Salthouse (1992, Figure 2) and Salthouse et al. (1990, Figure 3) for the control (young) groups show a different pattern, with little separation between the one-relevant and all-relevant problems. This suggests that the discrepancy between the findings in the current and earlier studies reflects an actual performance difference. The next step is to explore in more detail the pattern of performance in the current and earlier studies.

Comparing the results from the control group in the current experiment with the results of the control (young) group in Salthouse et al. (1992, Study 2) for both decision accuracy and response time reveals that Ss in the two experiments were equivalent in decision accuracy and response time on the one-relevant problems. There were differences associated with all-relevant problems, with Ss in the current experiment showing less decision accuracy and longer response times. This suggests that either the Ss in the current experiment were more sensitive to the increased processing demands of the all-relevant problems or that the all-relevant problems were in some way more difficult in the current experiment.

The possibility that the Ss in the current experiment differed from those in the Salthouse studies in some way that made them more sensitive to increased processing demands is considered first. This explanation is difficult to support, since both the current experiment and those carried out by Salthouse and his colleagues used Ss who were educated to degree level, and who would therefore be expected to perform at a similar level on the task. This is supported by the similarity in performance noted above between the control groups in the current experiment and in the experiment reported by Salthouse et al. (1990, Study 2) in relation to the effect of number of premises.

The possibility that the all-relevant problems in the current experiment were more difficult in some way is also difficult to support. Salthouse and his colleagues (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990) varied the Integrative Reasoning task in a number of ways, and the exact design of the current task was based on the reported outcomes associated with these differences. However, the actual parameters of the current task (e.g. presentation time, task material, number of premises) were based closely on those reported in the Salthouse studies, and did not differ in any way likely to influence

task performance. A major procedural difference did exist in that the current task was administered by the experimenter, while in the studies reported by Salthouse and his colleagues, the task was computer-administered. However, it is difficult to explain why experimenter rather than computer task administration should lead to the particular difference in task performance currently under consideration.

The evidence considered thus far suggests that the current experiment was very similar to those reported by Salthouse and his colleagues (Salthouse, 1992; Salthouse et al., 1989; Salthouse et al., 1990) in terms of both Ss and experimental task. Therefore, the current finding of a significant interaction between decision accuracy and number of relevant premises, which was not found by the earlier Salthouse studies, is rather perplexing. Since a plausible explanation cannot be extracted from the findings currently available, further study will be necessary. However, for the purposes of the current thesis, the important fact is that the dysphoric Ss were not found to be impaired relative to controls on any aspect of the Integrative Reasoning task. This may be because this task is not sensitive to depressive deficits. However, there is a possibility that ceiling effects on the easier problems (those with fewer premises, fewer relevant premises, and those in the memory-aid condition) may have reduced the sensitivity of the Integrative Reasoning task to any depressive deficit. Furthermore, the evidence is consistent with an explanation in terms of Ss in both groups adopting a guessing strategy on the more demanding problems, thus potentially obscuring any group differences.

Both the Integrative Reasoning task and the Discrimination Learning task reported in chapter IV were designed to investigate the possible role of storage and processing in depressive deficits on reasoning tasks. Taking the two experiments into consideration suggests that simply varying the storage and processing demands of an experimental task is not sufficient to elucidate the nature of any depressive deficits. Instead, the findings of both experiments highlight the role of strategy in the performance of reasoning tasks, with the suggestion that dysphoric Ss may have adopted simple, heuristic or even guessing strategies on the experimental tasks. It therefore seems important to consider task strategy more directly.

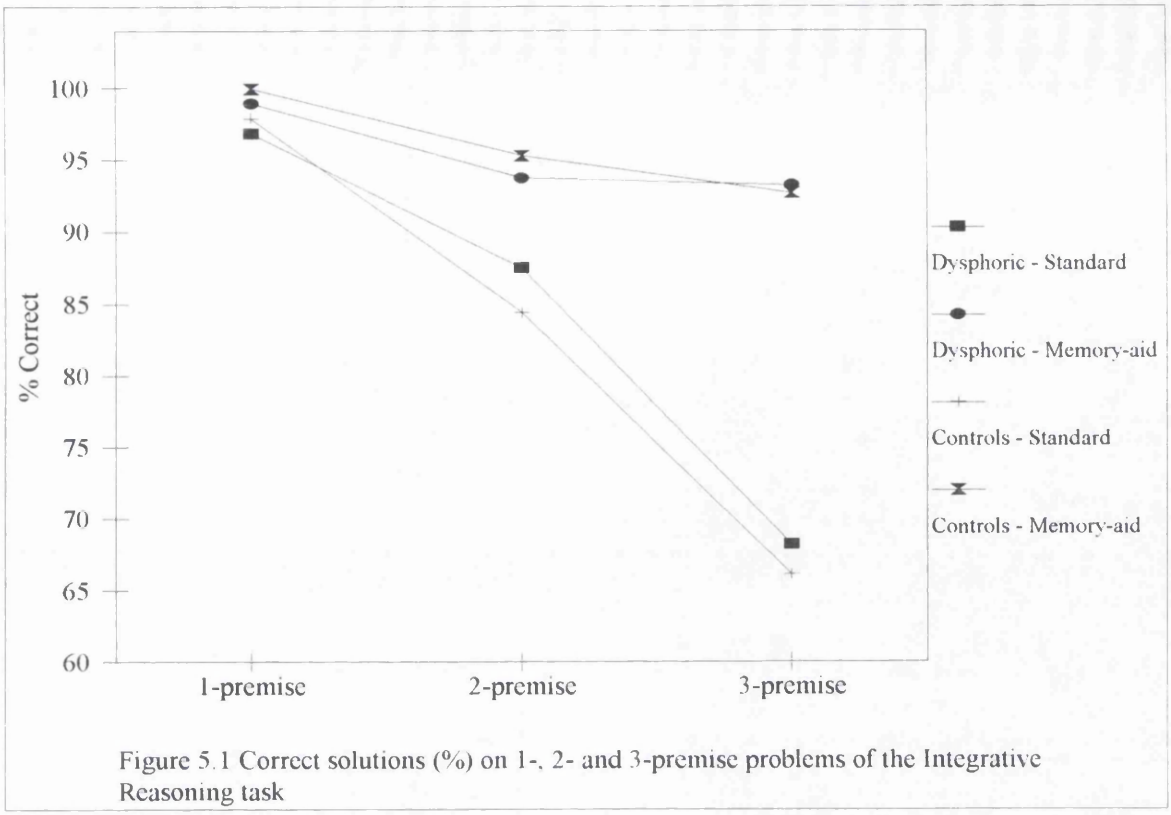

Figure 5.1 Correct solutions (%) on 1-, 2- and 3-premise problems of the Integrative Reasoning task

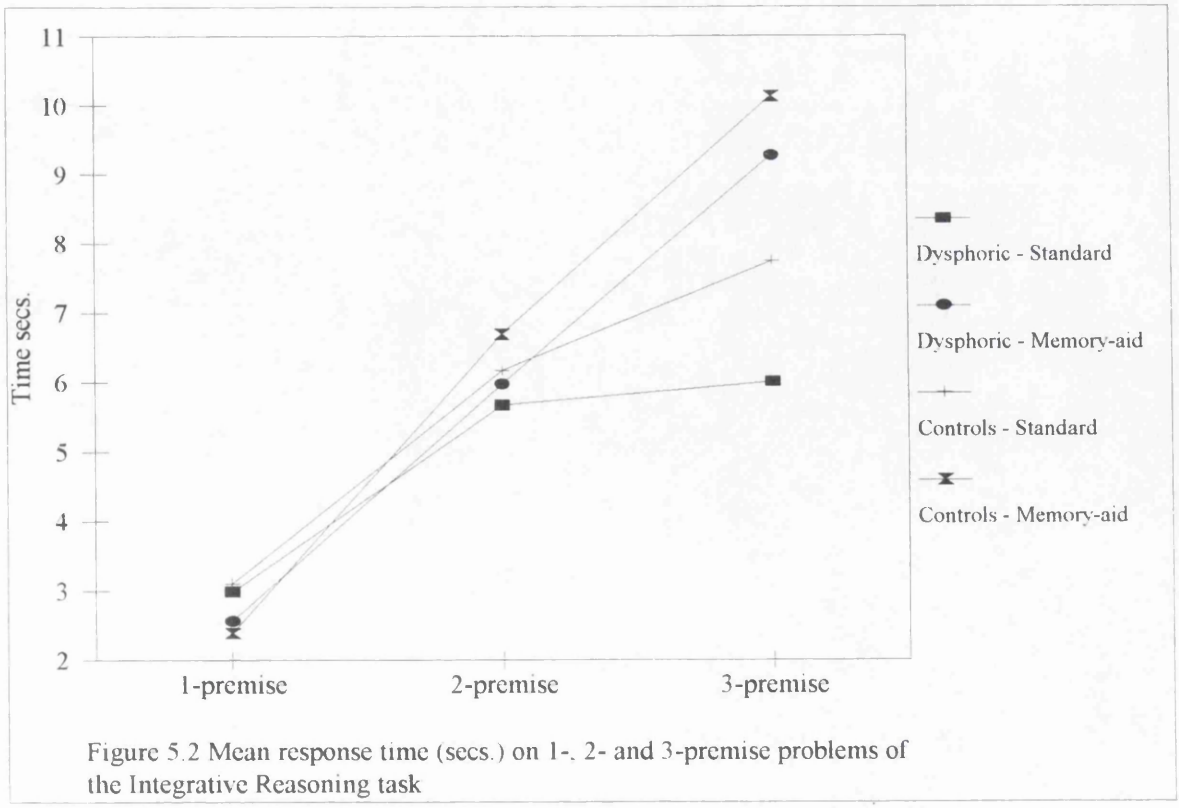

Figure 5.2 Mean response time (secs.) on 1-, 2- and 3-premise problems of the Integrative Reasoning task

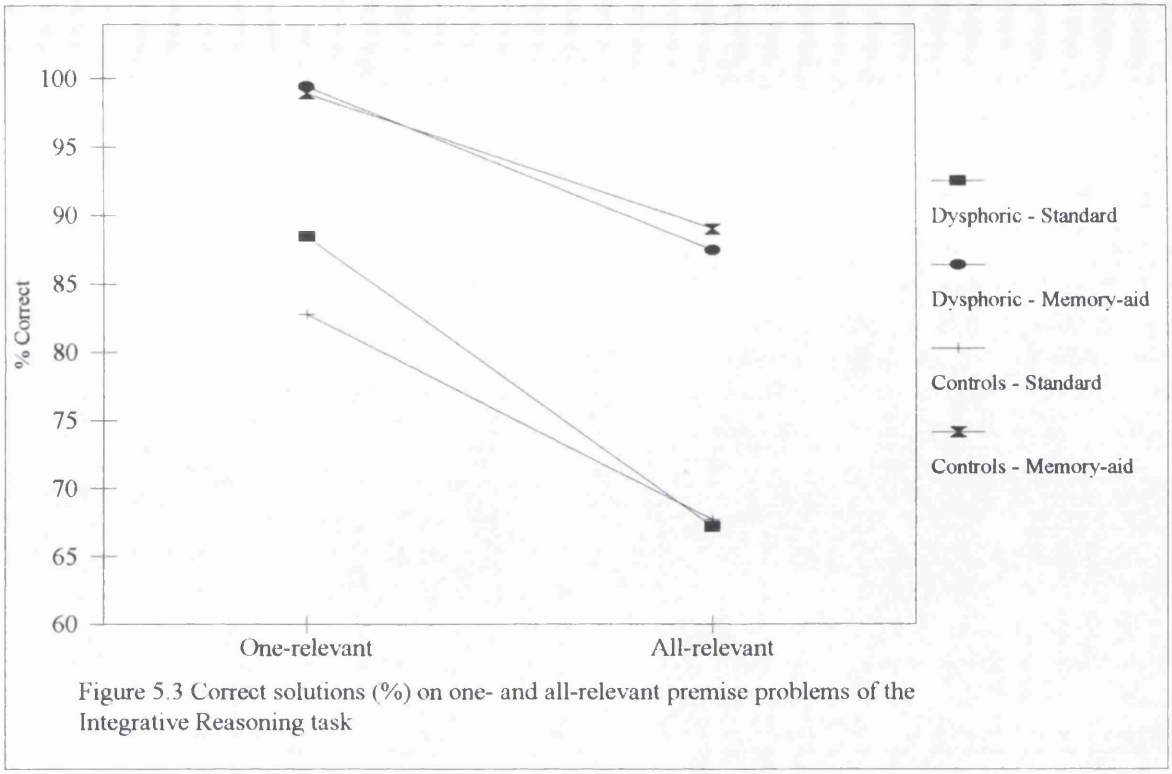

Figure 5.3 Correct solutions (%) on one- and all-relevant premise problems of the Integrative Reasoning task

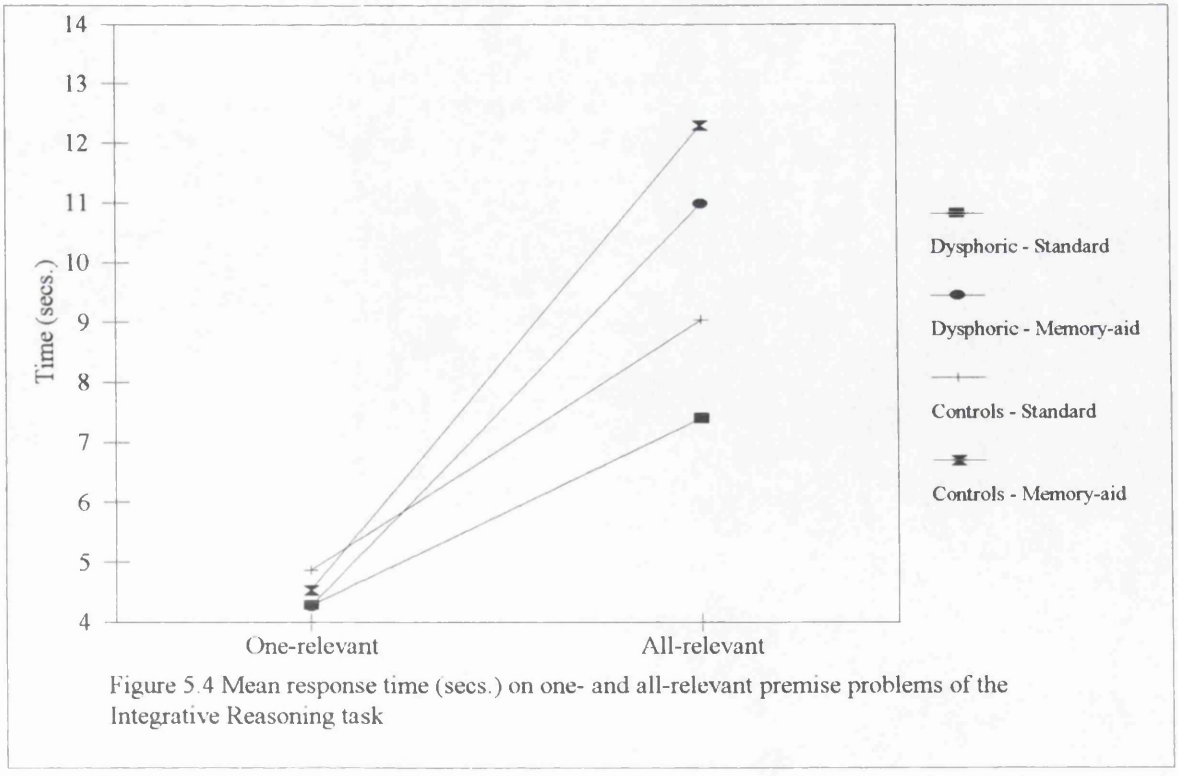

Figure 5.4 Mean response time (secs.) on one- and all-relevant premise problems of the Integrative Reasoning task

CHAPTER VI

EXPERIMENT 3

6.1 INTRODUCTION

The experiments reported in chapters IV and V compared dysphoric and control Ss on two reasoning tasks. They were designed to investigate the role of reduced storage and/or processing capacity in any depressive reasoning deficits by varying the storage and processing demands within each task. The Discrimination Learning task described in chapter IV required Ss to test and eliminate hypotheses in response to feedback in order to identify the correct solution. There was evidence that dysphoric Ss were impaired relative to controls on more complex Discrimination Learning problems. However, both the storage and processing demands of the task increased as problem complexity increased, so that depressive deficits on the more complex problems were difficult to interpret. The Integrative Reasoning task described in chapter V required Ss to draw inferences from one or more premises. This task was designed to vary storage and processing demands independently, but there were no significant group differences or interactions with group on this task.

Discussion of the findings in sections 4.6 and 5.6 suggest that the strategies adopted by Ss may play an important role in determing the pattern of findings. There was evidence to suggest that dysphoric Ss may not have used an appropriate hypothesis-testing strategy on the more complex Discrimination Learning problems. The fact that dysphoric Ss were impaired on every task measure was consistent with a general failure to carry out the task appropriately rather than a specific deficit. Furthermore, they were found to make more responses that suggested random or inconsistent responding, and also to test logically incorrect hypotheses more frequently than controls, suggesting the use of simple, heuristic strategies. On the Integrative Reasoning task, there was evidence that Ss in both groups may have adopted a guessing strategy on the more demanding problems. However, there was a trend, which did not reach significance, for dysphoric Ss to perform faster than controls as the number of premises increased, suggesting either that dysphoric Ss had greater available cognitive capacity than controls, an hypothesis which is not consistent with the bulk of evidence, or that they tended to use a guessing strategy more than controls. It was therefore decided to examine further the strategies used by dysphoric and control Ss whilst carrying out a reasoning task.

Both the Discrimination Learning and Integrative Reasoning tasks used a reception paradigm, that is, the experimenter determined what information the S received, thereby ensuring every S received all the information necessary to achieve the correct solution. This paradigm has the advantage that since all Ss have received all the relevant information, any errors can be attributed to a failure to use the information appropriately. However, it can also be valuable to see what information Ss would choose to have available for problem solution. This can be investigated using a selection paradigm in which Ss determine the information they are given.

The present study was designed to explore further the pattern of performance shown by dysphoric Ss in chapters IV and V. A Fault-Diagnosis task (Rouse, 1978) was selected which has been developed and widely used to study the real-life situations faced by people working with complex systems, such as aircraft pilots or industrial operators, who may have to locate a fault within a system. Ss are required to test potentially faulty units and use the feedback to identify the faulty one. The order of testing and the number of tests carried out before giving a final solution are determined by the S. This task is therefore similar to the Discrimination Learning and Integrative Reasoning tasks in that Ss are required to complete the task by storing and integrating information. In particular, it shares similarities with the Discrimination Learning task in that Ss must eliminate possible solutions successively in response to feedback. However, it differs from these two tasks in that the Fault Diagnosis task is presented in a selection rather than a reception paradigm, giving additional information about Ss' strategies. In particular, the design permitted examination of Ss' ability to use information available at the outset of each problem to identify possible solutions, and then to collect relevant information and eliminate possible solutions appropriately.

As noted above, the performance of dysphoric Ss on the Discrimination Learning task suggested that depressive deficits might be related to the complexity of the task. Although all the problems were based on the same logical principle, the dysphoric Ss showed minimal impairment on the simpler one- and two-dimensional problems, but marked impairment on the more complex four-dimensional problems. In the current study a series of pre-training problems of increasing complexity was devised to introduce the task components in a gradual way. This should also give useful information about the relationship between task complexity and sensitivity to depressive deficits.

The Fault Diagnosis task has two key components: the collection of relevant information; and the need to draw inferences. Thus, in the standard version of the task (selection condition), it was necessary for Ss to collect information and test hypotheses in order to identify the single faulty unit. It was predicted that dysphoric Ss would differ from controls in the strategies which they adopted. For example, a guessing strategy should result in giving solutions (correct or incorrect) without collecting adequate information. In order to assess Ss' ability to carry out both the collection of relevant information and the testing of hypotheses, a condition was included in which Ss were required only to draw inferences from the information available at the problem outset, without having to collect further relevant information. This condition was based on the reception paradigm, and was analogous to the Discrimination Learning and Integrative Reasoning tasks.

On the basis of the dysphoric deficits reported on the Discrimination Learning problems in chapter IV, it is predicted that dysphoric Ss will also be impaired on a Fault Diagnosis task which shares a number of similarities with the Discrimination Learning task. Pretraining trials of increasing complexity were included because the performance of dysphoric Ss on the Discrimination Learning task appeared to be sensitive to task complexity. It was predicted that dysphoric Ss would be unimpaired on the simplest pre-training trials, but would show deficits as complexity increased. The Fault Diagnosis problems were presented in two conditions which reflected reception and selection paradigms in order to permit examination of Ss' strategies and to identify which aspects of performance might be associated with depressive deficits.

6.2 EXPERIMENTAL HYPOTHESES

If dysphoric Ss are impaired in their ability to carry out reasoning tasks then they should show impairment relative to controls on a Fault Diagnosis task.

If dysphoric Ss are impaired in their ability to process information in WM then:

(i) they should show impairment relative to controls in a condition where they have to deduce which are the potentially faulty units;

(ii) they should be impaired in their ability to collect information and test hypotheses efficiently relative to controls;

(iii) they should show differentially greater impairment on more complex problems relative to control Ss.

If dysphoric Ss are more likely than controls to adopt simple heuristic or guessing strategies on complex reasoning tasks then they should show a different pattern of collecting information and testing hypotheses.

6.3 METHOD

6.3.1 EXPERIMENTAL MEASURES

6.3.1.1 Experimental Stimuli

This task was based on the fault-diagnosis task developed by Rouse (1978), and later modified by Duncan and colleagues (e.g. Brooke and Duncan, 1981; Morrison and Duncan, 1988). Each problem consisted of a network of 24 boxes or units in a 4x6 matrix connected in a wiring diagram (see Figure 6.1). For each problem, one of the units was faulty, and the aim of the task was to diagnose the faulty unit. Each individual unit had inputs from between one and three other units, and also had outputs to between one and three other units. Ss were told that unidirectional signals passed through the network from left to right, and ended at four indicators, each of which was either receiving a signal or not. Each unit in the network ultimately fed into between one and four of these indicators. At the outset of the problem, the only information available was the status of the four indicators, which each showed 'Yes' or 'No' to indicate whether they were receiving a signal or not. Ss were informed that an indicator could only receive a signal if all the units which fed into it, either directly or via other units, were in working order. If a unit was faulty then the signal could not pass through it, and this was propagated to all subsequent units, and resulted in one or more indicators which did not receive a signal.

The information available at the outset of each problem did not permit the deduction of the actual faulty unit, but only of the subset of units which was potentially faulty. The potential faulty units could be logically deduced to be those connected to all the 'No' indicators, and not connected to any 'Yes' indicators. In order to identify the actual faulty unit, the S had to test the connections of potentially faulty units until sufficient information was obtained. Feedback was given by the experimenter for each connection tested as to whether a signal was present or not. The faulty unit had a signal present in all of the connections entering the unit, but in none of those leaving it.

6.3.1.2 Pre-training Problems

Ss were initially shown single units with one input and a single output connection to an indicator, one with a 'Yes' signal and one with a 'No', to demonstrate how the status of the indicator reflected whether the unit was working. They were then shown two diagrams, each of four units, each with its own indicator, showing only one unit in each diagram to be faulty, again identifiable by the 'Yes'/'No' status of the indicators.

Ss were then introduced to the principles of the task using two problems at each of four levels in order of increasing complexity. Ss were initially asked to identify all the potentially faulty units on the basis of the status of the indicators, and then to discover which unit was the faulty one. They were instructed that they could test the connections between units by asking whether a signal was travelling from one unit to another, and that they should continue to do this until they thought they knew the answer. The first level diagrams consisted of six units connected to each other in a horizontal line, with a single indicator at the end of the line. This was designed to illustrate that if this indicator was not receiving a signal, then any of the units in the line could potentially be the faulty one. The second level diagrams introduced four parallel horizontal lines of six units, with an indicator for each line, and this introduced the concept of multiple indicators. The third level diagrams again showed four parallel horizontal lines, and this time these included connections crossing from one line to another, always in a left-right direction. Thus, the indicators were no longer necessarily connected to all the units in a single straight line, but had to be traced back along the crossing connections. Finally, the fourth level diagrams included units with more than one input or output, so that now some units were connected to more than one indicator. This therefore had to be taken into account when deciding which units were potentially faulty.

For each of these problems, Ss had to reach criterion before proceeding to the next problem. If Ss responded with only a subset of the potential faulty units at the beginning of each problem, they were told that they had not given the complete answer; if the answer included a unit which could not account for the pattern of the indicators, then they were told that their answer was wrong. This was repeated until the S identified the correct set of units. If they then gave the wrong answer when asked to test connections and identify the actual faulty unit, then they were told it was incorrect and that they should carry on asking questions. Again, this procedure was repeated until the S

identified the faulty unit. Ss were required to give four correct responses at each level, and a score of one was given for each correct or incorrect responses; therefore a S who gave a correct response each time would receive an optimal score of four at each level.

6.3.1.3 Experimental Conditions

After carrying out the pre-training problems, Ss then carried out a further six problems, three in each of two conditions. In one 'reception' condition, Ss were asked to deduce all the potentially faulty units from the information available in the wiring diagram alone. Both the units listed and the time to solution were recorded. In the other 'selection' condition, Ss were asked to test connections until they deduced which of the units was faulty. The connections tested, the solution given, and also the total time taken to solution were recorded. The two conditions were presented in counterbalanced order within each group, and Ss were not given feedback as to whether their answers were correct or wrong; each problem was presented only once.

6.3.1.4 Experimental Problems

The same three fault diagnosis diagrams were used for the three problems in both conditions, except that the pattern of indicators was different in each condition (see Figure 6.1 for an example). The diagrams were more complex than those used in the pre-training problems in terms of the number of connections between units. The difficulty of the problems was matched in terms of the number of potentially faulty units, so that in both conditions one diagram had three such units, one had five and one had eight, and these were associated with different diagrams in the two conditions.

Optimal performance in the reception condition was defined as giving the correct set of possible answers without any errors. Errors were recorded as the number of units omitted and the number of false positive units listed. Performance in the selection condition was assessed in terms of the tests made and the final solution given. Test errors were classified as omissions, false positives, or repetitions. Omission errors occurred when the S gave a solution without carrying out all the tests necessary to determine this. False positive errors were tests which were unnecessary because the answer could be deduced without carrying them out. The number of tests consisting of direct repetitions of earlier tests was recorded. Correct solutions were classified as error-free if the S reached the solution without carrying out any incorrect tests and collected sufficient information (i,e,

carried out all the necessary tests). Correct solutions which were not error-free were discarded unless Ss had carried out sufficient tests to reach a solution, rather than guessed.

In order to be certain whether a unit was faulty or not it was necessary to test all the inputs and the output. Once these tests had been carried out for the actual faulty unit then the S had enough information to stop testing. It was therefore possible to judge at what point in their sequence of tests Ss had actually collected sufficient information, and to score the number of redundant tests carried out beyond this point. This measure was only taken on problems where the correct solution was reached, since if a S failed to give the correct solution then it was unsurprising if they carried out incorrect tests.

6.3.2 CLINICAL MEASURES

The clinical measures were the same as those described in section 4.3.2.

6.3.3 SUBJECT SELECTION

6.3.3.1 Selection Criteria and Procedure

The selection criteria and procedure were the same as those described in section 4.3.3.1 and 4.3.3.2.

6.3.3.2 Subjects

Approximately 180 volunteers initially completed the BDI. Of the 42 who scored 11 or above, nine did not want to participate, and 33 carried out the experiment. Seven of these were later excluded since they did not score in the dysphoric range on the second administration of the BDI, and two others were excluded because they did not meet the selection criteria. Eighty-two Ss initially scored five or below, and individuals were randomly selected from this pool until a sample size which matched that of the dysphoric Ss was collected. Nine Ss did not want to take part, six scored above the criterion on the second administration of the BDI, and three did not meet the selection criteria. The final sample consisted of 24 Ss in each group.

T-tests confirmed that the dysphoric group scored significantly higher than the control group on both the first and second administrations of the BDI. The groups did not differ significantly in vocabulary or age (see Table 6.1).

Table 6.1 Means (and standard deviations) for age, vocabulary and BDI scores, and information on sex in the two groups

	Dysphoric	Controls
Sex	6m, 18f	7m, 17f
Age	21.4 (3.5)	20.7 (3.0)
Vocab	12.2 (2.0)	12.9 (1.6)
BDI 1	23.1 (6.5)	2.0 (1.9)
BDI 2	19.6 (7.3)	1.6 (1.7)

6.3.4 PROCEDURE

The procedure was the same as that described in section 4.3.4.

6.4 RESULTS

6.4.1 PRE-TRAINING PROBLEMS

Means and standard deviations for the performance measures on the pre-training problems are shown in Table 6.2 and these are illustrated in Figure 6.2.

Performance on the pre-training problems was measured by counting the number of trials required to reach criterion on each of the four problems at each of the four levels. Therefore, optimal performance at each level was reflected by a score of 4, while a score greater than this indicated that Ss had made errors and required extra trials to achieve the correct solution. Unacceptable skewness and kurtosis was detected in some of these variables. Two outliers were found and replaced in the level 3 data and one in the level

4 data, but this did not reduce skewness to acceptable levels. A log 10 transformation reduced skewness to acceptable levels in some, but not all of the variables, and a stronger transformation did not provide a better solution. Since a nonparametric was not available to compare the two groups on the four levels, the results were analysed using repeated measures ANOVA, and also nonparametric tests.

ANOVA was carried out with one between-subjects factor (group) and one within-subjects factor (level of problem). A Greenhouse-Geisser correction was applied to the group by level interaction, and this was not statistically significant (df=1.98, 138, p=0.92). There was a significant main effect of group (F=5.00, df=1,46, p=0.03), and the means indicated that the dysphoric Ss took more trials to reach criterion than the controls. There was a significant effect of level, which was also affected by sphericity. The effect remained significant after a Greenhouse-Geisser correction was applied (F=2.45, df=1.93,138, p<0.001). An inspection of the means showed that both groups took more trials at levels 1 and 4, than at 2 and 3 (see Figure 6.2).

Nonparametric analyses were also carried out. Mann-Whitney U-tests were carried out to compare the two groups at each level. The groups did not differ significantly at level 1 (U=261.0, p=0.56), or level 2 (U=263.0, p=0.28). The differences approached significance at level 3 (U=252.0, p=0.08) and level 4 (U=207.0, p=0.09), with dysphoric Ss taking more trials to reach criterion on these. The nonparametric analysis was therefore consistent with the parametric analysis in finding some evidence that the dysphoric Ss took more trials, although these did not reach significance in the nonparametric analysis.

Wilcoxon tests were used to compare the number of trials taken to achieve the correct solution by both groups at the different levels. These indicated that Ss took significantly more trials at level 1 than at level 2 (Z=-4.17, p<0.0001) or level 3 (Z=-4.35, p<0.0001), but fewer than at level 4 (Z=-2.30, p=0.02). There was no significant difference between levels 2 and 3 (Z=-0.85, p=0.40), and when compared with level 4, Ss took fewer trials at both Level 2 (Z=-4.48, P<0.0001) and level 3 (Z=-4.90, p<0.0001). This confirms the finding of the ANOVA that Ss took more trials at levels 1 and 4 than 2 and 3, and also suggests that Ss took more trials at level 4 than level 1.

Table 6.2 **Means (and standard deviations) for the number of trials needed to reach criterion on the four levels of the pre-training trials in the two groups (optimal = 4)**

	Dysphoric	Controls
No. of trials to criterion (optimal = 4)		
Level 1	5.67	5.04
	(2.24)	(1.20)
Log 10	0.81	0.78
	(0.14)	(0.09)
Level 2	4.33	4.04
	(1.01)	(0.20)
Log 10	0.73	0.70
	(0.09)	(0.02)
Level 3	4.13	4.00
	(0.34)	(0.00)
Log 10	0.71	0.70
	(0.03)	(0.00)
Level 4	7.75	5.46
	(4.18)	(1.72)
Log 10	0.90	0.80
	(0.19)	(0.11)

6.4.2 EXPERIMENTAL CONDITIONS

Since the measures used in the reception and selection conditions were not directly comparable, the results from the two conditions were considered separately.

6.4.2.1 Reception Condition

The means and standard deviations for performance measures in the reception condition are shown in Table 6.3.

A t-test was carried out to compare the groups on the time taken to give a response to the three problems in this condition. There was no significant difference between the groups on this measure ($t=1.37$, $df=46$, $p=0.18$).

The groups were then compared on the number of correct solutions achieved. As described in section 6.3.1.3, a correct solution was defined as listing all the possible faulty units, with no errors. Positive skewness was found in the data for this measure. An arcsine transformation was applied to the data, and this corrected skewness to an acceptable level. A t-test revealed no significant difference between the groups in the number of correct solutions given ($t=-1.63$, $df=46$, $p=0.11$), and inspection of the means revealed that both groups had a very low rate of success (see Figure 6.3).

The groups were compared for the number of potentially faulty units incorrectly omitted from those listed by Ss. Positive skewness was found in this variable, but a log 10 transformation of the data brought this within acceptable limits. There was no significant difference in the number of potentially faulty units omitted by the two groups ($t=0.83$, $df=46$, $p=0.41$) (see Figure 6.4).

The total number of units incorrectly listed by subjects as potentially faulty was compared for the two groups (see Fogire 6.4). As described in section 6.3.1.1, potentially faulty units were those that were connected to all the 'No' indicators, and none of the 'Yes' indicators. The false positive errors were first examined to see whether Ss had understood the principle that a faulty unit must be connected to a 'No' indicator. There were no errors involving a unit which was connected only to 'Yes' indicators, suggesting that Ss were able to apply this principle of the task.

The false positive errors were then inspected to see whether Ss had understood the principle that connection to a 'Yes' indicator eliminated a unit from being potentially faulty. It was found that Ss in both groups listed units that were connected to 'Yes' indicator(s) in addition to being connected to 'No' indicator(s). There was also evidence that Ss had not understood the principle that a unit must be connected to all the 'No' indicators in order to be a possible candidate. Again, Ss in both groups had listed units that were connected to only a subset of the 'No' indicators. The number of each error type was recorded. It was found that the data were positively skewed, but a square root

transformation brought this within acceptable limits. Repeated measures ANOVA was carried out to compare the groups on the total number of each error-type. There was no significant overall effect of group ($F=0.67$, $df=1,46$, $p=0.42$); nor a significant group x error type interaction ($F=0.64$, $df=1,46$, $p=0.43$). There was a significant effect of error type ($F=6.43$, $df=1,46$, $p=0.02$), and an inspection of the means revealed that Ss made more errors in listing units that were connected to a subset of 'No' indicators than listing units that were connected to a 'Yes' indicator (see Figure 6.5).

The pattern of false positive errors suggested that Ss understood the general principle that in order to be faulty, a unit must be connected to a 'No' indicator, but they seemed less able to apply the principles that faulty units were connected only to 'No' indicators, and to all the 'No' indicators. This raised the possibility that Ss were using simpler heuristic strategies rather than applying the more complex logical principles of the task. On the basis that Ss seemed to select units that were connected to 'No' indicators, two possible heuristic strategies seemed likely. One possibility was that Ss simply listed all the units that were connected to a 'No' indicator. The second possibility was that following from the pre-training trials in which units were connected in straight lines from left to right, Ss might list only those units that were in the same horizontal line as a 'No' indicator. For each of the three problems, the units that would be selected on the basis of the two strategies were identified, and the responses given by each S were compared with these. It was found that 15.3% of the solutions given by the dysphoric group and 11.1% of the solutions given by the control group were consistent with a strategy of listing all the units that were connected to a 'No' indicator. For the dysphoric group, 31.9%, and for the control group, 30.6% of responses were consistent with a strategy of listing those units that were connected to a 'No' indicator in a direct horizontal line. The number of solutions that were correct accounted for 11.1% of the solutions given by the dysphoric group, and 23.6% of those given by the control group. The remaining solutions (dysphoric: 41.7%; control: 34.7%) probably reflected either failed attempts to apply one of the above strategies, or the use of more idiosyncratic heuristic strategies that are not easily inferred from the data.

Table 6.3 Means (and standard deviations) for performance measures in the reception condition for the two groups

	Dysphoric	Controls
Total time taken (secs)	170.83	143.13
	(74.50)	(65.57)
Correct solutions /3	0.33	0.71
	(0.57)	(0.95)
Arcsine	0 .43	0.81
	(0.60)	(0.96)
Total omission errors	3.17	1.86
	(3.62)	(1.80)
Log 10	0.46	0.39
	(0.38)	(0.25)
Total false positive errors	12.33	11.00
	(7.78)	(9.18)
False positive errors connected to 'Yes' and 'No' indicators	4.58	3.42
	(4.29)	(4.64)
Square root	1.79	1.31
	(1.2)	(1.33)
False positive errors connected to a subset of 'No' indicators	7.58	7.58
	(6.98)	(7.05)
Square root	2.26	2.21
	(1.61)	(1.68)

6.4.2.2 Selection Condition

Means and standard deviations for performance measures in the selection condition are shown in Table 6.4.

A t-test was carried out to compare the groups on the time taken to achieve a solution for each of the problems. There was a significant difference in time to solution ($t=2.49$, $df=46$, $p=0.016$), and inspection of the means revealed that the dysphoric Ss took longer than the control Ss.

Unlike the reception condition, where solution required the listing of the potentially faulty units, performance in this condition required Ss to identify the actual faulty unit by testing connections between the units until they could deduce the answer. Performance could therefore be measured both in terms of the tests carried out by Ss, and by the solution given. Test-errors were classified as either false positives, omissions or repetitions. False positive tests were those which elicited information which could be deduced without carrying out the test. Omission errors were failures to test adequately the unit given as the final solution. Repetition errors were simple repeats of earlier tests. T-tests were used to compare the groups on the number of false positive, omission and repetition test-errors made (see Figure 6.6). Positive skewness was detected in the three variables. A square root transformation corrected skewness in the false positive and omission error variables, and t-tests showed there was no significant difference between the groups in the number of false positive ($t=1.51$, $df=46$, $p=0.14$), or omission ($t=-0.32$, $df=46$, $p=0.75$) errors made. The positive skewness in the repetition error variable was not corrected by any of the transformations attempted, and was therefore analysed using a nonparametric test. A Mann-Whitney U-test revealed a significant group difference ($U=209.5$, $p=0.04$), with dysphoric Ss making more repetition errors (see Figure 6.6).

The groups were compared for the number of correct solutions achieved without making test-errors, and for the number of correct solutions achieved which were not error-free (see Figure 6.7). There was no significant difference between the groups in the number of error-free solutions ($t=-0.75$, $df=46$, $p=0.46$), nor in the number of correct solutions which were not error-free ($t=-0.16$, $df=46$, $p=0.87$). Correct answers were common for both groups, but these were rarely error-free for either group.

Finally, on problems where a correct solution was achieved, the two groups were compared on the number of redundant tests performed when sufficient information had already been collected to deduce the correct answer (see Figure 6.6). The data showed unacceptable positive skewness and kurtosis, but a log 10 transformation brought both within acceptable limits. A t-test revealed a significant difference between the groups in the number of redundant tests carried out ($t=2.23$, $df=46$, $p=0.03$), showing that the dysphoric Ss carried out more than the controls. As the analysis above had already indicated that dysphoric Ss carry out more repetitive tests than controls, the redundant tests were inspected to see whether they were simply repetitions of earlier tests, or

whether they elicited new, albeit redundant, information. For the dysphoric group the proportion comprising repetitions of earlier tests was 23%, and this figure was 14% for the control group. The remaining tests elicited new information (77% for the dysphoric group, 86% for the control group).

Table 6.4 Means (and standard deviations) for performance measures in the selection condition for the two groups

	Dysphoric	Controls
Total time taken (secs)	145.58	114.29
	(50.52)	(35.17)
Total false positive tests	9.04	6.25
	(7.26)	(5.05)
Square root	2.78	2.31
	(1.17)	(0.98)
Total tests omitted	0.63	0.79
	(0.77)	(1.06)
Square root	0.53	0.59
	(0.60)	(0.68)
Total test repetition errors	0.75	0.21
	(1.07)	(0.51)
Error free solutions /3	0.63	0.79
	(0.77)	0.78
Correct solutions /3	2.25	2.29
	(0.90)	(0.86)
Total redundant tests	2.67	0.58
	(4.14)	(0.83)
Log 10	0.36	0.15
	(0.41)	(0.20)

Table 6.5 Correlation of BDI2[+] with performance measures on the Fault Diagnosis problems

	Dysphoric	Controls
No. of trials to criterion (optimal = 4)		
Level 1	-.11	-.14
Log 10	-.09	-.14
Level 2	.04	-.20
Log 10	.06	-.20
Level 3	.22	/
Log 10	.22	/
Level 4	-.16	-.19
Log 10	-.11	-.19
Reception Condition		
Total time taken (secs)	-.07	-.06
Correct solutions /3	-.06	-.08
Arcsine	.03	-.10
Total omission errors	.03	.05
Log 10	.15	.01
Total false +ve errors	-.03	.08
False +ve errors connected to 'Yes' & 'No' indicators	-.08	-.20
Square root	.09	-.17
False +ve errors connected to a subset of 'No' indicators	-.08	.23
Square root	-.11	.15
Selection Condition		
Total time taken (secs)	-.20	.26
Total false positive tests	-.16	-.14
Square root	-.20	-.24
Total tests omitted	-.29	-.17
Square root	-.31	-.17
Total test repetition errors	-.02	-.05
Error free solutions /3	.55*	.26
Correct solutions /3	.29	.15
Total redundant tests	-.10	.12
Log 10	-.03	.08

*p<.01; **p<.001, two-tail. "/" indicates that a coefficient could not be computed.
[+]BDI2 = Beck Depression Inventory given on the second occasion (see section 4.3.3.1).

6.4.3 Correlation of Performance with Severity of Depressive Symptomatology

Pearson correlation coefficients were calculated between each of the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2). This was carried out within each group because the selection criteria for the two groups (dysphoric >10; nondysphoric <6) ensured that, taken together, the scores on the BDI could not have a normal distribution. There were no significant correlations with the BDI1 and only one significant correlation with the BDI2: a positive correlation between score on the BDI2 and number of error-free solutions in the selection condition in the dysphoric group. Table 6.5 shows the results for the BDI2. This is the measure which was taken during the actual test session, and might therefore be expected to show a higher correlation with task performance than the BDI1 which was completed as part of the initial screening process (see section 4.3.3.2).

6.5 SUMMARY OF RESULTS

1. On the pretraining problems, the dysphoric Ss took more trials to reach criterion, but this did not interact significantly with the degree of complexity. All Ss tended to take more trials on levels 1 and 4 than on levels 2 and 3.

2. There were no significant group differences in performance in the reception condition, and both groups had a low rate of success on these problems.

3. The dysphoric group took significantly more time to complete the problems in the selection condition. There was no significant difference between the groups in the number of correct solutions, or in the rate of false positive or omission errors. The dysphoric group made significantly more repetition test errors, and they also carried out significantly more redundant tests.

4. There were no significant correlations between performance measures and the BDI1, and only one significant correlation between the performance measures and the BDI2.

6.6 DISCUSSION

The aim of the current experiment was to examine reasoning performance in dysphoric and control Ss on a Fault Diagnosis task in two conditions. In the reception condition Ss

were presented with information and had to draw inferences in a way which was analogous to the Discrimination Learning and Integrative Reasoning tasks reported in chapters IV and V. In the selection condition, Ss determined what information they collected and which hypotheses were tested before giving the solution, thus permitting more detailed examination of Ss' strategies. In addition, pretraining trials of increasing complexity were included to investigate further the effects of task complexity on depressive deficits.

On the basis of the findings reported in chapter IV of dysphoric deficits on more complex Discrimination Learning problems, it was predicted the dysphoric Ss would show impairment on the Fault Diagnosis task. On the pretraining problems, the dysphoric Ss took significantly more trials to reach criterion than the controls, although the predicted interaction with task complexity was not found to be significant. The experimental problems were presented in two conditions: reception and selection. In the reception paradigm, Ss were required to deduce which of the units were potentially faulty. Both groups of Ss performed poorly in using the information available at the outset of the problem to identify the potentially faulty units, and there was no significant difference between the groups on any of the experimental measures in this condition. In the selection condition, Ss were required to collect information in order to discover which unit was the faulty one. The dysphoric group was impaired relative to the control group on the experimental problems in this condition. There was no significant difference between the groups in the number of correct solutions achieved or in the number of false positive or omission test-errors made. However, the dysphoric Ss took significantly longer than the controls to reach a solution, made significantly more repetition errors, and also carried out significantly more redundant tests. Therefore, there is support for the prediction that dysphoric Ss would show deficits on the Fault Diagnosis task.

Unlike the Discrimination Learning task reported in chapter IV, there was no clear relationship between depressive deficits and task complexity on the Fault Diagnosis task. It is of interest that dysphoric Ss took more trials than the control Ss overall on the pretraining problems, since these problems were relatively simple. This indicates that dysphoric Ss can show impairment in carrying out simple logical operations which make relatively low demands on WM capacity.

The results for the pretraining trials were also analysed using nonparametric tests, and this showed a slightly different pattern in that there were no significant group differences, but the difference approached significance on the level 3 and 4 problems. As discussed in section 4.6, a discrepancy between parametric and nonparametric analyses is likely to reflect the lower power of the nonparametric analysis, and/or the possibility that the multiple comparisons in the ANOVA may reveal some differences at the expense of others. In the current experiment it seems most likely that the nonparametric tests lacked the necessary power to detect the group differences revealed by the ANOVA.

Both the parametric and nonparametric analyses of the pretraining trials found Ss in both groups took more trials at levels 1 and 4 than 2 and 3. While it was predicted that the more complex level 4 problems should take more trials than levels 2 and 3, it was anticipated that the simplest level 1 should take fewer trials. The finding that Ss took more trials at level 1 may be due to the initial novelty of the task rather than the complexity of the processing.

It was predicted that dysphoric Ss would be impaired relative to controls in their ability to draw inferences from information presented in a reception paradigm. This prediction was made on the basis that this condition was similar to the Discrimination Learning task in which Ss received all the information necessary to achieve the correct solution. In fact, there were no group differences on any of the measures in the reception condition. This failure to find group differences may be due to floor effects, since both groups of Ss had low success rates in identifying the possible faulty units correctly from the initial wiring diagrams. An examination of the error patterns sheds some light on this. Ss clearly found it difficult to integrate information from units which had multiple inputs or outputs. Their choices appeared to reflect a simplified strategy which involved listing all the units which were most obviously feeding into a 'No' indicator, regardless of whether or not the units were also connected to a 'Yes' indicator. For instance, Figure 6.1 shows an example in which five of the units were possibly faulty: units 11, 21, 22, 32, 33. Typical errors involved listing units 12 and 13, despite their ultimate connection with a 'Yes' indicator. The other difficulty centred around taking into account the need for faulty units to connect to all the 'No' indicators, where the diagram showed more than one. In Figure 6.1, typical errors of this nature involved listing units 24 to 26, 35 and 36, and 41 to 46, each of which were connected to only one or two of the three 'No' indicators.

Ss also carried out Fault Diagnosis problems in a selection paradigm in which they determined what information was made available to them, thus permitting an examination of their strategies. It was predicted that dysphoric Ss would show deficits in their ability to collect information and test hypotheses in this condition of the task. However, the errors made by Ss in listing the potentially faulty units in the reception condition were also reflected in their testing strategies in the selection condition. Their performance contained relatively high rates of false positive tests, i.e. testing units which could not be faulty, and the groups did not differ in this aspect of performance, again, possibly due to floor effects. The groups did appear to differ in the use they made of feedback given by the experimenter in response to the tests carried out by Ss. Dysphoric Ss were significantly more likely than controls to perform redundant tests, i.e. to carry on testing units beyond the point when they had collected sufficient information to be able to deduce the faulty unit correctly. One possible explanation for this would be in terms of storage deficits, since Ss had to retain the information collected from feedback. There was a significant difference between the groups in the number of repetition errors made overall, i.e. the total number of times in which the same connection was tested more than once, lending some support to an interpretation of this kind. However, the majority of redundant tests did not consist of repetition errors, but functioned to elicit new information, suggesting that the dysphoric Ss were still considering more than one potential solution at the point when only one viable solution remained.

It was predicted that dysphoric Ss would differ from controls in the strategy adopted in the selection condition. The dysphoric Ss took longer and made more redundant tests, but still achieved the correct solution as often as controls, consistent with an explanation in terms of a conservative response style which Johnson and Magaro (1987) suggested as a possible explanation for depressive deficits (see section 2.3.9.2). As discussed in section 2.3.4.3, studies have investigated response style in depressed Ss in recognition memory paradigms, with mixed findings. Thus, studies have reported evidence of a conservative response criterion rather than a true memory deficit (e.g. Miller & Lewis, 1977), while others have found the converse (e.g. Channon, Baker & Robertson, 1993); a more liberal response criterion has also been reported (Deptula, Manevitz & Yozawitz, 1991). In the current study, the fact that the majority of redundant tests made by dysphoric Ss elicited new information rather than repeating earlier tests is suggestive of a performance deficit, although it is possible that a conservative response bias may have contributed to this.

The results of the correlation coefficients calculated between the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2) need to be considered. Only one correlation reached significance, and this was a positive correlation between number of error-free solutions in the selection condition and the BDI2 in the dysphoric group. This correlation therefore goes against the experimental hypothesis since it suggests that greater depression was associated with better task performance on this measure. However, it should be noted that when this quantity of correlation coefficients is calculated then one or more would be expected to reach significance by chance, and therefore the implications of this single result will not be considered further.

In section 4.6 it was concluded that a failure to find significant correlations between task performance and depressive symptomatology (as measured by the BDI) could suggest either that there was no relationship between severity of depressive symptomatology and task performance, or that a relationship did exist, but that the measures used made it difficult to detect. In the current experiment there was evidence that the dysphoric group was impaired relative to the controls on the pretraining problems and on the problems in the selection condition, and therefore a relationship with the severity of depressive symptoms was predicted. It was noted in section 4.6 that correlations are more difficult to detect if the range of scores is restricted in some way. This is a particular problem for the control group, where there was a possible range of only 0-5 on the BDI, making it unlikely that any meaningful relationship with depressive symptomatology could be identified. The dysphoric group had a greater range of scores on the BDI (11-28 on BDI2), but there was a restricted range on many of the task measures. For example, the number of correct solutions had a range of 0-3 in each condition. It is possible the measures were not well-suited to identifying correlations, although it may be that a relationship between depressive symptomatology and task performance did not exist.

In summary, the dysphoric Ss were found to be impaired on the simple pretraining problems, and to show a different pattern of responding in the selection condition, making more repetition and redundant tests than controls. The groups did not differ in their ability to draw inferences in the reception paradigm, but the low rate of success in both groups suggests that this may have resulted from floor effects.

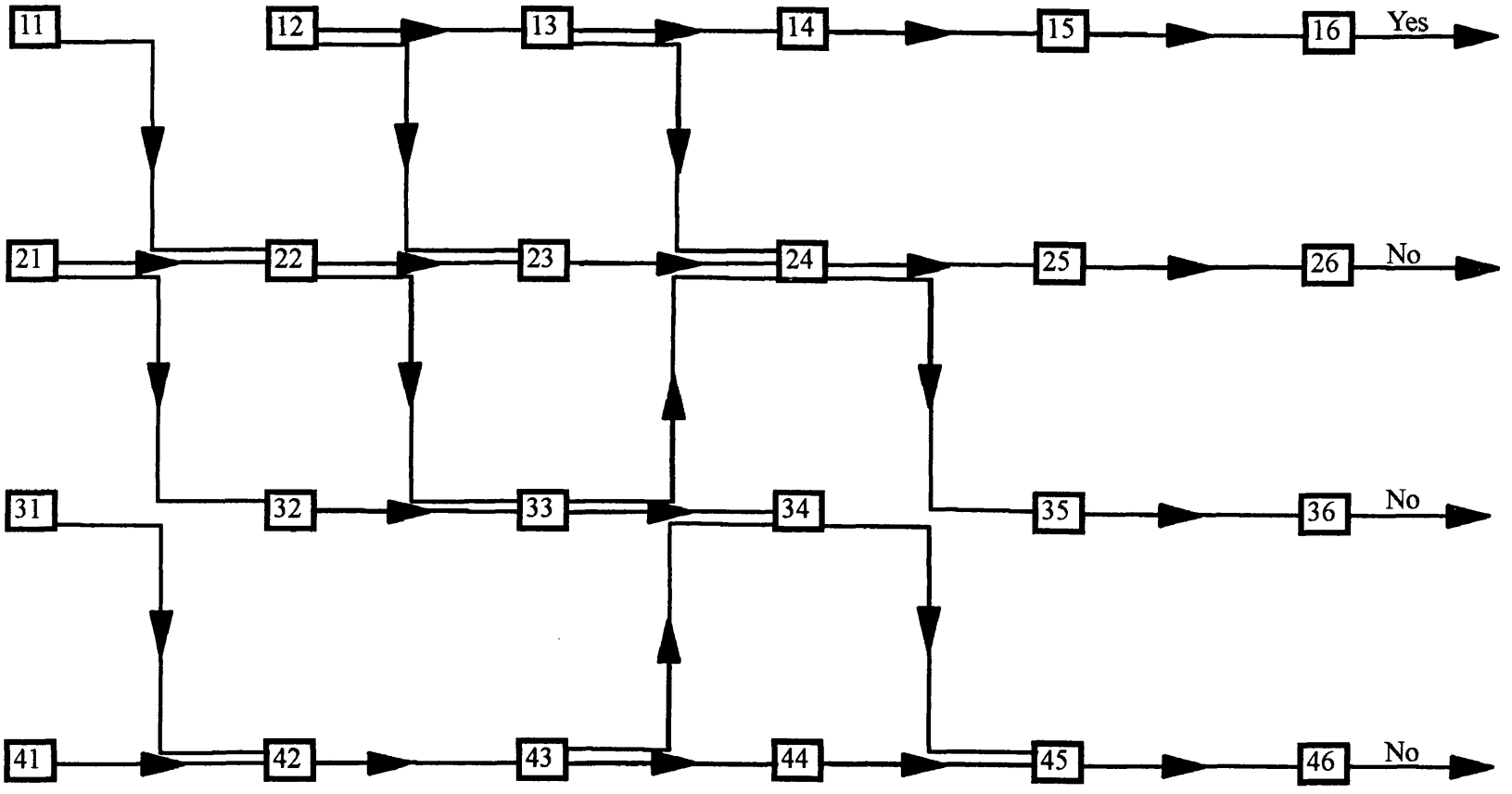

Figure 6.1 An example of a wiring diagram that has five possible faulty units, numbers 11, 21, 22, 32 and 33. The faulty unit can only be identified by testing. For instance, if 32 were the faulty unit, it would be neccessary to test the connections between 21 and 32, and between 32 and 33; any tests performed after these are redundant. Whether other tests are carried out before these depends upon the chosen starting point, since subjects might reasonably start by testing connections with inputs and outputs to any of the possible faulty units listed above.

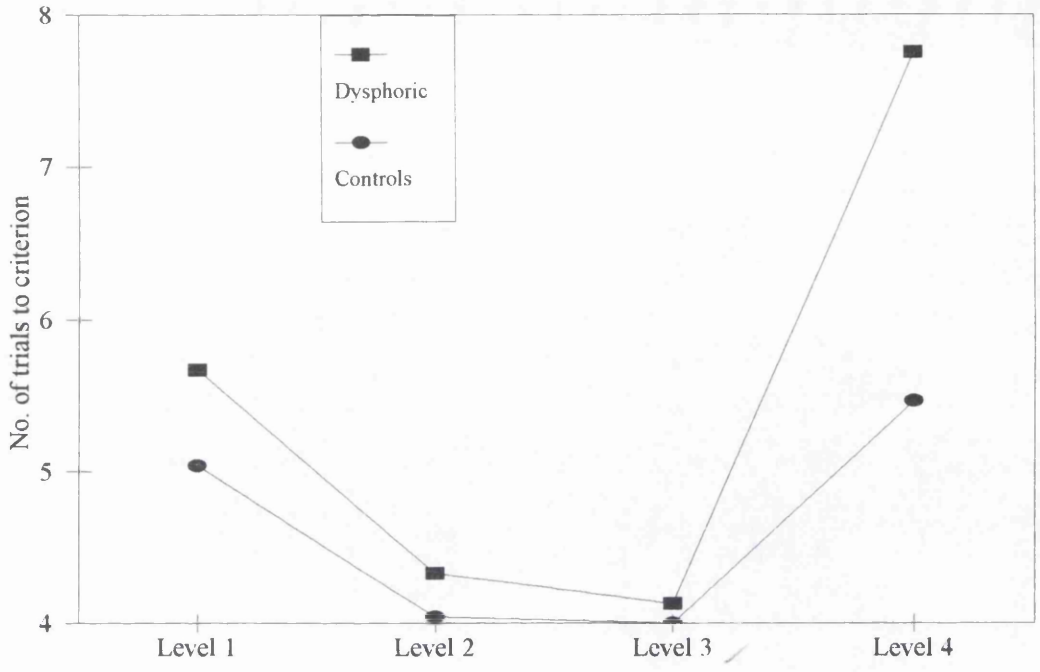

Figure 6.2 Number of trials to criterion (optimal = 4) on the pre-training trials of the Fault Diagnosis task

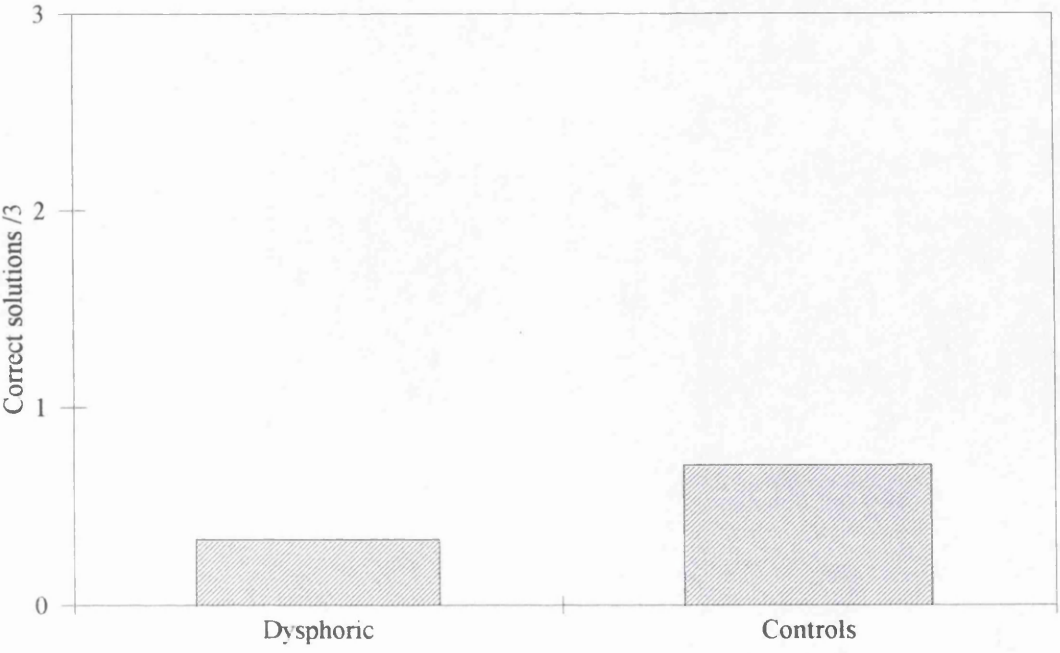

Figure 6.3 Correct solutions /3 in the reception condition of the Fault Diagnosis task

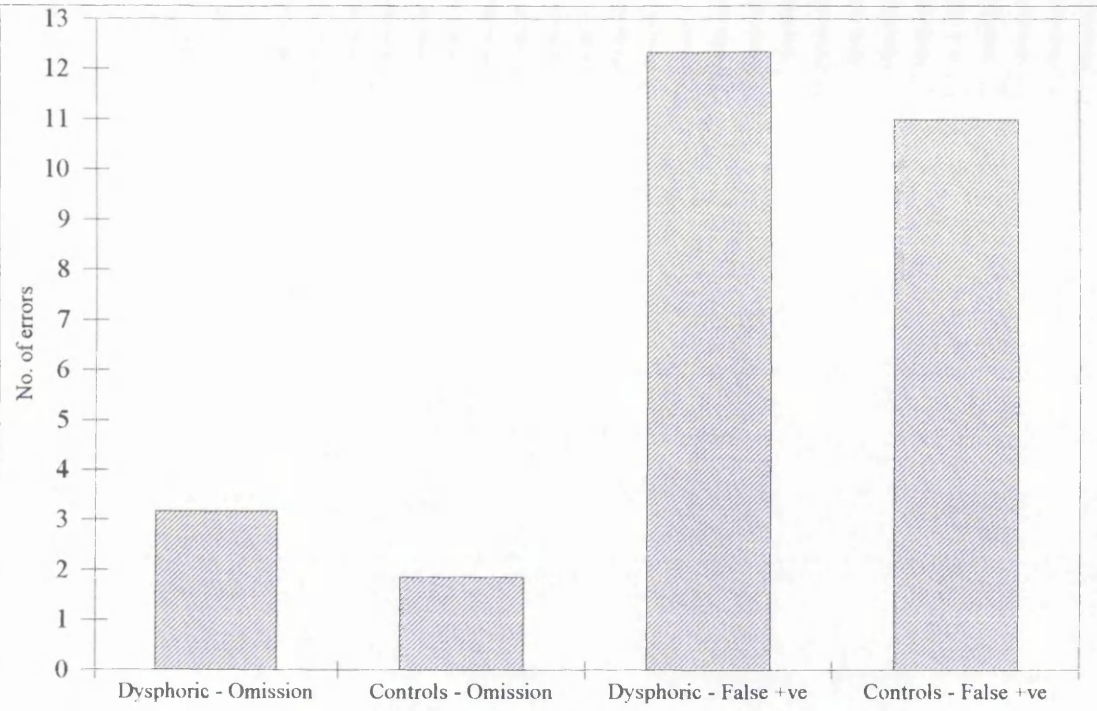

Figure 6.4 Total omission and false positive errors in the reception condition of the Fault Diagnosis task

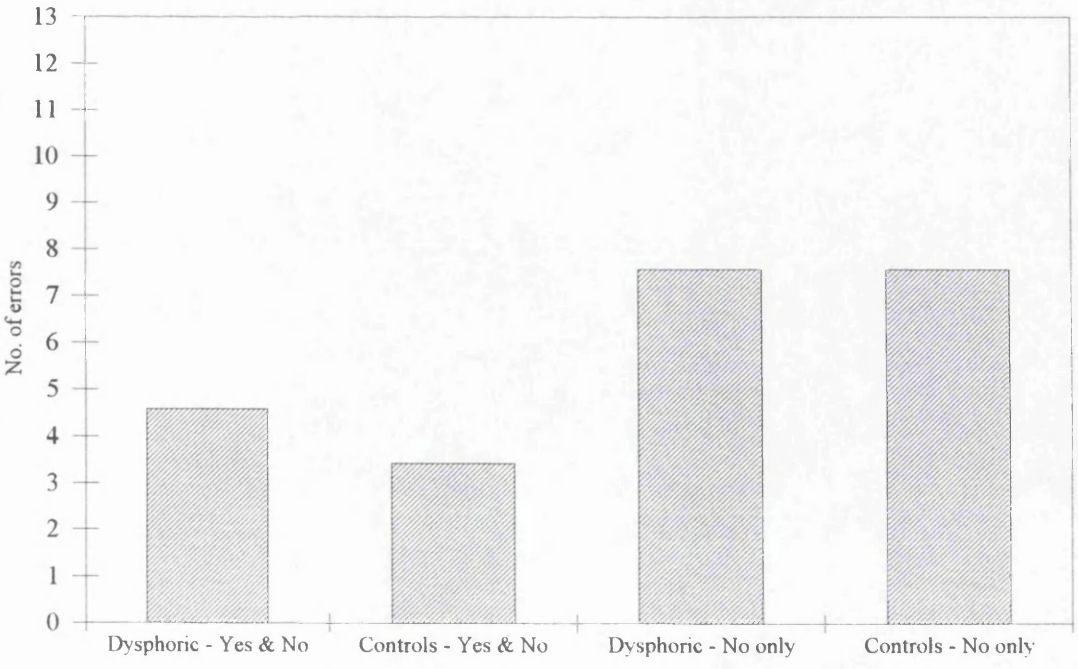

Figure 6.5 Number of false positive errors connected to 'Yes' and 'No' indicators and to a subset of 'No' indicators only in the reception condition of the Fault Diagnosis task

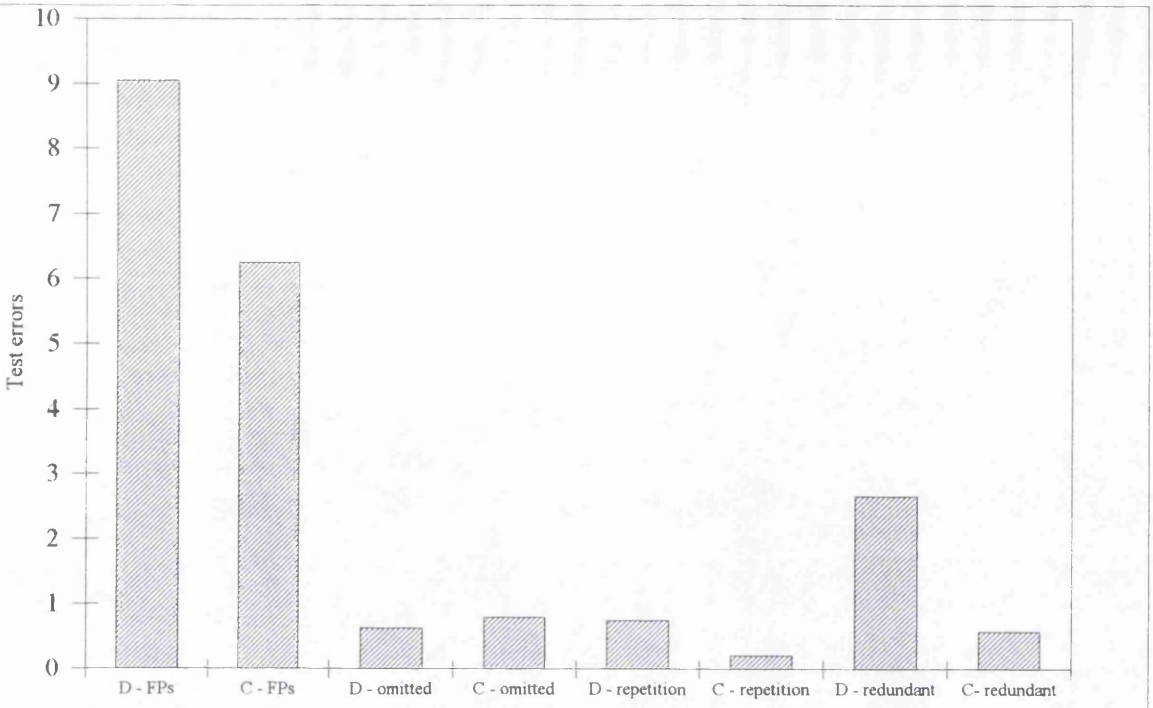

Figure 6.6 Number of false positive, omitted, repetition and redundant test errors in the selection condition of the Fault Diagnosis task

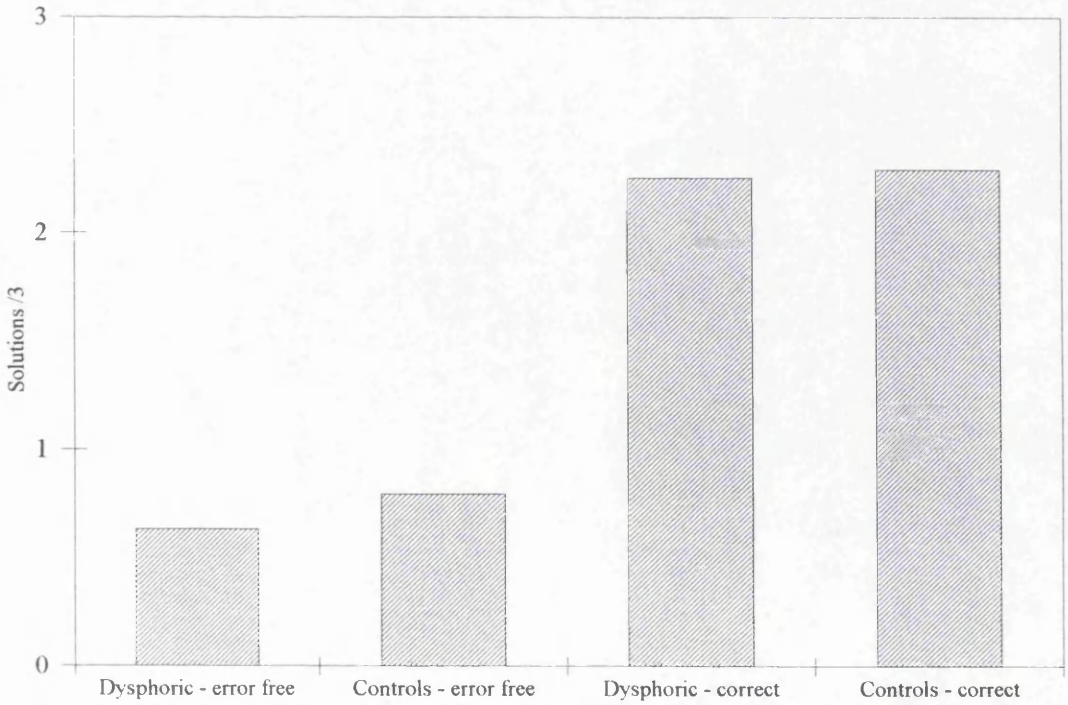

Figure 6.7 Number of error free and correct solutions /3 in the selection condition of the Fault Diagnosis task

CHAPTER VII

EXPERIMENT 4

7. INTRODUCTION

The experiments reported in chapters IV to VI compared dysphoric and control Ss on three reasoning tasks: Discrimination Learning, Integrative Reasoning and Fault Diagnosis. It was postulated that any depressive deficits might reflect reduced storage or processing capacity, or both. The experimental tasks were therefore designed to vary demands on storage and processing capacity. Dysphoric Ss were found to be impaired relative to controls on Discrimination Learning problems in chapter IV and reducing the storage demands of the task did not differentially facilitate the performance of the dysphoric Ss. Dysphoric Ss also showed deficits on the Fault Diagnosis task reported in chapter VI, in particular dysphoric Ss were more likely to collect redundant information beyond the point when the solution could be logically deduced. There were no significant group differences on the Integrative Reasoning task reported in chapter V, but there was evidence that ceiling effects on the easier problems may have reduced the sensitivity of the task to any depressive deficits. Furthermore, there was evidence Ss in both groups may have adopted a guessing strategy on more difficult problems, and there was some suggestion that dysphoric Ss may have had a greater tendency to do this than controls. Having explored the reasoning performance of dysphoric Ss on a range of tasks, it was decided to examine the WM performance of dysphoric Ss more directly.

In section 2.3.7, the existing work on attention and WM in depression was reviewed. It was concluded that there have been relatively few studies of attention and WM in depression. Studies of attention have found evidence of depressive deficits, with depressed Ss consistently poorer on measures of vigilance, and several studies finding deficits in selective attention. With regard to WM, the strongest evidence was in relation to CE function as measured by tasks sensitive to frontal lobe deficits. The effects of distraction on depression suggested the cognitive function of normal Ss given a secondary task may resemble that of depressed Ss without one. Therefore, there was support for the hypothesis that depression is associated with reduced and/or diverted processing resources. The paucity of studies meant the findings were far from being conclusive. It was notable that very few studies made use of paradigms developed within the field of cognitive psychology to measure attention and WM processes. It is possible these might provide valuable information about the nature of attentional and WM processes in depression.

The hypothesis that depression is associated with reduced and/or diverted cognitive resources has been put forward to explain the pattern of depressive cognitive deficits (see section 2.3.9). Earlier theories postulated an actual reduction in cognitive resources (e.g. Hasher & Zacks, 1979), possibly as the result of biological mechanisms (e.g. Roy-Byrne, Weingartner, Bierer, Thompson & Post, 1986). Later theories argue that total resources are not reduced, but limited by a narrowing of attentional focus whereby a proportion of resources is taken up with task-irrelevant or depression-relevant thoughts, thus reducing the resources available for task processing (e.g. Ellis & Ashbrook, 1988). A combination of both reduced resources and narrowed attentional focus has also been suggested (e.g. Watts, 1993). Other models have postulated a conservative response bias (Johnson & Magaro, 1987) or reduced cognitive initiative (e.g. Hertel & Hardin, 1990; Hertel & Rude, 1991b) as alternatives to reduced resources and/or narrowed attentional focus. It was noted in section 2.3.9 that the current evidence does not permit conclusions to be drawn about which model is best supported: the predictions of the different models are indistinguishable for many of the tasks which have been utilised and only a small number of studies have attempted to test hypotheses which would separate the models. Therefore, an important step in elucidating the mechanisms which underlie depressive deficits will be identifying and testing hypotheses that differentiate between these competing models.

One model that assumes narrowing of attentional focus to be the underlying causal mechanism was put forward by Hasher and Zacks (1988). They suggest that in both normal ageing and depression, normal inhibitory mechanisms may become less efficient. This permits irrelevant information to enter WM and receive sustained activation, leading to reduced resources. In the review of selective attention in depression reported in section 2.3.7.2 it was reported that two paradigms had been utilised, with mixed findings. Four studies (Frame & Oltmanns, 1982; Hemsley & Zawada, 1976; Knott et al., 1991; Pogue-Geile & Oltmanns, 1980) examined the ability of depressed Ss to 'filter' relevant from irrelevant aural material. Only Hemsley and Zawada (1976) reported a depressive deficit. The reason for the discrepany between the studies was unclear. Two studies tested depressed Ss on the Stroop Colour-Word Test (SCWT; Stroop, 1935). Both (Lemelin et al., 1996; Trichard et al., 1995) reported that depressed Ss showed significantly greater interference than controls on the standard SCWT task, and Lemelin et al. (1996) also reported a depressive deficit on a modified version of the task. These two studies provide strong evidence of a distractor inhibition disturbance in clinical depression.

Hasher and Zacks' (1988) model makes very specific predictions about the performance of depressed Ss on certain tasks. For example, it was tested in ageing by Gerard, Zacks, Hasher and Radvansky (1991) using a Fan Effect task. This is a task taken from the field of cognitive psychology. It is based on the paradigm described by Anderson (1974; 1983), and is thought to provide a sensitive index of WM. Ss learn sentences which are composed of associations between people and activities. These vary in the number of associations which have to be learnt for each item, for example:

The newsreader cut the apple pie into six pieces.

The judge took the car for a short test drive.
The judge decided to play chess with a friend.

The teacher got change from the ticket machine.
The teacher ran at least four miles a day.
The teacher paid a deposit on the new video.

Once the sentences have been learnt, Ss then have to distinguish under time pressure between study and foil sentences. The foil sentences consist of recombinations of the study sentences (e.g. The judge paid a deposit on the new video). During recognition testing a 'fan-effect' is observed on retrieval, such that sentences consisting of people or activities with only one associate are recalled faster and more accurately than those with more than one associate. Thus, for the above example, sentences concerning the newsreader would be judged faster and more accurately than those relating to the judge or the teacher. The difficulty of this task is increased when more than one person is associated with the same activity, known as a 'crossed fan', for example:

The judge took the car for a short test drive.
The judge decided to play chess with a friend.

The teacher took the car for a short test drive.
The teacher decided to play chess with a friend.
The teacher paid a deposit on the new video.

Gerard et al. (1991) predicted on the basis of Hasher and Zacks' (1988) model that reduced inhibitory mechanisms would result in older Ss experiencing more interference when trying to retrieve the experimental items under time pressure on the Fan Effect recognition task, which would in turn lead to an enhanced fan effect in the older Ss. Older Ss were found to take more trials to learn the study sentences, and they showed an enhanced fan effect relative to younger Ss on the recognition task. Furthermore, the older Ss were particularly prone to false positive recognition errors on foil items compared to younger Ss. Gerard et al. (1991) suggested that the older Ss took more trials to learn the material because reduced inhibitory mechanisms would allow more irrelevant material to become activated and to remain activated for longer, thus permitting spurious associations to develop between the items in the study sentences. Once these incorrect associations had been made, older Ss would be expected to have more difficulty in inhibiting them than younger Ss. The net result would be poorer encoding of the material, and in particular, any spurious associations which had been made would increase the likelihood of false recognition of foil sentences. These results were therefore interpreted by Gerard et al. (1991) as providing support for Hasher and Zacks' model in relation to ageing.

The current experiment was designed to investigate the Hasher and Zacks' (1988) model of cognitive deficits in depression using a Fan Effect task. If depression is characterised by inefficient inhibitory mechanisms (Hasher and Zacks, 1988) then dysphoric Ss should show a similar pattern of performance to the older Ss in the study carried out by Gerard et al. (1991), that is, they should take more trials to learn the study sentences, and show an enhanced fan effect relative to controls. Furthermore, the fan effect should be greater on foil than on studied items for dysphoric Ss.

7.2 EXPERIMENTAL HYPOTHESIS

If depression is associated with inefficient inhibitory mechanisms then:

The dysphoric Ss should take more trials to learn the study sentences than controls.

The dysphoric Ss should show an enhanced fan effect relative to controls.

The dysphoric Ss should make more false recognition errors on the foil sentences relative to the control group.

7.3 METHOD

7.3.1 EXPERIMENTAL MEASURES

7.3.1.1 Experimental Stimuli

The experimental task followed Gerard et al. (1991) and consisted of a series of learning trials followed by a recognition test, and a final recall test (see Appendix 8) .

7.3.1.1.1 Learning trials

The aim of the learning trials was for Ss to learn a list of eighteen study sentences which each described an association between one of nine people (e.g. the judge, the teacher) and one of nine activities (e.g. cut the apple pie into six pieces, took the car for a short test drive). In order to maximise the potential for interference between items, the associations were arbitrary and there was no obvious relationship between the person and the activity. The items used were taken from Gerard et al. (1991) except where they appeared culturally biased, and these were replaced (e.g. "the anchorman" became "the newsreader"). Within the eighteen study sentences, each person and each activity appeared in between one and three sentences in a 'crossed-fan' formation. The list was constructed to contain three critical fan levels, six at the 3-3 level, and three each at the 1-1, and 2-2 levels, where the first number refers to the number of associations with the person, and the second to the activity; three non-critical sentences at the 2-3 and 3-2 levels also had to be included in order to create the critical fan levels. The people and activities were randomly selected to appear at the different fan levels, and all Ss learnt the same list (see Appendix 8.1).

Five lists containing the study sentences in a different random order were created. These were presented to Ss on successive learning trials, and repeated if more than five learning trials were required. Five booklets were constructed, each containing the 18 sentences in a different random order. Each sentence was typed on a card 3 x 8 in., and the booklet was held together by a tag in the left-hand corner. There was a blank page at the front of each booklet to ensure that the first sentence could not be seen prior to presentation by the experimenter. On each learning trial, the 18 study sentences were presented by the experimenter one at a time for seven seconds each. The instructions given to Ss were as follows:

"In this booklet there are a number of pages, and on each page there is a sentence. I am going to show you each sentence one at a time, and I want you to look at them carefully and try to remember what they say because I am going to ask you questions about them afterwards."

Learning was then tested by asking questions of the form "What did the judge do?" or "Who took the car for a short test drive?" (see Appendix 8.2). Each activity and each person was tested once for a total of eighteen questions. The correct answer required the production of one, two or three people or activities, depending on the fan level of the item. If the S gave an incorrect answer, then the experimenter read out the entire correct answer. The questions were presented in a fixed random order, and five lists were constructed with the questions in a different order each time, and these were presented on successive learning trials, and repeated if necessary in the same way as the sentence lists. A learning criterion was used of two consecutive series of correct answers to all the questions.

7.3.1.1.2 Recognition test

Once the learning criterion was achieved, a seventy-item recognition test was given containing nine of the study sentences, three each from levels 1-1, 2-2, and 3-3 presented in the original list, and nine foils constructed by recombining items in the original list to produce three at each of the 1-1, 2-2, and 3-3 levels (see Apendix 8.3). Each of the nine study and nine foil sentences appeared three times during the list. Following Gerard et al. (1991), additional trials were included as fillers to vary the fan levels, consisting of both sentences from the 2-3 and 3-2 levels in the original list, and recombinations of the original items across fan levels. The list began with three randomly selected filler sentences, and the remaining sentences were presented in fixed random order. Each of the seventy sentences was typed onto a card 7 x 12 cm.. These were made up into booklets in the same way as the learning stimuli described in 7.3.1.1.1 with 10 sentences in each.

The recognition sentences were presented visually one at a time on cards, and the card following each recognition item gave feedback as to whether the previous item had appeared on the original study list using the words "present" or "absent". Ss were asked to respond "yes" or "no" as quickly as possible before turning each page to receive

feedback, and time taken from turning over the first card to responding to the last item was recorded. An example of the test sentences and feedback was given before the recognition test began. The instructions given to Ss were as follows:

"In this booklet are a number of pages, and on each page is a sentence. When I tell you to begin, I want you to turn the page to the first sentence and as quickly as you can tell me whether it appeared on the list you saw earlier. You should say 'yes' if it was on the list, and 'no' if it wasn't. Then, as quickly as you can, turn the page, and you will see the word 'present' or 'absent'. 'Present' means that the sentence you've just seen was on the list, and 'absent' means that it wasn't. This will tell you whether your answer was correct or not, and it is therefore very important that you say 'yes' or 'no' before you turn the page. Turn the to the next sentence and again say 'yes' or 'no'. Carry on in this way until you reach the end of this first booklet, and then as quickly as you can turn to the first page of this second booklet and carry on in the same way. When you reach the end of the second booklet, then turn to the first page of the third booklet which I will have place there for you, and carry on in the same way until I tell you to stop."

7.3.1.1.3 Final recall test

After completing the recognition task, Ss' retention of the original eighteen study sentences was tested by presenting the eighteen test questions again as described for the learning trials, in a different random order. Ss' responses were recorded by the experimenter, and no feedback about performance was given.

7.3.1.2 Experimental Manipulations

The critical manipulation in this study was the comparison of Ss' ability to correctly identify studied and foil sentences at the three fan levels (1-1, 2-2 and 3-3).

7.3.2 CLINICAL MEASURES

The clinical measures were the same as those described in section 4.3.2.

7.3.3 SUBJECT SELECTION

7.3.3.1 Selection Criteria and Procedure

The selection criteria and procedure were the same as those described in sections 4.3.3.1 and 4.3.3.2.

7.3.3.2 Subjects

Ss were recruited from a pool of approximately 190 first- or second-year university undergraduate volunteers who filled out the BDI as part of a lecture. Of the 36 who scored 11 or above, seven did not want to participate, and 28 carried out the experiment. Five of these were later excluded since they did not score in the dysphoric range on the second administration of the BDI, and one other was excluded since they did not meet the selection criteria. Seventy-six Ss initially scored five or below, and individuals were randomly selected from this pool until a sample size which matched that of the dysphoric Ss was collected. Fourteen Ss did not want to take part, one scored above the criterion on the second administration of the BDI, and one did not meet the selection criteria. The final sample therefore consisted of 23 Ss in each group.

T-tests confirmed that the dysphoric group scored significantly higher than the control group on both the first and second administrations of the BDI. The groups did not differ significantly in vocabulary or age (see Table 7.1).

Table 7.1 Means (and standard deviations) for age, vocabulary and BDI scores, and information on sex in the two groups

	Dysphoric	**Controls**
Sex	3m, 20f	6m, 17f
Age	21.0	21.5
	(3.2)	(3.0)
Vocab	12.7	12.3
	(1.4)	(1.6)
BDI 1	20.2	2.4
	(7.9)	(1.8)
BDI 2	17.8	1.8
	(7.0)	(1.5)

7.3.4 PROCEDURE

The procedure was the same as that described in section 4.3.4

7.4 RESULTS

7.4.1 LEARNING AND RETENTION OF THE STUDIED SENTENCES

Means and standard deviations for performance on the learning and retention measures of the Fan Effect task are shown in Table 7.2.

.

As described in section 7.3.1, Ss were first required to learn the associations between the items in the 18 study sentences. On each learning trial, the 18 study sentences were presented, and then retention was tested by asking Ss to recall the items associated with each person and activity. This was repeated until Ss produced all the associations correctly on two consecutive trials. The two groups were initially compared on the number of acquisition trials taken to achieve this learning criterion, and a t-test revealed no significant difference on this measure ($t=-1.19$, $df=44$, $p=0.24$).

Errors on the learning trials were classified as either omission errors, where an association was omitted from a response to an item, or as false positive errors, where an incorrect association was produced. The number of omission and false positive errors made by the two groups is shown in Figure 7.1. Unacceptable positive skewness and kurtosis was detected in the number of omission and false positive errors made on the learning trials, and this was corrected using a log 10 transformation. T-tests were carried out to compare the groups on these two measures, and it was found that the groups did not differ significantly in the number of false positive errors ($t=-1.52$, $df=44$, $p=0.14$). For both groups, the false positive errors were typically the result of incorrect associations between experimental items, rather than the introduction of new items by the subjects. The difference in the number of omission errors also did not reach significance ($t=-1.75$, $df=44$, $p=0.09$), but there was a tendency for the dysphoric group to make fewer omission errors (see Figure 7.1).

After Ss had carried out the recognition test (see section 7.4.2), their retention of the study sentences was tested on a final recall test in which they were required to recall the associations for each of the 18 items. The total number of items correctly recalled was found to have negative skewness, and this was corrected using a "reflex" transformation

(see section 3.2.2) followed by a log 10 transformation. A t-test revealed that there was no significant group difference in the number of correct responses (t=-0.78, df=44, p=0.44). The groups were also compared for the number of false positive errors made on the final recall test. Positive skewness and kurtosis were detected in the number of false positive errors made on the final recall test, and these were corrected with a log 10 transformation. A t-test found no significant group difference (t=-0.78, df=44, p=0.44).

Table 7.2 Means (and standard deviations) for performance measures on the learning trials of the Fan Effect task

	Dysphoric	Controls
Learning trials		
Trials to criterion	6.96	7.78
	(1.75)	(2.83)
Total omission	29.87	43.30
errors	(16.40)	(31.87)
Log 10	1.42	1.56
	(0.26)	(0.27)
Total FP errors	25.22	34.17
	(18.64)	(26.27)
Log 10	1.30	1.45
	(0.33)	(0.30)
Final recall test		
Total correct responses	35.22	34.96
/36	(1.28)	(1.33)
"Reflex" log 10	0.41	0.45
	(0.16)	(0.17)
Total FP errors	0.65	0.87
	(1.15)	(1.14)
Log 10	0.15	0.21
	(0.22)	(0.24)

7.4.2 FAN EFFECT IN RECOGNITION

Means and standard deviations for performance on the learning and retention measures of the Fan Effect task are shown in Table 7.3.

The time taken to complete the 70-item recognition test was measured by the experimenter. Positive skewness was detected in the time taken to complete the recognition test, and an outlier in the depressed group was replaced with the nearest score plus one (see section 3.2.2). This reduced skewness to an acceptable level, and a t-test found there was no significant difference between the groups in the amount of time taken to perform the recognition test (t=-1.07, df=44, p=0.29).

The critical analysis in this experiment was the comparison of the groups in their ability to classify correctly the sentences at the 1-1, 2-2 and 3-3 levels. Performance at the three critical fan levels was analysed separately for the studied and the foil sentences, excluding the filler items, but both are shown in Figure 7.2.. There was unacceptable skewness and kurtosis in the data. The nature of the data suggested that an arcsine transformation was the most appropriate, but this failed to reduce the skewness and kurtosis to acceptable levels. An analysis capable of carrying out the critical between-group comparison across the fan levels was necessary, and since such a non-parametric test was not available it was decided the analysis would be carried out using repeated measures ANOVA, and then repeated using non-parametric comparisons.

For the studied sentences, there was a significant effect of sentence level (F=4.38, df=2,88, p=0.015), and inspection of the means revealed more errors at the 3-3 level relative to the 2-2 and 1-1 levels (see Figure 7.2). The group by level interaction was not significant (F=0.55, df=2,88, p=0.58), indicating that the dysphoric Ss failed to show a differentially large fan effect, and there was no significant effect of group (F=1.70, df=1,44, p=0.20).

For the foil sentences there was a significant effect of sentence level (F=5.28, df=1.43,88, p=0.015), and inspection of the means revealed more errors at the 3-3 relative to the 2-2 and 1-1 levels. There was a significant group by level interaction (F=4.06, df=1.43, 88, p=0.032) (see Figure 7.2). T-tests were used to compare the groups at each level with a post-hoc significance level of 0.012. The groups did not differ significantly at the 1-1 (t=-

0.59, df=44, p=0.56) or 2-2 level (t=-0.03, df=44, p=0.974). There was a significant difference between the groups at the 3-3 level (t=2.75, df=44, p=0.009). The dysphoric group made more errors relative to the control group on this measure. There was no overall significant effect of group (F=3.65, df=1,44, p=0.063), although it approached significance.

Nonparametric analyses were also carried out. Mann-Whitney U-tests were used to compare the groups at each fan level on the studied and foil sentences. On the studied sentences the difference between the groups did not reach significance at the 1-1 (U=264.5, p=1.00), 2-2 (U=231.0, p=0.29) or 3-3 (U=241.0, p=0.55) levels. On the foil sentences, there was no significant difference at the 1-1 (U=253.0, p=0.56) or 2-2 (U=256.0, p=0.77) levels, but there was a significant group difference at the 3-3 level (U=164.5, p=0.009), thus confirming the findings of the ANOVA reported above.

7.4.3 Correlation of Performance with Severity of Depressive Symptomatology

Pearson correlation coefficients were calculated between each of the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2). This was carried out within each group because the selection criteria for the two groups (dysphoric >10; nondysphoric <6) ensured that, taken together, the scores on the BDI could not have a normal distribution. There were no significant results with either the BDI1 or the BDI2. Table 7.4 shows the results for the BDI2. This is the measure which was taken during the actual test session, and might therefore be expected to show a higher correlation with task performance than the BDI1 which was completed as part of the initial screening process (see section 4.3.3.2).

Table 7.3 Means (and standard deviations) for performance measures on the recognition test of the Fan Effect task

	Dysphoric	Controls
Recognition test		
Time taken (secs)	254.83	269.17
	(33.53)	(54.99)
Study sentences		
Fan 1-1 % errors	1.45	1.45
	(3.83)	(3.83)
Arcsine	0.14	0.14
	(0.21)	(0.21)
Fan 2-2 % errors	4.35	1.93
	(9.90)	(5.46)
Arcsine	0.26	0.16
	(0.37)	(0.25)
Fan 3-3 % errors	7.25	4.35
	(11.41)	(6.48)
Arcsine	0.38	0.29
	(0.43)	(0.32)
Foil sentences		
Fan 1-1 % errors	0.48	0.97
	(2.32)	(3.20)
Arcsine	0.09	0.12
	(0.13)	(0.18)
Fan 2-2 % errors	2.90	2.42
	(8.35)	(5.76)
Arcsine	0.18	0.18
	(0.32)	(0.27)
Fan 3-3 % errors	9.66	1.93
	(14.72)	(4.31)
Arcsine	0.48	0.17
	(0.48)	(0.24)

Table 7.4 Correlation of BDI2[+] with performance measures on the Fan Effect problems

	Dysphoric	Controls
Learning trials		
Trials to criterion	-.27	-.02
Total omission errors	-.20	-.18
Log 10	-.12	-.14
Total FP errors	-.16	-.15
Log 10	-.20	-.15
Final recall test		
Total correct responses /36	.20	.06
"Reflex" log 10	-.20	-.06
Total FP errors	-.17	-.20
Log 10	-.10	-.27
Recognition test		
Time taken (secs)	-.20	-.32
Study sentences		
Fan 1-1 % errors	.20	-.03
Arcsine	.20	-.03
Fan 2-2 % errors	.03	-.19
Arcsine	.18	-.16
Fan 3-3 % errors	-.19	.30
Arcsine	-.22	.27
Foil sentences		
Fan 1-1 % errors	.13	-.37
Arcsine	.13	-.37
Fan 2-2 % errors	.33	.01
Arcsine	.35	.03
Fan 3-3 % errors	.07	.45
Arcsine	.20	.45

*p<.01; **p<.001, two-tail. "/" indicates that a coefficient could not be computed.

[+]BDI2 = Beck Depression Inventory given on the second occasion (see section 4.3.3.1).

7.5 SUMMARY OF RESULTS

1. There was no significant difference between the groups in the number of trials to reach learning criterion, or in the number of false positive and omission errors made on these trials.

2. There was no significant difference between the groups in the number of correct responses or false positive errors made on the final recall test.

3. On the recognition test, for the study sentences, both groups showed a normal fan effect, and there was no significant difference between the groups in the strength of this effect.

4. For the foil sentences on the recognition test, both groups showed a normal fan effect, and the dysphoric Ss showed a significantly enhanced fan effect.

5. There were no significant correlations between the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2).

7.6 DISCUSSION

The aim of this experiment was to test the model postulated by Hasher and Zacks (1988) that depression is associated with reduced inhibitory mechanisms which permit irrelevant information to enter WM. Dysphoric and nondysphoric Ss were compared on a Fan Effect task known to be sensitive to deficits in normal ageing (Gerard et al., 1991).

The main prediction was that dysphoric Ss would show an enhanced fan effect, since depression is thought to be associated with reduced or diverted WM resources. This was confirmed in that dysphoric Ss showed an enhanced fan effect on the foil sentences relative to controls. On the studied sentences, both groups showed a normal fan effect, and did not differ significantly in terms of its size. There was no evidence of a significant difference between the groups in psychomotor speed in terms of time taken to complete the recognition test, suggesting that the dysphoric Ss maintained a normal speed of responding at the expense of response accuracy. This pattern therefore appears to reflect differences in the trade-off between speed and accuracy in the two groups.

There was no significant difference between the groups in learning or retention of the study sentences, ruling out the possibility that the enhanced fan effect shown by the dysphoric group on the foil sentences might be a function of impaired acquisition or retention of the relevant information in LTM. The learning trials for the study sentences were highly structured, and Ss were given cues for recall. Previous studies have found that depression is not invariably associated with deficits on memory tasks (see section 2.3.4); materials which need to be organised and structured may be more sensitive to depressive deficits (e.g. Channon et al., 1993a), and providing cues for such structuring tends to ameliorate such deficits (e.g. Hertel & Hardin, 1990). The learning procedure in the present experiment was therefore unlikely to produce deficits in the dysphoric group.

The finding that the dysphoric group showed an enhanced fan effect on the foil sentences is consistent with the findings of Lemelin et al. (1996) and Trichard et al. (1995) who both reported depressive deficits on the Stroop Colour-Word Test (SCWT; Stroop, 1935). It is also consistent with Hemsley and Zawada (1976) who reported a depressive deficit in 'filtering' relevant from irrelevant material. It is inconsistent with the findings of three studies (Frame & Oltmanns, 1982; Knott et al., 1991; Pogue-Geile & Oltmanns, 1980) which failed to find a depressive deficit on filtering tasks that were similar to that reported by Hemsley and Zawada (1976). Therefore, the findings of the current study add weight to the hypothesis that depression is associated with a distractor inhibition disturbance, although on the basis of the evidence currently available it is still not possible to accept this as unequivocal.

The finding that the dysphoric group showed an enhanced fan effect on the foil but not the study sentences is of interest, since it potentially contributes to our understanding of the nature of the underlying deficit in depression. The difference is unlikely to be attributable to a response bias, since the dysphoric group did not show a general tendency to say "yes" more often, but did this only at the 3-3 fan level, suggesting a more specific interpretation. Moreover, previous studies have more frequently described a conservative response bias associated with depression (see section 2.3.4.3), which should lead to more "no" rather than "yes" responses. This finding is inconsistent with the cognitive intiative model of depressive deficits (e.g. Hertel & Hardin, 1990; Hertel & Rude, 1991b) since there was no difference between the foil and study sentences in terms of cuing appropriate strategies.

The findings are consistent with Hasher and Zacks' (1988) model in that dysphoric Ss did show an enhanced fan effect on foil sentences which can be explained in terms of inefficient inhibitory mechanisms. Gerard et al. (1991) suggested this might allow irrelevant information to become activated and to remain so for longer periods. This would enable more spurious associations to develop between items, leading to an increased rate of false positive errors on the foil sentences during the recognition test. In line with this, the enhanced fan effect shown by older Ss in their study was stronger on foil than on studied sentences. Other models of reduced or diverted cognitive capacity do not assume dysphoric Ss have specific difficulties in inhibiting information, and therefore do not predict the elevated false recognition error rate for the foil sentences. However, it is possible the findings reflect the specific task characteristics of the fan effect paradigm, and that the foil sentences promote a tendency towards false positive errors. It would therefore be premature to attempt to distinguish definitively between the effects of reduced capacity and inefficient inhibitory mechanisms on the basis of this task alone.

With regard to the findings reported in chapters IV to VI, the Hasher and Zacks' (1988) proposal of inefficient inhibitory mechanisms would predict deficits on reasoning tasks that make demands on cognitive resources, such as the ones described. However, as is often the case (see section 7.1), it is difficult to generate specific predictions that would differentiate between Hasher and Zacks' (1988) model and other models that postulate reduced resources and/or narrowed attentional focus as the mechanism underlying depressive cognitive deficits for these tasks. More work is needed to test the competing models of depressive deficits by generating differential hypotheses.

As before., the failure to find any significant correlations between performance measures and scores on the BDI1 and BDI2 needs to be considered. As noted previously (e.g. section 4.6), this suggests either that there was no relationship between severity of depressive symptomatology and task performance, or that a relationship did exist, but that the measures used made it difficult to detect. In the current experiment, there was no significant difference between the groups on many of the performance measures, and it is perhaps unsurprising if there was no significant correlation between these variables and depressive symptomatology. The dysphoric group did show a significantly enhanced fan effect on the foil sentences. This measure might therefore be expected to show a relationship with severity of depressive symptoms, but none was found. It has been noted

that correlations are more difficult to detect if the range of scores is restricted in some way. This was a particular problem for the control group, where there was a possible range of only 0-5 on the BDI, making it unlikely that any meaningful relationship with depressive symptomatology could be identified. The dysphoric group had a greater range of scores on the BDI (11-35 on BDI2), but there was a restricted possible range on the relevant task variable (0-9). It is therefore possible that the measures were not well-suited to identifying correlations, although it is possible that a relationship between depressive symptomatology and task performance simply did not exist.

In summary, dysphoric Ss in the current experiment were found to show an enhanced fan effect on the foil sentences in the speeded recognition test. This finding is consistent with the model proposed by Hasher and Zacks (1988), which postulates inefficient inhibitory mechanisms in depression, although more work is needed to differentiate between the competing models of depressive cognitive deficits.

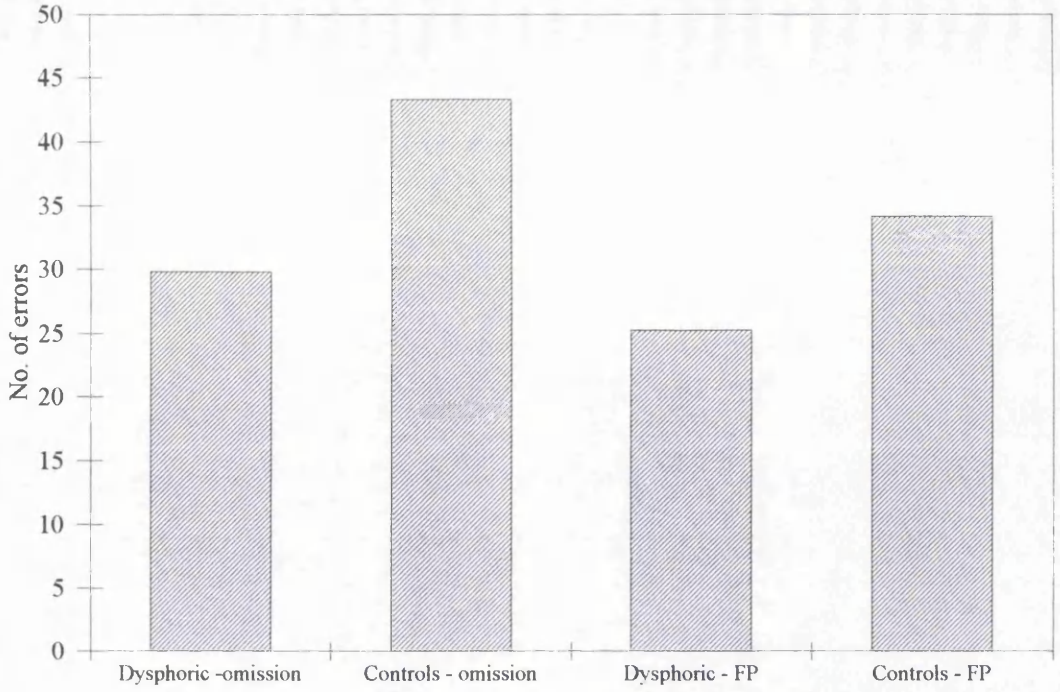

Figure 7.1 Total omission and false positive errors on the learning trials of the Fan Effect task

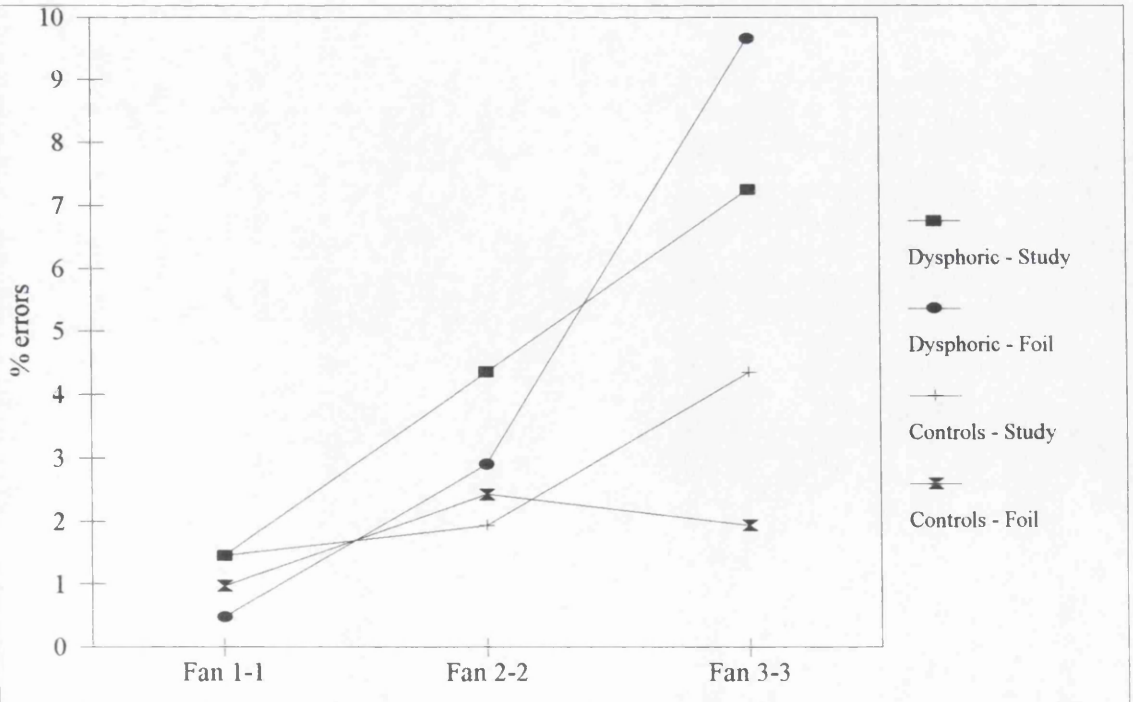

Figure 7.2 Percentage errors on the study and foil sentences of the recognition test - Fan Effect task

CHAPTER VIII

EXPERIMENT 5

8.1 INTRODUCTION

The experiments reported in chapters IV to VI compared the performance of dysphoric and nondysphoric Ss on three different reasoning tasks: Discrimination Learning, Integrative Reasoning, and Fault Diagnosis. The aim was to investigate the role of reduced storage and/or processing capacity in any dysphoric reasoning deficits by varying the storage and processing demands within each task. The experiment reported in chapter VII tested the hypothesis that WM capacity is reduced in dysphoric Ss in a more direct fashion using a Fan Effect task, and found evidence consistent with this. It also highlighted the value of carrying out work aimed at differentiating between the competing models of depressive cognitive function.

The hypothesis that reduced storage capacity might play a role in dysphoric reasoning deficits was explored in chapters IV to VI, with mixed results. The standard versions of the Discrimination Learning, Integrative Reasoning and Fault Diagnosis tasks all involve the serial presentation or collection of information, and therefore require Ss to store information for later integration. In the Discrimination Learning and Integrative Reasoning tasks, a memory-aid condition was introduced in which the information relevant to problem solution was left on view until the S responded, thereby reducing the storage demands of the task. This failed to differentially facilitate the performance of the dysphoric Ss in either case. In the Fault Diagnosis task the standard selection condition, in which Ss tested hypotheses and received feedback, placed greater demands on storage than the reception paradigm, in which Ss simply had to draw inferences about which units were potentially faulty. Dysphoric Ss were found to show deficits in the selection, but not the reception condition. However, the reception and selection conditions of the Fault Diagnosis task are not directly comparable in the same way as the standard and memory-aid conditions of the first two tasks, since Ss only had to identify the potentially faulty units in the reception condition, whilst they had to test their hypotheses and identify the actual faulty unit in the selection paradigm.

As a consequence of the pattern of findings reported in chapters IV and V, the importance of task strategy was identified. The pattern of performance on the Discrimination Learning task was consistent with dysphoric Ss adopting simple heuristic strategies. It

was also noted in section 5.6 that Ss in both groups may have adopted a guessing strategy on more complex Integrative Reasoning problems. It was therefore decided to examine the strategies adopted by Ss in more detail, and this was explored using the Fault Diagnosis task in chapter VI, with the selection condition of particular interest. As noted above, the dysphoric Ss did test hypotheses less efficiently than controls, carrying out a greater number of repetition tests, and also carrying out redundant tests beyond the point when the answer could be logically deduced. However, both groups performed relatively poorly in both conditions of the Fault Diagnosis task, and other group differences may have been obscured by floor effects. It was therefore decided to examine the role of strategy in depressive performance further in the current experiment.

Since the Discrimination Learning task reported in chapter IV appeared to be particularly sensitive to depressive deficits, it was decided to investigate this task further. Conditions were designed to investigate the role of strategy in performance on this task, and to examine further the role of storage.

The cognitive initiative model put forward by Hertel and colleagues (see section 2.3.9 for review) suggests that cognitive deficits associated with depression are not invariant, and they have reported work showing that giving instructions which cue appropriate strategies can facilitate the performance of depressed Ss to normal levels on memory tasks. In the current experiment, Ss were asked after each feedback trial to list the hypotheses which remained viable. This manipulation was similar to the 'report' condition used by Silberman et al. (1983; see section 4.1), and was intended to focus Ss' attention at the critical stages of the task.

The standard condition of the Discrimination Learning task described in chapter IV requires Ss to retain the feedback given on the three feedback trials across the intervening non-feedback trials. The memory-aid condition, also described in chapter IV, was designed to reduce the storage load of the task by removing the need to remember the feedback cards, the response made by the S, and the feedback given by the experimenter, since all this information was left on view to the S. However, the memory-aid condition did not remove all need to store information, since after each feedback trial Ss still had to remember the list of hypotheses which they were considering as possible solutions, in order to integrate this information with later feedback. In the current experiment, a

manipulation was introduced to reduce the storage load further. In one of the two conditions, Ss' lists of hypotheses were left on display for the remainder of the problem, to aid memory storage.

Thus, the current experiment was designed to investigate the model of cognitive deficits in depression put forward by Hertel and her colleagues (e.g. Hertel & Hardin, 1990; Hertel & Rude, 1991b) using the Discrimination Learning task, and to investigate further the role of storage in depressive reasoning deficits.

8.2 EXPERIMENTAL HYPOTHESES

If depression is associated with deficits in reasoning performance, then dysphoric Ss should be impaired relative to controls on Discrimination Learning problems.

If dysphoric Ss are impaired in their ability to adopt an appropriate strategy then an experimental manipulation which focuses attention on the task at the critical stages should differentially facilitate the performance of the dysphoric group.

If dysphoric Ss have reduced storage capacity then an experimental manipulation which reduces the necessity to retain information should differentially facilitate the performance of the dysphoric group.

8.3 METHOD

8.3.1 EXPERIMENTAL MEASURES

8.3.1.1 Experimental Stimuli

The experimental stimuli were the same as those described in section 4.3.1.1. As described in Section 4.3.1.1, each problem consisted of a series of trials, requiring a choice between two stimuli on each trial. The two stimuli varied on up to four dimensions (big-small, left-right, A-B, black-white), such that one of the two stimuli in each pair would be black, the other white, and so on. On some trials Ss received feedback as to whether their choice was correct, whilst other trials were non-feedback trials included to provide information about Ss' pattern of responding.

8.3.1.2 One- and Two-Dimensional Pretraining Problems

As described in section 4.3.1.2, Ss carried out a pretraining procedure consisting of four problems in which the stimuli varied on only one dimension, and four problems in which the stimuli varied on two dimensions.

8.3.1.3 Four-Dimensional Problems

The four-dimensional problems were the same as those described in section 4.3.1.3. The stimuli varied on all four dimensions simultaneously, giving eight possible solutions, and the task was to determine which was correct. Feedback was given every fifth trial. Optimal performance required Ss to discard four hypotheses after the first feedback, two more after the second feedback, and one more after the third feedback, leaving a single hypothesis which was the correct solution (see Figure 4.1).

8.3.1.4 Experimental Manipulations

Following the pretraining problems, Ss carried out the four-dimensional problems in two conditions: report and report-aid. In the report condition, following Silberman et al. (1983), after each feedback trial, Ss were asked to list all the hypotheses which they believed to be potentially correct, and these were recorded by the experimenter. In the report-aid condition, Ss' lists of hypotheses, as recorded by the experimenter, were displayed beside the appropriate stimuli (e.g. left, black, A, small). This manipulation removed the necessity for Ss to retain or regenerate the list of hypotheses produced after each feedback trial. Ss performed a total of eight problems, four in each of the two conditions. In both the report and report-aid conditions, feedback and non-feedback trials were used as described in section 4.3.1.3. In order to minimise the storage load, correct/wrong feedback about performance on the feedback trials was left in view as in the memory-aid condition described in section 4.3.1.4. To avoid possible practice effects, the order of presentation of the two conditions was counterbalanced within each group.

8.3.2 CLINICAL MEASURES

The clinical measures were the same as those described in section 4.3.2.

8.3.3 SUBJECT SELECTION

8.3.3.1 Selection Criteria and Procedure

The selection criteria and procedure were the same as those described in sections 4.3.3.1 and 4.3.3.2.

8.3.3.2 Subjects

The Ss were recruited from a pool of 190 undergraduate volunteers who initially completed the BDI. Of the 38 who scored 11 or above, nine did not want to participate, and 29 carried out the experiment. Four of these were excluded since they did not score in the dysphoric range on the second administration of the BDI. Seventy-eight Ss initially scored five or below, and individuals were randomly selected from this pool until a sample size which matched that of the dysphoric Ss was collected. Eight Ss did not want to take part, and three had to be excluded since they did not meet the selection criteria. The final sample consisted of 25 Ss in each group.

T-tests confirmed that the dysphoric group scored significantly higher than the control group on both the first and second administrations of the BDI. The groups did not differ significantly in vocabulary or age (see Table 8.1).

Table 8.1 Means (and standard deviations) for age, vocabulary and BDI scores, and information on sex in the two groups

	Dysphoric	Controls
Sex	6m, 19f	9m, 16f
Age	21.3 (3.8)	20.2 (2.4)
Vocab	12.1 (2.0)	12.8 (1.4)
BDI 1	22.3 (6.9)	2.1 (2.0)
BDI 2	18.6 (7.2)	1.6 (1.7)

8.3.4 PROCEDURE

The procedure was the same as that described in section 4.3.4.

8.4 RESULTS

8.4.1 ONE- AND TWO-DIMENSIONAL PROBLEMS

The means and standard deviations for the performance measures on the one- and two-dimensional problems are shown in Table 8.2.

On the one-dimensional training problems, all Ss in each group achieved the criterion of two consecutive correct responses on each trial, and were also able to verbalise the correct solutions.

On the two-dimensional training problems, most Ss again performed at ceiling level. All but six of the dysphoric and two of the control Ss achieved the learning criterion on each problem. All but two Ss in each group were able to verbalise the correct solution to each problem.

Table 8.2 Means (and standard deviations) for performance measures on the one- and two-dimensional problems for both groups

	Dysphoric	Controls
1-dimensional problems		
Reached learning criterion /4	4.00 (0.00)	4.00 (0.00)
Correct solutions /4	4.00 (0.00)	4.00 (0.00)
2-dimensional problems		
Reached learning criterion /4	3.76 (0.44)	3.92 (0.28)
Correct solutions /4	3.92 (0.28)	3.92 (0.28)

8.4.2 FOUR-DIMENSIONAL PROBLEMS

The means and standard deviations for the performance measures in the report and report-aid conditions of the four-dimensional problems are shown in Table 8.3.

8.4.2.1 Correct Responses on Non-feedback Trials

As described in section 4.3.1.2, for each set of non-feedback trials, there are sixteen possible response patterns. Half correspond to the selection of a single hypothesis (one pattern for each hypothesis), while the remaining eight are indeterminate and suggest that no single hypothesis is being tested. The data were scored as described in section 4.4.2.1 (using the keys shown in Appendix 5 and Appendix 6) in order to determine which hypotheses, if any, were being tested by Ss on each non-feedback trial, and whether they were logically correct or not.

The groups were first compared on the number of logically correct hypotheses tested on the non-feedback trials, i.e. those which were logically consistent with all the feedback available at that stage of the problem. Thus, a hypothesis was scored as correct if it was one of the four logically correct hypotheses after feedback 1, one of the two logically correct hypotheses after feedback 2, or the only remaining hypothesis after feedback 3. Since there were three sets of non-feedback trials, and the set of potentially correct hypotheses became smaller at each stage, any effect of stage of testing (after 1st, 2nd or 3rd feedback trial) was examined to see whether the groups differed in the efficiency with which they progressively ruled out incorrect hypotheses (see Figure 8.1).

Unacceptable skewness and kurtosis were detected in some of the variables, and an arcsine transformation was applied to the data. This reduced both skewness and kurtosis, but they remained above the acceptable limit in some of the variables. Since a test was needed to carry out a comparison of the data at the three test stages, and no such nonparametric multivariate test was available, the data were analysed using repeated measures ANOVA, and the analysis was repeated using non-parametric tests. ANOVA was carried out with two within-subject factors (condition: report vs. report-aid; stage of testing: after 1st, 2nd, or 3rd feedback trial). The degrees of freedom were corrected using the Greenhouse-Geisser test. The group x condition x stage of testing interaction was not significant ($F=1.15$; $df=1.97,96$, $p=0.32$). The group x condition ($F=3.23$, $df=1,48$, $p=0.08$) and group x stage ($F=2.84$, $df=1.56,96$, $p=0.08$) interactions both approached significance. There was no significant overall effect of group ($F=2.75$, $df=1,48$, $p=0.10$).

Post-hoc t-tests were carried out to explore the interactions between group and condition, and group and stage, using a post-hoc significance level of 0.01 (0.05/5). The number of correct responses obtained by each group in the two experimental conditions was compared. This data was found to be positively skewed, but an arcsine transformation brought this within acceptable levels. The groups did not differ in the report condition ($t=0.53$, $df=48$, $p=0.60$), but the dysphoric Ss achieved significantly fewer responses than controls in the report-aid condition ($t=2.64$, $df=48$, $p=0.01$). The two groups were then compared on the number of correct responses which they achieved after each feedback collapsed across conditions. T-tests revealed that the groups did not differ after feedback 1 ($t=0.22$, $df=48$, $p=0.83$), or feedback 2 ($t=1.45$, $df=48$, $p=0.16$), but that the dysphoric Ss achieved significantly fewer correct responses after feedback 3 ($t=2.64$, $df=48$, $p=0.01$).

There was no significant condition x stage of testing interaction ($F=0.66$, $df=1.97,96$, $p=0.52$), nor an overall effect of condition ($F=1.30$, $df=1,48$, $p=0.26$). There was a significant effect of stage of testing ($F=21.19$, $df=1.56,96$, $p<0.001$), and the means revealed that for both groups the proportion of correct hypotheses increased on successive trials.

For the nonparametric analysis of the data, Mann-Whitney U-tests were used to compare the performance of the dysphoric and control groups at each stage of testing in the two conditions. In the report condition, there was no significant group difference after the first ($U=243.0$, $p=0.17$) or second ($U=259.0$, $p=0.28$) feedback trials, but there was a significant group difference after the third feedback trial ($U=222.0$, $p=0.05$), and the dysphoric group tested fewer logically correct hypotheses than the control group. In the report-aid condition, there was no significant group difference after the first feedback trial ($U=274.0$, $p=0.44$), but the difference approached significance after the second feedback trial ($U=217.5$, $p=0.06$), and reached significance after the third feedback trial ($U=195.0$, $p=0.008$), with the dysphoric Ss again testing fewer logically correct hypotheses. The ANOVA reported above found that the group x condition and group x stage of testing interactions both approached significance, while on the nonparametric analysis, the comparisons which correspond with these interactions were found to reach significance.

The number of logically correct responses given after each feedback trial was collapsed across groups, and the report and report-aid conditions were compared using Wilcoxon

tests. There was no significant difference in the number of logically correct responses given in the two conditions after feedback 1 (Z=-0.19, p=0.85), feedback 2 (Z=-0.27, p=0.79), or feedback 3 (Z=-1.49, p=0.14), consistent with the parametric analysis. Wilcoxon tests were also used to compare performance at the different stages of testing collapsed across the groups. In the report condition, there was no significant difference in the number of logically correct responses given after the 1st and 2nd feedback trials (Z=12.53, p=0.12), but Ss achieved significantly more logically correct responses after the 3rd feedback trial relative to both the 1st (Z=-3.54, p=0.0004) and 2nd (Z=-3.37, p=0.0008) feedback trials. The pattern was similar in the report-aid condition, with no difference after the 1st and 2nd feedback trials (Z=-1.51, p=0.13), but significantly more logically correct responses after the 3rd feedback trial relative to the 1st (Z=-4.05, p=0.0001) and 2nd (Z=-3.82, p=0.0001) feedback trials. This is consistent with the ANOVA which found a significant effect of stage of testing.

8.4.2.2 Incorrect Responses on Non-feedback Trials

The groups were compared on the number of logically incorrect hypotheses tested overall on non-feedback trials (see Figure 8.2). The data were found to have unacceptable skewness and kurtosis, but an arcsine transformation brought these within acceptable limits. Repeated measures ANOVA with one within-subject factor (condition) revealed no significant group x condition interaction (F=1.52, df=1,48, p=0.22), but a significant overall effect of group (F=4.86, df=1,48, p=0.03), with the dysphoric group testing more incorrect hypotheses than the control group. There was no significant effect of condition (F=0.04, df=1,48, p=0.84).

8.4.2.3 Indeterminate Responses on Non-feedback Trials

The evidence described in section 8.4.2.2 suggests that dysphoric Ss tested significantly more incorrect hypotheses on non-feedback trials than control Ss. The groups were also compared on the number of indeterminate responses made. Assuming that by chance consistent response patterns would be generated on 50% (8/16) of trials, it is possible to compare the actual number of hypothesis patterns with a random responding rate. The rates were found to be 77% and 76% for the dysphoric group, and 76% and 81% for the control group in the report and report-aid conditions respectively, which was clearly above chance levels for both groups. Repeated measures ANOVA revealed no significant group

x condition interaction (F=1.81, df=1,48, p=0.19); nor a significant effect of group (F=0.23, df=1,48, p=0.63), or condition (F=0.51; df=1,48, p=0.19) (see Figure 8.3).

8.4.2.4 Correct Solutions

The correct solution to each problem was ascertained using the procedure described in section 4.4.2.1, and these are shown in Figure 8.4. The groups were compared on the number of correct solutions achieved in the two conditions (report vs. report-aid). Skewness and kurtosis were found in some of the variables, and an arcsine transformation failed to reduce these to acceptable levels. Since no suitable nonparametric multivariate test was available, repeated measures ANOVA was carried out in addition to a nonparametric analysis of the data. ANOVA revealed a significant group difference (F=10.99, df=1,48, p=0.002) which did not interact significantly with condition (F=0.01, df=1,48, p=0.93). Mean scores showed that the dysphoric Ss achieved fewer correct solutions than the control Ss in both conditions (see Figure 8.4). There was no significant effect of condition (F=0.01, df=1,48, p=0.91).

For the nonparametric analysis, the number of correct solutions achieved by the dysphoric and control groups in the two conditions was compared using Mann-Whitney U-tests. There was a significant group difference in both the report (U=218.0, p=0.02) and report-aid (U=204.5, p=0.01) conditions, confirming the significant overall group difference reported above. A Wilcoxon test was used to compare the number of correct solutions achieved in the two conditions collapsed across groups. There was no significant difference between the conditions (Z=-0.3429, p=0.73), again confirming the results of the ANOVA.

8.4.2.5 Correct Responding on Non-feedback Trials After Positive and Negative Feedback

As described in section 4.1, different types of processing are necessary following positive and negative feedback trials (Levine, 1966). Analyses were therefore carried out to examine the rate of logically correct responses on non-feedback trials immediately after positive and negative feedback. The results are illustrated in Figure 8.5. Repeated measures ANOVA with two within-subject factors (condition: report vs. report-aid; positive vs. negative feedback) was carried out for logically correct responses following positive and negative feedback. The three-way interaction involving group was not

significant (F=0.91, df=1,48, p=0.34), and nor was the group x feedback interaction (F=0.00, df=1,48, p=1.00) significant. The group x condition interaction did reach significance (F=3.96, df=1,48, p=0.05). Post-hoc t-tests with an adopted significance level of 0.025 were carried out for the number of correct responses collapsed across positive and negative feedback for the two groups in the report and report-aid conditions. There was no significant difference between the groups in the report condition (t=-0.53, df=48, p=0.60), but there was a significant group difference in the report-aid condition (t=-2.61, df=48, p=0.01). Inspection of the means revealed that the dysphoric Ss made fewer correct responses than the controls.

There was no significant overall effect of group (F=2.77, df=1,48, p=0.10), or condition (F=0.99, df=1,48, p=0.33). There was a significant overall effect of feedback (F=5.30, df=1,48, p=0.03), since Ss in both groups tended to make more correct responses after positive feedback than after negative feedback.

8.4.2.6 Consistency of Non-feedback Trials with Earlier Feedback

As described in section 4.4.2.5, Eimas (1970) reported on a measure which gives information about which of the feedback trials Ss actually took into account when responding on non-feedback trials. The groups were compared on the extent to which hypotheses were consistent with feedback trials that had 0, 1 or 2 intervening feedback trials (see Figure 8.6). The data were found to have skewness and kurtosis, and an arcsine transformation failed to reduce these to acceptable levels. Since no nonparametric multivariate test was available to analyse the data, repeated measures ANOVA was carried out in addition to nonparametric tests. ANOVA with two within-subject factors (condition; 0, 1 or 2 intervening feedback trials) was carried out. The degrees of freedom were adjusted using a Greenhouse-Geisser test. The three way group x condition x number of intervening feedback trials interaction was not significant (F=0.03, df=1.50,96, p=0.93), nor were the two way interactions of group x condition (F=1.73, df=1,48, p=0.20) or group x number of intervening feedback trials (F=1.19, df=1.48,96, p=0.30) significant. There was a significant effect of group overall (F=5.10, df=1,48, p=0.03), and the means showed that the hypotheses tested by dysphoric Ss were less likely to be consistent with feedback, regardless of the number of intervening feedback trials (see Figure 8.6).

There was no significant condition x number of intervening feedback trials interaction ($F=0.37$, $df=1.50,96$, $p=0.63$); nor was there a significant effect of condition ($F=2.07$, $df=1,48$, $p=0.16$). There was a significant effect of number of intervening feedback trials ($F=55.51$. $df=1.48,96$, $p<0.001$), and inspection of the means revealed that for both groups hypotheses were more likely to be consistent with earlier feedback trials (see Figure 8.6).

The analysis was repeated using nonparametric tests. The dysphoric and control groups were compared on the number of hypotheses that were consistent with feedback trials with 0, 1 and 2 intervening feedback trials in the two conditions using Mann-Whitney U-tests. In the report condition, there was no difference between the groups for either 0 ($U=298.5$, $p=0.87$), 1 ($U=240.5$, $p=0.15$) or 2 ($U=256.0$, $p=0.19$) intervening feedback trials. However, in the report-aid condition the group differences approached significance for 0 ($U=222.0$, $p=0.08$) intervening feedback trials, and reached significance for both 1 ($U=186.5$, $p=0.01$) and 2 ($U=222.0$, $p=0.02$) intervening trials. Inspection of the means revealed that the dysphoric group tested fewer consistent hypotheses in both cases.

The number of hypotheses that were consistent with feedback with 0, 1 and 2 intervening feedback trials in the two experimental conditions was compared using Wilcoxon tests with the data collapsed across groups. There was no significant difference between the two conditions for either 0 ($Z=-0.65$, $p=0.52$), 1 ($Z=-0.5$, $p=0.62$) or 2 ($Z=-1.03$, $p=0.30$) intervening trials. Wilcoxon tests were also used to compare the number of hypotheses that were consistent with feedback with 0, 1 and 2 intervening feedback trials collapsed across condition. Ss tested significantly more hypotheses that were consistent with feedback with 2 intervening feedback trials than with either 1 ($Z=-6.21$, $p<0.0001$) or 0 ($Z=-6.21$, $p<0.0001$) intervening trials. They also tested more hypotheses that were consistent with 1 than 0 intervening trials ($Z=-5.99$, $p<0.0001$).

The two analyses were consistent in finding that Ss were more likely to test hypotheses that were consistent with earlier feedback trials. When the two groups were compared, there was some inconsistency between the parametric and nonparametric analyses. The ANOVA revealed an overall deficit in the dysphoric group which did not interact with condition or number of intervening trials, but on the nonparametric analyses, the group difference reached significance only on the 2- and 3-intervening trials in the report-aid condition.

Table 8.3 **Means (and standard deviations) for performance measures in the report and report-aid conditions on the four-dimensional problems**

	Dysphoric		Controls	
	Report	Report-aid	Report	Report-aid
Correct on non-feedback trials:				
After feedback 1 /4	2.52	2.08	1.96	2.36
	(1.09)	(1.35)	(1.46)	(1.44)
Arcsine	1.90	1.61	1.55	1.82
	(0.73)	(0.95)	(1.04)	(1.04)
After feedback 2 /4	2.32	2.28	2.68	2.92
	(1.25)	(1.21)	(1.15)	(1.04)
Arcsine	1.81	1.83	2.04	2.22
	(0.89)	(0.81)	(0.81)	(0.74)
After feedback 3 /4	2.88	3.04	3.48	3.80
	(1.30)	(1.17)	(1.01)	(0.41)
Arcsine	2.20	2.32	2.66	2.88
	(0.96)	(0.86)	(0.78)	(0.40)
Total incorrect /12	1.52	1.76	1.04	0.68
	(1.85)	(1.81)	(1.70)	(1.11)
Arcsine	0.57	0.65	0.43	0.32
	(0.50)	(0.49)	(0.47)	(0.38)
Total indeterminate /12	2.76	2.92	2.84	2.32
	(2.20)	(2.29)	(1.97)	(1.93)
Correct solutions /4	3.28	3.32	3.80	3.84
	(0.94)	(0.85)	(0.50)	(0.37)
Arcsine	2.50	2.50	2.90	2.92
	(0.71)	(0.66)	(0.43)	(0.37)
After +ve feedback /6	4.20	3.68	4.28	4.64
	(1.53)	(1.35)	(1.31)	(1.15)
After -ve feedback /6	3.52	3.72	3.84	4.44
	(1.64)	(1.49)	(1.63)	(1.08)
Consistent with feedback:				
0 intervening feedback trials %	68.67	67.33	70.33	77.33
	(19.73)	(19.08)	(19.56)	(16.93)
Arcsine	2.01	1.99	2.10	2.24
	(0.50)	(0.52)	(0.55)	(0.51)
1 intervening feedback trial %	73.50	74.00	82.00	86.50
	(23.47)	(19.07)	(18.43)	(13.46)
Arcsine	2.16	2.15	2.40	2.53
	(0.70)	(0.54)	(0.56)	(0.49)
2 intervening feedback trials %	84.00	85.00	91.00	98.00
	(23.80)	(26.02)	(17.50)	(6.92)
Arcsine	2.52	2.58	2.78	3.00
	(0.75)	(0.79)	(0.56)	(0.27)

8.4.2.7 Hypothesis-listing

The means and standard deviations for the performance measures relating to hypothesis-listing performance are shown in Table 8.4.

In both the report and report-aid conditions, Ss listed the hypotheses which they believed to be correct after each feedback trial. These were scored by comparing each list with the hypotheses that were logically correct at that stage of the problem, i.e. four after feedback 1, two after feedback 2, and one after feedback 3. The data for the number of problems with correct hypothesis lists were found to be skewed, and an arcsine transformation failed to reduce this to acceptable levels. A multivariate test was needed to carry out the analysis, but since a nonparametric multivariate test was not available, the analysis was carried out using repeated measures ANOVA and then repeated using nonparametric tests. ANOVA revealed a significant group difference ($F=12.92$, $df=1,48$, $p=0.001$), and this did not interact with condition ($F=0.00$, $df=1,48$, $p=1.00$), nor was there a significant effect of condition ($F=1.20$, $df=1,48$, $p=0.28$). The dysphoric group achieved fewer correct lists of hypotheses (see Figure 8.7).

This analysis was repeated with nonparametric tests. Mann-Whitney U-tests were used to compare the dysphoric and control groups for the number of correct hypothesis lists in the two conditions. There were significant group differences in both the report ($U-171.0$, $p=0.0041$) and report-aid ($U=158.5$, $p=0.0016$) conditions, confirming the significant group difference reported above. The number of correct hypothesis lists given in the two conditions was compared collapsed across groups using a Wilcoxon test. There was no significant difference between the two conditions ($Z=-1.16$, $p=0.2478$), again consistent with the parametric analysis.

The errors made by Ss were coded as either errors of omission, where a correct hypothesis had not been listed, or errors of over-inclusion, where an incorrect hypothesis had been listed. Since the total number of possible omission errors (7 per problem) was smaller than the total possible over-inclusion errors (17 per problem), error rates were analysed in terms of proportions of the total possible. The number of omission and over-inclusion errors made in the hypothesis lists were found to have skewness and kurtosis. A log 10 transformation was not strong enough to reduce these, so a stronger transformation was applied ($-1/(x+0.01)$). This reduced them further, but both were still above an acceptable

level. Therefore, for the analysis of both variables, repeated measures ANOVA was carried out as a multivariate analysis of the data, but the analysis was repeated using nonparametric tests.

For errors of omission, ANOVA revealed a significant group difference ($F=13.04$, $df=1,48$, $p=0.001$) which did not interact with condition ($F=0.06$, $df=1,48$, $p=0.802$). The dysphoric Ss made more errors of omission than control Ss (see Figure 8.8). There was no significant effect of condition ($F=0.39$, $df=1,48$, $p=0.54$).

This analysis was repeated using nonparametric tests. Mann-Whitney U-tests were used to compare the number of omission errors made by the dysphoric and control groups in the report and report-aid conditions. There was a significant group difference in both the report ($U=196.0$, $p=0.0105$) and report-aid ($U=202.5$, $p=0.0195$) conditions, confirming the significant group difference reported above. A Wilcoxon test was applied to compare the number of omission errors made in the two conditions collapsed across the groups. There was no significant difference between the conditions ($Z=-0.36$, $p=0.72$) which was again consistent with the parametric analysis.

A similar pattern was found for the over-inclusion errors. Repeated measures ANOVA revealed a significant group difference ($F=13.92$, $df=1,48$, $p=0.001$), which did not interact with condition ($F=0.09$, $df=1,48$, $p=0.763$). The means revealed that the dysphoric Ss made more over-inclusion errors than the control Ss (see Figure 8.9). There was no significant effect of condition ($F=1.06$, $df=1,48$, $p=0.308$).

Mann-Whitney U-tests revealed significant group differences in the number of over-inclusion errors in both the report ($U=159.0$, $p=0.0021$) and report-aid ($U=166.0$, $p=0.0026$) conditions, confirming the significant group difference reported above. A Wilcoxon test was carried out to compare the number of over-inclusion errors made in the two conditions collapsed across the groups, and this revealed a significant difference ($Z=-2.0517$, $p=0.0402$). An inspection of the means revealed that there were more errors of over-inclusion in the report relative to the report-aid condition. This was not consistent with the finding of no significant effect of condition on the ANOVA.

Table 8.4 **Means (and standard deviations) for performance on hypothesis listing in the report and report-aid conditions on the four-dimensional problems**

	Dysphoric		Controls	
	Report	Report-aid	Report	Report-aid
Correct hypothesis lists				
after each feedback	1.72	1.92	3.08	3.28
/4	(1.70)	(1.61)	(1.32)	(1.24)
Arcsine	1.35	1.49	2.36	2.50
	(1.24)	(1.18)	(0.98)	(0.94)
Errors of omission	8.71	8.00	3.86	4.00
as % of possibles	(16.14)	(14.99)	(12.41)	(11.63)
-1(x+0.01)	-47.43	-44.75	-82.03	-75.72
	(44.13)	(42.69)	(36.80)	(39.96)
Errors of over-inclusion				
as % of total	21.12	15.53	4.29	2.94
possible	(25.14)	(22.20)	(9.24)	(6.66)
-1/(x+0.01)	-30.94	-35.73	-64.60	-73.38
	(40.68)	(41.49)	(41.83)	(40.12)

8.4.2.8 Correlation of Performance with Severity of Depressive Symptomatology

Pearson correlation coefficients were calculated between each of the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2). This was carried out within each group because the selection criteria for the two groups (dysphoric >10; nondysphoric <6) ensured that, taken together, the scores on the BDI could not have a normal distribution. There were no significant results with either the BDI1 or the BDI2. Table 8.5 shows the results for the BDI2. This is the measure which was taken during the actual test session, and might therefore be expected to show a higher correlation with task performance than the BDI1 which was completed as part of the initial screening process (see section 4.3.3.2).

Table 8.5 Correlation of BDI2[+] with performance measures on the four-dimensional Discrimination Learning problems

	Dysphoric		Controls	
	Report	Report-aid	Report	Report-aid
Correct on non-feedback trials:				
After feedback 1 /4	.27	.05	.22	.36
Arcsine	.28	.03	.19	.33
After feedback 2 /4	-.14	-.01	.32	.13
Arcsine	-.14	-.05	.36	.12
After feedback 3 /4	.34	.10	.22	-.12
Arcsine	.35	.06	.25	-.12
Total incorrect /12	-.21	-.18	-.36	-.03
Arcsine	-.18	-.16	-.46	-.10
Total indeterminate /12	.08	-.07	.15	.27
Correct solutions /4	.24	-.10	.25	-.24
Arcsine	.25	-.12	.25	-.24
After +ve feedback /6	.22	.24	.27	.27
After -ve feedback /6	.13	-.15	.35	.26
Consistent with feedback:				
0 intervening feedback trials %	.11	.06	.33	.27
Arcsine	.09	.03	.33	.19
1 intervening feedback trial %	-.00	.08	.27	.12
Arcsine	-.01	.00	.29	.08
2 intervening feedback trials %	.34	.05	.12	.20
Arcsine	.38	.01	.15	.20
Correct hypothesis lists after each feedback /4	.29	.17	.19	.02
Arcsine	.28	.19	.21	.02
Errors of omission as % of total possible	-.15	-.14	-.20	-.18
-1(x+0.01)	-.12	-.14	-.12	-.07
Errors of over-inclusion as % of total possible	-.25	-.16	-.10	.06
1/(x+0.01)	-.25	-.21	-.12	.03

*p<.01; **p<.001, two-tail. "/" indicates that a coefficient could not be computed.
[+]BDI2 = Beck Depression Inventory given on the second occasion (see section 4.3.3.1).

8.5 SUMMARY OF RESULTS

1. Both groups were at ceiling on the one-dimensional problems, and showed a similar degree of minimal impairment on the two-dimensional problems.

2. ANOVA revealed non-significant trends for dysphoric Ss to test fewer logically correct hypotheses than controls in the report-aid condition, and after the third feedback trial. When the comparisons were repeated using nonparametric tests, these differences reached significance.

3. The dysphoric Ss tested significantly more logically incorrect hypotheses on the non-feedback trials than controls.

4. Ss in both groups tested hypotheses on the majority of the non-feedback trials, and there was no significant group difference on this measure.

5. Dysphoric Ss achieved significantly fewer correct solutions than controls.

6. Dysphoric Ss were significantly less likely to test a logically correct hypothesis after feedback than controls in the report-aid condition, but there was no group difference in the report condition.

7. Dysphoric Ss were significantly less likely than controls to test hypotheses that were consistent with earlier feedback trials, regardless of the number of intervening feedback trials.

8. Dysphoric Ss were significantly less likely than controls to list the correct hypotheses after each feedback trial, and they made both more omission and more over-inclusion errors.

9. There were no significant correlations between the performance measures and scores on the Beck Depression Inventory (BDI1 and BDI2).

8.6 DISCUSSION

The aim of this experiment was to investigate the role of strategy and storage in the performance of dysphoric Ss on a Discrimination Learning task. It was predicted that dysphoric Ss would be differentially facilitated by manipulations which cued the use of an appropriate strategy and reduced the storage demands of the task.

The Discrimination Learning task was selected for further study because it had proved to be sensitive to depressive deficits both in the study carried out by Silberman et al. (1983) with clinically depressed patients, and in the experiment reported in chapter IV using dysphoric Ss. It was therefore predicted that dysphoric Ss would again be impaired relative to controls on the Discrimination Learning task in the current experiment. The findings reported in chapter 4 showed dysphoric Ss to be impaired relative to controls on the four-dimensional Discrimination Learning problems, with intact performance on the simpler one- and two-dimensional problems. On the four-dimensional problems, dysphoric Ss had difficulties in achieving correct solutions and listing the correct hypotheses in both the report and report-aid conditions. The dysphoric Ss were not differentially facilitated by the manipulations on any of the performance measures. For several of the measures parallel parametric and nonparametric analyses were carried out due to skewness and kurtosis in the data. Although there was discrepancy between the findings of the two analyses in several cases, there was no instance in which either analysis found evidence to support the prediction that dysphoric Ss would be differentially facilitated by the experimental manipulations.

Comparing the findings of the current experiment with those of the experiment reported in chapter IV using similar Discrimination Learning problems, there were differences in the pattern of performance. On the two-dimensional pre-training problems the control group were at ceiling in chapter IV, but performed at the same level as the dysphoric Ss in the current experiment, although the experimental paradigm was identical in both cases. The Ss in both studies were recruited using the same procedure, and had comparable BDI scores. It therefore seems likely that the differences are due to slight random fluctuation between the samples; the dysphoric groups were comparable in the two experiments on the two-dimensional problems.

On the four-dimensional problems, there was evidence that the deficits shown by the dysphoric group in the current experiment were less pervasive than those reported in chapter IV. Nevertheless, dysphoric Ss in the current experiment were still significantly less likely to give the correct solution, or to list the correct hypotheses after feedback than controls. The difference between the two experiments may be because the control group in the current experiment was close to ceiling on many of the measures, potentially masking group differences.

The complexity of the Discrimination Learning problems was found to be sensitive to dysphoric deficits in chapter IV. On the one- and two-dimensional problems in the pretraining phase of both the current experiment and that reported in chapter IV, the dysphoric Ss had minimal impairments, although these problems were based on the same logical principle as the four-dimensional problems which are reliably sensitive to depressive deficits. Increasing the complexity of the Discrimination Learning problems in this way increases task demands on both storage and processing capacity, both of which were investigated by direct manipulations in the current experiment.

It was predicted that reducing the storage demands of the task would differentially facilitate the dysphoric Ss. On the four-dimensional problems, the display of Ss' lists of hypotheses throughout the problem in the report-aid condition was designed to remove the necessity to hold in memory or regenerate the hypotheses which remained viable after each feedback trial. This did not appear to aid the dysphoric group differentially, since there was no significant interaction of group with condition in which depressive performance was facilitated. This is consistent with the findings reported in chapters IV and V that memory-aid manipulations involving the Discrimination Learning and Integrative Reasoning tasks respectively failed to facilitate differentially the performance of the dysphoric Ss. A comparison of the report and report-aid conditions in the current experiment revealed no significant effects involving condition alone. Only one measure showed a significant group by condition interaction, and another approached significance in the parametric analysis and reached it in the nonparametric analysis. However, contrary to prediction, it was the control group who was facilitated by the report-aid condition. This may indicate that failure to retain the list of potential hypotheses does not play a significant role in any performance deficits on this task for either group, although it is possible that the control Ss were so close to ceiling in the report condition that the report-

aid manipulation could have only a limited impact on performance. Nevertheless, the failure of the report-aid manipulation in this experiment, and the memory-aid manipulations in chapters IV and V to differentially facilitate the performance of dysphoric Ss suggests that deficits associated with depression are unlikely to be explained solely in terms of an underlying storage deficit.

It was also predicted that cuing an appropriate strategy by focusing Ss' attention on the task at the critical stages would differentially facilitate the performance of the dysphoric Ss. The fact that the dysphoric group remained significantly impaired relative to the control Ss in both conditions suggests that this manipulation failed to facilitate depressive performance to normal levels, although without a direct comparison with performance on the standard condition the possibility that dysphoric Ss were differentially facilitated to some lesser extent cannot be excluded. Silberman et al. (1983) compared clinically depressed and normal control Ss in the standard and report conditions, and found that depressed Ss were differentially aided on some aspects of task performance, but remained below normal levels overall.

The cognitive initiative model of cognitive function in depression put forward by Hertel and her colleagues (see section 2.3.9.2) suggests that deficits associated with depression are not invariant, and they have reported work showing that giving instructions which cue appropriate strategies can facilitate the performance of depressed Ss to normal levels on memory tasks. In the current experiment, although the report condition was designed to focus attention at the critical stages of the task, it did not provide direct cues to assist Ss to work out or implement an adequate strategy for task performance. It is difficult to generate a means of cuing an adequate performance strategy more directly in a task of this type, whilst retaining the reasoning difficulty of the task. Further cues for performance could have been provided, but these would have also served to make the reasoning components of the task easier.

The hypothesis-listing manipulation provided useful insight into the depressive deficits observed in task performance. The dysphoric group made significantly more omission errors than controls, showing an impaired ability to list the relevant correct hypotheses after each feedback trial. The dysphoric Ss were also significantly more likely than controls to make over-inclusion errors, signifying a failure to eliminate hypotheses. It is

of interest that the dysphoric Ss made significantly more of both error-types than controls. If the group differences had occurred on only one type of error then this would have suggested a specific deficit associated with depression, while the actual pattern of findings confirms the finding reported in chapter IV that the performance of dysphoric Ss on this task is suggestive of more general difficulties in using feedback appropriately on the four-dimensional problems.

Consistent with the experiments reported in chapters IV to VII, there were no significant correlations between performance measures and scores on the BDI (BDI1 and BDI2). It has been suggested before (e..g. section 4.6) that this could be due either to a lack of any significant relationship between severity of depressive symptomatology and task performance, or to a failure of the measures in detecting an existing relationship. In the current experiment there was strong evidence that the dysphoric group was impaired relative to the controls on many of the task measures, and therefore a relationship with the severity of depressive symptoms was predicted. With regard to the possibility that the measures may have obscured a relationship, it should be noted that correlations are more difficult to detect if the range of scores is restricted in some way. This was a particular problem for the control group, where there was a possible range of only 0-5 on the BDI, making it unlikely that any meaningful relationship with depressive symptomatology could be identified. The dysphoric group had a greater range of scores on the BDI (11-39 on BDI2), but there was a restricted range on many of the task measures. For example, the number of correct solutions had a range of 0-4 in each condition. It is therefore possible that the measures were not well-suited to identifying correlations, although it is possible that a relationship between depressive symptomatology and task performance simply did not exist.

In summary, dysphoric Ss were found to be impaired relative to controls in both the report and report-aid conditions of the Discrimination Learning problems. These deficits appeared to reflect impaired processing and manipulation of information, and the dysphoric Ss were not facilitated to normal levels by manipulations which reduced storage demands or cued an appropriate strategy.

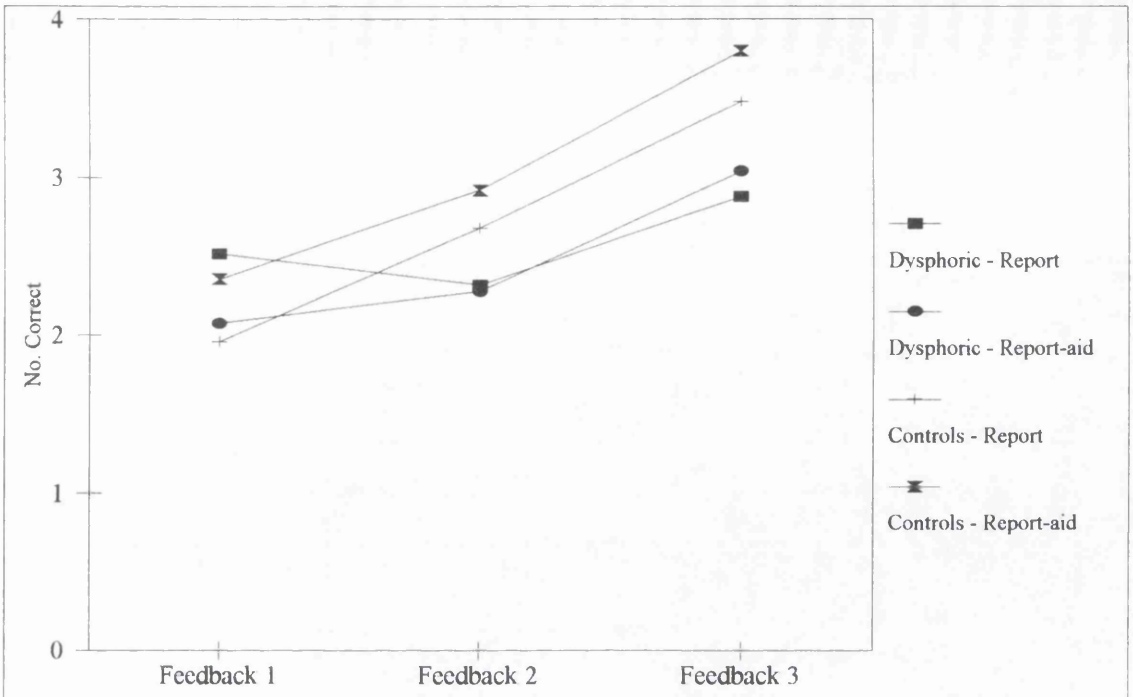

Figure 8.1 Responding on Non-feedback trials - No. Correct /4 After Feedbacks 1-3 in the report and report-aid conditions of the Discrimination Learning task

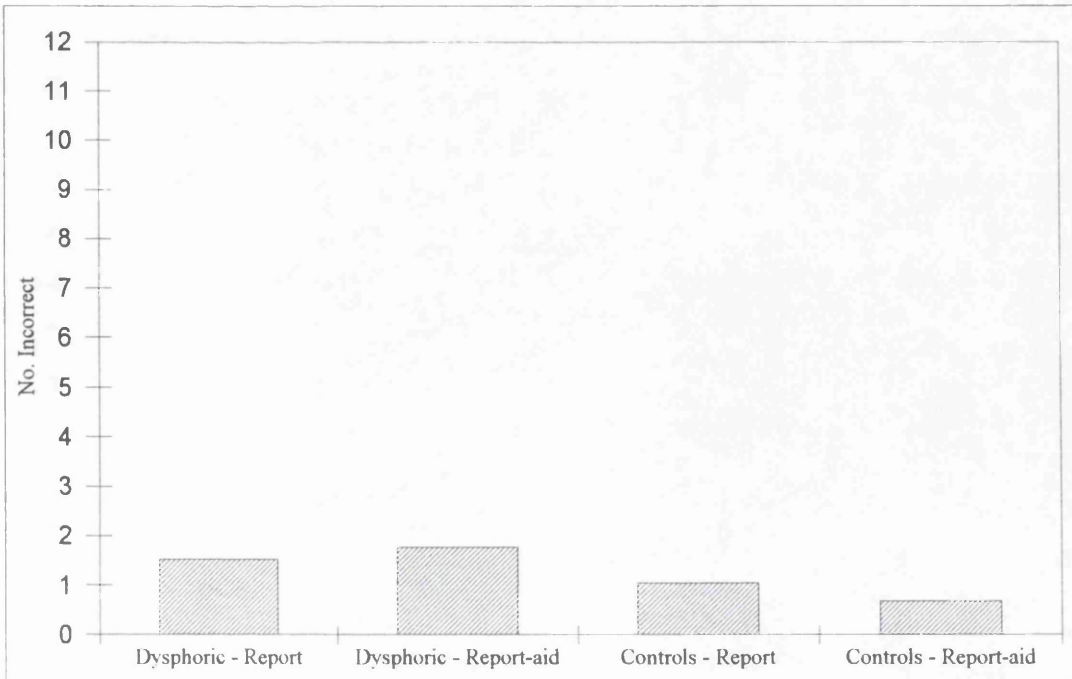

Figure 8.2 Total Number of Logically Incorrect Responses /12 in the report and report-aid conditions of the Discrimination Learning task

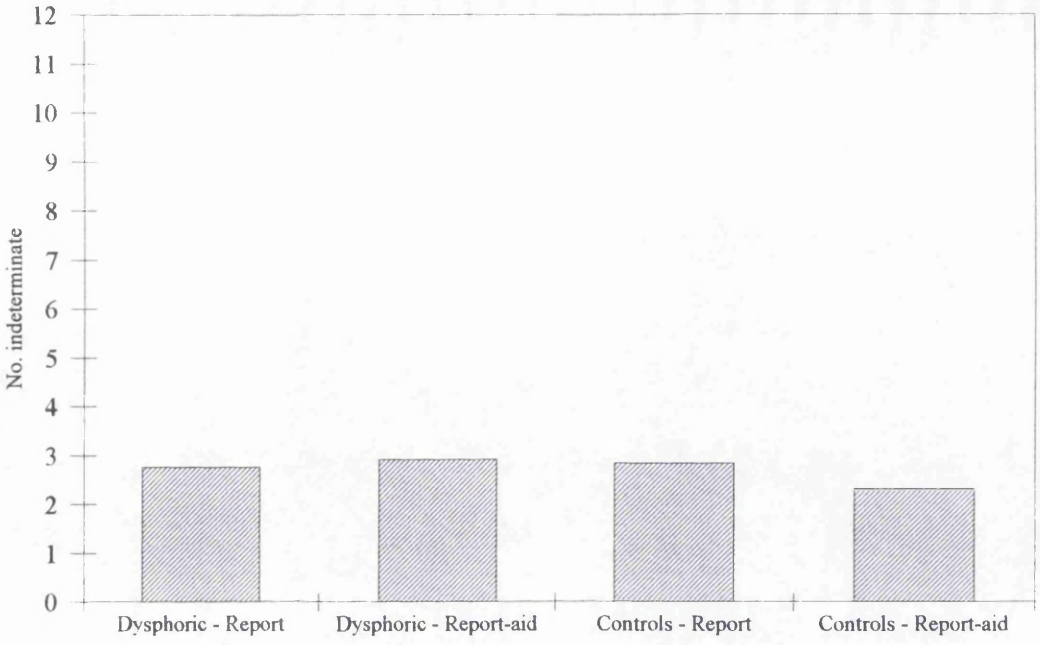

Figure 8.3 Total number of indeterminate responses /12 in the report and report-aid conditions of the Discrimination Learning task

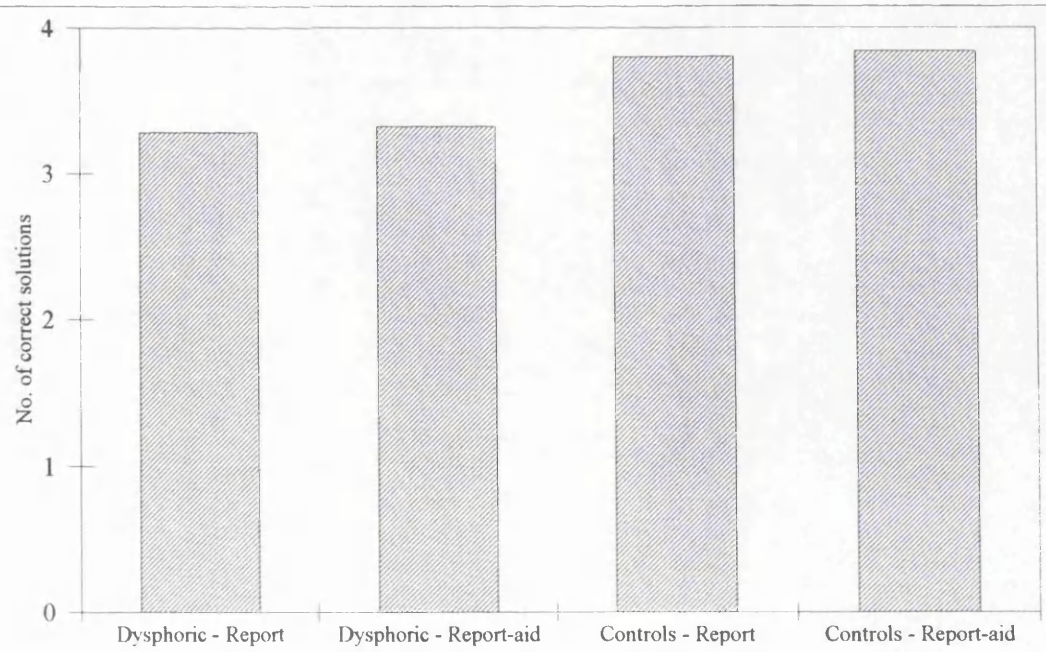

Figure 8.4 Total number of correct solutions /4 in the report and report-aid conditions of the Discrimination Learning task

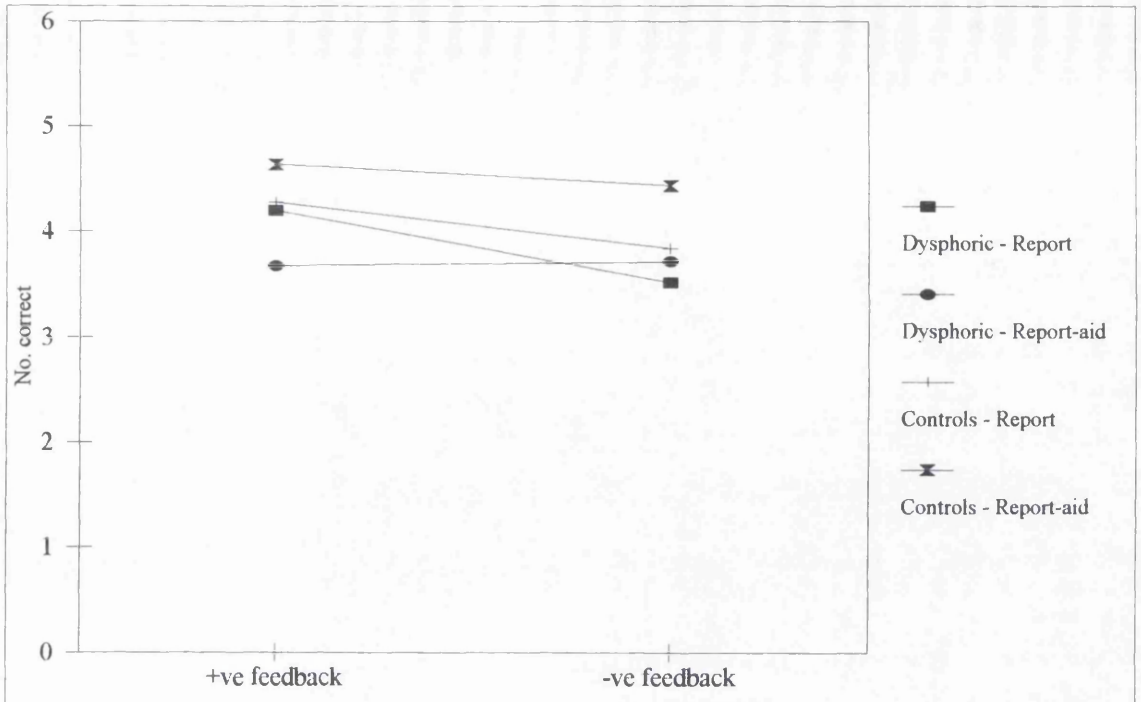

Figure 8.5 Logically correct responding /6 on non-feedback trials after positive and negative feedback in the report and report-aid conditions of the Discrimination Learning task

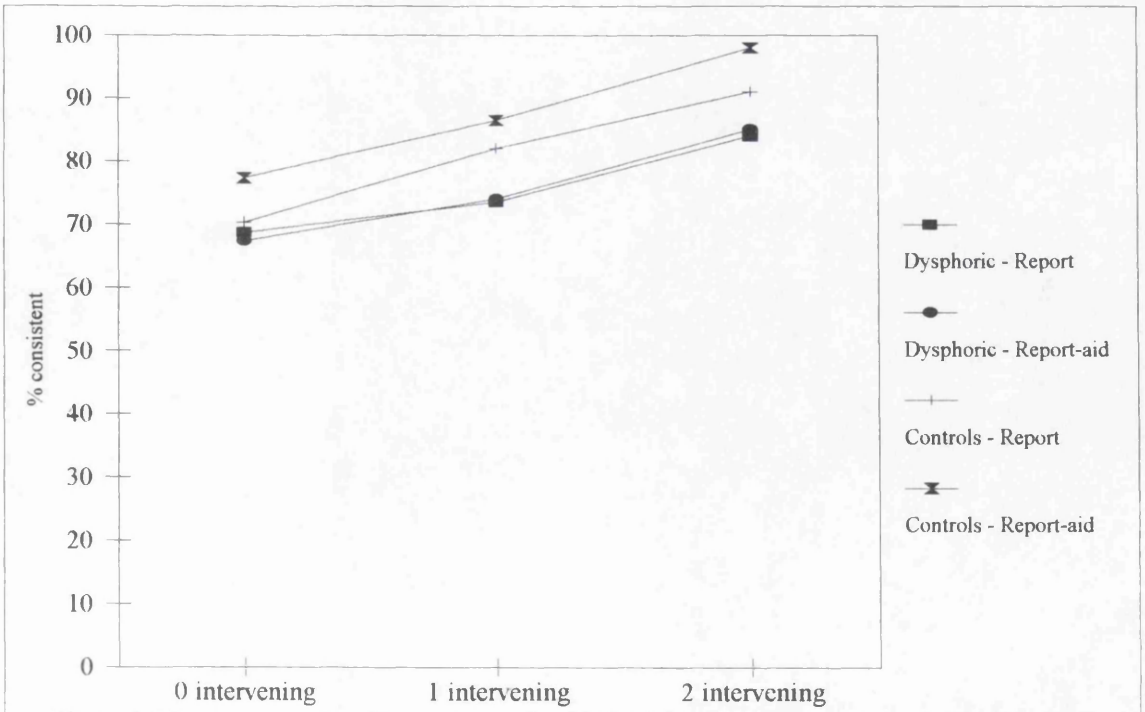

Figure 8.6 Percentage of hypotheses on non-feedback trials consistent with prior feedback in the report and report-aid conditions of the Discrimination Learning task

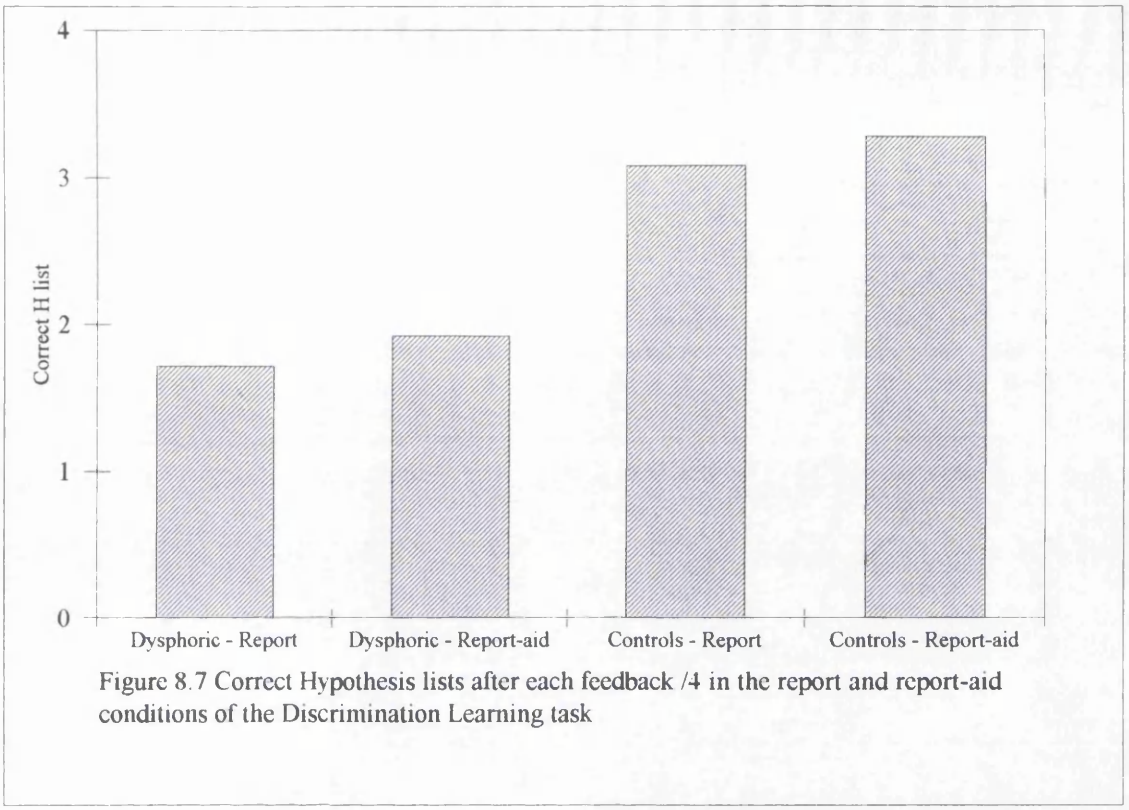

Figure 8.7 Correct Hypothesis lists after each feedback /4 in the report and report-aid conditions of the Discrimination Learning task

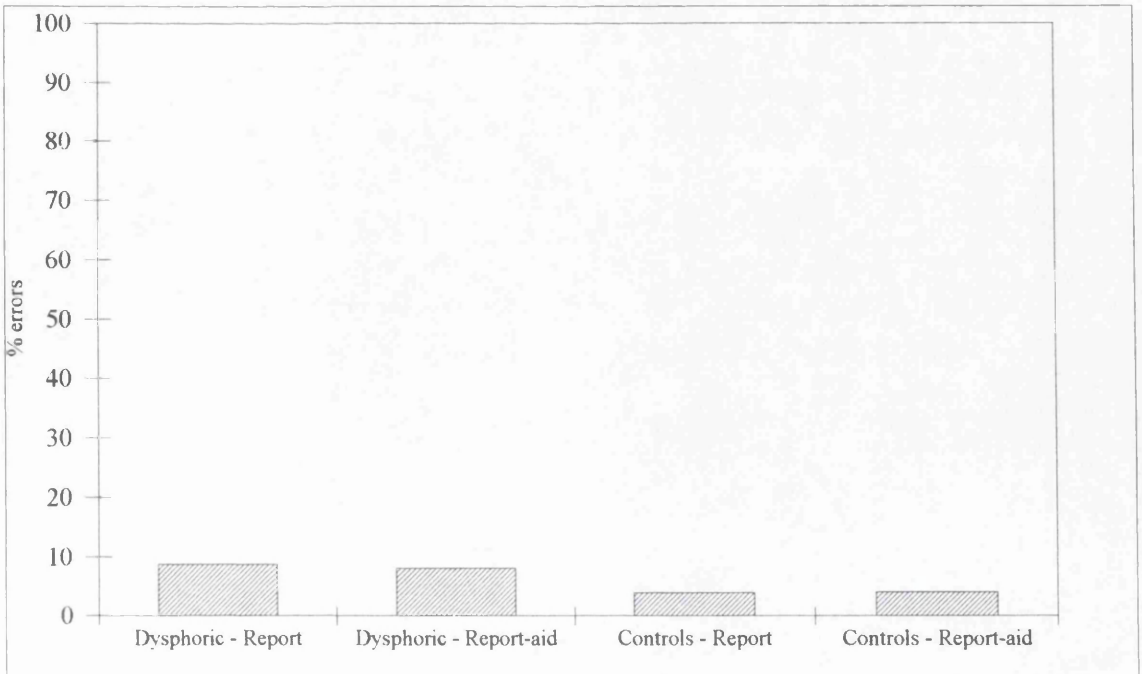

Figure 8.8 Errors of omission as % of possible total in the report and report-aid conditions of the Discrimination Learning task

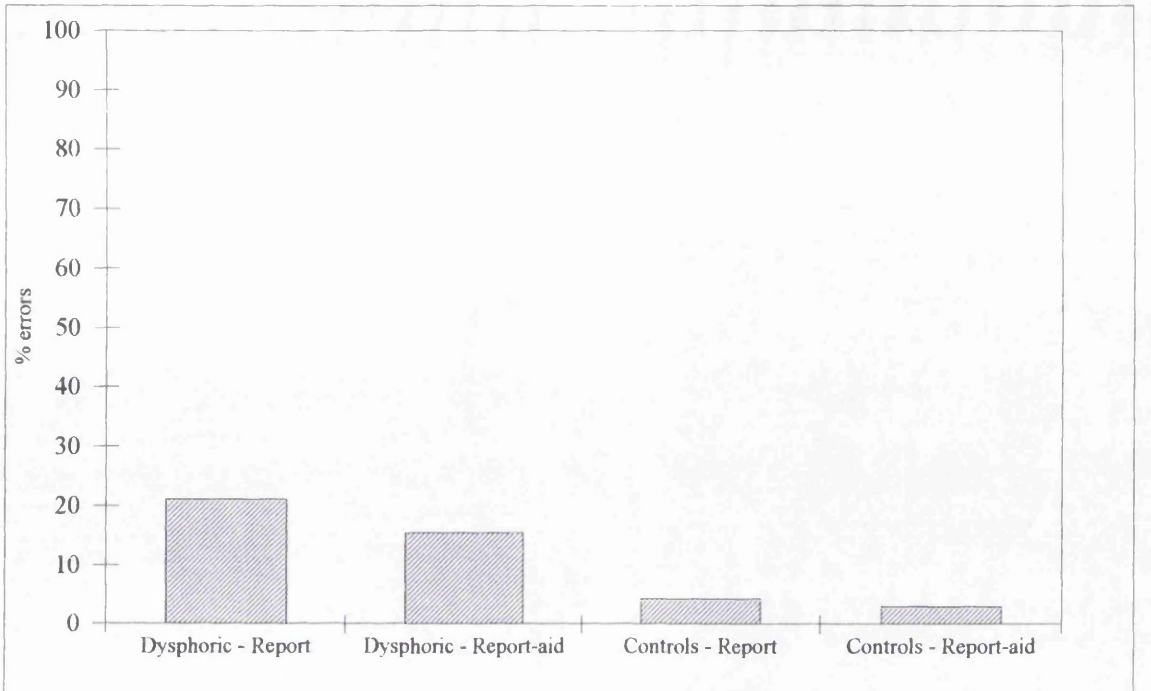

Figure 8.9 Errors of over-inclusion as % of possible total in the report and report-aid conditions of the Discrimination Learning task

CHAPTER IX

GENERAL DISCUSSION & CONCLUSIONS

9.1 CRITIQUE OF METHODOLOGY

The aim of this series of experiments was to explore the nature and extent of any reasoning deficits in depression; to investigate the mechanisms that might underlie them; to relate the findings to current models of cognitive function in depression; and to consider any implications for the development of remedial strategies. Before reviewing these issues, it is necessary to evaluate whether the methodology of the current series of experiments was appropriate.

First, it is necessary to consider whether the experimental tasks were adequate measures of reasoning performance. All the tasks in the current series of experiments had previously been reported in the cognitive psychology and/or neuropsychology literature. They were novel to the Ss taking part in each experiment and used material that was abstract and neutral. The disadvantage of using tasks of this type is that in everyday life people often have to reason with material which relates to their past experience and knowledge. This means that care must be taken when trying to make generalisations from the current series of experiments. The findings may be most relevant to other situations where problems that are novel, abstract and neutral are being addressed.

The advantage of using tasks that are novel, abstract and neutral is that the impact of any individual variation in past experience and existing knowledge is reduced, which is desirable when the aim is to measure performance abilities. The review of the literature in section 2.1.3.3 regarding whether depression is associated with reasoning deficits on tasks containing personally-relevant or emotionally-salient material illustrated the profound effects that pre-existing beliefs and knowledge can have on performance. This literature suggested qualitative differences between depressed and nondepressed Ss in their performance on these tasks, with significant group differences reported by the majority of studies. However, on the basis of this literature it was not possible to reach any firm conclusions with regard to depressive reasoning *deficits*.

In the current thesis, five experiments (see chapters IV to VIII) were carried out with students identified as dysphoric or nondysphoric on the basis of their scores on the BDI. This is a standard paradigm which has been used widely over the past twenty-five years

(see section 3.2.1.1). The BDI was given on two occasions to minimise the chances of selecting Ss with a transient elevation in their depressive symptomatology. In each of the five experiments reported in chapters IV to VIII, the dysphoric group scored significantly higher than the control group on both test occasions. An inspection of the mean BDI scores for the dysphoric group in each of the five experiments reveals that these ranged from 18.7 to 23.1 on BDI1 and from 17.7 to 19.6 on BDI2. This is higher than one might expect with a cut-off of 11, but is probably a result of the recruitment procedure (see section 4.3.3.1). After administering the BDI on the first occasion (BDI1), the dysphoric samples contained Ss with a full range of scores from 11 upwards, but in each experiment some were excluded because their score fell below 11 on the second occasion (BDI2). It is unsurprising that Ss with higher scores on the BDI were less likely to be excluded, and has the benefit of increasing the validity of this approach. In fact during the course of the data collection, a number of Ss had to be referred to their General Practitioner for treatment of their depression.

The main aim of the current series of experiments was to study the effects of dysphoric mood (the IV) on reasoning task performance (the DV). Before drawing any conclusions about the effects of dysphoric mood, it is important to establish that the experimental groups did not differ on any other variable which might influence reasoning task performance, thereby potentially confounding the results. Studies of cognitive function in depression reviewed in section 2.3 have identified age, sex, educational level and IQ as potential confounding factors, and steps were taken in the current studies to control for these. In each of the experiments reported in chapters IV to VIII there were no significant differences between the dysphoric and nondysphoric groups in age, sex, or IQ (score on the Vocabulary sub-test of the WAIS-R). Educational level was controlled for by the fact that all Ss were undergraduate students. Therefore, the results of the experiments in chapters IV to VIII cannot be explained in terms of the confounding effects of age, sex, IQ or educational level. Furthermore, a brief history was taken from all Ss, and any with significant sensory impairments, physical illness or injury, dyslexia or previous psychiatric history (other than depression in the dysphoric group) were excluded from the analysis. This should reduce further the likelihood that any group differences on the DVs resulted from variables other than the IV.

The main criticism of the methodology used in the current series of experiments is that while the BDI gives a measure of the degree of depressive symptomatology, it does not permit a diagnosis of depression to be made. This means that some caution is necessary when attempting to extrapolate the findings to clinical depression. However, as noted in section 3.2.1.1, comparison of studies using clinical and nonclinical samples has generally revealed a similar pattern of performance on a range of tasks. Furthermore, as noted above, the validity is increased by the relatively high mean scores of the dysphoric Ss on the BDI, and by screening Ss on the basis of a brief history taken during the test session.

Another issue which needs to be addressed is the specificity of the findings to depression. On the basis of the current experiments the possibility that similar deficits are associated with other psychiatric conditions or that they are mediated by a factor common to psychiatric conditions in general cannot be excluded. The methodology of the current work would have been improved by the inclusion of a control group with some other psychiatric disorder. However, the use of analogue populations makes this difficult to achieve since most other psychiatric disorders do not occur in high enough proportions in student populations to make this a viable proposition. The exception is anxiety; however, it is difficult to differentiate successfully between anxiety and depression even in clinical populations since the two disorders frequently co-exist (see section 2.1.1.1).

A common approach with analogue populations has been to use measures of depression and anxiety to identify a depressed group and an anxious group. However, this frequently results in the identification of a group of Ss who score highly on both the depression and anxiety measures (depressed/anxious), and a group who have low scores on the depression measure, but only mild/moderate scores on the anxiety measure (non-depressed/mildly anxious). Comparison of these groups therefore confounds severity of symptomatology with group membership. This problem may result in part from the measures used. For example, it has been common practice to use the BDI and the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch & Lushene, 1970) in combination, but research suggests that the STAI correlates very highly with the BDI, and may not measure a separate construct (see e.g. Endler, Cox, Parker & Bagby, 1991). The fact that depression and anxiety frequently co-exist adds to the problem, and at lower levels of severity the overlap between them may be greater. Studies conducted with clinical samples are better able to address the question of specificity of deficits, with separation of groups based on

diagnostic criteria. The difficulty of collecting large samples of patients, and the added complications of psychotropic medication (see section 2.3.1) mean that clinical samples may be most effectively used to test precise hypotheses generated from work with analogue populations.

Overall, it can be concluded that the study of cognitive function in depression is associated with a number of methodological issues which are difficult to resolve. The methodology of the current series of experiments has addressed these issues as far as possible, and is appropriate for testing whether dysphoric Ss show deficits on neutral, abstract and novel measures of reasoning, and for exploring the mechanisms which might underlie any deficits. Taken within the context of existing knowledge, these results can then be considered with regard to the wider issue of reasoning processes in depression.

9.2 SUMMARY OF FINDINGS

Before considering the results of the current experiments with regard to reasoning processes in depression, the predictions which arose as a result of the review of the existing literature in chapter II will be briefly summarised.

While there are a number of competing models of reasoning, common themes can be extracted. First, there is a consensus that reasoning draws on a number of cognitive processes, most notably attention, WM, and LTM, and that these may limit reasoning performance (see section 2.2). For example, distracted attention, inadequate WM resources, or the non-availability of relevant knowledge may all impair performance on a reasoning task. The review in section 2.2 highlighted the fact that some authors (e.g. Johnson-Laird, 1993; Shallice, 1982; Sternberg, 1986) have suggested that the storage and strategic processing of information in WM constitute the essential features of reasoning, making WM capacity an important limiting factor for performance on reasoning tasks. Second, it is generally accepted that humans do not always reason in the way that might be expected on the basis of pure logic. Instead, human reasoning often seems to rely on heuristics and to be overly influenced by existing beliefs (see section 2.2.2.3.5). It is likely that these points are related, with heuristics used to compensate for the constraints placed on performance by the limits of the cognitive system.

In chapter II, evidence regarding cognitive function in depression was considered in relation to processing of both emotional (see section 2.1.3.3) and neutral (see section 2.3) material. These two areas have developed as relatively separate fields of study, but both make similar predictions regarding cognitive function in depression. Evidence relating to depressive processing of emotional material suggests that depression is associated with biases towards negative material (e.g. Beck, 1967; Williams et al., 1988). This is likely to take up processing resources, leaving fewer available for other purposes. A review of depressive performance on neutral tasks suggests depressed Ss are often, but not always, impaired in their performance on measures of psychomotor speed, LTM, attention, WM and reasoning. Current models of depressive cognitive function cite reduced and/or diverted processing resources (e.g. Ellis & Ashbrook, 1988; Hasher & Zacks, 1979, 1988), reduced initiative to generate appropriate strategies (e.g. Hertel & Hardin, 1990; Hertel & Rude, 1991b), or a conservative response bias (e.g. Johnson & Magaro, 1987) as possible explanations for depressive deficits. Finally, there is increasing evidence that depression may be associated with frontal lobe dysfunction. The review of frontal lobe function in section 2.2.3 concluded that the frontal lobes play an important role in performance of reasoning tasks, and this is likely to be mediated by frontal lobe involvement in WM processes (Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997).

In the light of the review carried out in chapter II it was predicted that depression would be associated with reasoning deficits and it was hypothesised that depressive deficits in WM function might underlie impaired depressive reasoning performance. Therefore, the experiments reported in chapters IV to VIII were designed to investigate the role of storage and the use of strategies during reasoning. Within-task manipulations were designed to vary the storage load, and to vary the strategic demands of the tasks, and the effect on task performance for the dysphoric and nondysphoric groups was compared.

9.2.1 DO DYSPHORIC SUBJECTS SHOW DEFICITS ON REASONING TASKS?

The experiments reported in chapters IV, V, VI and VIII all compared dysphoric and nondysphoric Ss on reasoning tasks. There was clear evidence of dysphoric performance impairment on the Discrimination Learning problems reported in chapters IV and VIII, and also evidence of impaired performance by the dysphoric Ss on the Fault Diagnosis task reported in chapter VI. There were no significant differences between the dysphoric and nondysphoric groups in their performance on the Integrative Reasoning task reported

in chapter V, although there were problems with the methodology of this task which may have reduced its sensitivity to dysphoric deficits. Overall, there was evidence of dysphoric reasoning deficits in the current studies consistent with the findings of previous studies reviewed in section 2.3.8.

9.2.2 STORAGE OF INFORMATION

Storage of information in both WM and LTM plays a crucial role in reasoning operations (see section 2.2.6). The possibility that reduced WM storage capacity might underlie depressive reasoning deficits was investigated in the Discrimination Learning and Integrative Reasoning tasks described in chapters IV, V and VIII. In these three experiments manipulations were introduced to vary the storage demands of the reasoning task. This was achieved by leaving some of the information essential to task completion on view rather than requiring Ss to retain it. In this sense the memory-aid and report-aid conditions of the Discrimination Learning task equated with the memory-aid condition of the Integrative Reasoning task. The findings of these three experiments have been discussed at length in chapters IV, V and VIII. In summary there was no evidence from these three experiments that reducing storage demands differentially facilitated the performance of dysphoric Ss, suggesting that any dysphoric reasoning deficits are unlikely to be explicable solely in terms of reduced capacity to store information.

9.2.3 STRATEGY USE

Strategy refers to the order and nature of the processing steps carried out during task performance, and, as such, is heavily dependent on WM. The possibility that deficits in implementing an appropriate strategy might underlie depressive reasoning impairments was investigated using the Discrimination Learning, Integrative Reasoning and Fault Diagnosis tasks described in chapters IV, V, VI and VIII. These have all been studied extensively by other authors, and information was therefore available to inform predictions about the likely nature of strategy-use on these tasks. For each task it was possible to identify a strategy which would lead to logically perfect task performance. It was also possible to measure and describe the way in which Ss actually carried out the task, and to compare this with the logically optimal strategy. Finally, conditions were included which were designed to elucidate the type of strategy being used, and possible limiting factors which influenced this.

The Discrimination Learning task (chapters IV and VIII) and the Fault Diagnosis task (chapter VI) were consistent in finding dysphoric Ss to be impaired relative to nondysphoric Ss in their task performance. On both tasks there was evidence that dysphoric Ss were impaired in their use of an appropriate strategy. It is possible to speculate that on the Discrimination Learning task the dysphoric Ss may have adopted a simple heuristic strategy. On the Fault Diagnosis task the dysphoric Ss collected more redundant information suggesting either a performance deficit or the adoption of a more cautious strategy. While there was no significant difference between the groups on the Integrative Reasoning task (chapter V), the task design was such that group differences in strategy may have gone undetected, and therefore no firm conclusions could be drawn in relation to this task.

Overall, there was evidence that dysphoric Ss were impaired relative to the controls in their performance on reasoning tasks that were neutral, novel and abstract. These deficits could be explained in terms of a failure to adopt appropriate performance strategies by the dysphoric Ss. There was little evidence that reduced capacity to store information provided an adequate explanation of dysphoric deficits.

9.3 RELATIONSHIP TO MODELS

As noted in section 9.1, one of the aims of the current series of experiments was to consider the findings in relation to current models of cognitive function in depression. In this section each model is taken in turn and considered in the light of the experiments reported in chapters IV-VIII.

9.3.1 REDUCED CAPACITY

The hypothesis that depression is associated with reduced cognitive resources forms the basis of several models (e.g. Ellis & Ashbrook, 1988; Hasher & Zacks, 1979, 1988), although the models vary with regard to whether this is thought to be fixed and invariant, or responsive to environmental factors. Hasher and Zacks (1979) first suggested that the pattern of cognitive deficits associated with depression was compatible with reduced attentional resources, that is, deficits are more likely on tasks which make high demands on these resources. Ellis and Ashbrook (1988) introduced the notion that increased processing of irrelevant aspects of the task, or processing of non-task material, such as personal concerns, might be one mechanism by which depression reduces available

resources. This means resources may be diverted as well as reduced, but the overall effect will be one of reduced availability of resources.

The models of cognitive function in depression have lacked precision regarding the nature of the cognitive resources thought to be involved. As outlined in section 2.2.5.1.4, there is longstanding debate regarding the nature of cognitive resources in general, with particular disagreement as to whether resources are unitary or multiple in nature. For the most part, current models of cognitive resources assume multiple resources (e.g Baddeley, 1986; Baddeley & Hitch, 1974; Wickens, 1980). If a unitary system is assumed, then reducing the demands of one aspect of task processing should, theoretically, free capacity for other aspects. However, if multiple systems are assumed, then reducing the demands of one aspect of task processing may have no effect on processes which depend on different resources. The current lack of precision and consensus regarding the nature of cognitive resources, combined with the failure of models of cognitive function in depression to specify their assumptions about the nature of resources, makes it difficult to identify specific predictions. Therefore, the issue of unitary versus multiple resources will not be considered further in relation to the current findings, although it is an issue that theorists and researchers should consider in the future.

Hypotheses of reduced or diverted cognitive resources in depression make identical predictions in most instances (Hartlage, Alloy, Vazquez & Dykman, 1993). The main prediction is that depressive performance should be impaired by increasing demands on capacity. Thus, it is predicted that depressed Ss should not show deficits on tasks that make no, or only minimal demands on cognitive resources, but that performance should suffer as task demands increase. Since reasoning tasks are demanding of cognitive resources, depressive deficits would generally be predicted. The dysphoric deficits observed on the Discrimination Learning (chapters IV and VIII) and Fault Diagnosis (chapter VI) tasks are therefore in line with this prediction. More specifically, depressive deficits should be related to task demands, with greater deficits on more demanding tasks. Thus, group differences are predicted on tasks which are within the capacity of the nondysphoric group, but which exceed the capacity of the dysphoric group; ceiling or floor effects may obscure group differences. There are considerable practical difficulties in choosing a task which places an appropriate level of demand on capacity. If several different tasks are selected to vary the demands on capacity then there may be

confounding effects due to other differences between the tasks. For example, in section 2.3.4.2 it was noted that it has been difficult to discern whether free recall is more sensitive to depressive deficits than recognition memory. The optimal approach is to vary capacity demands within a task.

The Discrimination Learning (chapters IV and VIII), Integrative Reasoning (chapter V) and Fault Diagnosis (chapter VI) tasks all included manipulations of task complexity which should vary demands on cognitive resources and thereby provide the means to test the hypothesis that depressive deficits are related to task demands. In each task, complexity was manipulated by varying the number of processing operations required to carry out the task while maintaining the same logical principle. The assumption that these manipulations would increase task complexity was based on theoretical models of the tasks, but there was no a priori way to guarantee they would influence performance in the expected manner. This is a common problem, as noted in section 2.2.5.1.3, and the level of task demands must usually be determined by empirical observation of performance levels. This approach was adopted in the current series of experiments.

In summary, the findings of two experiments with the Discrimination Learning task (chapters IV and VIII) suggested that while dysphoric Ss were able to apply the logical principle of the task on the simpler one- and two-dimensional problems, they were impaired relative to controls when required to carry out additional processing on more complex problems. In contrast, the findings on the Fault Diagnosis task (chapter VI) were not consistent with a simple relationship between depressive deficits and task complexity, with the finding that dysphoric Ss were impaired relative to controls on the relatively simple pre-training problems, but not on the more complex problems in the reception condition. This constitutes a direct contradiction of the prediction. The failure to find a group difference on the Integrative Reasoning task (chapter V) may be explained by the ceiling effects on most of the problem types, and the possibility that Ss adopted guessing as a rewarding strategy on this task.

Models that argue depression is associated with reduced capacity (e.g. Hasher & Zacks, 1979) predict that deficits will arise when a task exceeds available capacity, and theoretically this should be robust to environmental influences. This approach predicts a strong relationship between task complexity and sensitivity to depressive deficits, with

no depressive deficits on tasks that fall within the capacity of the depressed Ss. There was some support for this prediction in the current series of experiments. Models that postulate diverted resources in depression, such as processing of task-irrelevant or depression-relevant material (e.g. Ellis & Ashbrook, 1988), predict that depressive capacity should be more variable than if resources were actually reduced (e.g. Hasher & Zacks, 1979). If resources are diverted, depressed Ss have cognitive capacity available, but the extent to which this is applied to the experimental task will vary. Predictions are similar to those of resource-reduction models in suggesting that depressed Ss should show deficits on tasks that exceed their capacity. However, if depressed Ss have resources available that they are not applying to the task, an increase in task difficulty may cue depressed Ss to switch from task-irrelevant to task-relevant processing, and the result may be either no change or an actual improvement in performance. This has been labelled the distraction effect (see section 2.3.7.3.1). In the current series of experiments there was no strong evidence to suggest the performance of dysphoric Ss improved differentially on more difficult tasks, and the reverse was more often true.

As noted above, while models of reduced capacity in depression disagree regarding the mechanism by which capacity is reduced and/or diverted, a lack of clarity and precision in the formulation of these models makes it difficult to test hypotheses which successfully distinguish between them. Hasher & Zacks (1988) proposed a model which explained cognitive deficits in both ageing and depression in terms of reduced inhibitory mechanisms allowing irrelevant material access to WM. Irrelevant material occupies capacity and can interfere with successful processing. This model generates specific hypotheses about the likely nature of depressive deficits on certain tasks. In chapter VII the Fan Effect task was used to test the predictions of Hasher & Zacks' (1988) model in dysphoric Ss. The results were consistent with the hypothesis that depression is characterised by inefficient inhibitory mechanisms which allow irrelevant material to enter WM. While it would be premature to draw any firm conclusions about the contribution of inefficient inhibitory mechanisms to dysphoric deficits on the basis of this task alone, the results do suggest the Hasher and Zacks' (1988) model merits further study.

As noted in section 2.3.9, some later models of cognitive function in depression have combined hypotheses from two or more of the earlier models. Hartlage et al. (1993) postulate that depression is associated with reduced capacity, and that the remaining

capacity may be taken up by task-irrelevant processing. This model predicts that while it may be possible to facilitate the performance of depressed Ss to the extent that capacity is taken up by irrelevant thoughts, when task demands exceed capacity then any fixed reduction in capacity will result in depressive deficits regardless of task manipulations. However, this model is difficult to test since it can account for almost any pattern of performance observed in depressed Ss.

9.3.2 RESPONSE STYLE

One of the earliest hypotheses regarding depressive deficits was the proposal that depressed individuals are simply not motivated to do well when given tasks to perform or are unable to sustain motivation (McAllister, 1981; Miller, 1975). The problem with this hypothesis is that it is difficult to derive specific predictions about performance from it. In the current series of experiments, no attempt was made to manipulate motivation directly, and this hypothesis will not be considered further. The response bias model (Johnson & Magaro, 1987) predicts depressed Ss are capable of carrying out the processing necessary to perform tasks, but will show a conservative response bias. There was some support for this hypothesis on the Fault Diagnosis task reported in chap VI. In the selection condition the dysphoric Ss were found to carry out significantly more repetitive and redundant tests, but achieved the correct solution as often as controls, suggesting a more cautious approach. However, the performance of the dysphoric Ss on the Discrimination Learning task in both chaps IV and VIII was not consistent with the response bias model. Dysphoric Ss used a less efficient strategy than controls, and achieved fewer correct solutions. Furthermore, on the Fan Effect task reported in chapter VII, a response bias model would predict that dysphoric Ss should make more false negative and fewer false positive errors than controls. In fact, the groups did not differ in the number of false negative errors, but dysphoric Ss made significantly more false positive errors relative to controls. Therefore, the weight of evidence in the current series of experiments goes against the predictions of the response bias model.

Hertel and her colleagues (e.g. Hertel, 1994; Hertel & Hardin, 1990; Hertel & Rude, 1991b) have put forward a more sophisticated version of the motivational hypothesis. They argue that depressed Ss are characterised by reduced initiative, leading to failure to use strategies spontaneously or engage in elaborative thinking, but that they are capable of this when directed. This is the strongest version of the hypothesis that cognitive

capacity in depression is flexible and open to environmental influence. The model proposed by Hertel and her colleagues implies that any deficiency in cognitive capacity associated with depression is not fixed, but modifiable according to task demands. Hertel and colleagues reported a series of studies (e.g. Hertel, 1994; Hertel & Hardin, 1990; Hertel & Rude, 1991a,b) with results consistent with the notion that depressed Ss can carry out the processes involved, but are less likely than controls to do so spontaneously. It should be noted that impairment in initiating a strategy could be due either to a motivational impairment or to a cognitive impairment.

Hertel and her colleagues have hypothesised that tasks sensitive to depressive deficits will be those which permit the spontaneous use of strategies, rather than those which direct or bypass the use of strategies. The experiments reported by Hertel and colleagues were carried out with memory tasks. Their findings show that task manipulations which ensure all Ss carry out the same processing of task material can facilitate the performance of depressed Ss to normal levels. It was important to test this hypothesis on a reasoning task, and this was carried out in chapter VIII using the Discrimination Learning task.

The results of the experiment reported in chapter VIII indicated that the performance of the dysphoric group was not facilitated to normal levels by a task manipulation designed to cue Ss to carry out the necessary processing at the appropriate time, that is, to carry out the necessary elimination of hypotheses in response to each feedback. The failure of the manipulation to facilitate the performance of dysphoric Ss to normal levels can be explained in two ways: either the manipulation was not effective in cuing an appropriate strategy; or the manipulation was successful in cuing an appropriate strategy, but the dysphoric Ss were unable to adopt it due to reduced cognitive capacity. The manipulation was selected because it focused Ss' attention on the task at the critical stages, but it is not directly comparable to those implemented by Hertel and colleagues in that it only prompted a strategy, rather than requiring Ss to carry out the necessary processing. Consideration of the task procedure suggests that the inclusion of more specific instructions to Ss regarding their task strategy would be likely to change the essential nature of the task, and therefore it seems that this type of task may be less amenable to attempts to cue strategy than the memory tasks reported by Hertel and her colleagues.

Overall, the results of the current series of experiments were consistent with the hypothesis of reduced availability of cognitive resources in depression. As noted in section 9.2, dysphoric deficits were not explicable simply in terms of a reduced ability to store information, rather the dysphoric Ss were characterised by a failure to use appropriate task strategies. According to the model put forward by Hertel and her colleagues (see above) failure to use an appropriate strategy might be caused by a failure to initiate a strategy, due to either motivational or cognitive factors, rather than an inability to implement an appropriate strategy due to reduced resources. It was reported in chapter VIII that dysphoric Ss showed deficits in implementing an appropriate strategy when this was cued by the experimental procedure, although it was not clear whether this was because the manipulation was unsucessful in prompting an appropriate strategy or because the dysphoric Ss did not have sufficient resources to implement it.

9.4 IMPLICATIONS

As noted in section 2.3.2, depressed patients frequently complain of difficulties with concentration, memory and 'thinking'. These symptoms can give rise to great distress, and depressed patients may fear they are suffering from a dementing illness. Furthermore, poor cognitive function can make performance of everyday tasks more difficult and may reinforce depressed patients' negative beliefs about themselves. For these reasons at least, it is of great importance to understand the influence of depression on cognitive function so that information and reassurance can be given to patients, and help given in terms of developing remedial strategies. In addition, information about cognitive function may inform theories of how depression is caused and the mechanisms by which it is maintained, with implications for treatment.

The current series of experiments found that dysphoric Ss often, but not always, showed deficits on measures of reasoning. This is consistent with the studies reported in section 2.3.8 that found depressed Ss to be impaired on reasoning tasks. These findings mean that when depressed patients in clinical settings report that their 'thinking' is impaired, they may be experiencing an objective rather than a subjective deficit. However, with regard to making predictions about depressive reasoning performance in everyday life, the tasks used in the current studies are limited in that they are abstract, novel and neutral, whereas the tasks with which people are engaged in real life are more likely to be concrete, familiar and emotionally-salient or personally-relevant. Furthermore,

performance on the tasks in the current studies was compared with ideal performance based on logic, whereas the evidence reviewed in section 2.2 suggested that humans seldom reason using logically perfect strategies, and that they often adopt heuristics.

In terms of explaining reasoning performance in depression and making predictions about which tasks, both in the laboratory and in real life, are likely to show depressive deficits it may be more helpful to consider the likely source of any depressive deficits. While the details of the models of cognitive function reviewed in section 2.3.9 vary, the basic tenet is the same: depression seems to be associated with reduced and/or diverted WM capacity. Depressive deficits are predicted on tasks that fall within the WM capacity of nondepressed Ss and outside the capacity of depressed Ss. In situations where nondepressed Ss rely on simple heuristics which do not tax their WM capacity to the limit, depressed Ss may not be disadvantaged. Similarly, in situations where the emotionally-salient or personally-relevant nature of the material results in biased processing by either nondepressed Ss, depressed Ss, or both, the difference between depressed and nondepressed Ss may be more complex than a simple depressive deficit.

Based on the above hypotheses, prediction of whether depressed Ss will be impaired on a particular reasoning task will depend on careful analysis of how the task is performed by nondepressed Ss. If task performance is thought to place heavy demands on WM capacity, such as the Discrimination Learning task reported in chapters IV and VIII, then depressive deficits may be predicted. If a task exceeds the WM capacity of nondepressed Ss, or if a simpler heuristic strategy such as guessing is adopted by nondepressed Ss, such as the Integrative Reasoning task reported in chapter V, then depressive deficits are less likely. If adoption of an effortful strategy actually seems to interfere with task performance by nondepressed Ss, as was the case in the study reported by Hertel and Knoedler (1996; see section 2.3.8.2.4.2), then depressed Ss may show a performance advantage. Tasks that have personally-relevant or emotionally-salient material have seldom been subject to this type of scrutiny. For example, depressive deficits have been reported on measures of social problem-solving such as the MEPS and SPSI (see section 2.1.3.3.6.3), but it is not clear what cognitive processes nondepressed Ss draw on when carrying out these tasks. However, it should be noted that tasks with personally-relevant or emotionally-salient material are also likely to be the subject of biased processing by both depressed and nondepressed Ss which will further complicate the issue.

With regard to treatment of depression, the findings of the current series of experiments have several implications. The first, as mentioned above, is that patient reports of 'thinking' difficulties need to be taken seriously. It is important for the clinician to have an understanding of cognitive function in depression so that s/he can appreciate the fact that depressed clients may be experiencing objective rather than subjective cognitive impairment. The clinician may then be able to empathise with the client regarding his/her difficulties, give reassurance that this is a normal part of depressive disorders, and provide information on the likely nature, extent and course of these deficits. For the latter, it is important that good quality research evidence is available to provide this information. The review of cognitive function in depression contained in section 2.3 made it clear that while it may be safe to conclude that depression is associated with cognitive deficits, the true nature, extent and course of these is unclear. For example, the findings are mixed with regard to whether depressive cognitive impairment recovers to normal levels with clinical recovery. There is still a need to carry out further research to provide answers to these questions and it is important that future work takes account of the methodological issues outlined in section 2.3.1.

In addition to being able to provide depressed clients with information about depressive cognitive impairments, it would also be desirable for clinicians to be able to direct remediation of deficits as part of the therapeutic intervention, for example by directing clients towards appropriate strategies or compensatory aids. In the current studies, task manipulations were included that were designed to identify the processes underpinning depressive reasoning deficits. It was predicted that the performance of dysphoric Ss might be facilitated by reducing the storage demands or cuing appropriate strategies, but in fact the performance deficits of dysphoric Ss proved resilient to these interventions. Further work is needed to identify ways in which the performance of depressed Ss might be facilitated so that this information can be used for therapeutic intervention.

Recently, Channon and Green (in press) reported on a study which investigated spontaneous strategy use in clinically depressed Ss, and the effectiveness of providing hints about performance strategies. Three tasks were used, and while these did not measure reasoning directly, they were all known to be sensitive to executive funtion and required the generation of a performance strategy. Both depressed and control Ss were randomly allocated to a stategy aid or no strategy aid condition on the three tasks, with

those in the strategy aid condition receiving a hint regarding task strategy. The results indicated that overall depressed Ss were impaired relative to controls on all three tasks, and were found to be less likely to use appropriate performance strategies. Provision of strategy hints increased the use of performance strategies in two of the three tasks, but did not significantly improve performance for either group.

The results of the study reported by Channon and Green (in press) are consistent with the findings of the current thesis that depressive deficits seem explicable in terms of a failure to use appropriate task strategies to the same extent as controls, and that it seems difficult to remediate these deficits simply by prompting an appropriate strategy. Channon and Green suggest that remediation of depressive deficits on measures of executive function, such as reasoning tasks, may require specific performance cues, practice, or possibly specific training in the use of appropriate strategies, and this requires further investigation.

Depressive deficits in 'thinking' become particularly important when there is patient and/or clinician concern that cognitive deficits may be due to a dementia. Understanding the degree and pattern of cognitive deficits in depression can aid the clinician in making a differential diagnosis between depression and dementia. The recognition and acceptance by clients (and clinicians) that depression can be associated with measurable cognitive deficits is the first step. A desirable aim for the future is the development of screening tests that reliably differentiate between depression and dementia. Increased understanding of cognitive function in depression (and in dementia) will help to achieve this goal.

Another relevant issue with regard to treating depression is that clinicians need to be aware that depressed clients may have an impaired ability to understand or process complex information. This is important for all clinicians working with depressed clients, but may be particularly relevant to clinicians carrying out psychological therapies such as cognitive-behaviour therapy. Cognitive-behaviour therapy (CBT) is based on the principles outlined in the section on Beck's theory of depression (see section 2.1.3.3.1). As part of the therapy, the clinician explains the principles of CBT to the client, so that the client can understand and apply these principles to their own thoughts and behaviour (Beck, Rush, Shaw & Emery, 1979). For example, the client must learn to identify and challenge his/her own automatic negative thoughts (see section 2.1.3.3.1). In order to do this the client must understand the concept of automatic negative thoughts, be able to

judge whether a thought reaches the criteria for being an automatic negative thought, record the thought, and generate alternatives to it. This requires complex processing of information and is likely to place heavy demands on WM capacity. The clinician needs to be aware that engaging in this type of therapy is going to be difficult for the depressed client in terms of the demands placed on cognitive processing, quite apart from any other difficulties associated with entering into therapy. If a client seems to be having difficulty engaging in CBT, cognitive deficits should be considered as a possible explanation. Research into how depressed clients come to learn, remember and implement the complex concepts associated with CBT might have the dual benefits of both making CBT more accessible to clients, and in increasing understanding of cognitive function in depression.

Understanding the biological aspects of depression such as cognitive or brain function has important implications for understanding the nature of depression, and also has implications for somatic treatments of depression. Currently, the work being done with functional imaging (see section 2.1.3.2.2) is providing strengthening evidence of frontal lobe dysfunction in depression. A review of frontal lobe function is contained within section 2.2.3 of the current thesis. In essence, the frontal lobes are implicated in a range of cognitive functions including planning, organising behaviour, and WM function. The findings of the current series of experiments, that depression is associated with deficits on reasoning tasks, and that this is likely to be associated with reduced and/or diverted WM capacity, is entirely consistent with the hypothesis that depression is associated with frontal lobe dysfunction. Further work in the area of functional brain imaging in depression, and exploration of the depressive performance on tasks known to be sensitive to frontal lobe function are likely to be extremely fruitful areas of research in the future.

9. 5 CONCLUSIONS

1. Dysphoric Ss sometimes, but not always, show deficits on reasoning tasks that are abstract, novel and neutral.

2. Deficits shown by dysphoric Ss on reasoning tasks are not explicable solely in terms of reduced capacity to store information.

3. The performance of dysphoric Ss on reasoning tasks is characterised by deficits in implementing an appropriate performance strategy.

4. Depressive deficits on abstract, novel and neutral tasks are likely to be underpinned by reduced and/or diverted WM capacity.

5. Depressive reasoning deficits are most likely to occur on reasoning tasks which are solved by nondepressed Ss using a strategy that places heavy demands on WM capacity and are less likely to occur on reasoning tasks which nondepressed Ss solve using strategies which do not place heavy demands on WM capacity.

9.6 RECOMMENDATIONS FOR FURTHER WORK

On the basis of the current thesis, the following areas are recommended for study in the future:

1. The current series of experiments has highlighted that investigation of reasoning in depression is under-researched compared with other aspects of cognitive function in depression. This is surprising in the light of the importance of reasoning in everyday life. Further work could examine neutral, abstract and novel tasks of the type described in the current studies with regard to elucidating further the mechanisms underlying depressive deficits, for example by use of dual task methodology to investigate the role of WM more directly. The current series of experiments highlighted the importance of strategy when carrying out reasoning tasks, and this is an area that would merit future investigation using a range of laboratory and real life tasks. Work in this field should aim to identify remedial strategies for facilitating the performance of depressed Ss.

2. The area of reasoning in depression needs to be investigated further using clinically depressed, rather than dysphoric, Ss. In particular, the inclusion of groups with other psychiatric conditions will be important in establishing the specificity of reasoning deficits to depression. As outlined in section 9.1, one of the most serious disadvantages of using analogue Ss is the difficulty in recruiting Ss with a psychiatric disorder other than depression for comparison. This is more appropriately done using clinical populations who have received a formal diagnosis. In general, the use of clinical Ss who are not receiving psychotropic medication will greatly improve the validity of studies of cognitive function in depression and other psychiatric disorders in the future.

3. Future work investigating the question of whether depression is associated with reasoning deficits should also focus on real life tasks. One possible starting point would

be to use existing measures of social problem-solving. These could be subjected to task analysis to identify the strategies used by nondepressed Ss and the likely demands on WM capacity. Predictions about depressive performance could then be made on the basis of this, and within-task demands on capacity could be varied. This approach should start to identify the areas of everyday life that are most likely to present difficulties to depressed Ss. A further aim would be to identify remedial strategies for use by depressed Ss, and to discover whether performance returns to normal levels on recovery.

4. In spite of the fact that a large body of work already exists, there is still a need to investigate the nature, extent and course of all areas of cognitive function in depression, since many important questions remain unanswered. Future experiments should be designed carefully to avoid the confounding variables that have made interpretation of the existing work so difficult. In particular, the influence of psychotropic medication on cognitive function is deserving of more attention than has been paid in the past.

5. There is a need to identify hypotheses that will distinguish between current models of cognitive function in depression and to focus future work on testing these.

6. There is a need to develop existing models of cognitive function in depression to reflect current models from cognitive psychology and neuropsychology that suggest processing resources are multiple rather than unitary in nature.

7. Future work on cognitive function in depression should focus on the role of the frontal lobes in depression, with the performance of depressed Ss measured on tasks that are known to be sensitive to the frontal lobes, and their pattern of performance compared with that of patients with frontal lobe lesions. The availability of brain imaging techniques has already facilitated this area of work. In the future it will increasingly be possible to relate information about brain function in depression to models of normal and abnormal brain function that will be far more sophisticated than those currently available.

8. Research into how depressed clients come to learn, remember and implement the complex concepts associated with CBT might have the dual benefits of both making CBT more accessible to clients, and in increasing understanding of cognitive function in depression.

CHAPTER X

REFERENCES

Abas, M.A., Sahakian, B.J. & Levy, R. (1990). Neuropsychological deficits and CT scan changes in elderly depressives. Psychological Medicine, 20, 507-520.

Abramson, L.Y., Garber, J., Edwards, N.B. & Seligman, M.E.P. (1978). Expectancy changes in depression and schizophrenia. Journal of Abnormal Psychology, 87, 102-109.

Abramson, L.Y., Seligman, M.E.P. & Teasdale, J.D. (1978). Learned helplessness in humans: Critique and reformulation. Journal of Abnormal Psychology, 87, 49-74.

Ackermann, R. & DeRubeis, R.J. (1991). Is depressive realism real? Clinical Psychology Review, 11, 565-584.

Alloy, L.B. & Abramson, L.Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? Journal of Experimental Psychology: General, 108, 441-485.

Alloy, L.B., Abramson, L.Y. & Viscusi, D. (1981). Induced mood and the illusion of control. Journal of Personality and Social Psychology, 41, 1129-1140.

Alloy, L.B. & Clements, C.M. (1992). Illusion of control: Invulnerability to negative affect and depressive symptoms after laboratory and natural stressors. Journal of Abnormal Psychology, 101, 234-245.

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

American Psychiatric Association (1987). Diagnostic and statistical manual of mental disorders (3rd ed., rev; DSM-III-R). Washington DC: Author.

Anderson, J.R. (1974). Retrieval of propositional information from long-term memory. Cognitive Psychology, 6, 451-474.

Anderson, J.R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Anderson, J.R. (1993). Problem solving and learning. American Psychologist, 48, 35-44.

Anderson, J.R. & Bower, G.H. (1973). Human associative memory. Washington, DC: V. H. Winston.

Andreasen, N.C. (1976). Do depressed patients show thought disorder? The Journal of Nervous and Mental Disease, 163, 186-192.

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory, Vol. 2. New York: Academic Press.

Austin, M.P., Ross, M.., Murray, C., O'Carroll, R.E., Ebmeier, K.P. & Goodwin, G.M. (1992). Cognitive function in major depression. Journal of Affective Disorders, 25, 21-30.

Backman, L. & Forsell, Y. (1994). Episodic memory functioning in a community-based sample of old adults with Major Depression: Utilisation of cognitive support. Journal of Abnormal Psychology, 103, 361-370.

Baddeley, A.D. (1986). Working memory. Oxford: Oxford University Press.

Baddeley, A. (1990). Human memory: Theory and practice. Hove: Lawrence Erlbaum Associates.

Baddeley, A.D. & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), Recent advances in learning and motivation, Vol. 8 (pp. 47-89). New York: Academic Press.

Barkley, R.A. (1996). Critical issues in research on attention. In G.R. Lyon & N.A. Krasnegor (Eds.) Attention, memory, and executive function (pp. 45-56). Baltimore: Paul H. Brookes Publishing Co.

Barnard, P.J. & Teasdale, J.D. (1991). Interacting cognitive subsystems: A systemic approach to cognitive-affective interaction and change. Cognition and Emotion, 5, 1-39.

Bazin, N., Perruchet, P., De Bonis, M. & Feline, A. (1994). The dissociation of explicit and implicit memory in depressed patients. Psychological Medicine, 24, 239-245.

Beats, B.C., Sahakian, B.J. & Levy, R. (1996). Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. Psychological Medicine, 26, 591-603.

Beatty, W.W., Wonderlich, S.A., Staton, R.D. & Ternes, L.A. (1990). Cognitive functioning in bulimia: Comparison with depression. Bulletin of the Psychonomic Society, 28, 289-292.

Bebbington, P., Katz, R., McGuffin, P., Tennant, C. & Hurry, J. (1989). The risk of mnor depression before age 65: Results from a community study. Psychological Medicine, 19, 393-400.

Beck, A.T. (1967). Depression: Causes and treatment. Philadelphia: University of Pennsylvania Press.

Beck, A.T. (1976). Cognitive therapy and the emotional disorders. New York: International Universities Press.

Beck, A.T., Feshbach, S. & Legg, D. (1962). The clinical utility of the digit symbol test. Journal of Consulting Psychology, 26, 263-268.

Beck, A.T., Kovacs, M. & Weissman, A. (1979). Assessment of suicidal ideation: The Scale for Suicidal Ideators. Journal of Consulting and Clinical Psychology, 47, 343-352.

Beck, A.T., Rush, A.J., Shaw, B.F. & Emery, G. (1979). Cognitive therapy of depression. New York: John Wiley.

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

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

Benassi, V.A. & Mahler, H.I.M. (1985). Contingency judgments by depressed college students: Sadder but not always wiser. Journal of Personality and Social Psychology, 49, 1323-1329.

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.

Berndt, D.J. & Berndt, S.M. (1980). Relationship of mild depression to psychological deficit in college students. Journal of Clinical Psychology, 36, 868-874.

Bernstein, A.S., Riedel, J.A., Graae, F., Seidman, D., Steele, H., Connolly, J. & Lubowsky, J. (1988). Schizophrenia is associated with altered orienting activity; depression with electrodermal (cholinergic?) deficit and normal orienting response. Journal of Abnormal Psychology, 97, 3-12.

Bernstein, A.S., Riedel, J.A., Graae, F., Seidman, D., Steele, H., Lubowsky, J., Yeager, A., Wrable, J. & Margolis, R. (1990). The effects of prolonged stimulus repetition with repeated switching of target status on the orienting response in schizophrenia and depression. The Journal of Nervous and Mental Disease, 178, 96-104.

Bernstein, A.S., Schnur, D.B., Bernstein, P., Yeager, A., Wrable, J. & Smith, S. (1995). Differing patterns of electrodermal and finger pulse responsivity in schizophrenia and depression. Psychological Medicine, 25, 51-62.

Berry, D.C. & Dienes, Z. (1993). Implicit learning: Theoretical and empirical issues. Hove: Lawrence Erlbaum Associates.

Blackburn, I.M. (1975). Mental and psychomotor speed in depression and mania. British Journal of Psychiatry, 126, 329-335.

Blaney, P.H. (1986). Affect and memory: A review. Psychological Bulletin, 99, 229-246.

Blaney, P.H., Behar, V., & Head, R. (1980). Two measures of depressive cognitions: Their association with depression and with each other. Journal of Abnormal Psychology, 89, 678-682.

Blankstein, K.R., Flett, G.L. & Johnston, M.E. (1992). Depression, problem-solving ability, and problem-solving appraisals. Journal of Clinical Psychology, 48, 749-759.

Boone, K.B., Lesser, I.M., Miller, B.L., Whol, M., Berman, N., Lee, A., Palmer, B. & Back, C. (1995). Cognitive functioning in older depressed outpatients: Relationship of presence and severity of depression to neuropsychological test scores. Neuropsychology, 9, 390-398.

Bower, G.H. (1981). Mood and memory. American Psychologist, 36, 129-148.

Bower, G.H. (1987). Commentary on mood and memory. Behaviour Research and Therapy, 25, 443-455.

Bowlby, J. (1978). Attachment and loss, Vol. 2: Separation: Anxiety and anger. Harmondsworth, Middlesex: Penguin.

Bowlby, J. (1981). Attachment and loss, Vol 3: Loss: Sadness and depression Harmondsworth, Middlesex: Penguin.

Boyd, J.H. & Weissman, M.M. (1981). Epidemiology of Affective Disorders: A reexamination and future directions. Archives of General Psychiatry, 38, 1039-1046.

Bradley, B. & Mathews, A. (1983). Negative self-schemata in clinical depression. British Journal of Clinical Psychology, 22, 173-181.

Bradley, B.P., Mogg, K. & Millar, N. (1996). Implicit memory bias in clinical and non-clinical depression. Behaviour Research & Therapy, 34, 865-879.

Bradley, B.P., Mogg, K. & Williams, R. (1994). Implicit and explicit memory for emotional information in non-clinical subjects. Behaviour Research and Therapy, 32, 65-78.

Bradley, B.P., Mogg, K. & Williams, R. (1995). Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety. Behaviour Research and Therapy, 33, 755-770.

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

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. Psychological Review, 85, 1-21.

Brand, N. & Jolles, J. (1987). Information processing in depression and anxiety. Psychological Medicine, 17, 145-153.

Brand, A.N., Jolles, J. & Gispen-de Wied, C. (1992). Recall and recognition memory deficits in depression. Journal of Affective Disorders, 25, 77-86.

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

Brewin, C.R. & Furnham, A. (1986). Attributional versus preattributional variables in self-esteem and depression: A comparison and test of learned helplessness theory. Journal of Personality and Social Psychology, 50, 1013-1020.

Broadbent, D.E. (1958). Perception and communication. London: Pergamon Press.

Brooke, J.B. & Duncan, K.D. (1981). Effects of system display format on performance in a fault location task. Ergonomics, 24, 175-189.

Brown, G.W., Adler, Z. & Bifulco, A. (1988). Life events, difficulties and recovery from chronic depression. British Journal of Psychiatry, 152, 487-498.

Brown, G.W., Andrews, B., Harris, T., Adler, A. & Bridge, L. (1986). Social support, self-esteem and depression. Psychological Medicine, 16, 813-831.

Brown, G.W. & Harris, T. (1978). Social origins of depression: A study of psychiatric disorder in women. London: Tavistock.

Brown, R.G., Scott, L.C., Bench, C.J. & Dolan, R.J. (1994). Cognitive function in depression: Its relationship to the presence and severity of intellectual decline. Psychological Medicine, 24, 829-847.

Brown, W.A., Johnston, R. & Mayfield, D. (1979) The 24-hour dexamethasone suppression test in a clincial setting: Relationship to diagnosis, symptoms, and response to treatment. American Journal of Psychiatry, 136, 543-547.

Bruner, J.S., Goodnow, J.J. & Austin, G.A. (1956). A study of thinking. New York: Wiley.

Burgess, P.W. & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. Neuropsychologia, 34, 263-273.

Byrne, D.G. (1976a). Choice reaction times in depressive states. British Journal of Social and Clinical Psychology, 15, 149-156.

Byrne, D.G. (1976b). Vigilance and arousal in depressive states. British Journal of Social and Clinical Psychology, 15, 267-274.

Byrne, D.G. (1977). Affect and vigilance performance in depressive illness. Journal of Psychiatric Research, 13, 185-191.

Calev, A. & Erwin, P.G. (1985). Recall and recognition in depressives: Use of matched tasks. 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.

Calev, A., Nigal, D. & Chazan, S. (1989). Retrieval from semantic memory using meaningful and meaningless constructs by depressed, stable bipolar and manic patients. British Journal of Clinical Psychology, 28, 67-73.

Cane, D.B. & Gotlib, I.H. (1985). Depression and the effects of positive and negative feedback on expectations, evaluations, and performance. Cognitive Therapy and Research, 9, 145-160.

Channon, S. & Baker, J. (1994). Reasoning strategies in depression: Effects of depressed mood on a syllogism task. Personality and Individual Differences, 17, 707-711.

Channon, S., Baker, J.E. & Robertson, M.M. (1993a). Effects of structure and clustering on recall and recognition memory in clinical depression. Journal of Abnormal Psychology, 102, 323-326.

Channon, S., Baker, J.E. & Robertson, M.M. (1993b). Working memory in clinical depression: An experimental study. Psychological Medicine, 23, 87-91.

Channon, S. & Green, P.S.S. (in press). Executive functioning in depression: the role of performance strategies in aiding depressed and non-depressed participants. Journal of Neurology, Neurosurgery and Psychiatry.

Cheng, P.W. & Holyoak, K.J. (1985). Pragmatic reasoning schemas. Cognitive Psychology, 17, 391-416.

Cicerone, K.D., Lazar, R.M. & Shapiro, W.R. (1983). Effects of frontal lobe lesions on hypothesis sampling during concept formation. Neuropsychologia, 21, 513-524.

Clark, D.A., Beck, A.T. & Beck, J.S. (1994). Symptom differences in major depression, dysthymia, panic disorder, and generalized anxiety disorder. American Journal of Psychiatry, 151, 205-209.

Clark, L.A. & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. Journal of Abnormal Psychology, 100, 316-336.

Cohen, G. (1983). The psychology of cognition. 2nd edition. Academic Press: London.

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.

Colby, C.A. & Gotlib, I.H. (1988). Memory deficits in depression. Cognitive Therapy and Research, 12, 611-627.

Cornell, D.G., Suarez, R. & Berent, S. (1984). Psychomotor retardation in melancholic and nonmelancholic depression: Cognitive and motor components. Journal of Abnormal Psychology, 93, 150-157.

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

Costello, C.G. (1978). A critical review of Seligman's laboratory experiments on learned helplessness and depression in humans. Journal of Abnormal Psychology, 87, 21-31.

Costello, C.G. (1993). The advantages of the symptom approach to depression. In C.G. Costello (Ed), Symptoms of depression (pp.1-21). New York: John Wiley.

Coughlan, A.K. & Hollows, S.E. (1984). Use of memory tests in differentiating organic disorder from depression. British Journal of Psychiatry, 145, 164-167.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychological Bulletin, 104, 163-191.

Coyne, J.C. (1976). Toward an interactional description of depression. Psychiatry, 39, 28-40.

Coyne, J.C. & Gotlib, I.H. (1983). The role of cognition in depression: A critical appraisal. Psychological Bulletin, 94, 472-505.

Cummings, J.L. (1993). The neuroanatomy of depression. Journal of Clinical Psychiatry, 54, 14-20.

Dalgleish, T. & Watts, F.N. (1990). Biases of attention and memory in disorders of anxiety and depression. Clinical Psychology Review, 10, 589-604.

Dalla Barba, G., Parlato, V., Iavarone, A. & Boller, F. (1995). Anosognosia, intrusions and 'frontal' functions in Alzheimer's disease and depression. Neuropsychologia, 33, 247-259.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behaviour, 19, 450-466.

Danion, J.M., Kauffmann-Muller, F., Grange, D., Zimmermann, M.A. & Greth, P. (1995). Affective valence of words, explicit and implicit memory in clinical depression. Journal of Affective Disorders, 34, 227-234.

Danion, J.M., Willard-Schroeder, D., Zimmermann, M.A., Grange, D., Schlienger, J.L. & Singer, L. (1991). Explicit memory and repetition priming in depression: Preliminary findings. Archives of General Psychiatry, 48, 707-711.

Davidson, J., Turnbull, C., Strickland, R. & Belyea, M. (1984). Comparative diagnostic criteria for melancholia and endogenous depression. Archives of General Psychiatry, 41, 506-511.

Davis, H. & Unruh, W.R. (1980). Word memory in non-psychotic depression. Perceptual and Motor Skills, 51, 699-705.

de Silva, P. (1994). Obsessions and compulsions: Investigation. In S.J.E. Lindsay & G.E. Powell (Eds.), The Handbook of Clinical Adult Psychology (2nd edition). London: Routledge.

Deardorff, W.W. & Funabiki, D. (1985). A diagnostic caution in screening for depressed college students. Cognitive Therapy and Research, 9, 277-284.

Deijen, J.B., Orlebeke, J.F. & Rijsdijk, F.V. (1993). Effect of depression on psychomotor skills, eye movements and recognition memory. Journal of Affective Disorders, 29, 33-40.

Delgado, P.L., Price, L.H., Heninger, G.R. & Charney, D.S. (1992). Neurochemistry. In E.S. Paykel (Ed.). Handbook of affective disorders (2nd edition) (pp. 219-253). Edinburgh: Churchill Livingstone.

Dempster, F.N. (1991). Inhibitory processes: A neglected dimension of intelligence. Intelligence, 15, 157-173.

Denny, E.B. & Hunt, R.R. (1992). Affective valence and memory in depression: Dissociation of recall and fragment completion. Journal of Abnormal Psychology, 101, 575-580.

Deptula, D., Manevitz, A. & Yozawitz, A. (1991). Asymmetry of recall in depression. Journal of Clinical and Experimental Neuropsychology, 13, 854-870.

Deptula, D. & Pomara, N. (1990). Effects of antidepressants on human performance: A review. Journal of Clinical Psychopharmacology, 10, 105-111.

Dobson, D.J.G. & Dobson, K.S. (1981). Problem-solving strategies in depressed and nondepressed college students. Cognitive Therapy and Research, 5, 237-249.

Doerfler, L.A., Mullins, L.L., Griffin, N.J., Siegel, L.J. & Richards, C.S. (1984). Problem-solving deficits in depressed children, adolescents, and adults. Cognitive Therapy and Research, 8, 489-500.

Dolan, R.J., Bench, C.J., Brown, R.G., Scott, L.C. & R.S.J. Frackowiak, R.S.J. (1994). Neuropsychological dysfunction in depression: The relationship to regional cerebral blood flow. Psychological Medicine, 24, 849-857.

Donnelly, E.F., Waldman, I.N., Murphy, D.L., Wyatt, R.J. & Goodwin, F.K. (1980). Primary affective disorder: Thought disorder in depression. Journal of Abnormal Psychology, 89, 315-319.

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

Dunning, D. & Story, A.L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? Journal of Personality and Social Psychology, 61, 521-532.

Dykman, B.M., Abramson, L.Y. & Albright, J.S. (1991). Effects of ascending and descending patterns of success upon dysphoric and nondysphoric subjects' encoding, recall, and predictions of future success. Cognitive Therapy and Research, 15, 179-199.

Dykman, B.M., Abramson, L.Y., Alloy, L.B. & Hartlage, S. (1989). Processing of ambiguous and unambiguous feedback by depressed and nondepressed college students: Schematic biases and their implications for depressive realism. Journal of Personality and Social Psychology, 56, 431-445.

Dykman, B.M., Horowitz, L.M., Abramson, L.Y. & Usher, M. (1991). Schematic and situational determinants of depressed and nondepressed students' interpretation of feedback. Journal of Abnormal Psychology, 100, 45-55.

D'Zurilla, T.J. & Nezu, A.M. (1990). Development and preliminary evaluation of the Social Problem-Solving Inventory. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2, 156-163.

Eimas, P.D. (1970). Effects of memory aids on hypothesis behaviour and focusing in young children and adults. Journal of Experimental Child Psychology, 10, 319-336.

Elliott, C.L. & Greene, R.L. (1992). Clinical depression and implicit memory. Journal of Abnormal Psychology, 101, 572-574.

Elliott, R., Sahakian, B.J., McKay, A.P., Herrod, J.J., Robbins, T.W. & Paykel, E.S. (1996). Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance. Psychological Medicine, 26, 975-989.

Ellis, H.C. & Ashbrook, P.W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J. Forgas (Eds.), Affect, cognition and social behaviour (pp. 25-43). Toronto: Hogrefe.

Ellis, H.C. & Ashbrook, P.W. (1989). The "state" of mood and memory research: A selective review. In D. Kuiken (Ed.), Mood and memory: Theory, research and applications. (Special Issue). Journal of Social Behaviour and Personality, 4, 1-21.

Engle, R.W., Cantor, J. & Carullo, J.J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 972-992.

Evans, J., Williams, J.M.G., O'Loughlin, S. & Howells, K. (1992). Autobiographical memory and problem-solving strategies of parasuicide patients. Psychological Medicine, 22, 399-405.

Evans, J.St.B.T. (1984). Heuristic and analytic processes in reasoning. British Journal of Psychology, 75, 451-468.

Evans, J.St.B.T. (1991). Theories of human reasoning: The fragmented state of the art. Theory and Psychology, 1, 83-105.

Evans, J.St.B.T., Barston, J.L. & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. Memory & Cognition, 11, 295-306.

Evans, J.St.B.T., Newstead, S.E. & Byrne, R.M.J. (1993). Human reasoning: The psychology of deduction. Hove: Lawrence Erlbaum Associates.

Farmer, A. & McGuffin, P. (1989). The classification of the depressions: Contemporary confusion revisited. British Journal of Psychiatry, 155, 437-443.

Farmer, E.W., Berman, J.V.F. & Fletcher, Y.L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. The Quarterly Journal of Experimental Psychology, 38A, 675-688.

Fennell, M.J.V. & Campbell, E.A. (1984). The Cognitions Questionnaire: Specific thinking errors in depression. British Journal of Clinical Psychology, 23, 81-92.

Fennell, M.J.V., Teasdale, J.D., Jones, S. & Damle, A. (1987). Distraction in neurotic and endogenous depression: An investigation of negative thinking in major depressive disorder. Psychological Medicine, 17, 441-452.

Foulds, G.A. (1952). Temperamental differences in maze performance. Part II. The effect of distraction and of electroconvulsive therapy on psychomotor retardation. British Journal of Psychology, 43, 33-41.

Frame, C.L. & Oltmanns, T.F. (1982). Serial recall by schizophrenic and affective patients during and after psychotic episodes. Journal of Abnormal Psychology, 91, 311-318.

Freud, S. (1917/1965). Mourning and melancholia. The complete psychological works of Sigmund Freud, Vol. 14 (standard edition.) (pp.239-260). Translated and editied by J. Strachey. London: Hogarth Press.

Friedman, A.S. (1964). Minimal effects of severe depression on cognitive functioning. Journal of Abnormal and Social Psychology, 69, 237-243.

Frith, C.D., Stevens, M., Johnstone, E.C., 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.

Fuster, J.M. (1989). The prefrontal cortex: Anatomy, physiology and neuropsychology of the frontal lobe, (2nd ed.). New York: Raven Press.

Galotti, K.M. (1989). Approaches to studying formal and everyday reasoning. Psychological Bulletin, 105, 331-351.

Gangadhar, B.N, Ancy, J., Janakiramaiah, N. & Umapathy, C. (1993). P300 amplitude in non-bipolar, melancholic depression. Journal of Affective Disorders, 28, 57-60.

Garber, J. & Hollon, S.D. (1980). Universal versus personal helplessness in depression: Belief in uncontrollability or incompetence? Journal of Abnormal Psychology, 89, 56-66.

Gardiner, J.M. & Java, R.I. (1993). Recognising and remembering. In A.F. Collins, S.E. Gathercole, M.A. Conway & P.E. Morris (Eds.) Theories of memory (pp.163-188). Hove: Lawrence Erlbaum Associates.

Garnham, A. & Oakhill, J. (1994). Thinking and reasoning. Oxford: Blackwell

Gass, C.S. & Russell, E.W. (1986). Differential impact of brain damage and depression on memory test performance. Journal of Consulting and Clinical Psychology, 54, 261-263.

Gaylin, W. (1968). The meaning of despair: Psychoanalytic contributions to the understanding of depression. New York: Jason Aronson.

George, M.S., Ketter, T.A. & Post, R.M. (1993). SPECT and PET imaging in mood disorders. Journal of Clinical Psychiatry, 54:11 (suppl), 6-13.

Gerard, L., Zacks, R.T., Hasher, L. & Radvansky, G.A. (1991). Age deficits in retrieval: The fan effect. Journal of Gerontology: Psychological Sciences, 46, P131-136.

Gilbert, P. (1992). Depression: The evolution of powerlessness. Hove: Lawrence Erlbaum Associates.

Gilhooly, K.J., Logie, R.H., Wetherick, N.E. & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. Memory and Cognition, 21 115-124.

Glass, R.M., Uhlenhuth, E.H., Hartel, F.W., Matuzas, W. & Fischman, M.W. (1981). Cognitive dysfunction and imipramine in outpatient depressives. Archives of General Psychiatry, 38, 1048-1051.

Goddard, L., Dritschel, B. & Burton, A. (1996). Role of autobiographical memory in social problem solving and depression. Journal of Abnormal Psychology, 105, 609-616.

Goldstein, K. & Scheerer, M. (1941). Abstract and concrete behaviour: An experimental study with special tests. Psychological Monographs, 53 (Whole No. 239).

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

Gorham, D.R. (1956). A proverbs test for clinical and experimental use. Psychological Reports, 2, 1-2. Monograph Supplement.

Gotlib, I.H. (1984). Depression and general psychopathology in university students. Journal of Abnormal Psychology, 93, 19-30.

Gotlib, I.H. & Asarnow, R.F. (1979). Interpersonal and impersonal problem-solving skills in mildly and clinically depressed university students. Journal of Consulting and Clinical Psychology, 47, 86-95.

Gotlib, I.H. & Hammen, C.L. (1992). Psychological aspects of depression: Towards a cognitive-interpersonal integration. Chichester: John Wiley.

Gotlib, I.H. & Olson, J.M. (1983). Depression, psychopathology, and self-serving attributions. British Journal of Clinical Psychology, 22, 309-310.

Graf, P. & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. Journal of Verbal Learning and Verbal Behaviour, 23, 553-568.

Grant, D.A. & Berg, E.A. (1948). A behavioural analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. Journal of Experimental Psychology, 38, 404-411.

Greden, J.F., Albala, A.A., Smokler, I.A., Gardner, R. & Carroll, B.J. (1981). Speech pause time: A marker of psychomotor retardation among endogenous depressives. Biological Psychiatry, 16, 851-859.

Greden, J.F. & Carroll, B.J. (1980). Decrease in speech pause times with treatment of endogenous depression. Biological Psychiatry, 15, 575-587.

Greenhouse, S.W. & Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24, 95-112.

Grove, W.M. & Andreasen, N.C. (1992). In E.S. Paykel (Ed.) Handbook of Affective Disorders (pp. 25-41). Edinburgh: Churchill Livingstone.

Gruzelier, J., Seymour, K., Wilson, L., Jolley, A., & Hirsch, S. (1988). Impairments on neuropsychologic tests of temporohippocampal and frontohippocampal functions and word fluency in remitting schizophrenia and affective disorders. Archives of General Psychiatry, 45, 623-629.

Gunther, V. & Kryspin-Exner, I. (1991). Ergopsychometry in depressive patients. Journal of Affective Disorders, 23, 81-92.

Haaga, D.A.F. & Beck, A.T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. Behaviour Research and Therapy, 33, 41-48.

Hall, K.R.L. & Stride, E. (1954). Some factors affecting reaction times to auditory stimuli in mental patients. Journal of Mental Science, 100, 462-477.

Halstead, W.C. (1947). Brain and intelligence. Chicago: University of Chicago Press.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. British Journal of Social and Clinical Psychology, 6, 278-296.

Hammen, C.L. (1978). Depression, distortion, and life stress in college students. Cognitive Therapy and Research, 2, 189-192.

Hammen, C.L. & Krantz, S. (1976). Effect of success and failure on depressive cognitions. Journal of Abnormal Psychology, 85, 577-586.

Hargreaves, I.R. (1985). Attributional style and depression. British Journal of Clinical Psychology, 24, 65-66.

Hart, R.P. & Kwentus, J.A. (1987). Psychomotor slowing and subcortical-type dysfunction in depression. Journal of Neurology, Neurosurgery, and Psychiatry, 50, 1263-1266.

Hartlage, S., Alloy, L.B., Vazquez, C. & Dykman, B. (1993). Automatic and effortful processing in depression. Psychological Bulletin, 113, 247-278.

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

Hasher, L. & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), The psychology of learning and motivation Vol. 22 (pp. 193-225). San Diego, CA: Academic Press.

Hatzenbuehler, L.C., Parpal, M. & Matthews, L. (1983). Classifying college students as depressed or nondepressed using the Beck Depression Inventory: An empirical analysis. Journal of Consulting and Clinical Psychology, 51, 360-366.

Hayes, B. & Hesketh, B. (1989). Attribution theory, judgmental biases, and cognitive behaviour modification: Prospects and problems. Cognitive Therapy and Research, 13, 211-230.

Hemsi, L.K., Whitehead, A. & Post, F. (1968). Cognitive functioning and cerebral arousal in elderly depressives and dements. Journal of Psychosomatic Research, 12, 145-156.

Hemsley, D.R. & Zawada, S.L. (1976). 'Filtering' and the cognitive deficit in schizophrenia. British Journal of Psychiatry, 128, 456-461.

Henry, G.M., Weingartner, H. & Murphy, D.L. (1973). Influence of affective states and psychoactive drugs on verbal learning and memory. American Journal of Psychiatry, 130, 966-971.

Henriques, J.B., Glowacki, J.M. & Davidson, R.J. (1994). Reward fails to alter response bias in depression. Journal of Abnormal Psychology, 103, 460-466.

Herskovic, J.E., Kietzman, M.L. & Sutton, S. (1986). Visual flicker in depression: response criteria, confidence ratings and response times. Psychological Medicine, 16, 187-197.

Hertel, P.T. (1994). Depressive deficits in word identification and recall. Cognition and Emotion, 8, 313-327.

Hertel, P.T. & Hardin, T.S. (1990). Remembering with and without awareness in a depressed mood: Evidence of deficits in initiative. Journal of Experimental Psychology: General, 119, 45-59.

Hertel, P.T. & Knoedler, A.J. (1996). Solving problems by analogy: The benefits and detriments of hints and depressed mood. Memory & Cognition, 24, 16-25.

Hertel, P.T. & Milan, S. (1994). Depressive deficits in recognition: Dissociation of recollection and familiarity. Journal of Abnormal Psychology, 103, 736-742.

Hertel, P.T. & Rude, S.S. (1991a). Recalling in a state of natural or experimental depression. Cognitive Therapy and Research, 15, 103-127.

Hertel, P.T. & Rude, S.S. (1991b). Depressive deficits in memory: Focusing attention improves subsequent recall. Journal of Experimental Psychology: General, 120, 301-309.

Hill, A.B., Kemp-Wheeler, S.M. & Jones, S.A. (1986). What does the Beck Depression Inventory measure in students? Personality and Individual Differences, 7, 39-47.

Hinchcliffe, M.K., Lancashire, M. & Roberts, F.J. (1971). Depression: Defence mechanisms in speech. British Journal of Psychiatry, 118, 471-472.

Hiroto, D.S. & Seligman, M.E.P. (1975). Generality of learned helplessness in man. Journal of Personality and Social Psychology, 31, 311-327.

Hirschfeld, R.M.A. (1981). Situational depression: Validity of the concept. British Journal of Psychiatry, 139, 297-305.

Hirschfeld, R.M.A. & Cross, C.K. (1982). Epidemiology of affective disorders: Psychosocial risk factors. Archives of General Psychiatry, 39, 35-46.

Hitch, G.J. & Baddeley, A.D. (1976). Verbal reasoning and working memory. Quarterly Journal of Experimental Psychology, 28, 603-621.

Howell, D.C. (1987). Statistical methods for psychology, (2nd ed.). Boston: PWS-Kent Publishing Co.

Hull, C.L. (1920). Quantitative aspects of the evolution of concepts: An experimental study. Psychological Monographs, 28 (Whole No. 123).

Ilsley, J.E., Moffoot, A.P.R. & O'Carroll, R.E. (1995). An analysis of memory dysfunction in major depression. Journal of Affective Disorders, 35, 1-9.

Ingram, R.E., Kendall, P.C., Smith, T.W., Donnell, C. & Ronan, K. (1987). Cognitive specificity in emotional distress. Journal of Personality and Social Psychology, 53, 734-742.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory and Language, 30, 513-541.

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

Johnson, O. & Crockett, D. (1982). Changes in perceptual asymmetries with clinical improvement of depression and schizophrenia. Journal of Abnormal Psychology, 91, 45-54.

Johnson-Laird, P.N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (1993). Human and machine thinking. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Johnson-Laird, P.N. & Bara, B.G. (1984). Syllogistic inference. Cognition, 16, 1-62.

Johnson-Laird, P.N. & Byrne, R.M.J. (1991). Deduction. Hove, East Sussex: Lawrence Erlbaum Associates.

Judd, F.K. & Burrows, G.D. (1992). Anxiety disorders and their relationship to depression. In E. Paykel (Ed.) Handbook of Affective Disorders (pp. 77-87). Edinburgh: Churchill Livingstone.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, New Jersey: Prentice-Hall.

Karno, M., Hough, R.L., Burnam, A., Escobar, J.I., Timbers, D.M., Santana, F. & Boyd, J.H. (1987). Lifetime prevalence of specific psychiatric disorders among Mexican Americans and non-hispanic whites in Los Angeles. Archives of General Psychiatry, 44, 695-701.

Katon, W. & Roy-Byrne, P.P. (1991). Mixed anxiety and depression. Journal of Abnormal Psychology, 100, 337-345.

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

Kimberg, D.Y. & Farah, M.J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organised behaviour. Journal of Experimental Psychology: General, 122, 411-428.

King, S. & Phillips, S. (1985). Problem-solving characteristics of process and reactive schizophrenics and affective-disordered patients. Journal of Abnormal Psychology, 94, 17-29.

Klein, D.C., Fencil-Morse, E. & Seligman, M.E.P. (1976). Learned helplessness, depression, and the attribution of failure. Journal of Personality and Social Psychology, 33, 508-516.

Klein, M. (1934/1968). A contribution to the psychogenesis of manic-depressive states. In W. Gaylin (Ed.), The meaning of despair: Psychoanalytic contributions to the understanding of depression. New York: Jason Aronson

Knott, V., Lapierre, Y., Griffiths, L., de Lugt, D. & Bakish, D. (1991). Event-related potentials and selective attention in major depressive illness. Journal of Affective Disorders, 23, 43-48.

Koh, S.D., Kayton, L. & Berry, R. (1973). Mnemonic organisation in young nonpsychotic schizophrenics. Journal of Abnormal Psychology, 81, 299-310.

Koh, S.D. & Wolpert, E.A. (1983). Memory scanning and retrieval in affective disorders. Psychiatry Research, 8, 289-297.

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

Kovacs, M. & Beck, A.T. (1978). Maladaptive cognitive structures in depression. American Journal of Psychiatry, 135, 525-533.

Krames, L. & MacDonald, M.R. (1985). Distraction and depressive cognitions. Cognitive Therapy and Research, 9, 561-573.

Krantz, S.E. & Gallagher-Thompson, D. (1990). Depression and information valence influence depressive cognition. Cognitive Therapy and Research, 14, 95-108.

Krantz, S. & Hammen, C. (1979). Assessment of cognitive bias in depression. Journal of Abnormal Psychology, 88, 611-619.

Kuiper, N.A. (1978). Depression and causal attributions for success and failure. Journal of Personality and Social Psychology, 36, 236-246.

Kwiatkowski, S.J. & Parkinson, S.R. (1994). Depression, elaboration, and mood congruence: Differences between natural and induced mood. Memory and Cognition, 22, 225-233.

Kwon, S. & Oei, T.P.S. (1994). The roles of two levels of cognitions in the development, maintenance, and treatment of depression. Clinical Psychology Review, 14, 331-358.

Kyllonen, P.C. & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?!. Intelligence, 14, 389-433.

Lakey, B. (1988). Self-esteem, control beliefs, and cognitive problem-solving skill as risk factors in the development of subsequent dysphoria. Cognitive Therapy and Research, 12, 409-420.

Langer, E.J. (1975). The illusion of control. Journal of Personality and Social Psychology, 32, 311-328.

Larner, S. (1977). Encoding in senile dementia and elderly depressives: A preliminary study. British Journal of Social and Clinical Psychology, 16, 379-390.

Lefebvre, M.F. (1981). Cognitive distortion and cognitive errors in depressed psychiatric and low back pain patients. Journal of Consulting and Clinical Psychology, 49, 517-525.

Lemelin, S., Baruch, P., Vincent, A., Laplante, L., Everett, J. & Vincent, P. (1996). Attention disturbance in clinical depression: Deficient distractor inhibition or processing resource deficit? The Journal of Nervous and Mental Disease, 184, 114-121.

Levin, H.S., Eisenberg, H.M. & Benton, A.L. (1991). Frontal lobe function and dysfunction. New York: Oxford University Press.

Levine, M. (1966). Hypothesis behaviour by humans during discrimination learning. Journal of Experimental Psychology, 71, 331-338.

Levy, R. & Maxwell, A.E. (1968). The effect of verbal context on the recall of schizophrenics and other psychiatric patients. British Journal of Psychiatry, 114 311-316.

Lewinsohn, P.M. (1974). A behavioural approach to depression. In R.J. Friedman & M.M. Katz (Eds.), The psychology of depression: Contemporary theory and research. New York: Winston-Wiley.

Lewinsohn, P.M. (1975). The behavioural study and treatment of depression. In M. Hersen, R.M. Eisler & P.M. Miller (Eds.), Progress in behaviour modification, Vol. 1. New York: Academic Press.

Lewinsohn, P.M., Mischel, W., Chaplin, W. & Barton, R. (1980). Social competence and depression: The role of illusory self-perception. Journal of Abnormal Psychology, 89, 203-212.

Lewinsohn, P.M., Youngren, M.A. & Grosscup, S.J. (1979). Reinforcement and depression. In R.A. Depue (Ed.), The psychobiology of the depressive disorders: Implications for the effects of stress. New York: Academic Press.

Linehan, M.M., Camper, P., Chiles, J.A., Strosahl, K. & Shearin, E. (1987). Interpersonal problem solving and parasuicide. Cognitive Therapy and Research, 11, 1-12.

Logie, R.H., Gilhooly, K.J. & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. Memory and Cognition, 22, 395-410.

Luria, A.R. (1973). The frontal lobes and the regulation of behaviour. In K.H. Pribram & A.R. Luria (Eds.), Psychophysiology of the frontal lobes (pp. 3-26). New York: Academic Press.

Lyon, G.R. & Krasnegor, N.A. (1996). Attention, memory, and executive function. Baltimore: Paul H. Brookes Publishing Co.

MacLeod, C. (1990). Mood disorders and cognition. In M.W. Eysenck (Ed.), Cognitive Psychology: An International Review (pp.9-55). New York: Wiley.

MacLeod, C., Mathews, A. & Tata, P. (1986). Attentional bias in emotional disorders. Journal of Abnormal Psychology, 95, 15-20.

Manschreck, T.C., Maher, B.A., Rosenthal, J.E. & Berner, J. (1991). Reduced primacy and related features in schizophrenia. Schizophrenia Research, 5, 35-41.

Marcos, T., Salamero, M., Gutierrez, F., Catalan, R., Gasto, C. & Lazaro, L. (1994). Cognitive dysfunctions in recovered melancholic patients. Journal of Affective Disorders, 32, 133-137.

Marshalek, B., Lohman, D.F. & Snow, R.E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. Intelligence, 7, 107-127.

Martin, D.J., Abramson, L.Y. & Alloy, L.B. (1984). Illusion of control for self and others in depressed and nondepressed college students. Journal of Personality and Social Psychology, 46, 125-136.

Martin, I. & Rees, L. (1966). Reaction times and somatic reactivity in depressed patients. Journal of Psychosomatic Research, 9, 375-382.

Marx, E.M., Williams, J.M.G. & Claridge, G.C. (1992). Depression and social problem-solving. Journal of Abnormal Psychology, 101, 78-86.

Mathews, A. & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. Annual Review of Psychology, 45, 25-50.

Mathews, A., Ridgeway, V. & Williamson, D. (1996). Evidence for attention to threatening stimuli in depression. Behaviour Research and Therapy, 34, 695-705.

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

McCarthy, R.A. & Warrington, E.K. (1990). Cognitive Neuropsychology: A clinical introduction. San Diego: Academic Press.

Metalsky, G.I., Halberstadt, L.J. & Abramson, L.Y. (1987). Vulnerability to depressive mood reactions: Toward a more powerful test of the diathesis-stress and causal mediation components of the reformulated theory of depression. Journal of Personality and Social Psychology, 52, 386-393.

Metalsky, G.I. & Joiner, T.E. Jr (1992). Vulnerability to depressive symptomatology: A prospective test of the diathesis-stress and causal mediation components of the hopelessness theory of depression. Journal of Personality and Social Psychology, 63, 667-675.

Metcalfe, M. & Goldman, E. (1965). Validation of an inventory for measuring depression. British Journal of Psychiatry, 111, 240-242.

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.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Miller, W.R. (1975). Psychological deficit in depression. Psychological Bulletin, 82, 238-260.

Miller, W.R. & Seligman, M.E.P. (1975). Depression and learned helplessness in man. Journal of Abnormal Psychology, 84, 228-238.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J.M. Warren & K. Akert (Eds.), The frontal granular cortex and behaviour (pp. 313-334). New York: McGraw Hill.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. British Medical Bulletin, 27, 272-277.

Milner, B. (1982). Some cognitive effects of frontal-lobe lesions in man. Philosophical Transactions of the Royal. Society of London B, 298, 211-226.

Moffoot, A.P.R., O'Carroll, R.E., Bennie, J., Carroll, S., Dick, H., Ebmeier, K.P. & Goodwin, G.M. (1994). Diurnal variation of mood and neuropsychological function in major depression with melancholia. Journal of Affective Disorders, 32, 257-269

Mogg, K., Bradley, B.P. & Williams, R. (1995). Attentional bias in anxiety and depression: The role of awareness. British Journal of Clinical Psychology, 34, 17-36.

Mogg, K., Bradley, B.P., Williams, R. & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. Journal of Abnormal Psychology, 102, 304-311.

Morrison, D.L. & Duncan, K.D. (1988). Strategies and tactics in fault diagnosis. Ergonomics, 31, 761-784.

Myers, J.K., Weissman, M.M., Tischler, G.L., Holzer, C.E., Leaf, P.J., Orvaschel, H., Anthony, J.C., Boyd, J.H., Burke, J.D., Kramer, M. & Stoltzman, R. (1984). Six-month prevalence of psychiatric disorders in three communities: 1980 to 1982. Archives of General Psychiatry, 41, 959-967.

Navon, D. (1984). Resources - A theoretical soup stone? Psychological Review, 91, 216-234.

Navon, D. & Gopher, D. (1979). On the economy of the human-processing system. Psychological Review, 86, 214-255.

Nelson, H.E. (1976). A modified card sorting test sensitive to frontal lobe defects. Cortex, 12, 313-324.

Nelson, J.C. & Charney, D.S. (1981). The symptoms of major depressive illness. American Journal of Psychiatry, 138, 1-13.

Newell, A. & Simon, H.A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Nezu, A.M. (1987). A problem-solving formulation of depression: A literature review and proposal of a pluralistic model. Clinical Psychology Review, 7, 121-144.

Nezu, A.M., Nezu, C.M. & Perri, M.G. (1989). Problem-solving therapy for depression: Theory, research and clinical guidelines. New York: John Wiley.

Nisbett, R. & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgement. Englewood Cliffs, NJ: Prentice-Hall.

Nolen-Hoeksema, A. (1987). Sex differences in unipolar depression: Evidence and theory. Psychological Bulletin, 101, 259-282.

Norman, D.A. & Bobrow, D.G. (1975). On data-limited and resource-limited processes. Cognitive Psychology, 7, 44-64.

Norman, D.A. & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R.J. Davidson, G.E. Schwartz & D.E. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory, Vol. 4 (pp. 1-18). New York: Plenum Press.

Oakhill, J.V. & Johnson-Laird, P.N. (1984). Representation of spatial descriptions in working memory. Current Psychological Research and Reviews, 3, 52-62.

Oatley, K. & Johnson-Laird, P.N. (1987). Towards a cognitive theory of emotions. Cognition and Emotion, 1, 29-50.

Oliver, J.M. & Simmons, M.E. (1984). Depression as measured by the DSM-III and the Beck Depression Inventory in an unselected adult population. Journal of Consulting and Clinical Psychology, 52, 892-898.

Owen, A.M., Downes, J.J., Sahakian, B.J., Polkey, C.E. & Robbins, T.W. (1990). Planning and spatial working memory following frontal lobe lesions in man. Neuropsychologia, 28, 1021-1034.

Parkin, A.J & Russo, R. (1990). Implicit and explicit memory and the auotmatic/effortful distinction. European Journal of Cognitive Psychology, 2, 71-80.

Paykel, E.S. (1989a). The background: Extent and nature of the disorder. In K.R. Herbst & E.S. Paykel (Eds.), Depression: An integrative approach (pp.3-17). Oxford: Heinneman Medical Books.

Paykel, E.S. (1989b). Treatment of depression: The relevance of research for clinical practice. British Journal of Psychiatry, 155, 754-763.

Pennington, B.F., Bennetto, L., McAleer, O. & Roberts, R.J. (1996). Critical issues in research on attention. In G.R. Lyon & N.A. Krasnegor (Eds.) Attention, memory, and executive function (pp. 327-348). Baltimore: Paul H. Brookes Publishing Co.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia, 12, 323-330.

Perrig, W.J. & Perrig, P. (1988). Mood and memory: Mood-congruity effects in absence of mood. Memory and Cognition, 16, 102-109.

Perris, C. (1992). Bipolar-unipolar distinction. In E.S. Paykel (Ed.) Handbook of Affective Disorders (pp. 57-75). Edinburgh: Churchill Livingstone.

Peselow, E.D., Corwin, J., Fieve, R.R., Rotroson, J. & Cooper, T.B. (1991). Disappearance of memory deficits in outpatient depressives responding to imipramine. Journal of Affective Disorders, 21, 173-183.

Peterson, C. & Seligman, M.E. (1984). Causal explanations as a risk factor for depression: Theory and evidence. Psychological Review, 91, 347-74.

Peterson, C.R., Semmel, A., von Baeyer, C., Abramson, L.Y., Metalsky, G.I. & Seligman, M.E.P. (1982). The Attributional Style Questionnaire. Cognitive Therapy and Research, 6, 287-299.

Petrides, M. (1982). Motor conditional associate-learning tasks after selective prefrontal lesions in the monkey. Behaviour Brain Research, 5, 407-413.

Petrides, M & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. Neuropsychologia, 20, 249-262.

Platt, J.J. & Spivack, G. (1972). Problem-solving thinking of psychiatric patients. Journal of Consulting and Clinical Psychology, 39, 148-151.

Pogue-Geille, M.F. & Oltmanns, T.F. (1980). Sentence perception and distractibility in schizophrenia, manic, and depressed patients. Journal of Abnormal Psychology, 89, 115-124.

Pollard, P. (1982). Human reasoning: Some possible effects of availability. Cognition, 12, 65-96.

Pope, B., Blass, T., Siegman, A.W. & Raher, J. (1970). Anxiety and depression in speech. Journal of Consulting and Clinical Psychology, 35, 128-133.

Posner, M.I. & Snyder, C.R.R. (1975). Attention and cognitive control. In R.L. Solso (Ed.), Information processing and cognition: The Loyola Symposium. Hillsdale, New Jersey: Erlbaum.

Potts, R., Camp, C. & Coyne, C. (1989). The relationship between naturally occurring dysphoric moods, elaborative encoding, and recall performance. Cognition and Emotion, 3, 197-205.

Power, M.J. & Champion, L.A. (1986). Cognitive approaches to depression: A theoretical critique. British Journal of Clinical Psychology, 25, 201-212.

Prabhakaran, V., Smith, J.A.L., Desmond, J.E., Glover, G.H. & Gabrieli, J.D.E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. Cognitive Psychology, 33, 43-63.

Pritchard, W.S. (1981). Psychophysiology of P300. Psychological Bulletin, 89, 506-540.

Raven, J.C. (1965). Advanced Progressive Matrices Sets I and II: Plan and use of the scale with a report of experimental work. London: H.K. Lewis.

Richards, P.M. & Ruff, R.M. (1989). Motivational effects on neuropsychological functioning: Comparison of depressed versus nondepressed individuals. Journal of Consulting and Clinical Psychology, 57, 396-402.

Richardson-Klavehn, A. & Bjork, R.A. (1988). Measures of memory. Annual Review of Pscyhology, 39, 475-543.

Rips, L.J. (1989). The psychology of knights and knaves. Cognition, 31, 85-116.

Rips, L.J. (1990). Reasoning. Annual Review of Psychology, 41, 321-353.

Rizley, R. (1978). Depression and distortion in the attribution of causality. Journal of Abnormal Psychology, 87, 32-48.

Robbins, T.W., Joyce, E.M. & Sahakian, B.J. (1992). Neuropsychology and imaging. In E. Paykel (Ed.) Handbook of Affective Disorders (pp. 289-309). Edinburgh: Churchill Livingstone.

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

Robinson, R.G., Kubos, K.L., Starr, L.B., Rao, K. & Price, T.R. (1984). Mood disorders in stroke patients: Importance of location of lesions. Brain, 107, 81-93.

Roediger, H.L. (1990). Implicit memory: Retention without remembering. American Psychologist, 45, 1043-1056.

Roediger, H.L. & McDermott, K.B. (1992). Depression and implicit memory: A commentary. Journal of Abnormal Psychology, 101, 587-591.

Rogers, D., Lees, A.J., Smith, E., Trimble, M. & Stern, G.M. (1987). Bradyphrenia in Parkinson's Disease and psychomotor retardation in depressive illness: An experimental study. Brain, 110, 761-776.

Rohling, M.L. & Scogin, F. (1993). Automatic and effortful memory processes in depressed persons. Journal of Gerontology: Psychological Sciences, 48, P87-P95.

Rosch, E.H. (1973). Natural categories. Cognitive Psychology, 4, 328-350.

Rotheram-Borus, M.J., Trautman, P.D., Dopkins, S.C. & Shrout, P.E. (1990). Cognitive style and pleasant activities among female adolescent suicide attempters. Journal of Consulting and Clinical Psychology, 58, 554-561.

Rouse, W.B. (1978). Human problem solving performance in a fault diagnosis task. IEEE Transactions on Systems, Man, and Cybernetics, 8, 258-271.

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.

Ruiz-Caballero, J.A. & Gonzalez, P. (1994). Implicit and explicit memory bias in depressed and nondepressed subjects. Cognition and Emotion, 8, 555-569.

Rush, A.J., Weissenburger, J., Vinson, D.B. & Giles, D.E. (1983). Neuropsychological dysfunctions in unipolar nonpsychotic major depressions. Journal of Affective Disorders, 5, 281-287.

Russell, P.N. & Beekhuis, M.E. (1976). Organization in memory: A comparison of psychotics and normals. Journal of Abnormal Psychology, 85. 527-534.

Sacco, W.P. & Hokanson, J.E. (1978). Expectations of success and anagram performance of depressives in a public and private setting. Journal of Abnormal Psychology, 87, 122-130.

Sackeim, H.A., Freeman, J., McElhiney, M., Coleman, E., Prudic, J. & Devanand, D.P. (1992). Effects of major depression on estimates of intelligence. Journal of Clinical and Experimental Neuropsychology, 14, 268-288.

Sackeim, H.A., Prohovnik, I., Moeller, J.R., Brown, R.P., Apter, S., Prudic, J., Devanand, D.P. & Mukherjee, S. (1990). Regional cerebral blood flow in mood disorders. I. Comparison of major depressives and normal controls at rest. Archives of General Psychiatry, 47, 60-70.

Sadowski, C. & Kelley, M.L. (1993). Social problem solving in suicidal adolescents. Journal of Consulting and Clinical Psychology, 61, 121-127.

Salthouse, T.A. (1992). Working-memory mediation of adult age differences in integrative reasoning. Memory and Cognition, 20, 413-423.

Salthouse, T.A. (1993). Influence of working memory on adult age differences in matrix reasoning. British Journal of Psychology, 84, 171-199.

Salthouse, T.A., Legg, S., Palmon, R. & Mitchell, D. (1990). Memory factors in age-related differences in simple reasoning. Psychology and Aging, 5, 9-15.

Salthouse, T.A., Mitchell, D.R.D., Skovronek, E. & Babcock, R.L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 507-516.

Salzman, L.F., Goldstein, R.H., Atkins, R. & Babigian, H. (1966). Conceptual thinking in psychiatric patients. Archives of General Psychiatry, 14, 55-59.

Sara, G., Gordon, E., Kraiuhin, C., Coyle, S., Howson, A. & Meares, R. (1994). The P300 ERP component: An index of cognitive dysfunction in depression? Journal of Affective Disorders, 31, 29-38.

Savard, R.J., Rey, A.C. & Post, R.M. (1980). Halstead-Reitan Category Test in bipolar and unipolar affective disorders. The Journal of Nervous and Mental Disease, 168, 297-304.

Schneider, W. & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G.H. Bower (Ed.), The psychology of learning and motivation Vol. 21 (pp. 53-119). New York: Academic Press.

Schneider, W. & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Schotte, D.E. & Clum, G.A. (1982). Suicide ideation in a college population: A test of a model. Journal of Consulting and Clinical Psychology, 50, 690-696.

Schotte, D.E. & Clum, G.A. (1987). Problem-solving skills in suicidal psychiatric patients. Journal of Consulting and Clinical Psychology, 55, 49-54.

Schotte, D.E., Cools, J. & Payvar, S. (1990). Problem-solving deficits in suicidal patients: Trait vulnerability or state phenomenon? Journal of Consulting and Clinical Psychology, 58, 562-564.

Schustack, M.W. (1988). Thinking about causality. In R.J. Sternberg & E.E. Smith (Eds.), The psychology of human thought (pp.92-115). Cambridge: Cambridge University Press.

Seligman, M.E.P. (1975). Helplessness: On depression, development and death. San Francisco: Freeman.

Seligman, M.E.P, Abramson, L.Y., Semmel, A. & von Baeyer, C. (1979). Depressive attributional style. Journal of Abnormal Psychology, 88, 242-247.

Shallice, T. (1982). Specific impairments of planning. Philosophical Transactions of the Royal Society of London B, 298, 199-209.

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge: Cambridge University Press.

Shallice, T. & Burgess, P.W. (1991). Deficits in strategy application following frontal lobe damage in man. Brain, 114, 727-741.

Shallice, T. & Burgess, P.W. (1993). Supervisory control of action and thought selection. In A.D. Baddeley & L. Weiskrantz (Eds.), Attention: Selection, awareness and control: A tribute to Donald Broadbent. Oxford: Oxford University Press.

Shallice, T. & Evans, M.E. (1978). The involvement of the frontal lobes in cognitive estimation. Cortex, 14, 294-303.

Shapiro, M.B., Campbell, D., Harris, A. & Dewsbery, J.P. (1958) Effects of E.C.T. upon psychomotor speed and the "distraction effect" in depressed psychiatric patients. Journal of Mental Science, 104, 681-695.

Shaw, B.F. & Dobson, K.S. (1981). Cognitive assessment of depression. In T.V. Merluzzi, C.R. Glass & M. Genest (Eds.), Cognitive Assessment (pp.361-387). New York: The Guildford Press.

Shiffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Shipley, J.E., Kupfer, D.J., Spiker, D.G., Shaw, D.H., Coble, P.A., Neil, J.F. & Cofsky, J. (1981). Neuropsychological assessment and EEG sleep in affective disorders. Biological Psychiatry, 16, 907-918.

Shipley, W.C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. Journal of Psychology, 9, 371-377.

Silberman, E.K., Weingartner, H., Laraia, M., Byrnes, S. & Post, R.M. (1983). Processing of emotional properties of stimuli by depressed and normal subjects. The Journal of Nervous and Mental Disease, 171, 10-14.

Silberman, E.K., Weingartner, H. & Post, R.M. (1983). Thinking disorder in depression. Archives of General Psychiatry, 40, 775-780

Silberman, E.K., Weingartner, H., Targum, S.D. & Byrnes, S. (1985). Cognitive functioning in biological subtypes of depression. Biological Psychiatry, 20, 654-661.

Smith, J.D., Tracy, J.I. & Murray, M.J. (1993). Depression and category learning. Journal of Experimental Psychology: General, 122, 331-346.

Snaith, R.P. (1987). The concepts of mild depression. British Journal of Psychiatry, 150, 387-393.

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs, 74 (Whole No. 498), 1-29.

Spielberger, C.D., Gorsuch, R.L. & Lushene, R.E. (1970). Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire). Palo Alto: Consulting Psychologists Press.

Spitzer, R.L., Williams, J.B.W., Gibbon, M. & First, M.B. (1992). The Structured Clinical Interview for DSM-III-R (SCID): I. History, rationale, and description. Archives of General Psychiatry, 49, 624-629.

Sprock, J., Braff, D.L., Saccuzzo, D.P. & Atkinson, J.H. (1983). The relationship of depression and thought disorder in pain patients. British Journal of Medical Psychology, 56, 351-360.

Squire, L.R. & Zouzounis, J.A. (1988). Self-ratings of memory dysfunction: Different findings in depression and amnesia. Journal of Clinical and Experimental Neuropsychology, 10, 727-738.

Standish-Barry, H.M.A.S, Bouras, N. Bridges, P.K. & Bartlett, J.R. (1982). Pneumo-encephalographic and computerized axial tomography scan changes in affective disorder. British Journal of Psychiatry, 141, 614-617.

Steif, B.L., Sackeim, H.A., Portnoy, S., Decina, P. & Malitz, S. (1986). Effects of depression and ECT on anterograde memory. Biological Psychiatry, 21, 921-930.

Sternberg, D.E. & Jarvik, M.E. (1976). Memory functions in depression: Improvement with antidepressant medication.. Archives of General Psychiatry, 33, 219-224.

Sternberg, R.J. (1982). A componential approach to intellectual development. In R.J. Sternberg (Ed.), Advances in the psychology of human intelligence Vol. 1 (pp. 413-463). Hillsdale, NJ: Erlbaum.

Sternberg, R.J. (1986). Toward a unified theory of human reasoning. Intelligence, 10, 281-314.

Sternberg, R.J. (1988). Intelligence. In R.J. Sternberg & E.E. Smith (Eds.), The psychology of human thought (pp.267-308). Cambridge: Cambridge University Press.

Stromgren, L.S. (1977). The influence of depression on memory. Acta Psychiatrica Scandinavica, 56, 109-128.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662.

Surtees, P.G. & Sashidharan, S.P. (1986). Psychiatric morbidity in two matched community samples: A comparison of rates and risks in Edinburgh and St. Louis. Journal of Affective Disorders, 10, 101-113.

Szabadi, E., Bradshaw, C.M. & Besson, J.A.O. (1976). Elongation of pause-time in speech: A simple, objective measure of motor retardation in depression. British Journal of Psychiatry, 129, 592-597.

Tabachnick, B.G. & Fidell, L.S. (1983). Using multivariate statistics. New York: Harper & Row.

Tabachnick, B.G. & Fidell, L.S. (1989). Using multivariate statistics (2nd edition). New York: Harper & Row.

Tancer, M.E., Brown, T.M., Evans, D.L., Ekstrom, D., Haggerty, J.J., Pedersen, C. & Golden, R.N. (1990). Impaired effortful cognition in depression. Psychiatry Research, 31, 161-168.

Taylor, S.E. & Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. Psychological Bulletin, 103, 193-210.

Teasdale, J.D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? Advances in Behaviour Research and Therapy, 5, 3-25.

Teasdale, J.D. (1988). Cognitive vulnerability to persistent depression. Cognition and Emotion, 2, 247-274.

Teasdale, J.D. & Barnard, P.J. (1993). Affect, cognition and change: Re-modelling depressive thought. Hove: Lawrence Erlbaum Associates.

Teadale, J.D. & Dent, J. (1987). Cognitive vulnerability to depression: An investigation of two hypotheses. British Journal of Clinical Psychology, 26, 113-126.

Teasdale, J.D. & Fogarty, S.J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant memories from episodic memory. Journal of Abnormal Psychology, 88, 248-257.

Teasdale, J.D. & Russell, L. (1983). Differential effects of induced mood on the recall of positive, negative and neutral words. British Journal of Clinical Psychology, 22, 163-171.

Teasdale, J.D. & Taylor, R. (1981). Induced mood and accessibility of memories: An effect of mood state or of induction procedure? British Journal of Clinical Psychology, 20, 39-48.

Teasdale, J.D., Taylor, R. & Fogarty, S.J. (1980). Effects of induced elation-depression on the accessibility of memories of happy and unhappy experiences. Behaviour, Research and Therapy, 18, 339-346.

Tirre. W.C. & Pena, C.M. (1993). Components of quantitative reasoning: General and group ability factors. Intelligence, 17, 501-521.

Toms, M., Morris, N. & Ward, D. (1993). Working memory and conditional reasoning. The Quarterly Journal of Experimental Psychology, 46A, 679-699.

Trichard, C., Martinot, J.L., Alagille, M., Masure, M.C., Hardy, P., Ginestet, D. & Feline, A. (1995). Time course of prefrontal lobe dysfunction in severely depressed in-patients: A longitudinal neuropsychological study. Psychological Medicine, 25, 79-85.

Tucker, D.M. & Liotti, M. (1989). Neuropsychological mechanisms of anxiety and depression. In F. Boller & J. Grafman (Eds.), Handbook of Neuropsychology, Vol. 3. Amsterdam: Elsevier.

Turner, M.L. & Engle, R.W. (1989). Is working memory capacity task dependent? Journal of Memory and Language, 28, 127-154.

Turvey, M.T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. Psychological Review, 80, 1-52.

Tversky, A. & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. Psychological Review, 90, 293-315.

Tyler, S.W., Hertel, P.T., McCallum, M.C. & Ellis, H.C. (1979). Cognitive effort and memory. Journal of Experimental Psychology: Human Learning and Memory, 5, 607-617.

Velten, E. (1968). A laboratory task for induction of mood states. Behaviour Research and Therapy, 6, 473-482.

Vilkki, J. & Holst, P. (1989). Deficient programming in spatial learning after frontal lobe damage. Neuropsychologia, 27, 971-976.

Vredenburg, K., Flett, G.L. & Krames, L. (1993). Analogue versus clinical depression: A critical reappraisal. Psychological Bulletin, 113, 327-344.

Warren, E.W. & Groome, D.H. (1984). Memory test performance under three different waveforms of ECT for depression. British Journal of Psychiatry, 144, 370-375.

Wason, P.C. (1966). Reasoning. In B.M. Foss (Ed.), New horizons in psychology, Vol. 1. Harmondsworth: Penguin.

Watkins, J.T. & Rush, A.J. (1983). Cognitive Response Test. Cognitive Therapy and Research, 7, 425-436.

Watkins, P.C., Mathews, A., Williamson, D.A. & Fuller, R.D. (1992). Mood-congruent memory in depression: Emotional priming or elaboration? Journal of Abnormal Psychology, 101, 581-586.

Watkins, P.C., Vache, K. Verney, S.P., Muller, S. & Mathews, A. (1996). Unconscious mood-congruent memory bias in depression. Journal of Abnormal Psychology, 105, 34-41.

Watts, F.N. (1992). Applications of current cognitive theories of the emotions to the conceptualization of emotional disorders. British Journal of Clinical Psychology, 31, 153-167.

Watts, F.N. (1993). Problems of memory and concentration. In C.G. Costello (Ed.), Symptoms of depression (pp. 113-140). New York: John Wiley.

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., Dalgleish, T., Bourke, P. & Healy, D. (1990). Memory deficit in clinical depression: Processing resources and the structure of materials. Psychological Medicine, 20, 345-349.

Watts, F.N., MacLeod, A.K. & Morris, L. (1988). Associations between phenomenal and objective aspects of concentration problems in depressed patients. British Journal of Psychology, 79, 241-250.

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

Watts, F.N. & Sharrock, R. (1985). Description and measurement of concentration problems in depressed patients. Psychological Medicine, 15, 317-326.

Watts, F.N. & Sharrock, R. (1987). Cued recall in depression. British Journal of Clinical Psychology, 26, 149-150.

Waugh, N.C. & Norman, D.A. (1965). Primary memory. Psychological Review, 72, 89-104.

Wechsler, D. (1945). A standardised memory scale for clinical use. Journal of Psychology, 19, 87-95.

Wechsler, D. (1955). Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation.

Wechsler, D. (1981). Wechsler adult intelligence scale - revised Manual. New York: Psychological Corporation.

Weckowicz, T.E., Tam, C.I., Bay, K.S., Collier, G. & Beelen, L. (1981). Perception of reinforcement and psychomotor retardation in depressed patients. Canadian Journal of Behavioural Science, 13, 129-143.

Weckowicz, T.E., Tam, C.I., Mason, J. & Bay, K.S. (1978). Speed in test performance in depressed patients. Journal of Abnormal Psychology, 87, 578-582.

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

Whitehead, A. (1973). Verbal learning and memory in elderly depressives. British Journal of Psychiatry, 123, 203-208.

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N.J.: Erlbaum.

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & D.R. Davies (Eds.), Varieties of Attention (pp. 63-102). Orlando: Academic Press.

Wickens, C.D. (1989). Attention and skilled performance. In D.H. Holding (Ed.), Huma n Skills, Second Edition (pp.71-105). Chichester: John Wiley.

Wiegersma, S., van der Scheer, E. & Hijman, R. (1990). Subjective ordering, short-term memory, and the frontal lobes. Neuropsychologia, 28, 95-98.

Wilkinson, I.M. & Blackburn, I.M. (1981). Cognitive style in depressed and recovered depressed patients. British Journal of Clinical Psychology, 20, 283-292.

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

Williams, J.M.G., Watts, F.N., MacLeod, C. & Mathews, A. (1988). Cognitive psychology and emotional disorders. Chichester: Wiley.

Willis, M.H. & Blaney, P.H. (1978). Three tests of the learned helplessness model of depression. Journal of Abnormal Psychology, 87, 131-136.

Winer, B.J. (1971). Statistical principles in experimental design. Second Edition. New York: McGraw-Hill.

Wittchen, H.U. & Essau, C.A. (1993). Comorbidity and mixed anxiety-depressive disorders: Is there epidemiologic evidence? Journal of Clinical Psychiatry, 54:1 (suppl), 9-15.

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.

Wolpe, J. (1979). The experimental model and treatment of neurotic depression. Behaviour Research and Therapy, 17, 555-565.

Yee, C.M., Deldin, P.J. & Miller, G.A. (1992). Early stimulus processing in dysthymia and anhedonia. Journal of Abnormal Psychology, 101, 230-233.

Yee, C.M. & Miller, G.A. (1994). A dual-task analysis of resource allocation in dysthymia and anhedonia. Journal of Abnormal Psychology, 103, 625-636.

Zarantonello, M., Slaymaker, F., Johnson, J. & Petzel, T. (1984). Effects of anxiety and depression on anagram performance, ratings of cognitive interference, and the negative subjective evaluation of performance. Journal of Clinical Psychology, 40, 20-25.

Zemore, R. & Dell, L.W. (1983). Interpersonal problem-solving skills and depression-proneness. Personality and Social Psychology Bulletin, 9, 231-235.

Zemore, R. & Johansen, L.J. (1980). Depression, helplessness, and failure attributions. Canadian Journal of Behavioural Science, 12, 167-174.

Zuroff, D.C. (1981). Depression and attribution: Some new data and a review of old data. Cognitive Therapy and Research, 5, 273-281.

APPENDIX 1

BECK DEPRESSION INVENTORY (BDI)

On this questionnaire are groups of statements. Please read each group of statements carefully. Then pick out the one statement in each group which best describes the way you have been feeling in the PAST WEEK, INCLUDING TODAY. Circle the number beside the statement you picked. If several statements in the group seem to apply equally well, circle each one. Be sure to read all the statements in each group before making your choice.

1.	I do not feel sad	0
	I feel sad	1
	I am sad all the time and can't snap out of it	2
	I am so sad or unhappy that I can't stand it	3
2.	I am not particularly discouraged about the future	0
	I feel discouraged about the future	1
	I feel I have nothing to look forward to	2
	I feel that the future is hopeless and that things cannot improve	3
3.	I do not feel like a failure	0
	I feel I have failed more than the average person	1
	As I look back on my life, all I can see is a lot of failures	2
	I feel I am a complete failure as a person	3
4.	I get as much satisfaction out of things as I used to	0
	I don't enjoy things the way I used to	1
	I don't get real satisfaction out of anything anymore	2
	I am really dissatisfied and bored with everything at the moment	3
5.	I don't feel particularly guilty	0
	I feel guilty a good part of the time	1
	I feel quite guilty most of the time	2
	I feel guilty all of the time	3

6.	I don't feel I am being punished	0
	I feel I may be punished	1
	I expect to be punished	2
	I feel I am being punished	3
7.	I don't feel disappointed in myself	0
	I am disappointed in myself	1
	I am disgusted with myself	2
	I hate myself	3
8.	I don't feel I am worse than anybody else	0
	I am critical of myself for my weaknesses or mistakes	1
	I blame myself for all my faults	2
	I blame myself for anything bad that happens	3
9.	I don't have any thoughts of killing myself	0
	I have thoughts of killing myself, but I would not carry them out	1
	I would like to kill myself	2
	I would kill myself if I had the chance	3
10.	I don't cry any more than usual	0
	I cry more than I used to	1
	I cry all the time now	2
	I used to be able to cry, but now I can't cry even though I want to	3
11.	I am no more irritated now than I ever am	0
	I get annoyed or irritated more easily than I used to	1
	I feel irritated all the time now	2
	I don't get irritated at all by things that used to irritate me	3
12.	I have not lost interest in other people	0
	I am less interested in other people than I used to be	1
	I have lost most of my interest in other people	2
	I have lost all of my interest in other people	3

13.	I make decisions about as well as I ever could	0
	I put off making decisions more than I used to	1
	I have greater difficulty in making decisions than before	2
	I can't make decisions at all anymore	3
14.	I don't feel I look any worse than I used to	0
	I am worried that I am looking old or unattractive	1
	I feel that there are permanent changes in my appearance that make me look unattractive	2
	I believe that I look ugly	3
15.	I can work about as well as before	0
	It takes extra effort to get started at doing something	1
	I have to push myself very hard to do anything	2
	I can't do any work at all	3
16.	I can sleep as well as usual	0
	I don't sleep as well as I used to	1
	I wake 1-2 hours earlier than usual and find it hard to get back to sleep	2
	I wake up several hours earlier than I used to and cannot get back to sleep	3
17.	I don't get more tired than usual	0
	I get tired more easily than I used to	1
	I get tired from doing almost anything	2
	I am too tired to do anything	3
18.	My appetite is no worse than usual	0
	My appetite is not as good as it used to be	1
	My appetite is much worse now	2
	I have no appetite at all now	3

19.	I haven't lost much weight, if any, lately	0
	I have lost more than 5 pounds	1
	I have lost more than 10 pounds	2
	I have lost more than 15 pounds	3

I am purposely trying to lose weight by eating less YES/NO

20.	I am no more worried about my health than usual	0
	I am worried about my physical problems such as aches and pains; or upset stomach; or constipation	1
	I am very worried about physical problems and its hard to think of much else	2
	I am so worried about my physical problems, that I cannot think about anything else	3

21.	I have not noticed any recent changes in my interest in sex	0
	I am less interested in sex than I used to be	1
	I am much less interested in sex now	2
	I have lost interest in sex completely	3

APPENDIX 2

SUBJECT INITIAL SELECTION INFORMATION SHEET

We are hoping to complete the study within the next few weeks. We would be very grateful if you would fill out the information below and complete the questionnaire. These will be collected at the end of the lecture. All information will of course be kept completely confidential.

Some of you will be contacted and asked to take part in the rest of the study.

NAME:

SEX: DATE OF BIRTH:

ARE YOU RIGHT OR LEFT HANDED?

NAME OF COURSE:

CONTACT ADDRESS:

CONTACT TEL NO:

APPENDIX 3

INFORMATION SHEET FOR Ss PARTICIPATING IN STUDY

Information Sheet

The effects of mood on learning, memory and reasoning

Director of project: Dr. S. Channon
Telephone 0171 387 7050 ext. 5931

This study is concerned with the ways in which people learn, remember and solve problems. In particular, we are looking at the ways in which these processes may be affected by moods. The study is designed to improve our understanding of the relationship between mood and cognitive processes. This has relevance for everyday living where learning and problem-solving play an important role.

You will be given a series of psychological tests which measure aspects of learning, memory and problem-solving. These will be arranged to take place at a time convenient to you, and you will be able to take breaks if you feel tired. You will also be asked a series of questions concerned with the way you are feeling and any difficulties you have been having, and asked to fill out a set of questionnaires. The study does not include any blood tests or other medical procedures. You will be paid a small sum to cover your expenses.

You will be asked to sign a consent form, and any information you give will be treated in strict confidence. You do not have to take part in this study if you do not want to. If you decide to take part you may withdraw at any time without giving a reason.

Should this experiment prompt any concern on your part in relation to your own mood, you are welcome to discuss this further with us.

APPENDIX 4

CONSENT FORM

The effects of mood on learning, memory and reasoning

This study will look at how people learn, remember and reason, to help understand the nature and extent of any difficulties associated with mood.

Director of project: Dr. Shelley Channon

To be completed by the volunteer: **Delete as necessary**:

Have you read the information sheet about the study? Yes/No

Have you had an opportunity to ask questions and discuss the study? Yes/No

Have you received satisfactory answers to all your questions? Yes/No

Have you received enough information about this study? Yes/No

Which researcher have you spoken to about this study? .....

Do you understand that you are free to withdraw from this study at any time, and without giving a reason for withdrawing? Yes/No

Do you agree to take part in this study? Yes/No

Signature of volunteer .....

Name.....

Date .....

Address.....

Signature of researcher.....

Name.....

Date.....

APPENDIX 5

SCORING PROTOCOL FOR RESPONSE PATTERNS ON EACH SET OF FOUR NON-FEEDBACK TRIALS ON THE DISCRIMINATION LEARNING TASK

	Left	Small	White	B/F/S/T	A/O/H/X	Black	Big	Right
Problem 1	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
Problem 2	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x

N.B. Any other pattern indicates that the subject is not sorting to a consistent hypothesis.

	Left	Small	White	B/F/S/T		A/O/H/X	Black	Big	Right
Problem 3	x	x	x	x		x	x	x	x
	x	x	x	x		x	x	x	x
	x	x	x	x		x	x	x	x
	x	x	x	x		x	x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
Problem 4	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x
	x	x	x	x	x		x	x	x

N.B. Any other pattern indicates that the subject is not sorting to a consistent hypothesis.

APPENDIX 6

SOLUTIONS AFTER FEEDBACK TRIALS ON THE DISCRIMINATION LEARNING TASK

W=wrong C=correct

Feedback WCW

Outcome 1	Left choice	4 of	Big	Black	AFHX	Right
	Right choice	4 of	Small	White	BOST	Left
Outcome 2	Left choice	2 of	Big	Black	BOST	Left
	Right choice	2 of	Small	White	AFHX	Right
Outcome 3	Left choice	1 of	Small	Black	BOST	Right
	Right	1 of	Big	White	AFHX	Left

Feedback CWC

Outcome 1	Left choice	4 of	Small	White	BOST	Left
	Right choice	4 of	Big	Black	AFHX	Right
Outcome 2	Left choice	2 of	Small	White	AFHX	Right
	Right choice	2 of	Big	Black	BOST	Left
Outcome 3	Left choice	1 of	Big	White	AFHX	Left
	Right choice	1 of	Small	Black	BOST	Right

Feedback CWW

Outcome 1	Left choice	4 of	Small	White	BOST	Left
	Right choice	4 of	Big	Black	AFHX	Right
Outcome 2	Left choice	2 of	Small	White	AFHX	Right
	Right choice	2 of	Big	Black	BOST	Left
Outcome 3	Left choice	1 of	Small	Black	BOST	Right
	Right	1 of	Big	White	AFHX	Left

Feedback WCC

Outcome 1	Left choice	4 of	Big	Black	AFHX	Right
	Right choice	4 of	Small	White	BOST	Left
Outcome 2	Left choice	2 of	Big	Black	BOST	Left
	Right choice	2 of	Small	White	AFHX	Right
Outcome 3	Left choice	1 of	Big	White	AFHX	Left
	Right choice	1 of	Small	Black	BOST	Right

APPENDIX 7

INTEGRATIVE REASONING TEST MATERIALS

Set A Practice problems

1. B and C do the opposite.

   If B increases, what happens to C?

2. K and L do the same.

   M and N do the opposite.

   If M decreases, what happens to N?

3. V and W do the same.

   U and V do the opposite.

   If U increases, what happens to W?

4. F and G do the same.

   D and E do the opposite.

   B and C do the same.

   If F decreases, what happens to G?

5. R and S do the opposite.

   P and Q do the same.

   Q and R do the opposite.

   If P increases, what happens to S?

Set A Experimental problems

1. B and C do the same.

   D and E do the opposite

   If B increases, what happens to C?

2. M and N do the opposite.

   L and M do the same.

   If L decreases, what happens to N?

3. A and B do the same.

   C and D do the opposite.

   B and C do the same.

   If A decreases, what happens to D?

4. G and H do the opposite.

   F and G do the same.

   If F increases, what happens to H?

5. U and V do the opposite.

   S and T do the same.

   T and U do the same.

   If S decreases, what happens to V?

6. G and H do the opposite.

   S and T do the same.

   T and U do the same.

   If G decreases, what happens to H?

7. R and S do the opposite.

   T and U do the opposite.

   V and W do the opposite.

   If V increases, what happens to W?

8. E and F do the same.

   G and H do the opposite.

   F and G do the same.

   If E increases, what happens to H?

9. R and S do the opposite.

   P and Q do the same.

   If P increases, what happens to Q?

10. A and B do the opposite.

    If A increases, what happens to Q?

11. L and M do the opposite.

    J and K do the same.

    If J decreases, what happens to K?

12. C and D do the same.

    A and B do the opposite.

    E and F do the same.

    If C decreases, what happens to D?

13. L and M do the same.

    If L increases, what happens to M?

14. E and F do the same.

    If E decreases, what happens to F?

15. B and C do the same.

    A and B do the opposite.

    If A increases, what happens to C?

16. Q and R do the opposite.

    If Q decreases, what happens to R?

17. G and H do the same.

    E and F do the opposite.

    If G decreases, what happens to H.

18. R and S do the opposite.

    Q and R do the same.

    If Q decreases, what happens to S?

19. F and G do the same.

    B and C do the same.

    D and E do the same.

    If F increases, what happens to G?

20. W and X do the opposite.

    U and V do the opposite.

    V and W do the opposite.

    If U increases, what happens to X?

Set B Practice problems

1. G and H do the opposite.

   If G decreases, what happens to H?

2. V and W do the same.

   T and U do the opposite.

   If T increases, what happens to U?

3. K and L do the opposite.

   J and K do the opposite.

   If J decreases, what happens to L?

4. P and Q do the same.

   T and U do the opposite.

   R and S do the same.

   If P increases, what happens to Q?

5. D and E do the same.

   F and G do the opposite.

   E and F do the same.

   If D increases, what happens to G?

Set B Experimental problems

1. C and D do the opposite.

   A and B do the same.

   E and F do the same.

   If E decreases, what happens to F?

2. T and U do the same.

   R and S do the same.

   If T decreases, what happens to U?

3. E and F do the opposite.

   G and H do the same.

   If E decreases, what happens to F?

4. C and D do the same.

   If C decreases, what happens to D?

5. S and T do the same.

   Q and R do the opposite.

   R and S do the same.

   If Q decreases, what happens to T?

6. K and L do the opposite.

   M and N do the opposite.

   If M increases, what happens to N?

7. W and X do the same.

   U and V do the opposite.

   S and T do the opposite.

   If W increases, what happens to X?

8. C and D do the same.

   E and F do the same.

   D and E do the opposite.

   If C increases, what happens to F?

9 S and T do the opposite.

   Q and R do the same.

   If Q increases, what happens to R?

10. U and V do the opposite.

    If U increases, what happens to V?

11. L and M do the same.

K and L do the opposite.

If K decreases, what happens to M?

12. T and U do the opposite.

If T decreases, what happens to U?

13. W and X do the opposite.

V and W do the same.

If V increases, what happens to X?

14. E and F do the same.

D and E do the same.

If D increases, what happens to F?

15. S and T do the opposite.

Q and R do the same.

U and V do the same.

If S increases, what happens to T?

16. F and G do the same.

If F increases, what happens to G?

17. V and W do the opposite.

T and U do the same.

X and Y do the opposite.

If X decreases, what happens to Y?

18. B and C do the opposite.

A and B do the opposite.

If A decreases, what happens to C?

19. M and N do the same.

K and L do the opposite.

L and M do the opposite.

If K decreases, what happens to N?

20. R and S do the opposite.

T and U do the same.

S and T do the opposite.

If R increases, what happens to U?

APPENDIX 8

TEST MATERIALS FOR FAN EFFECT TASK

LEARNING TRIALS

18 experimental 'facts' with fan level in brackets:

The clerk nervously watched the tightrope walker (1-1)

The pharmacist arrived at the train station early (1-1)

The newsreader cut the apple pie into six pieces (1-1)

The vicar paid a deposit on the new video (2-2)

The author found a spot to sunbathe at the beach (2-2)

The judge decided to play chess with a friend (2-2)

The teacher got change from the ticket machine (3-3)

The executive ran at least four miles a day (3-3)

The executive took the car for a short test drive (3-3)

The doctor got change from the ticket machine (3-3)

The teacher ran at least four miles a day (3-3)

The doctor took the car for a short test drive (3-3)

The judge got change from the ticket machine (2-3)

The author took the car for a short test drive (2-3)

The vicar ran at least four miles a day (2-3)

The executive paid a deposit on the new video (3-2)

The doctor found a spot to sunbathe at the beach (3-2)

The teacher decided to play chess with a friend (3-2)

QUESTIONS

18 questions with correct answer in brackets:

1. Who decided to play chess with a friend?

(teacher, judge)

2. The newreader did what?

(cut the apple pie into six pieces)

3. Who paid a deposit on the new video?

(vicar, executive)

4. Who nervously watched the tightrope walker?

(clerk)

5. The vicar did what?

(ran at least four miles a day, paid a deposit on the new video)

6. The executive did what?

(ran at least four miles a day, took the car for a short test drive, paid a deposit on the new video)

7. The judge did what?

(decided to play chess with a friend, tot change from the ticket machine)

8. Who got change from the ticket machine?

(teacher, doctor, judge)

9. Who took the car for a short test drive?

(executive, doctor, author)

10. The clerk did what?

(nervously watched the tightrope walker)

11. Who arrived at the train station early?

(pharmacist)

12. The pharmacist did what?

(arrived at the train station early)

13. The author did what?

(found a spot to sunbathe at the beach, took the car for a short test drive)

14. Who ran at least four miles a day?

(executive, teacher, vicar)

15. Who cut the apple pie into six pieces?

(newsreader)

16. Who found a spot to sunbathe at the beach?

(author, doctor)

17. The teacher did what?

(got change from the ticket machine, ran at least four miles a day, decided to play chess with a friend)

18. The doctor did what?

(took the car for a short test drive, got change from the ticket machine, found a spot to sunbathe at the beach)

RECOGNITION TEST

Practice sentence: The fish swam in the lake

Absent

Seventy trials with feedback presented as seven booklets (fan levels in brackets):

The clerk found a spot to sunbathe at the beach (1-2)

Absent

The teacher arrived at the train station early (3-1)

Absent

The vicar took the car for a short test drive (2-3)

Absent

The pharmacist nervously watched the tightrope walker (1-1)

Absent

The executive got change from the ticket machine (3-3)

Absent

The doctor paid a deposit on the new video (3-2)

Absent

The clerk nervously watched the tightrope walker (1-1)

Present

The judge got change from the ticket machine (2-3)

Present

The executive took the car for a short test drive (3-3)

Present

The pharmacist arrived at the train station early (1-1)

Present

The newsreader cut the apple pie into six pieces (1-1)

Present

The author found a spot to sunbathe at the beach (2-2)

Present

The judge paid a deposit on the new video (2-2)

Absent

The teacher ran at least four miles a day (3-3)

Present

The author decided to play chess with a friend (2-2)

Absent

The teacher took the car for a short test drive (3-3)

Absent

The vicar paid a deposit on the new video (2-2)

Present

The teacher decided to play chess with a friend (3-2)

Present

The judge decided to play chess with a friend (2-2)

Present

The vicar found a spot to sunbathe at the beach (2-2)

Absent

The newreader arrived at the train station early (1-1)

Absent

The clerk cut the apple pie into six pieces (1-1)

Absent

The doctor ran at least four miles a day (3-3)

Absent

The pharmacist took the car for a short test drive (1-3)

Absent

The doctor got change from the ticket machine (3-3)

Present

The newsreader found a spot to sunbathe at the beach (1-2)

Absent

The vicar ran at least four miles a day (2-3)

Present

The teacher took the car for a short test drive (3-3)

Absent

The teacher ran at least four miles a day (3-3)

Present

The newsreader arrived at the train station early (1-1)

Absent

The doctor ran at least four miles a day (3-3)

Absent

The doctor found a spot to sunbathe at the beach (3-2)

Present

The vicar paid a deposit on the new video (2-2)

Present

The pharmacist arrived at the train station early (1-1)

Present

The executive took the car for a short test drive (3-3)

Present

The judge cut the apple pie into six pieces (2-1)

Absent

The author got change from the ticket machine (2-3)

Absent

The doctor got change from the ticket machine (3-3)

Present

The author decided to play chess with a friend (2-2)

Absent

The executive took the car for a short test drive (3-3)

Present

The pharmacist arrived at the train station early (1-1)

Present

The clerk nervously watched the tightrope walker (1-1)

Present

The author took the car for a short test drive (2-3)

Present

The pharmacist nervously watched the tightrope walker (1-1)

Absent

The executive got change from the ticket machine (3-3)

Absent

The doctor got change from the ticket machine (3-3)

Present

The clerk cut the apple pie into six pieces (1-1)

Absent

The author found a spot to sunbathe at the beach (2-2)

Present

The doctor ran at least four miles a day (3-3)

Absent

The judge paid a deposit on the new video (2-2)

Absent

The clerk cut the apple pie into six pieces (1-1)

Absent

The judge paid a deposit on the new video (2-2)

Absent

The vicar paid a deposit on the new video (2-2)

Present

The executive got change from the ticket machine (3-3)

Absent

The judge decided to play chess with a friend (2-2)

Present

The newreader ran at least four miles a day (1-3)

Absent

The executive nervously watched the tightrope walker (3-1)

Absent

The clerk nervously watched the tightrope walker (1-1)

Present

The author found a spot to sunbathe at the beach (2-2)

Present

The vicar found a spot to sunbathe at the beach (2-2)

Absent

The newsreader cut the apple pie into six pieces (1-1)

Present

The judge decided to play chess with a friend (2-2)

Present

The pharmacist nervously watched the tightrope walker (1-1)

Absent

The executive decided to play chess with a friend (3-2)

Absent

The newsreader cut the apple pie into six pieces (1-1)

Present

The newsreader arrived at the train station early (1-1)

Absent

The teacher took the car for a short test drive (3-3)

Absent

The teacher ran at least four miles a day (3-3)

Present

The author decided to play chess with a friend (2-2)

Absent

The vicar found a spot to sunbathe at the beach (2-2)

Absent