PARKINSON'S DISEASE:
MEMORY DEFICITS

A dissertation presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
in Psychology
at
Massey University,
New Zealand

Craig J. Whittington

1999

Abstract

Parkinson's disease (PD) has long been thought of as a debilitating motor disease: only relatively recently has research focused on cognitive functioning. It is now widely accepted that memory processes are among the primary cognitive functions to deteriorate in PD. However, less is known about the role that task variables (e.g., difficulty) and participant characteristics (e.g., gender, disease progression) play in this relationship. In addition, few studies in the PD literature have looked at the important issues of statistical power and the magnitude of memory deficits. The present investigation addressed some of these concerns. The first stage involved conducting a power analysis, based on 48 studies, followed by a meta-analysis. The meta-analysis included 32 effect sizes from studies assessing recognition memory in PD. This analysis paved the way for a large-scale study examining recognition memory in 41 nondemented PD participants compared to 41 age- and education-matched healthy controls. Both verbal and nonverbal recognition tasks were specifically designed for this purpose, the latter employing two levels of difficulty. In addition, prospective memory (remembering to remember) was examined with two event-based tasks because no study to date has looked at this issue in PD. The results of the power analysis indicated that past research has typically had insufficient statistical power to detect all but the largest memory deficits. Integrating the data from many studies, the meta-analysis suggested that nondemented, medicated PD participants may suffer from small recognition deficits. Support was provided by the subsequent primary study. In addition, it was found that the progression of memory impairment operates in parallel with the progression of motor symptoms. Moreover, task demands interacted with disease stage, such that nonverbal recognition deficits were only seen in early-stage PD participants when the task was made more difficult. Conversely, advanced-stage participants produced deficits irrespective of the level of difficulty. With respect to prospective memory, only the advanced-stage participants showed clear evidence of memory impairment. The main outcome of this research is that recognition memory impairment does occur in PD. However, low levels of statistical power in previous

research and moderating factors such as symptom severity and task difficulty have likely obscured true deficits.

Table Of Contents

List of Tables

List of Figures

Preface

The work presented here evolved from a review of the literature on cognitive deficits in Parkinson's disease (PD) focusing on memory deficits. The review indicated that it is widely recognised that people with PD often suffer from cognitive dysfunction including memory deficits. However, the exact nature of these deficits is still under debate. It also became clear that there are considerable inconsistencies in the research findings coupled with little consideration for issues of statistical power, effect size (ES), and cumulative research synthesis. In the PD literature, it is often reported that recall memory is impaired, whereas recognition memory is relatively intact. However, these reports are often based on studies with inadequate statistical power.

Statistical power is important because studies with low statistical power will have little chance of detecting true deficits (effects) when using conventional statistical significance testing (Cohen, 1988). Moreover, statistical significance provides little information about the magnitude of any deficit. To interpret an individual study's results, an alternative to statistical significance testing is the use of point estimates of ES and confidence intervals (Schmidt, 1992, 1996), the approach adopted in this thesis.

To interpret the results of multiple studies, an alternative to the traditional narrative review is the meta-analysis. Meta-analysis is a perspective that uses objective techniques for research synthesis. It is generally used to produce a quantitative summary of a set of studies based on ES rather than significance tests (Bangert-Drowns, 1986; Cooper & Hedges, 1994). In situations where a field of research contains many inconsistent results, as judged by significance tests, meta-analysis is potentially most effective. In these circumstances, the statistical techniques developed by Hunter and Schmidt (1990) allow a meta-analyst to decide whether the variance in the results is due to sampling error or possible moderating factors. Without quantification of the results, it is difficult, if not

impossible in most situations, for a narrative review to achieve this. Above all, meta-analysis maintains all of the qualities of the traditional literary format, but also provides a process for judging whether a set of studies produced consistent results.

The main purpose of the present research, then, was to examine memory deficits, more specifically recognition memory deficits, in PD. A meta-analytical review and a power analysis paved the way for a large-scale experimental investigation of recognition memory impairment in PD. In addition, the study was designed to look at prospective memory, a topic that has not been examined in people with PD. This investigation also assessed a range of possible moderating factors that might interact with the disease process itself to bring about any observed memory deficits.

It should be noted that the meta-analytical data reported here are only a subset of what were originally collected. The complete data set includes ESs calculated from over 100 studies that used recall and/or recognition memory tasks. However, to limit the focus of the present investigation, only the recognition memory data were analysed and reported here.

This dissertation is unconventional in the sense that statistical power and ES are used extensively in the interpretation of both past and present research findings, rather than statistical significance. Further justification for this approach is given both in chapter 2 and in chapter 4. In my opinion, because the present work (like most current psychological research) is exploratory, nothing is lost by taking this approach. In fact, as demonstrated in chapter 4, interesting effects would have been discarded had this approach not been taken.

Finally, a planned comparison approach with simple effects tests was used to examine the influence of potential moderating variables at two different experimental sessions. In retrospect, a multivariate approach, such as multiple regression, could have been taken instead. However, as discussed in chapter 4,

the composition of the sample varied as a function of time. Therefore, a multivariate analysis would have been complicated by differences between sessions.

Organisation of this Dissertation

This dissertation consists of four chapters. Chapter 1 includes six major sections that provide a general overview of PD and the known cognitive deficits, recognition memory, prospective memory, memory systems, and concludes with a summary and rationale for conducting a power analysis and a meta-analysis on the PD literature.

Chapter 2 contains the statistical power analysis and the meta-analysis. It has four major sections detailing the issues and rationale, the methods used, the results, and a discussion. (Note that substantive parts of the method, results, and discussion have been accepted for publication. A copy of the article (in press) is provided in Appendix U.)

Chapter 3 has four major sections detailing the experimental rationale, the methods used, the results, and a discussion of the results.

Finally, chapter 4 provides a general discussion of the present research, including its limitations and future directions, and a general conclusion.

Acknowledgements

I have been most fortunate to work with John Podd, an amazing supervisor and mentor.

My appreciation goes out to all those people who helped me recruit participants, and most of all, to those who gave their time freely to take part in the study.

Thanks to O'Reilly Tucker for his help as an independent judge of coding reliability, and Harvey Jones for his help with the computer programming.

A special thanks to Hayley, for her unwavering support and encouragement. In addition, I am grateful to all the members of my family for always being there when I needed them.

Financial support for this research came from the Massey University Research Fund (1-0575-67302A), the Massey University Research Equipment Fund (1-0575-98025), and from the School of Psychology, Massey University.

All procedures involving human participants were approved by the Massey University Human Ethics Committee, and depending on region, the Manawatu-Whanganui Ethics Committee, the Hawkes Bay Ethics Committee, or the Wellington Ethics Committee.

CHAPTER 1

INTRODUCTION

Overview of Parkinson's Disease

Parkinson's disease (PD) takes its name from James Parkinson who provided the following clinical description of this neurodegenerative disorder that he labelled *Paralysis Agitans*, over 180 years ago:

> Involuntary tremulous motion, with lessened muscular power, in parts not in action and even when supported; with a propensity to bend the trunk forwards, and to pass from a walking to a running pace: the senses and intellect being uninjured. (Parkinson, 1817/1955, p. 153)

Despite Parkinson's claim that the disease affected motor functions but left the intellect intact, a number of early authors noted that some patients suffered from cognitive disturbances (e.g., Ball, 1882; Charcot & Vulpian, 1861; Gowers, 1899; all cited in Dubois, Boller, Pillon, & Agid, 1991). However, over the intervening years since Parkinson's description of PD the major research focus has been on the motor symptoms. In this respect, a major breakthrough was made in the 1960s with the discovery of greatly reduced dopamine levels in the putamen, caudate nucleus, and substantia nigra in PD patients (Schneider, 1993) (see Figure 1). This led to the treatment of PD with dopamine replacement therapy, and eventually to other drugs such as dopamine agonists (Schneider, 1993).

Another important discovery occurred in the 1980s when it was found that a meperidine analogue called 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)

reproduced the major pathology of PD and resulted in parkinsonian-type symptoms (Schneider, 1993). This finding stimulated a resurgence in research activity that has continued to this day. However, it has become clear that the treatment of PD has not been successful in alleviating the cognitive symptoms of this disorder. These cognitive deficits will be the focus of the review that follows. However, the cognitive symptoms do not occur in complete isolation from the physical symptoms, so the neurological profile of PD will be briefly introduced first. The review also provides an overview of memory systems, and a summary of recognition memory models and measures.

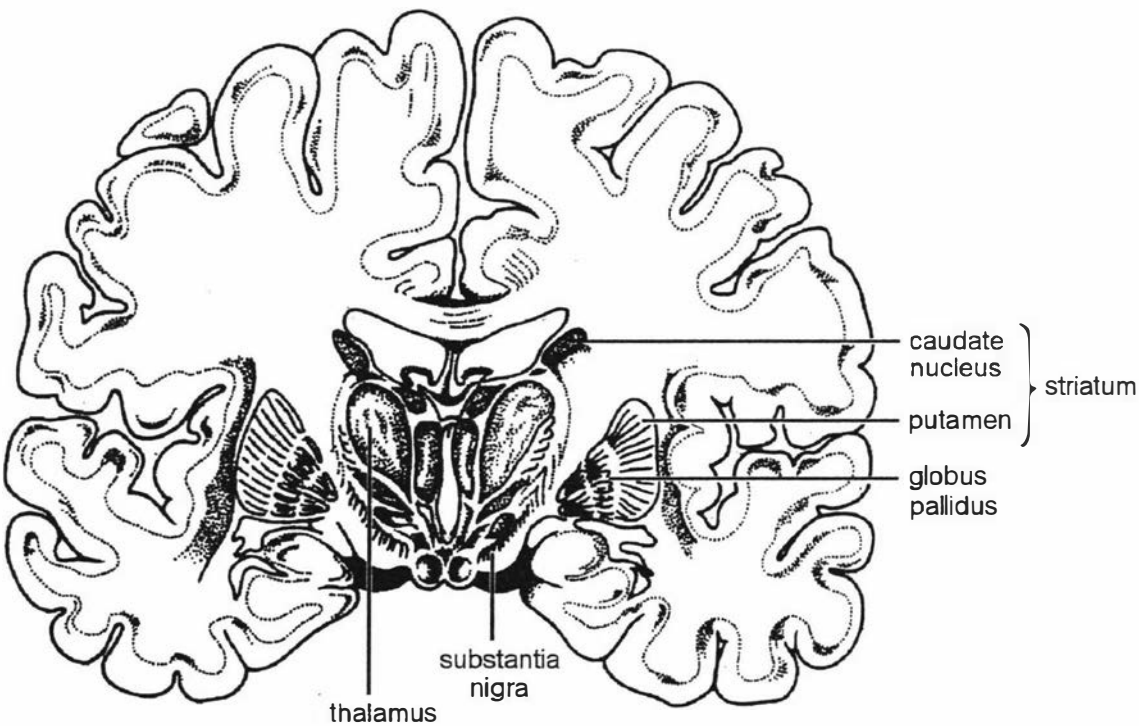


Figure 1. A coronal section through the normal human brain showing the location of the major brain structures involved in Parkinson's disease. From "The Brain: Degeneration, Damage and Disorder" by J. Metcalfe, 1998, p. 110. Copyright 1998 by The Open University.

Parkinsonism

Currently, it is recognised that PD includes a number of diseases and syndromes that share some of the same physical symptoms (Hopfensperger & Koller, 1991). Hopfensperger and Koller categorised parkinsonism into several different types: (a) Parkinson's disease (idiopathic parkinsonism), (b) "parkinsonism plus" syndromes (neuronal system degeneration with variable parkinsonian features, but minimal tremor), such as progressive supranuclear palsy, and (c) secondary parkinsonism (disorder other than PD, but with parkinsonian features), such as postinfectious diseases (e.g., Von Economo's disease), drug induced (e.g., from neuroleptics), toxin induced (e.g., MPTP), accompanying hereditary diseases (e.g., Huntington's disease), accompanying other conditions (e.g., Alzheimer's disease), and in the elderly (some parkinsonian features such as stooped posture, shuffling gait, and postural instability). Although these different types of parkinsonism share some common traits, there are important differences. The present dissertation will focus on only the idiopathic form of Parkinson's disease.

Symptoms

The cardinal motor signs of PD are tremor, rigidity, akinesia /bradykinesia, and postural instability (Bakheit, 1995; Stacy & Jankovic, 1992; Wooten, 1984). Hoehn and Yahr (1967) examined 183 cases of idiopathic PD and found that in 70.5% tremor was the initial symptom; 10% were free from tremor and 10% were free from rigidity. The tremor of PD is most prominent at rest and often involves a "pill-rolling" movement of the thumb and forefinger, with a frequency of about 3 to 6 Hz (Hopfensperger & Koller, 1991). Tremor may also be present in the limbs, jaw, and tongue, but often stops when a person makes a voluntary movement (Lezak, 1995). Rigidity of the limbs and trunk is common and may take on a consistent feel (*leadpipe rigidity*) or an inconsistent, jerky feel (*cogwheel rigidity*) (Knight, 1992). Muscular rigidity produces the characteristic stooped posture and expressionless (masklike) face (Knight, 1992). Bradykinesia is characterised by

slowness and poverty of voluntary movement and akinesia involves a difficulty initiating movement (Lezak, 1995). Postural instability manifests as poor balance and a slowed shuffling gait (Wooten, 1984). A number of nonmotor symptoms also occur in PD. Koller (1992a) suggested that such problems include subjective sensory and olfactory dysfunction, depression, dementia, and autonomic nervous system disturbances (e.g., constipation, sexual dysfunction, dermatitis, sweating, and urinary problems).

Symptom onset is insidious; motor signs are initially irregular and progression can be slow (Koller, 1992b). Generally, resting tremor and rigidity and/or bradykinesia present early, while mid-line features, disturbances of gait, speech, phonation, and righting reflexes develop later (Bakheit, 1995). However, the clinical course is highly variable and there is substantial heterogeneity of the clinical features (Koller, 1992a). For example, Koller (1992a) listed age at onset, presence of motor manifestations, presence of nonmotor manifestations, clinical course, responsiveness to drug therapy, and adverse reactions to drug therapy as features contributing to the heterogeneity of PD.

Disease stage or severity is commonly rated by means of Hoehn and Yahr's (1967) 5-stage scale (H&Y), or, increasingly, by the Unified Parkinson's Disease Rating Scale (van Hilten, van der Zwan, Zwinderman, & Roos, 1994). The latter is more comprehensive, consisting of both a motor exam scored by a physician, and a self-report disability scale.

Epidemiology

The symptoms of PD usually appear in the sixth decade – between the ages of 51 and 60 (Schwab, 1960); however, about 5% of patients present before age 40 (Koller, 1993). Marttila and Rinne (1981), summarising the results of 11 studies that examined prevalence and incidence rates of PD, found that among white races there is an average of 100 affected per 100,000 population. The prevalence

rate is approximately equal in males and females. After the age of 50 the prevalence of PD rapidly increases up to the age group of 70 to 79 years. Prevalence in this older group ranges from 300 to 1800 per 100,000 population. Similarly, incidence rates increase rapidly after the age of 40 years. The highest figures were found in the 70 to 79 year old age group with a wide range of estimates, from 53 to 229 per 100,000 population.

The relative risk of mortality in PD is three times that of the general population and increases markedly after age 65 (Hoehn & Yahr, 1967; Kalat, 1998). The death rate from causes related to PD is also greater in males than in females (Kurtzke & Murphy, 1990). Finally, the rate of mortality is less in those patients who develop tremor as their initial symptom as opposed to rigidity or bradykinesia.

Pathophysiology

Recent evidence suggests that PD begins some 5 years before the cardinal symptoms first appear (Markham & Diamond, 1993), after the loss of 70-80% of the dopaminergic neurons in the substantia nigra pars compacta (Brown & Marsden, 1990; Kalat, 1998) (see Figure 1). This loss of neurons causes a disruption of the dopaminergic nigrostriatal pathway. Symptoms then develop after approximately a 85% reduction of dopamine levels in the putamen (the motor part of the striatum), and around a 75% reduction of dopamine in the caudate nucleus (one of the "cognitive" areas of the striatum) (Testa, Brumback, Baik, Leech, & Cannon, 1998). This in turn disrupts the output from the globus pallidus to the thalamus, which leads to decreased excitation from the thalamus to the cerebral cortex (Kalat, 1998). Markham and Diamond (1993) also described other changes which include: (a) the presence of rounded eosinophilic intracytoplasmatic inclusions, called Lewy bodies; (b) the loss of the melanin-containing neurons in the locus coeruleus and the dorsal motor nucleus of the vagus nerve; and, (c) the loss of neurons in the ventral tegmental area of the

midbrain, the raphe nuclei of the brainstem, the substantia innominata of Meynert, and the hypothalamus.

It is less clear which changes to the CNS are responsible for each of the motor signs of PD. While degeneration of the nigrostriatal dopamine neurons is thought to underlie symptoms such as rigidity and akinesia, other pathways may also be involved in the pathophysiology of tremor (Rinne, 1991; but see DeLong & Wichman, 1993; Graybiel, 1993; Latash & Anson, 1996). In contrast, reduced functioning of the *complex loop* between the caudate nucleus and prefrontal cortex, and disruption of the mesocortical system may contribute to cognitive dysfunction (Brown & Marsden, 1990; Sagar, Atchison, Doherty, Ball, & Cooper, 1995). The complex loop is thought to have a number of functional subdivisions each projecting to a different cortical site (i.e., dorsolateral prefrontal, lateral orbitofrontal, and anterior cingulate circuits) (Brown & Marsden, 1990). The mesocortical system has projections to certain cortical areas, the amygdala, septal nuclei, and olfactory tubercle (De Keyser, Herregodts, & Ebinger, 1990). In addition, disruption of nondopaminergic pathways, such as the ascending noradrenergic and cholinergic projections to the frontal cortex, may also be related to cognitive deficits (Growdon, Corkin, & Rosen, 1990; Perry et al., 1985; Pillon et al., 1989; Sagar et al., 1995). Specifically, cognitive deficits appear to be associated with cholinergic neuronal loss in the nucleus basalis of Meynert and medial septal area of the basal forebrain (Testa et al., 1998). Dubois et al. (1987) demonstrated that a subthreshold dose of scopolamine produced specific memory deficits in nondemented PD participants. They suggested that this indicates problems with the cholinergic pathway. Finally, a number of studies have revealed nondopaminergic abnormalities involving serotonergic neurons in the raphe nucleus and GABAergic biochemical pathways, the consequence of which remains to be clarified (Growdon et al., 1990; Mahurin, Feher, Nance, Levy, & Pirozzolo, 1993).

Aetiology

The aetiology of PD is unknown, but there are four prominent hypotheses: the accelerated ageing theory, the toxin theory, the genetic theory, and the oxidative mechanism theory (Jankovic, 1992). According to the accelerated ageing theory, PD is caused by an abnormal acceleration of the normal age-related attrition of neurons (Jankovic, 1992). This acceleration could be due to genetic or environmental factors, or a combination of both. The toxin theory holds that dopaminergic neurons are selectively destroyed by an extrinsic or intrinsic toxin (Jankovic, 1992). Support comes from epidemiological studies that have shown that exposure to rural living, well water, pesticides, herbicides, and wood pulp mills are associated with an increased risk of PD (Olanow & Koller, 1998). According to the genetic theory, there could be a genetic deficiency resulting in an abnormally small number of substantia nigra neurons, and the mechanisms responsible for DNA repair may also be defective (Jankovic, 1992). Finally, according to the oxidative mechanism theory, abnormally high levels of oxygen-free radicals are produced during the oxidation of dopamine, as a result of increased iron, and due to the loss of normal protective anti-oxidant mechanisms (Jankovic, 1992). These free radicals can then potentially disrupt normal structures in the brain (for a recent review see Youdim & Riederer, 1997). Whether idiopathic PD results from an endogenous defect within the nervous system or has an environmental cause is open to speculation, but it is likely that there is a multifactorial aetiology (Markham & Diamond, 1993; Marsden, 1990; Jankovic, 1992).

Diagnosis

Currently PD must be diagnosed using clinical criteria because a biologic marker has not yet been found (Koller, 1992b). Idiopathic PD accounts for some 60 to 80% of parkinsonism cases, but difficulty of diagnosis results in 20 to 30% of patients being misdiagnosed (Larsen, Dupoint, & Tandberg, 1994). Diagnosis is

complicated by the lack of a universal definition of PD. Typically, researchers have diagnosed PD based on the presence of at least two of the motor signs (tremor, bradykinesia, and rigidity), and the exclusion of other forms of parkinsonism (Koller, 1992a). To improve matters, several investigators (e.g., Calne, Snow, & Lee, 1992; Larsen et al., 1994) have proposed clinical diagnostic criteria that can be used to classify patients into one of three categories: clinically definite, clinically probable, or clinically possible idiopathic PD. Larsen et al. proposed that to be classified as having clinically definite PD a patient must fulfil all of the following criteria:

1. Presence of resting tremor and at least two of the following signs: (a) akinesia/bradykinesia, (b) rigidity, and (c) postural abnormality.

2. Unilateral onset of signs and asymmetrical development of the disease.

3. Good to excellent response to dopamine agonism.

4. At the onset of the disease, there should be an absence of significant changes on computerised axial tomography or magnetic resonance imaging (MRI) other than mild diffuse cortical atrophy or mild hyperintense periventricular foci on MRI. There should also be an absence of clinical exclusion criteria like dementia, pyramidal and cerebellar signs, and autonomic failure, which may indicate other neurodegenerative disorders. There must also be an absence of environmental factors, like drugs and toxic substances, and the absence of a history of encephalitis that may cause a symptomatic parkinsonism.

The validity of the suggested diagnostic system has yet to be established, and future evidence will mostly likely result in revised criteria (Larsen et al., 1994). Moreover, the availability and cost of neuroimaging facilities will limit the use of this aspect of the diagnosis, although this will change as accessibility increases. Returning to Calne et al.'s (1992) criteria for clinically definite PD, it can be noted that they do not include evidence from neuroimaging or pharmacological

response, although the authors suggest that failure to respond to levodopa may be grounds for exclusion. For a patient to be classified with clinically definite idiopathic PD, they suggest that a combination of any three of the following features is sufficient: resting tremor, rigidity, bradykinesia, or impairment of postural reflexes.

Treatment

The advent of levodopa therapy remains the greatest advance in the treatment of PD (Markham & Diamond, 1993), although new drugs and surgical procedures, such as deep-brain stimulation and neural transplantation, show promise (Ali & Morley, 1999). Nonpharmacological treatment (e.g., education, exercise, and nutrition) is also important (Olanow & Koller, 1998). Levodopa, used in combination with a peripheral decarboxylase inhibitor, is effective in treating many of the motor signs of PD, particularly the most disabling symptoms of akinesia and bradykinesia (Marsden, 1990). Levodopa works by restoring dopamine levels in the striatum (i.e., putamen and caudate nucleus). After 3 to 5 years of levodopa therapy, approximately 50% of patients experience negative side-effects (Poewa, 1994), including drug-induced abnormal involuntary movements (e.g., chorea, dystonia), fluctuations in response (e.g., "end-of-dose" akinesia), and neuropsychiatric complications (Olanow & Koller, 1998; Poewa, 1994; Sage & Mark, 1994). To counter these fluctuations, sustained release levodopa preparations and/or different combinations of drug therapy can be used (Koller, Silver, & Lieberman, 1994). Other drug therapy involves the use of anticholinergics, dopamine agonists, catechol-o-methyltransferase (COMT) inhibitors, and monoamine oxidase (MAO)-B inhibitors, such as selegiline hydrochloride (deprenyl).

Anticholinergic drugs work by restoring the balance between acetylcholine and the reduced supply of dopamine, helping to control early rigidity and tremor, but not akinesia or impaired postural reflexes (Koller et al., 1994). However, the use

of anticholinergics is limited due to common adverse side-effects, such as hallucinations and memory impairment (Olanow & Koller, 1998). Dopamine agonists (e.g., bromocriptine) were developed to mimic the effects of dopamine by directly stimulating the normal striatal dopamine receptors (Ali & Morley, 1999). COMT inhibitors (e.g., tolcapone) were designed to increase the amount of levodopa available to cross the blood-brain barrier by limiting the breakdown of levodopa in the stomach and liver (Ali & Morley, 1999). Other drugs (e.g., selegiline) inhibit MAO-B, thus prolonging the duration of the action of dopamine (Marsden, 1990). Several lines of evidence suggest that selegiline may also have neuroprotective effects because it is able to block the development of oxygen-free radicals generated by dopamine metabolism (Olanow & Koller, 1998). Similarly, recent research has shown that dopamine agonists may have antioxidant properties (Olanow & Koller, 1998). In contrast, levodopa may accelerate neuronal degeneration during its metabolism to dopamine (Olanow et al., 1995).

Nevertheless, levodopa therapy remains the mainstay of treatment for PD, although other treatment options are continually being developed that may be used alone or in combination with levodopa. The effectiveness of drug therapy depends on many factors and must be tailored to suit each individual. An algorithm for the management of PD has been developed (Olanow & Koller, 1998; Koller et al., 1994) and demonstrates the complexity of treatment options. Current drug therapy concentrates on correcting the chemical imbalance related to the motor symptoms of PD and is not specifically designed to alleviate cognitive impairment (Mayeux, 1994). Nonetheless, some 30% of PD patients show a global impairment in cognition usually described as dementia (Olanow & Koller, 1998). Moreover, subtle cognitive deficits are frequently reported in PD patients who otherwise show no signs of generalised dementia.

Dementia

Recent evidence suggests that PD patients have a 4 to 5 times higher probability of developing dementia than would be expected for healthy people of the same age and socioeconomic status (Mayeux, 1994). However, the concept of dementia in PD is currently under debate with little agreement about its nature, severity, or prevalence (Huber & Bornstein, 1992). Much of the debate revolves around whether the dementia of PD should be classified as one of the so-called *subcortical dementias*, or attributed to cortical pathology. It remains to be seen whether the dementia is specific to PD or due to a different disorder, such as concomitant Alzheimer's disease (AD), diffuse Lewy body disease, or frontotemporal dementia (Olanow & Koller, 1998). Another important issue is whether selective cognitive deficits seen in PD are distinct from global dementia or whether they occur on the same continuum (Growdon et al., 1990).

The dementia debate is important but difficult to resolve because of a lack of comparability across studies. For example, the criteria used to define dementia have varied between studies, and samples have often contained mixed etiologies of parkinsonism (Levin, Llabre, & Weiner, 1989). Furthermore, many investigators have collapsed their data across possible moderator variables (e.g., disease stage, disease duration, age at disease onset, and medication), thus obscuring cognitive changes. With these caveats in mind, the next section will provide a general overview of the literature on specific cognitive deficits in PD participants without dementia. The impact of moderator variables will be covered in detail in a later section on Recognition Memory Deficits.

Cognitive Deficits

Within the experimental literature, cognitive deficits in PD are frequently described in terms of *executive function*, *bradyphrenia*, *visuospatial function*,

attention, and *memory*. In addition, behavioural similarities between PD patients and patients with frontal lobe (FL) dysfunction (Farina et al., 1994) have led to the use of the term *frontal deficit* when describing impaired performance on a number of tasks thought to be under the control of the frontal lobes (i.e., measures of executive function and bradyphrenia). Nevertheless, it is important to note that differences do exist between PD and FL patients (Faglioni, Botti, Scarpa, Ferrari, & Saetti, 1997; Owen, Sahakian, & Robbins, 1998). Furthermore, verbal abilities remain intact in nondemented participants (Levin et al., 1989), as does thinking and reasoning (Lezak, 1995), while depression (although common) appears to play little part in the observed deficits (Bieliauskas, Klawans, & Glantz, 1986; Taylor, Saint-Cyr, & Lang, 1988).

It is unclear what particular neural pathways are responsible for the cognitive deficits seen in PD, but disruption of the complex loops, the ascending dopaminergic mesocortical pathway, and/or nondopaminergic pathways is suspected. At this point it will be useful to look at the theoretical models used to explain cognitive impairment, before going on to describe each of the major deficits observed in PD.

Theoretical Models

Brown & Marsden's (1990) *theory of processing resources* holds that cognitive impairments will be seen when task demands exceed a participant's central processing resources. The theory is based on the attentional model of Norman and Shallice (1980, cited in Brown & Marsden, 1988a). Norman and Shallice proposed that a *supervisory attentional system* (SAS) with limited capacity is "called upon in situations that involve planning and decision making, in novel or poorly learned tasks, or in situations where some habitual response has to be overcome as may occur in switching set" (Brown & Marsden, 1990, p. 26). A disruption of the processing resources within the SAS (through frontal lobe dysfunction) is thought to underlie many of the cognitive impairments seen in PD.

For example, Brown and Marsden (1990) used their theory to explain a dissociation in performance that they found between PD and control participants on a modified Stroop task (e.g., Brown & Marsden, 1988a). Specifically, the PD participants were impaired when they had to rely on internal cues to perform the task, but were unimpaired when external cues were available. Brown and Marsden (1990) suggested that the internally cued task placed greater demands on the SAS than the externally cued task, and these demands exceeded the resources available to the PD participants.

Filoteo et al. (1997) argued against a single unifying theory of cognitive dysfunction in PD. They pointed to several studies (e.g., Downes, Sharp, Costall, Sagar, & Howe, 1993) that failed to find support for Brown and Marsden's (1990) theory of processing resources. Filoteo et al. suggested that impaired performance on some tasks (e.g., verbal fluency, selective and shifting attention) may be due to a deficit in inhibitory processes and not in depleted central processing resources.

Other theories used to explain the pattern of cognitive deficits seen in PD have made a distinction between effort-demanding (internal or active) tasks and automatic (passive) tasks (Taylor, Saint-Cyr, & Lang, 1986; Weingartner, Burns, Diebel, & LeWitt, 1984). For instance, Weingartner et al. proposed a theory where task demands are placed on a continuum from automatic to effortful. The theory proposes that deficits in the early stage of PD are limited to effort-demanding tasks. The results of their study provide some support; PD participants performed normally on tasks requiring automatic processing (e.g., frequency monitoring), but abnormally on effort-demanding tasks (e.g., serial list learning). Further support was provided by Appollonio et al. (1994) who found that PD deficits could be categorised according to the level of effort required by the task. Taylor and colleagues have frequently made a distinction between tasks requiring internal resources for successful completion and those that use external cues, only the former are impaired in PD (Taylor et al., 1986; Taylor & Saint-Cyr, 1992).

Finally, Bondi, Kaszniak, Bayles, and Vance (1993) found that PD participants were impaired on tasks requiring more effort, self-directed planning, efficient organisation, and sustained attention. Furthermore, their analysis revealed that poor performance on "frontal" tasks accounted for deficits in memory, and in both visuoperceptual and visuoconstructive skills.

Executive Functioning

The cognitive processes thought to be under executive control include anticipation, planning, initiation, and monitoring of goal-directed behaviours, as well as the ability to use feedback and modify behaviour accordingly (Levin, Tomer, & Rey, 1992b). In nondemented PD participants there appears to be a consistent impairment in performance on the Wisconsin Card Sorting Test (WCST), which is commonly attributed to a deficit in *set-shifting* ability (e.g., Bondi et al., 1993; Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991; Lees & Smith, 1983; Levin et al., 1989; Taylor et al., 1986). However, Cooper et al. noted that a deficit in set-formation may underlie this impairment. Moreover, Owen et al. (1998) noted that performance on the WCST involves many different aspects of cognitive function, which makes it difficult to determine what function is impaired. Deficits have been noted on the Color and Color-Word forms of the Stroop Test and on the Trail Making Test, also attributable to a set-shifting deficit (Hietanen & Teravainen, 1986; but see Brown & Marsden, 1988a). Furthermore, deficits on measures of categorisation (e.g., Cooper et al., 1991; Farina et al., 1994), have been attributed to impaired cognitive sequencing or working memory (Cooper et al., 1991). There is also some evidence of impaired temporal structuring as demonstrated by poor performance on the Spatial Delayed Response Test (Taylor et al., 1986) and the Verbal Temporal Ordering Test (Bondi et al., 1993). Finally, deficits have been found in semantic and letter fluency (e.g., Azuma et al., 1997; Bondi et al., 1993; but see Hanley, Dewick, Davies, Playfer, & Turnball, 1990).

Bradyphrenia

Bradyphrenia refers to the slowing of thought (Brown & Marsden, 1990), and is the cognitive analogue of bradykinesia (Levin et al., 1992b). Within the literature there appears to be little agreement over whether bradyphrenia is real or not (e.g., Brown & Marsden, 1990; Lezak, 1995). Moreover, bradyphrenia is difficult to assess because processing speed and motor demands are confounded on many tasks (Levin et al., 1992b). Brown and Marsden have rejected the common use of the term, arguing instead that any observed slowing is probably task-specific rather than a generalised slowing of cognition.

Visuospatial Functioning

Visuospatial functioning involves a number of processes and a precise definition is rarely given in the PD literature (Waterfall & Crowe, 1995). More generally, the term is taken to involve the ability to appreciate the relative position of stimuli and objects in space, integrate those objects into a coherent spatial framework, and perform mental operations involving spatial concepts (Brown & Marsden, 1988b; Dubois et al., 1991). Visuospatial deficits have been frequently reported in PD, although early research often used tasks that confounded visuospatial and motor performance (Growdon & Corkin, 1986). Nevertheless, researchers using visuospatial tasks that minimised motor responses (i.e., *visuo-perceptual* tasks) have still reported deficits in PD (Dubois et al., 1991; Growdon & Corkin, 1986; Mohr, Mendis, & Grimes, 1995). However, other investigators have argued against a generalised visuospatial deficit (Brown & Marsden, 1990; Knight, 1992), pointing instead to deficits in short-term memory and mental flexibility (e.g., Taylor et al., 1988) caused primarily by frontal lobe dysfunction (Bondi et al., 1993; Levin et al., 1992b; Ogden, Growdon, & Corkin, 1990), or to an impairment in maintaining visual attention. The latter is believed to be mediated by the basal ganglia (e.g., Filoteo et al., 1994).

Waterfall and Crowe (1995) attempted to integrate the findings from 70 PD studies by categorising tasks that involve visuospatial abilities into 13 components. They used meta-analysis to provide a quantitative estimate of the magnitude of visuospatial deficits. Unfortunately, they made no distinction between studies that excluded participants with dementia and those that did not. Nevertheless, their results do not rule out the possibility that visuospatial deficits are associated with frontal lobe dysfunction.

More recently, Cronin-Colomb and Braun (1997) argued that several lines of evidence now suggest that neither frontal lobe dysfunction nor attentional deficits underlie the wide range of visuospatial deficits in PD. They found visuospatial deficits in nondemented PD participants that were not associated with executive function or verbal memory. This finding supports an earlier study (Mohr et al., 1990) that demonstrated that high functioning PD participants, with no evidence of executive dysfunction, had impaired performance on memory independent visuospatial tasks.

Attention

A number of studies have shown that PD participants have deficits in visual, auditory, and tactile selective attention tasks, and this has been interpreted as an impairment in the maintenance of attention (e.g., Bradshaw et al., 1993; Filotea et al., 1994; Maddox, Filotea, Delis, & Salmon, 1996; Sharpe, 1992; Wright, Burns, Geffen, & Geffen, 1990). More recently, Filotea et al. (1997) suggested that PD patients do not have a general deficit in the maintenance of attention, but rather have impaired inhibitory processes. That is, they suffer from a "specific impairment in the ability to inhibit the movement of their attention from one cognitive process to another," and this underlies an observed deficit in selective attention (Filotea et al., 1997, p. 339). Several lines of evidence suggest the basal ganglia may be directly involved in inhibitory attentional functions (Filotea et al., 1997); thus, damage to these brain structures in PD would produce a selective

attention deficit. However, it is also possible that more widespread damage involving the frontal regions, the superior colliculus, the thalamic structures, as well as the basal ganglia is implicated (Maddox et al., 1996).

Memory Functioning

It has long been recognised that memory is not a unitary process. As well as memory being divided into short- and long-term, there now is a bewildering array of other proposed memory types (Schacter & Tulving, 1994). Many of these, such as declarative and procedural memory, have been investigated in PD patients (e.g., Harrington, Haaland, Yeo, & Marder, 1990; Saint-Cyr, Taylor, & Lang, 1988). Other researchers have made the distinction between explicit and implicit memory tests (e.g., Bondi & Kaszniak, 1991; Ferraro, Balota, & Connor, 1993; Heindel, Salmon, Shults, Walicke, & Butters, 1989; Huberman, Moscovitch, & Freedman, 1994). Within the framework of declarative memory, distinctions have been made between short-term (working) memory and long-term memory (e.g., Sullivan, Sagar, Cooper, & Jordan, 1993), and between verbal and nonverbal memory (e.g., Beatty, 1992).

With respect to working memory, the available evidence suggests that PD may disrupt verbal and nonverbal recall, independent of antiparkinsonian medication (e.g., Cooper et al., 1991; Cooper, Sagar, & Sullivan, 1993; Levin et al., 1989; Sahakian et al., 1988; Sullivan et al., 1993; Taylor, Saint-Cyr, & Lang, 1990). It has also been suggested that performance may return to normal after a delay period (Taylor et al., 1986). To explain this pattern of performance some authors have suggested that short-term storage may be intact, but slowed information processing (i.e., bradyphrenia) produces the deficits in immediate recall (Sagar, Sullivan, Gabriel, Corkin, & Growdon, 1988; Sullivan & Sagar, 1989; Taylor et al., 1986).

Procedural memory tasks generally involve testing the ability to acquire a motor or cognitive skill. Although few studies have specifically excluded demented PD patients, those that do have shown deficits in procedural memory tasks thought to be subserved by the frontal lobes (Bondi & Kaszniak, 1991). For example, impaired performance was found using the Tower of Toronto task (Saint-Cyr et al., 1988), the skill learning component of the fragmented pictures test (Bondi & Kaszniak, 1991), and the serial reaction time task (Ferraro et al., 1993). In contrast, normal performance has been found using a mirror-reading test (Bondi & Kaszniak, 1991; Harrington et al., 1990; Huberman et al., 1994) and a pursuit-rotor tracking task (Bondi & Kaszniak, 1991; Heindel et al., 1989). In addition, normal performance on the perceptual memory component of the fragmented pictures test (Appollonio et al., 1994; Bondi & Kaszniak, 1991) and on a word stem-completion repetition priming task (Bondi & Kaszniak, 1991; Huberman et al., 1994), indicates that the perceptual representation system is not damaged in PD.

Turning to long-term memory, the distinction between internal or effort-demanding tasks and passive or automatic tasks has received considerable attention from memory researchers. For example, after finding a dissociation between effort-demanding and automatic memory tasks, Appollonio et al. (1994) pointed to deficits involving "the conscious, effortful strategic aspects of searching long-term memory" (p. 366). Other investigators (e.g., Bondi & Kaszniak, 1991; Mohr et al., 1990) have noted that their results are consistent with Weingartner et al.'s (1984) theory of automatic versus effortful processing differences in PD. In particular, Mohr et al. showed that in PD participants with normal executive functioning, deficits were only seen on the more effort-demanding memory tasks. Memory task demands were also specifically manipulated in a study of incidental and intentional recall carried out by Cooper and Sagar (1993b). These authors found that PD participants were impaired only on the more effort-demanding tasks, and suggested that their results support Brown and Marsden's (1990) processing resources theory.

In addition, recall deficits are more common in tasks requiring the use of internal strategies for semantic clustering (e.g., Buytenhuijs et al., 1994; Taylor et al., 1990) or in tasks that involve learning nonsemantically related items (e.g., Taylor et al., 1986). Recently, Pillon et al. (1996) found that PD may cause a primary deficit in memory for spatial locations. This visuospatial learning deficit was seen when memory was tested using free recall, cued recall, and recognition, but not when a comparable verbal memory task was used. The authors suggested that task complexity may contribute to this dissociation. That is, the visuospatial memory task was more complex because attention was divided between the picture and its location, both of which had to be recalled.

Most reviews of the PD literature note a dissociation between free recall and recognition memory. That is, PD participants generally have impaired recall while recognition memory remains relatively normal (e.g., Beatty, 1992; Brown & Marsden, 1988b, 1990; Cummings & Benson, 1988; Dubois et al., 1991; Growdon & Corkin, 1986; Karayanidis, 1989; Knight, 1992; Knight, Godfrey, & Shelton, 1988; Mahurin et al., 1993; Sagar & Sullivan, 1988; Saint-Cyr & Taylor, 1993; Taylor et al., 1988). Most of the evidence for this dissociation comes from studies that have not examined recall and recognition memory in the same participants, although those that have have also found deficits in recall, but not recognition (e.g., Breen, 1993). This dissociation in memory functioning has been explained in terms of the effort-demanding/automatic distinction (e.g., Knight et al., 1988; Taylor et al., 1988). In addition, it has been used as evidence that the deficit lies with retrieval strategies rather than encoding (e.g., Beatty, 1992; Brown & Marsden, 1988b; Dubois et al., 1991; Mahurin et al., 1993). However, recent evidence calls this into question; different aspects of both retrieval and encoding may be at risk (Faglioni et al., 1997). In addition, Faglioni et al. highlighted the fact that the dissociation between recall and recognition memory seen in PD is not seen in FL patients. This finding raises the possibility that frontal dysfunction alone does not fully account for memory deficits. Figlioni et al. instead point to a dysfunction in the loop between the medial frontal cortex and the hippocampal system.

Faglioni et al. (1997) have also challenged the view that PD participants should necessarily perform poorly on effort-demanding tasks but normally on automatic memory tasks. They used the selective reminding procedure (Buschke & Fuld, 1974) and stochastic modelling (Brainerd, Howe, & Desrochers, 1982) to demonstrate that PD participants may have specific deficits restricted to those abilities that operate at an automatic level.

The conclusion that recognition memory is relatively normal in PD is not universal, and may depend on the modality of the material to be remembered (Sagar & Sullivan, 1988). For example, a number of researchers have noted a face recognition memory deficit in contrast to normal verbal recognition (e.g., Blonder, Gur, Gur, Saykin, & Hurtig, 1989; Dewick, Hanley, Davies, Playfer, & Turnbull, 1991). This is of interest given that face recognition could be distinct from other forms of recognition as it may involve at least some face-specific processing mechanisms (Bruce & Young, 1986; Geschwind, 1979; but see Ellis, 1981). It has also been noted that recognition memory for nonverbal material involving patterns or abstract designs may be at greater risk than verbal material (Beatty, 1992). For example, Sahakian et al. (1988) found PD participants were impaired on tests of spatial and pattern recognition memory.

In summary, nondemented PD patients may have specific cognitive deficits involving executive functions, visuospatial abilities, attention, and memory. These deficits share some similarities with those seen in FL patients, but differences do exist. This indicates that a disruption to the frontal lobe does not fully account for all of the deficits observed in PD participants. Moreover, whether these deficits can be explained by a single unifying theory or represent unique abnormalities is not known at this time. In this respect, memory dysfunction is particularly difficult to assess. The issue of whether or not recognition memory is impaired in PD is paramount to the present dissertation and so will be explored in greater depth in the next section.

Recognition Memory Deficits

The common assertion made in the literature is that PD participants have impaired recall but relatively intact recognition memory. Three studies (Flowers, Pearce, & Pearce, 1984; Lees & Smith, 1983; Taylor et al., 1986) appear to have been instrumental in establishing this view; here, each will be reviewed in some detail.

Lees and Smith (1983) used a sample of 30 *de novo* participants with idiopathic PD matched to an equal number of healthy controls according to gender and age. In the PD group there were 19 males with a mean age of 59.4[1] years (*SD* = 6.8) and 11 females with a mean age of 58.0 years (*SD* = 7.9). Mean disease duration was 2.4 years (range 6 months to 5 years). Their H&Y stage ranged from I to III (I: *n* = 12; II: *n* = 13; III: *n* = 5), and no participant presented with depression or with prior exposure to antiparkinsonian medications.

Lees and Smith (1983) used a two-alternative, forced-choice (2AFC) recognition memory paradigm. Two separate tasks were used to test recognition memory for words and for unknown faces. Each task used 50 target items each presented for 3 s with recognition being tested without delay. The authors argued that the results of their nonparametric analysis indicated no specific memory deficits in the PD participants. Since this conclusion appeared to be based largely on the outcome of statistical significance tests, a reanalysis[2] was made of Lees and Smith's data in terms of statistical power and observed effect size (ES) (see chapter 2 for more information about statistical power and ES).

In the Lees and Smith (1983) study, comparisons between PD participants and controls were made with two-tailed Mann-Whitney U tests that did not have satisfactory statistical power. Given an alpha level of .05 and 30 participants per

[1] Where an author only reported a statistic to one decimal place, that convention has been followed here.
[2] When ever reference is made to a reanalysis of another author's data, the reanalysis was done by the present author, unless otherwise stated.

group, a conservative estimate of statistical power indicates that the test had no more than 12% power to detect a small effect (Cohen, 1988). Thus, there was very little chance of rejecting the null hypothesis (Type II error probability = 100 - power = 88%) even if it was in fact false. Power to detect medium effects (48%) was also inadequate, but rose to 86% for large effects. The observed ESs calculated from the reported data were very small both for word recognition memory (d = .06, -.46–.57)[3] and for face recognition memory (d = .17, -.35–.68). Thus, apart from the very low level of statistical power, there was little evidence that this sample of PD participants had impaired recognition memory. However, the participants in this study were by no means representative of the entire population of PD suffers. That is, they were unmedicated and early in the course of the disease; so no generalisation beyond these types of patients can be made.

Flowers et al. (1984) used a more representative sample of PD participants. The sample size varied between tasks, but ranged from 34 to 50. The mean age ranged from 61.1 to 63.7 years. Little information was given on the severity of PD in this sample, except that participants ranged from 3 to 24 on the Webster scale of symptom severity and clinical disability. Disease duration ranged from less than a year to 33 years. All participants were stabilised on some form of antiparkinsonian medication. Participants were not screened for possible dementia. A similar number of healthy participants acted as controls matched approximately by gender, age, occupation, and background.

Flowers et al. (1984) assessed both immediate- and delayed-recognition memory with the aid of verbal and nonverbal tasks. The first verbal recognition memory task (pictures of common objects) used a 5AFC paradigm, total number correct being the dependent variable. Half of the 32 items were tested after a 1 min delay (immediate test) and the other half were tested after a 45 min delay (delayed test). The other verbal recognition memory task (words and numbers) used a 2AFC paradigm, total number correct again being the dependent variable. Half of the 16 items were tested after a 1 min delay (immediate test) and the other half were

[3] Here and in all future presentations of the d statistic, the 95% confidence interval for d will be given following the statistic (see Appendix D for formulae).

tested after a 15 min delay (delayed test). Two nonverbal recognition memory tasks (black and white histogram shapes, and coloured abstract pictures) were conducted, each following a yes–no paradigm. Recognisability was assessed using the signal detection theory (SDT) measure, *d'*. The first task consisted of 24 items, half of which were tested after a 1 min delay (immediate test) and the other half after a 45 min delay (delayed test). The other nonverbal test consisted of 20 items, half of which were tested after a 1 min delay (immediate test) and the other half after a 30 min delay (delayed test).

Flowers et al. (1984) claim that recognition memory was poorer for the PD participants on all tasks relative to the controls, but that the differences were neither "large" nor "reliable". Furthermore, they suggest that their results indicate that PD participants have normal registration and ability to retain information. Any memory problems must involve "retrieval or some higher level processing stage" (p. 1180). As with the study of Lees and Smith (1983), a reanalysis of some of Flowers et al.'s data was made in terms of statistical power and observed ESs.

Flowers et al. (1984) made comparisons between PD and control participants using two-tailed *t* tests. Given an alpha level of .05 and the number of participants used, these tests had no more than 17% power to detect small effects (Cohen, 1988). Power to detect a medium effect was also inadequate, ranging from 53 to 69%. When ESs were calculated from the reported data, it was found that the observed delayed recognition memory deficits ranged from small-to-medium in size (Cohen, 1988). However, there was no evidence of impaired immediate recognition memory except for the words and numbers task. It has been suggested that several of the tasks used by Flowers et al., may have produced performance close to floor and ceiling levels, thus reducing the observed ESs further (Sahakian et al., 1988). In any event, the important point to note is that their nonsignificant findings may well have been due to having too little statistical power to detect the effect.

Taylor et al. (1986) used a sample of 40 (15 female, 25 male) nondemented, idiopathic PD participants matched to an equal number of healthy controls according to gender, age, education, and verbal IQ. The PD participants had a mean age of 60.53 years (*SD* = 12.19) and a mean disease duration of 6.62 years (*SD* = 5.26). Their H&Y stages ranged from I to IV (I: *n* = 10; II: *n* = 13; III: *n* = 14; IV: *n* = 3), and 70% of the sample were taking antiparkinsonian medications.

Taylor et al. (1986) examined recognition memory with the Rey Auditory Verbal Learning Test (RAVLT). Participants were given a 15-word list presented as a set repeated over 5 trials in a spontaneous recall paradigm. Recognition memory was assessed (after recall) by presenting a paragraph in which all target words were embedded with an equal number of distractors (batch testing). Scores were taken as the difference between correctly identified words minus false positives. Nonparametric statistical tests were used because of frequent ceiling effects. The authors reported that there was no statistically significant difference between the PD and control groups. As before, a reanalysis of these data was made in terms of statistical power and observed ES.

As with the two studies described earlier, Taylor et al. (1986) had insufficient statistical power to detect small or medium effects. Given an alpha level of .05 and the number of participants used, the test of RAVLT recognition memory had no more than 22% and 72% power to detect small and medium effects, respectively (Cohen, 1988). Therefore, there was little chance of rejecting the null hypothesis for small ESs. The observed ES (*d* = .26, -.19–.71, estimated from the Z statistic) represents a possible, albeit small, recognition memory deficit.

Despite the widespread belief that recognition memory is relatively normal in PD patients, several studies using nondemented participants have reported some form of impaired recognition memory (e.g., Bondi et al., 1993; Cooper et al., 1993; O'Sullivan, 1998; Owen et al., 1993; Sahakian et al., 1988). Studies using PD participants unselected for dementia have also reported recognition deficits (Allain et al., 1995; El-Awar, Becker, Hammond, Nebes, & Boller, 1987; Massman,

Delis, Butters, Levin, & Salmon, 1990; Reid et al., 1989; Sagar, Cohen, Sullivan, Corkin, & Growdon, 1988; Sullivan & Sagar, 1989; Tweedy, Langer, & McDowell, 1982). Yet studies such as Flowers et al. (1984) and Taylor et al. (1986), appear to have been instrumental in establishing the conclusion that recognition memory is relatively unimpaired in PD (Whittington, Podd, & Kan, in press). However, this conclusion appears unwarranted given the limited statistical power of these studies, the small-to-medium observed ESs, and the number of subsequent studies detecting recognition memory deficits. Thus, recognition memory may be impaired in PD depending on circumstances. Cognitive deficits may vary as a function of many factors such as task difficulty, disease stage, age at onset of motor symptoms, depression, and medication (Beatty, 1992; Levin & Katzen, 1995; Levin, Tomer, & Rey, 1992a; Starkstein et al., 1989). Each of these potential moderators will be covered in greater detail below.

Task Difficulty

"Effortfulness" has sometimes been used to explain why some studies find recall deficits but relatively intact recognition memory in PD (e.g., Knight, 1992; Taylor et al., 1988). This reasoning rests on the assumption that recognition tests require less effort, less self-initiated activity, or less processing resources than recall (Craik, & McDowd, 1987; Hasher & Zacks, 1979). The net outcome is that participants perform better on recognition memory tests than recall tests. With respect to PD, deficits are only detected by the more sensitive free recall tasks rather than by tests of recognition. A similar issue has been discussed in the literature on age-related decline in memory (e.g., Craik & McDowd, 1987) and memory deficits in schizophrenia (e.g., Calev, 1984). To examine this issue, some researchers have used methodological procedures designed to equate recall and recognition performance. For instance, Parkin and Lawrence (1994) compared 22 (4 male, 18 female) neurologically normal older adults (mean age = 71.90 years) with 23 younger adults (mean age = 48.45 years). To equate recall performance with that of recognition, recall of 24 target words was compared with 24 matched

words from a yes-no recognition test that originally consisted of 40 target words and 38 distractors. Parkin and Lawrence reported a statistically significant difference between young and old participants with respect to both recall and recognition memory. A reanalysis of their results in terms of ES indicated large memory deficits in the older group relative to the younger group, both in recall (d = 2.15, 1.40-2.90) and in recognition (hit rate less false alarm rate: d = 2.00, 1.27-2.73). Thus, the usual finding of an age-related decline in recall, but not recognition memory was not supported when the two tasks were equated in terms of effort.

Craik and McDowd (1987) questioned whether scores on recall and recognition memory tests are really comparable. Brown (1976) discussed this issue at some length and concluded that unless strong assumptions are made, recall and recognition can only be compared under highly restrictive conditions. To overcome this problem, Craik and McDowd used statistical methods (analysis of covariance and regression), which allowed recognition and recall scores to be compared. However, any dissociation between recall and recognition will be confounded to the extent that factors other than effortfulness (e.g., procedural differences) vary between the two measures of memory. One approach to overcoming this problem is to study the affect of task difficulty by varying this factor within a single recognition task.

It has been suggested that if a recognition task was made more difficult (effortful) PD participants should begin to show a deficit relative to controls (Breen, 1993). The results of Weingartner et al. (1984) suggest that the relationship between effort and performance may be linear, but others have argued that performance decrements may only become apparent after a threshold is crossed (Brown & Marsden, 1988a). To my knowledge, no one has tested this theory directly in PD by specifically varying the difficulty level within one recognition memory task. Nonetheless, some research suggests that recognition task difficulty is an important factor.

Albert, Butters, and Levin (1979) used a recognition memory test with two levels of task difficulty to assess alcoholic patients with Korsakoff's disease. The sample consisted of 11 patients with a mean age of 59.5 years, matched to 15 healthy controls on age, education, IQ, and socioeconomic status. The recognition memory test followed a 3AFC format with 66 "easy" questions and 66 "hard" questions. The easy questions concerned "people and events of extended fame" (p. 213), whereas the hard questions involved "incidents and people whose prominence was transient" (p. 213); care was taken to avoid ceiling effects. The authors reported a statistically significant difference between the two groups on the hard questions only. A reanalysis of their results in terms of ES indicated that the Korsakoff patients had a relatively small deficit in recognition memory for the easy questions (d = .37, -.44–1.17), but were severely impaired on the hard items (d = 1.98, 1.00–2.96). In addition, Craik and Jennings (1992) noted that an age-related decline in recognition memory for pictures is small, unless the recognition test is made difficult by using very similar target and distractor items.

Dujardin, Bourriez, and Guieu (1995) studied the effects of ageing and task difficulty on verbal and nonverbal recognition memory. The sample consisted of 10 (5 males, 5 females) younger participants ranging in age from 18 to 21 years (M = 19.1, SD = 1.2), and 10 (4 males, 6 females) older participants ranging in age from 56 to 71 years (M = 62.8, SD = 4.3). The two groups had similar IQ, and no participant reported a history of neurological disorder or current use of medication. Both the verbal recognition task and the nonverbal recognition task used 12 target words or 12 target figures, respectively. All figures were drawn so as to be difficult to verbalise. A yes–no test was given in which the 12 target words or 12 target figures were mixed with 24 distractors. Two levels of task difficulty were created by using distractors that were either easy to distinguish from the targets (easy level) or hard to distinguish from the targets (hard level). Results were reported for hits and false alarms separately. The authors reported that there were no statistically significant interactions between age group and task difficulty. However, a reanalysis of their data in terms of ES was made by subtracting the mean false alarm rate from the mean hit rate and then dividing the

difference by the standard deviation of the hit rate. The results of the reanalysis indicated that the older participants suffered from verbal recognition memory deficits relative to the younger participants on the hard level of the task (d = 1.58, .52–2.64), but not on the easy level (d = -.20, .73– -1.13). With respect to the nonverbal task, the older group's performance was inferior to that of the controls on both levels of difficulty, although more so on the hard level (d = 2.27, 1.08–3.46) than the easy level (d = 1.72, .63–2.80).

Although no research has specifically examined the effect of task difficulty on PD patients using a recognition task, the results of two studies indicate that it may be important. Owen et al. (1993) examined recognition memory in 13 (8 male, 5 female) late-stage idiopathic PD participants and 42 healthy controls matched on age and premorbid IQ. All PD participants were medicated and had a mean age of 66.54 years (SD = 7.03) and a mean disease duration of 10.6 years (SD = 4.36). Their H&Y stage ranged from III to IV (III: n = 7; IV: n = 6). The authors used a 2AFC pattern recognition memory task with two sets of 12 stimuli each. While the study was not designed specifically to examine the effects of task difficulty, it was reported that controls found the first set of items less difficult than the second set. A reanalysis of these results in terms of ES indicated that the late-stage medicated PD participants showed some signs of impaired recognition, but only with regard to the more difficult Set 2 items (d = .37, -.17–.91).

Dewick et al. (1991) studied recognition memory in 19 (10 male, 9 female) PD participants aged between 59 and 84 years (M = 72.9, SD = 7.5). None were demented and all but one were taking antiparkinsonian medication. Twenty-one (9 male, 12 female) healthy elderly served as controls (matched on age, education, verbal ability, and premorbid IQ). A 2AFC task was used to assess recognition memory for 50 targets (common words). In addition, an identical procedure was used with 50 unfamiliar male faces. Dewick et al. reported that the PD patients had a statistically significant deficit in face recognition, but not in word recognition. A reanalysis of their results in terms of ES indicated that the PD participants had impaired memory performance in both word recognition and face

recognition, although the impairment was far greater for the more difficult task (d = .22, -.42–.85, and d = 1.87, 1.11–2.63, respectively). Interestingly, both groups found the face recognition task more difficult than the word recognition task. Other researchers have noted this distinction, although there is some debate over whether face and word recognition share the same mechanism (Glass, Holyoak, & Santa, 1979). While it is possible that the PD participants suffer from a more general difficulty in the structural coding of facial stimuli (Dewick et al., 1991), it is also possible that the difference in impairment may be related to task difficulty.

Taken together, these results indicate that the magnitude of the recognition deficit in PD may depend on the difficulty of the task. Nevertheless, a direct investigation of the effects of task difficulty on recognition memory is yet to be conducted.

Gender

There is some evidence of gender-related deviations in brain anatomy (Kalat, 1998; Markowitsch, 1998). For instance, Johnson, Pinkston, Bigler, and Blatter (1996) found that women generally have a larger corpus callosum relative to total brain size compared to men, and they suggest that there may be differences in brain organisation. In neurologically intact older adults, women tend to have superior verbal memory (Bleecker, Bolla-Wilson, Agnew, & Meyers, 1988; Herlitz, Nilsson, & Backman, 1997; Hultsch, Hertzog, & Dixon, 1990; Hultsch, Hertzog, Small, McDonald-Miszczak, & Dixon, 1992; Kramer, Delis, & Daniel, 1988; Zelinski, Gilewski, & Schaie, 1993; Zelinski & Stewart, 1998), whereas men generally have better spatial memory (Kramer et al., 1988; Rosselli & Ardila, 1991). Gender may have a differential affect on the memory abilities of neurologically impaired adults. For example, Henderson and Buckwalter (1994) reported the usual finding of superior verbal recall in female control participants, but the opposite result in patients with Alzheimer's disease.

More generally, research looking at gender differences in memory functioning has focused on recall tasks. For example, in a study of verbal learning in 136 (68 males, 68 females) neurologically intact adults, Kramer et al. (1988) demonstrated that women had better verbal memory skills. That is, greater immediate recall, delayed recall, and semantic clustering, but similar recognition memory to that of males. The superior performance of women could not be attributed to differences in verbal intelligence because the two groups were matched for age, education, and age-adjusted WAIS-R verbal subtest score. The authors suggested that women had better performance because they organised the material more effectively during encoding and this resulted in more efficient retrieval. In contrast, males used serial clustering as a recall strategy, which is less efficient, and they tended to recall items from the beginning and end of the list.

Kramer et al. (1988) found that recognition memory was similar for males and females arguing that efficient retrieval is not as important as it is for tests of recall. Nevertheless, a reanalysis of their results in terms of ES provided some evidence that the female group had superior recognition memory (d = .33, -.14–.80). Similar results were obtained by Bleecker et al. (1988) and Herlitz et al. (1997). Moreover, Herlitz et al. noted that the gender-related ES was similar in recognition and recall, and that gender differences were similar in magnitude for participants aged 35 to 80 years. In addition, although the female participants had better verbal ability, this did not fully account for the results, and neither did differences in interest or familiarity with the material.

With respect to PD, little information is available about gender differences in memory dysfunction, although Jacobs and Schofield (1998) claim that gender generally has not been associated with cognitive deficits. However, recently, Lyons, Hubble, Troster, Pahwa, and Koller (1998) examined cognitive and motor performance in 315 male and 315 female PD participants and found that the female group had better Mini-Mental State Examination (MMSE) scores. In terms of motor performance, the male group generally had more severe symptoms, whereas the female group experienced more levodopa-induced dyskinesia. It

appears that only one study of recognition memory (O'Sullivan, 1998) has examined gender effects in relation to PD.

O'Sullivan (1998) examined 22 (17 males, 5 females) nondemented, medicated PD participants ranging in age from 48 to 82 years ($M = 72.18$, $SD = 7.93$) matched to 22 controls on gender, age, education, and occupation. Recognition memory (discriminability) was assessed using the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987). O'Sullivan found that there was no statistically significant difference in performance between the male and female PD participants. However, a reanalysis of the results in terms of ES indicated that the female PD participants had somewhat better recognition memory than the male participants ($d = .43$, -.46–1.31). Moreover, this difference is likely to be attenuated by age because older participants had poorer performance than younger ($d = .57$, -.33–1.46), and the female participants were seven years older, on average, than the male participants. Unfortunately, no analysis was reported that examined the difference between PD and control participants as a function of gender.

In summary, gender differences exist with respect to memory functioning in normals and neurologically impaired patients. However, at this stage it is unclear whether such differences occur in PD.

Disease Stage

Although the relationship between motor disability and cognitive decline is not straightforward (Starkstein & Robinson, 1991), there is some suggestion that disease stage or severity is related to memory dysfunction (Huber, Freidenberg, Shuttleworth, Paulson, & Christy, 1989; Mortimer, Pirozzolo, Hansch, & Webster, 1982). Furthermore, participants in the early stages of PD may profit more from memory cues than those in the advanced stages (Scholz & Sastry, 1985). Few studies have examined this issue with respect to recognition memory, although

there is some evidence that participants in the advanced stages of PD show greater recognition deficits than those in the early stages (e.g., Lees & Smith, 1983; Owen et al., 1992). Sahakian et al. (1988) assessed recognition memory in two samples of nondemented idiopathic PD participants matched to an equal number of healthy controls according to age and verbal IQ. One PD group contained 13 (5 female, 8 male) *de novo* participants with a mean age of 61.3 years (*SD* = 6.3). Disease duration was 1.7 years (range 0.5–4.5) and H&Y stage ranged from I to III (I: n = 7; II: n = 4; III: n = 2). The other PD group consisted of 14 (6 female, 8 male) medicated participants with a mean age of 64.2 years (*SD* = 6.8). Disease duration ranged from 3 to 16 years (*M* = 7.6 years) and H&Y stage ranged from I to III (I: n = 2; II: n = 3; III: n = 9).

Recognition memory was examined by Sahakian et al. (1988) using a 2AFC pattern recognition memory task. The authors reported a statistically significant deficit in the medicated PD group only. A reanalysis of their data in terms of ES confirmed a small pattern recognition memory deficit in the *de novo* early-stage PD participants (*d* = .17, -.63–.98), but a larger impairment in the more advanced-stage medicated participants (*d* = 1.44, .58–2.30).

Owen et al. (1992) used a sample of 44 nondemented idiopathic PD participants matched to an equal number of healthy controls according to age and premorbid IQ. For the purposes of the study, the PD group was divided into three subgroups. One subgroup contained 15 *de novo* participants with a mean age of 55.73 years (*SD* = 11.23). Mean disease duration was 1.5 years (*SD* = 16.27) and H&Y stage ranged from I to III (I: n = 3; II: n = 10; III: n = 2). Another subgroup consisted of 15 early-stage medicated participants with a mean age of 58.86 years (*SD* = 10.96). Mean disease duration was 7.07 years (*SD* = 4.92) and H&Y stage ranged from I to II (I: n = 3; II: n = 12). The third subgroup of 14 advanced-stage medicated participants had a mean age of 65.85 years (*SD* = 6.47) and a mean disease duration of 10.21 years (*SD* = 5.87). Their H&Y stage ranged from III to IV (III: n = 8; IV: n = 6).

Using a 2AFC pattern recognition memory task, Owen et al. (1992) reported that there was no statistically significant difference in performance between any of the groups. A reanalysis of their data indicated only a small pattern recognition memory deficit in the early-stage *de novo* and early-stage medicated PD participants (d = .06, -.46–.58, and d = .19, -.33–.71, respectively), but a larger impairment in the advanced-stage medicated PD participants (d = .56, .02–1.09).

Finally, as described earlier, Lees and Smith (1983) found no statistically significant difference in recognition memory between *de novo* PD participants with mild symptoms and controls. The magnitude of the ES (d = .06, -.46–.57) confirms this finding. Taken together, the data from early- and late-stage PD patients suggest that recognition memory functioning may decline as PD progresses.

Age at Onset

Differences between PD patients with early- and late-onset of symptoms have been used to support the theory that there may be more than one form of PD (e.g., Biggins et al., 1992; Horiguchi, Nishimatsu, Inami, Sukegawa, & Shoda, 1991; Lieberman et al., 1979; Reid et al., 1989; Starkstein, Bolduc, Mayberg, Presziosi, & Robinson, 1990; Zetusky, Jankovic, & Pirozzolo, 1985). Late-onset PD has been associated with more widespread cognitive decline relative to early-onset PD (e.g., Biggins et al., 1992; Caparros-Lefebvre, Pecheux, Petit, Duhamel, & Petit, 1995; Dubois, Pillon, Sternic, Lhermitte, & Agid, 1990; Horiguchi et al., 1991; Hietanen & Teravainen, 1988; Katzen, Levin, & Llabre, 1998; Mahieux et al., 1998; Reid et al., 1989). Moreover, evidence suggests that this is not due to concomitant Alzheimer's disease pathology in late-onset participants (Dubois et al., 1990). In addition, longitudinal studies have shown that late-onset participants have more rapid progression of PD (Goetz, Tanner, Stebbins, & Buchman, 1988), more severe motor disability, poorer response to dopaminergic therapy, and a higher incidence of gait disturbance (Biggins et al., 1992; Caparros-Lefebvre et al., 1995).

In terms of memory, there appears to be less distinction between early- and late-onset participants; both groups show similar deficits relative to age-matched healthy controls (Dubois et al., 1990; Haeske-Dewick, 1996; Hietanen & Teravainen, 1988). Dubois et al. suggested that late age at onset compounds only the frontal-lobe related symptoms rather than those that rely on temporo-parietal lobe functioning (e.g., memory). Thus, they argue that the differences between early- and late-onset participants are due to age-related factors and not the disease process itself.

In an investigation of recognition memory, Haeske-Dewick (1996) examined 13 (8 male, 5 female) early-onset (< 55 years old at onset) and 13 (9 male, 4 female) late-onset (> 55 years old at onset) PD participants. The early- and late-onset participants were compared to equal numbers of age-matched controls. The early-onset PD group had a mean age of 55.80 years (*SD* = 6.24) and a mean disease duration of 8.23 years (*SD* = 5.30). Mean Age at onset of symptoms was 47.57 years. Their H&Y stage ranged from I to V (I: n = 1; II: n = 3; III: n = 6; IV: n = 2; V: n = 1). The late-onset PD group had a mean age of 73.88 years (*SD* = 6.15) and a mean disease duration of 8.38 years (*SD* = 4.97). Mean age at onset was 65.50 years. Their H&Y stage ranged from II to V (II: n = 3; III: n = 6; IV: n = 3; V: n = 1). All PD participants were taking antiparkinsonian medication. The two PD groups were similar with regard to clinical symptoms, disease duration, premorbid IQ, and depression.

Haeske-Dewick (1996) used a 2AFC face recognition memory task. They concluded that both the early- and the late-onset PD participants had significantly poorer recognition memory relative to their respective controls. A reanalysis of their data in terms of ES confirmed that both early- and late-onset participants had similar deficits relative to controls (*d* = 1.12, .25–1.98, and d = 1.24, .37–2.12, respectively).

In an earlier study, Dewick et al. (1991) used a sample of participants, which included the late-onset participants used by Haeske-Dewick (1996), to test recognition memory for words and numbers. Dewick et al. reported that there was no statistically significant differences between the PD and control groups. However, a reanalysis of the Dewick et al. data provided evidence of small recognition memory deficits for two immediate recognition tasks (d = .22, -.42–.85; d = .30, -.37–.98) and for a delayed recognition task (d = .17, -.51–.84).

Taken together, these data suggest that while recognition memory deficits may exist in PD participants, age at onset is unlikely to be a moderator. Nevertheless, because of the paucity of findings on this issue, a firm conclusion awaits further study.

Depression

Depression has also been related to the severity of cognitive impairment in PD (Starkstein et al., 1989). The issue is complicated because it is difficult to distinguish between the influence of depression and that of PD (Lieberman, 1998). Furthermore, it has been suggested that the mechanisms of depression and cognitive impairment in PD may interact, accelerating the progression of impairment (Starkstein et al., 1990). On the other hand, Starkstein et al. suggested the possibility of two different forms of PD. One form presents with depression and rapid cognitive decline, whereas the other presents without depression and a gradual cognitive decline. These issues are important because depression is common in PD; approximately half of all patients suffer from mild to moderate depressive symptoms (Lieberman, 1998; Levin et al., 1992b). Nevertheless, what little evidence there is suggests that depression contributes little to the memory dysfunction in PD (Owen et al., 1993; Taylor et al., 1988) or in healthy elderly people (Boone et al., 1995; Burt, Zembar, & Niederehe, 1995). However, as with age at onset, more data are needed to clarify the role depression plays in the memory deficit seen in PD.

Medication Effects

In general, there is little agreement regarding the affect of antiparkinsonian medication on the cognitive function of PD patients (Levin & Katzen, 1995). However, with respect to memory, the available evidence suggests that dopaminergic therapy has no effect, or may actually improve it (Cooper et al., 1993; El-Awar et al., 1987; Growdon et al., 1990; Growdon et al., 1998; Lange et al., 1992; Mohr et al., 1989). In addition, Huber, Shulman, Paulson, and Shuttleworth (1987) found that delayed recognition memory was not affected by the absolute dopamine level, but rather was state dependent. That is, changes in dopamine level between learning and recognition phases produced recognition deficits in PD participants that were not evident when dopamine was held constant.

The issue is more complicated for anticholinergic medication since it has been found to have both no effect (Appollonio et al., 1994; Bondi & Kaszniak, 1991; Levin, Llabre, Reisman, Weiner, & Brown, 1991; Sahakian et al., 1988) and a detrimental affect (van Spaendonck, Berger, Horstink, Buytenhuijs, & Cools, 1993) on recognition memory. These studies provide no compelling evidence that medication per se brings about recognition memory deficits.

Prospective Memory

To this point, the present review has focused on retrospective memory (i.e., memory for past events). However, the process of remembering to remember, or *prospective memory*, also deserves some attention. Prospective memory tasks require the participant to execute a planned action at some point in the future (Einstein & McDaniel, 1990; Morris, 1992). Cockburn and Smith (1994) constructed a simple model of performance on a prospective memory task, which involved remembering to ask about a future appointment upon hearing a timer ring. To perform the task correctly a participant must encode and store the

instructions, activate the intention to respond when the timer rings, retrieve the appropriate response, and execute this response.

The term, prospective memory, although widely used, has been criticised on the grounds that it implies a memory system distinct from retrospective memory (Mayes, 1997). Mayes argued that the only distinctive aspect of prospective memory is that it concerns the "realisation of delayed intentions" (p. 1423). (However, in keeping with the literature, the term will be used here.) Nevertheless, after failing to find a relationship between retrospective and prospective memory task performance, other researchers have argued instead that these tasks may utilise different neurological processes or structures (Huppert & Beardsall, 1993; Kvavilashvili, 1987; Maylor, 1990, 1993). Still other researchers (e.g., Dobbs & Rule, 1987; Tombaugh, Grandmaison, & Schmidt, 1995) have differentiated between two components of prospective memory: *remembering-to-remember* and *remembering-what-to-remember* (or *content*). This second component is what is measured by retrospective memory tasks. Therefore, it has been argued that it is important to consider the relative weight of each component within a prospective memory task when trying to make a distinction between prospective and retrospective memory (Tombaugh et al., 1995). The importance of the remembering-to-remember component may vary depending on the balance between internal and external processes.

In this regard, Einstein and McDaniel (1990) made a distinction between *time-based* tasks and *event-based* tasks, the former depending more on self-initiated retrieval processes than the latter. Einstein and McDaniel defined time-based prospective memory tasks as those requiring a person to remember to perform some action at a particular time or after a particular amount of elapsed time (e.g., remembering to catch a 1 p.m. bus). In contrast, an environmental event signals the appropriateness of the intended action in an event-based task (e.g., remembering an appointment when the alarm rings).

Prospective memory plays an important role in daily activities (Cockburn & Smith, 1994), yet no research has looked at whether PD affects prospective memory (Knight, 1998). More generally, research shows that there is a decline in prospective memory performance with age, but only in situations that require more effort, or processing resources, to execute. That is, on complex tasks (Einstein, Holland, McDaniel, & Guynn, 1992) or tasks where participants must use self-initiated cues (Einstein, McDaniel, Richardson, Guynn, & Cunfer, 1995; Maylor, 1990). Support for this line of reasoning comes from research that shows an age-related decline in performance on time-based tasks but not event-based tasks.

For example, Einstein et al. (1995) conducted a series of experiments using 36 younger participants aged 18 to 22 years (M = 20.2), 28 middle-aged participants aged 35 to 49 years (M = 42.5), and 26 older participants aged 61 to 76 years (M = 66.3). Einstein et al.'s first experiment established that the elderly group had impaired prospective memory on a time-based task relative to younger participants. In the second experiment they failed to find an age-related prospective memory deficit using an event-based task. However, their final experiment provided the most compelling evidence that prospective memory is impaired in older participants when the task requires self-initiated processes. Type of task was manipulated in a single experiment and care was taken to ensure that the tasks were identical, except that one task was event-based and the other was time-based. Each prospective memory task was embedded in a general knowledge and problem-solving task. For the time-based task, participants were instructed to press a response key every 5 min, whereas in the event-based task the response key was to be pressed whenever the word *president* appeared in a question. The authors reported that there were no reliable differences between the three groups in terms of event-based prospective memory. With the time-based task, however, older participants had significantly poorer prospective memory than either the middle-aged or the younger participants. A reanalysis of Einstein et al.'s results in terms of ESs supported these findings. That is, there was little evidence of an important difference in event-based prospective memory. Nevertheless, the older participants had a large

deficit in time-based prospective memory relative to both the middle-aged and the younger participants (d = 1.29, .69–1.89, and d = .85, .32–1.38, respectively).

Task complexity may also moderate the relationship between age and prospective memory. Einstein et al. (1992) demonstrated an age-related decline in prospective memory using a complex event-based task, but not using a simple event-based task. They used a sample of 48 younger participants aged 19 to 22 years (M = 20.56) and 48 older participants aged 60 to 80 years (M = 69.13). The groups were matched in terms of the WAIS-R Vocabulary Subscale and the number of years of education beyond high school. The prospective memory task was embedded in a short-term memory measure. Participants were instructed to press a response key whenever a particular target word occurred. In the simple prospective memory task, a single target word was used, whereas in the complex task four target words were used. The investigators concluded that the older participants had significantly poorer performance than the younger participants on the complex task only. A reanalysis of their results in terms of ESs confirmed that the two groups had similar performance on the simple prospective memory task (d = .17, -.41–.75), but a large difference in performance on the complex task (d = 1.06, .44–1.67). The differences in performance on the complex task were replicated in a second experiment using a different sample of participants.

Prospective memory tasks may also be sensitive to the onset of dementia. Huppert and Beardsall (1993) studied 12 minimally demented (mean age = 87.3 years), nine mildly/moderately demented (mean age = 80.6 years), and 27 healthy elderly (mean age = 81.1 years) participants. The groups were similar with respect to premorbid IQ and school leaving age. Prospective memory was assessed with three event-based tasks. The first task, *remembering an appointment*, required participants to remember to ask the interviewer to do something when a buzzer sounded. The second task, *remembering a belonging*, required participants to provide the interviewer with a personal belonging at the start of the session and then remember to ask for it back when the session was complete. The final task, *remembering to deliver a message*, was embedded

within another task that involved walking a specific route around the room. The prospective memory aspect of the route walking task involved picking up an envelope at the start of the route and then putting it back at the end. The task was performed both immediately and after a 20 min delay.

The results of Huppert and Beardsall's (1993) study indicated that both the minimally demented group and the mildly/moderately demented group had statistically significant prospective memory deficits relative to controls. A reanalysis of the data in terms of ES showed that the largest deficit occurred for remembering a belonging, in both the minimally demented group (d = 2.25, 1.43–3.07) and the mildly/moderately demented group (d = 2.55, 1.64–3.45). Interestingly, verbal recall was less impaired than prospective memory, despite the recall task being more difficult, particularly in the minimally demented patients. Furthermore, an analysis of covariance reported by Huppert and Beardsall showed that prospective memory deficits remained after the variance due to retrospective memory impairment was partialled out.

More recent research indicates that prefrontal functioning, and to a lesser extent hippocampal functioning, are involved in prospective memory performance. McDaniel, Glisky, Rubin, Guynn, and Routhieaux (1999) studied event-based prospective memory in 41 elderly participants. Participants were subdivided into one of four groups based on whether they were high- or low-functioning on frontal tasks and whether they were high- or low-functioning on hippocampal tasks. McDaniel et al. concluded that the low-frontal-functioning group had significantly poorer prospective memory than the high-functioning group, suggesting that frontal processes play a key role. In addition, although there was no statistically significant difference in performance between the hippocampal groups, the low-hippocampal participants were reported as having poorer prospective memory. A reanalysis of their results in terms of ESs revealed that there was a large prospective memory deficit in low-frontal-functioning participants relative to high-frontal-functioning participants (d = .89, .23–1.55). There was also a small deficit in low-hippocampal participants relative to high-hippocampal participants (d = .43,

-.21–1.07), but no interaction between frontal and hippocampal status. The authors also suggested that their results are consistent with the view that executive functions, such as organising the execution of an intended action (served by frontal processes), are involved in prospective memory. Furthermore, hippocampal functioning may be involved in the retrieval process of prospective memory.

Memory Systems: An Overview

The memory models developed in cognitive psychology suggest that memory is not a homogenous entity, but is multifaceted and highly versatile (Richardson-Klavehn & Bjork, 1988). "Memory seems to comprise several distinct yet mutually interacting systems and subsystems, each differing either in the type of information stored, the processes acting on that information, or both" (Heindel, Salmon, & Butters, 1993, p. 740). In Schacter and Tulving's (1994) view, "memory systems are large, elaborate, and complex. They have fuzzy boundaries, have overlapping constituent processes, and interact with one another in intricate ways" (p. 18). In addition, there is widespread agreement that these systems are anatomically distinct, although there is less agreement on what constitutes a memory system or the exact neural networks that mediate their processing (Gabrieli, Fleischman, Keane, Reminger, & Morrell, 1995).

Over a number of years, Tulving and Schacter formulated a comprehensive theory of human memory, which now consists of five major systems: *procedural memory*, *perceptual representation* (PRS), *semantic memory*, *working memory*, and *episodic memory* (Schacter & Tulving, 1994). The history of their multiple memory system approach, variations, and contrasting theories have been reviewed extensively (e.g., Schacter & Tulving, 1994). I focus here on Schacter and Tulving's five major systems and what is currently known about the neural structures that subserve these systems.

Procedural memory is involved in learning various kinds of behavioural and cognitive skills and algorithms. Schacter and Tulving proposed several subsystems of procedural memory: motor skills, cognitive skills, simple conditioning, and simple associative conditioning. The procedural memory system operates at the subconscious level and its output is noncognitive, whereas the other four major systems are concerned with cognition (Schacter & Tulving, 1994). Several lines of evidence suggest that the procedural memory system involves processes mediated by the primary motor and sensory cortices, the basal ganglia, the cerebellum, and the dorsolateral and the midlateral frontal lobes (Moscovitch, 1994; Oscar-Berman & Bardenhagen, 1998; Salmon, Lineweaver, & Heindel, 1998).

The second major memory system described by Schacter and Tulving (1994), working memory, is responsible for temporarily holding and processing information. Baddeley (1992a, 1992b) established a multiple component model of working memory, comprising of a central executive and two active slave systems: the visuospatial sketchpad and the phonological loop. The central executive is responsible for directing attention and determining which items will be stored in working memory (Baddeley, 1994). The visuospatial sketchpad and phonological loop are responsible for holding and manipulating visuospatial information and for maintaining speech-based information, respectively (Baddeley, 1994). Evidence from a variety of sources, including neuropsychological and neuroimaging studies, suggests that structures in the prefrontal cortex underlie the central executive system (Owen et al., 1998).

The remaining three memory systems can be characterised as long-term systems. The PRS system is involved in the identification of words and objects, and is typically associated with priming (Schacter & Tulving, 1994). Schacter and Tulving proposed three subsystems: visual-word-form, auditory-word-form, and structural-description. The PRS appears to be mediated by structures in the posterior cortex (Mayes, 1998; Moscovitch, 1994), or more specifically, the right occipital cortex (Fleischman et al., 1995; Gabrieli et al., 1995).

The fourth major system is semantic memory, which involves factual information about the world and includes such facts as "Robins are birds," and "Chairs have legs" (Shoben, 1992). Schacter and Tulving (1994) differentiate two subsystems: spatial and relational. Evidence from lesion and functional neuroimaging studies suggests that the temporal cortex, particularly the left anterolateral temporal lobe, subserves semantic memory (Fink & Randolph, 1998). Vandenberghe, Price, Wise, Josephs, and Frackowiak (1996), using positron-emission tomography (PET), noted parietal and frontal cortical involvement in addition to the left temporal lobe. There was also modality-specific activation in the left posterior inferior temporal sulcus for picture tasks and in the left superior temporal sulcus, left anterior middle temporal gyrus, and left inferior frontal sulcus for words.

Lastly, the episodic memory system is an extension of semantic memory that allows the recollection of personal experiences through the use of multifeature representations involving spatial, temporal, or contextual information (Schacter & Tulving, 1994). Several lines of evidence indicate that a number of neural structures are associated with episodic memory, including the "hippocampus and related limbic structures in the medial temporal lobes and diencephalon....dorsolateral and ventromedial frontal lobes, [and] cingulate cortex" (Moscovitch, 1994, p. 272). The neuroanatomy of episodic memory has recently been reviewed extensively (e.g., Eichenbaum, 1997; Gabrieli, 1998; Squire, 1994; Ungerleider, 1995; Woodruff-Pak, 1997).

Schacter and Tulving (1994) differentiate their concept of memory systems from "forms of memory," "memory processes," "memory tasks," and "expressions of memory." They suggested that while these terms are related to the concept of a memory system they do not in themselves constitute a memory system. According to these authors, terms such as recognition memory and olfactory memory describe forms of memory and should not be thought of as specific memory systems. Furthermore, memory processes (e.g., encoding, rehearsal, and retrieval) concern the operation of memory, possibly in more than one system.

Also, care should be taken when relating task performance to the functioning of a memory system. Memory tasks may tap a single system, but frequently more than one system is involved. Moreover, task performance generally involves an interaction between cognitive functions (Markowitsch, 1998).

Schacter and Tulving (1994) went on to suggest that expressions of memory such as *explicit* and *implicit* are not memory systems, but refer to forms of memory that can be distinguished both on psychological and behavioural grounds. The question of whether they are subserved by the same underlying memory system or discrete systems is still open to debate. Commonly, the terms explicit and implicit are used to distinguish between two broad classes of memory tests. Moscovitch (1994) defines explicit tests as those that require conscious recollection of past events, whereas implicit tests are those in which memory for the past is inferred from changes in performance with experience or practice. A similar distinction has been made between *declarative* (explicit) and *nondeclarative* (implicit) memory (e.g., Cohen & Squire, 1980; Squire, 1992). Under this scheme, Schacter and Tulving's (1994) procedural and PRS system are described as nondeclarative, while the other systems are declarative.

Recent findings from memory studies using functional neuroimaging techniques have generally supported neuropsychological research results, but have also lead to new insights. Berent and Giordani (1998) after an extensive review of the literature drew the following conclusions:

1. Human memory involves brain regions believed to be involved in motor control and sensory processing.

2. Memory encoding, storage and retrieval processes are likely mediated by different brain regions.

3. Distinctions between declarative, procedural and other types of memory have utility in cognitive research and these various classes of memory appear to be mediated by different neural networks.

4. The normal brain may work differently than the diseased brain and the young brain may work differently than the old brain in how each changes physiologically in response to experiential stimuli.

5. Short-term memory traces appear to be represented differently than are long term-term memories and these differences may be systematically dynamic.

6. Memory may be mediated simultaneously by different brain regions.

7. A localizing model that consists solely of brain mapping will not suffice to answer the complex questions involved in understanding how the brain responds to experience.

8. Memory subtypes may be mediated by different neurochemical systems and these distinctions may prove to be more important than are differences in structural regions.

9. Neural substrates of memory are heterogeneous and distributed widely in the brain. (pp. 135-136)

While both recall and recognition memory fit within the episodic memory system, theoretical models are needed to explain the processes involved. Generally, models of recognition have either grown out of retrieval models of recall or have been contrasted with these models. In addition to the interest in recognition memory models, the way recognition is measured has received considerable attention. The next two sections briefly summarise the models and the measures of recognition memory.

Models of Recognition Memory

Theoretical accounts of recognition memory have a relatively long history and as a consequence much has been written on the topic. Most models have been developed over a number of years and subtle changes have been made to account for subsequent experimental findings. Generally, recognition memory has been modelled by either a single-process or a dual-process. Single-process models have been developed under a number of guises and include discrete-state models (e.g., Bernbach, 1967; Kintsch, 1967), strength models (e.g., Anderson & Bower, 1972, 1974; Kintsch, 1970), scanning models (e.g., Murdock, 1974), multiple-trace models (e.g., Hintzman, 1988), the Search of Associative Memory (SAM) model (Gillund & Shiffrin, 1984), spreading activation models (e.g., Anderson, 1983), and connectionist models (e.g., Ratcliff, 1990). These models all share the assumption that recognition judgements are only based on an assessment of familiarity (or strength). In contrast, dual-process models (e.g., Atkinson & Juola, 1974; Mandler, 1980; Yonelinas, 1994) assume that recognition is based on familiarity and retrieval mechanisms. Familiarity is a rapid, direct-access process that is independent of the slower recall-like search process which involves associative or elaborative processing (Johnson & Hasher, 1987).

Signal detection theory (Green & Swets, 1966; McNicol, 1972) has played a prominent role in memory theory since Egan (1958, cited in McNicol, 1972) first applied SDT to the study of recognition memory. With regard to recognition memory, SDT is applied to the familiarity component and therefore its use can be found in both single- and dual-process models. Anderson and Bower (1972) described the application of SDT to recognition memory as follows:

> [The SDT] approach assumes that there is one normal distribution of strengths for those items which have been studied recently, and a different normal distribution for those items which have not. The first distribution is called the "old" and the second the "new." The average distance between

> the two distributions reflects the amount of strength that was added to the studied items by their recent presentation. It is assumed that the subject chooses some criterion point, *C*, along the dimension of strength....If the strength of a particular item exceeds *C*, the subject will decide it is an old item and thus "recognize" it. If the strength is less than *C*, the subject will decide it is a new item and reject it. (pp. 98-99)

In summary, early models of recognition memory assumed that recognition judgements were based on a single-process (usually an assessment of familiarity). Later models have used a dual-process mechanism where recognition judgements are thought to be based on both a familiarity component and a recall-like search process. Regardless of the type of model, the familiarity component has most often been described in terms of SDT.

Measures of Recognition Memory

Measures of recognition can be characterised by the type of information the participant must recognise, by the type of test used, by the type of procedure used, and by the experimental measure (Murdock, 1982). Drawing mainly from Murdock's work, each characteristic is described below in more detail.

Recognition memory tests generally involve *item information*, but they can also be designed to test *associative information* and *serial-order information.* Item information concerns the fact that some item, object, or event has occurred in the past. It is often context specific and may be in the form of words, numbers, common objects, abstract drawings, or pictures of faces. As noted earlier, face recognition could be distinct from other forms of recognition as it may involve at least some face-specific processing mechanisms (Bruce & Young, 1986; Geschwind, 1979), although others argue that there is no compelling evidence to support this view (e.g., Ellis, 1981). Associative information concerns the relationship between two items, objects, or events (e.g., names and faces, words

and their meanings, and artists and their works). Serial-order information concerns the order of items or sequence of events.

Three main types of test are used to assess recognition memory: *yes–no tests, forced-choice tests,* and *batch testing*. A yes–no test involves a binary response to a question (e.g., is Wellington the capital of New Zealand?) Forced-choice tests involve the presentation of a single target along with one or more distractors. The more similar the distractor is to the target, the harder the task and, thus, the lower the recognition score (Singh & Cole, 1985). Forced-choice tests are also the only type of test nearly free from response bias (i.e., biases in favour of a particular response when certain fixed alternatives are available) (Cronbach, 1950, p. 3). Finally, in batch testing, all targets and distractors are presented at once. The participant must then work through the items discriminating targets from distractors.

There are two main types of procedure: the *study–test procedure* and the *continuous-task procedure*. In the study–test procedure a list of items (e.g., words, drawings, or sentences) is shown to the participant, followed immediately, or after some delay, by one of the recognition tests described above. In the continuous-task procedure there is no clear separation between the study and test phase. Stimuli are presented one after another and the participant must respond to each item as it appears. That is, "on the first presentation of an item, the correct response is 'no' (*I have not seen this item before in this list*) and on subsequent presentations [of the same item] the correct response is 'yes' (*I have seen this item before in this list*) (Murdock, 1982, p. 5).

There are three main measures used in the assessment of recognition memory: *accuracy*, *latency*, and *confidence*. Accuracy is usually expressed in SDT terminology. In a yes–no procedure, scores can be calculated for the *hit rate* (proportion of "yes" responses to old items) and for the *false alarm rate* (proportion of "yes" responses to new items) separately. However, these measures of accuracy are not free from *response bias*. The SDT model of

recognition memory is important because it provides means of dealing with response bias. That is, a measure of the average strength of the old items, known as *d*-prime (*d'*). This statistic is a function of both the hit rate and the false alarm rate. In a forced-choice procedure, percentage correct (PC) is usually used, but it is possible to convert the data to *d'*. Latency can be defined as the amount of time elapsing between stimulus presentation and response initiation. Confidence judgements are the participant's assessment of his or her own accuracy, and are often integrated with yes-no responses or given separately.

General Summary

Parkinson's disease is a neurodegenerative disorder that is primarily associated with motor symptoms, but can also produce specific cognitive deficits in individuals without dementia. Whether or not a single unifying theory, such as Brown and Marsden's (1990) processing resources theory, can account for all cognitive dysfunction in PD is open to debate. Many issues also need clarification, particularly with regard to memory. One issue highlighted in this review concerned the apparent dissociation between recall and recognition memory.

Several studies (e.g., Flowers et al., 1984; Lees & Smith, 1983; Taylor et al., 1986) appear to have been instrumental in establishing the view that recognition memory is relatively normal in PD. However, this notion appears unwarranted given that these studies had only limited statistical power. Moreover, other studies have reported impaired recognition memory in nondemented PD participants (e.g., O'Sullivan, 1998; Sahakian et al., 1988). The evidence reviewed in this chapter indicates that deficits in recognition memory may be moderated both by task variables (e.g., task difficulty) and by participant variables (e.g., disease stage or severity). The extent to which either is involved may be better estimated by using *meta-analytic* procedures than the narrative approach so far adopted. At the most basic level, meta-analysis refers to a set of statistical techniques for summarising a body of literature (see chapter 2 for full details). Although the use of meta-analysis is becoming increasingly common in the psychological literature, no such

analysis has been conducted within the PD literature with respect to recognition memory.

A meta-analysis will not usually provide answers to questions that have not been explored at the primary research level. However, a meta-analysis may be very effective in terms of highlighting both the strengths and the weaknesses of existing studies, and the directions for future primary research (Eagly & Wood, 1994).

In the present investigation, a meta-analysis was conducted before designing a primary level study. The meta-analysis was designed to isolate possible sources of variability among the studies and to estimate the magnitude of any memory deficit. In addition, a power analysis was conducted first in order to quantify the level of statistical power in previous research. Low statistical power would further justify the use of the meta-analysis, because low power means the risk of a Type II statistical error (erroneously concluding there is no effect) is high (power = 1 - probability of a Type II error). It will be seen that many of the studies reviewed have very poor levels of power (< 30%). One way of overcoming this problem is to combine the results in a meta-analysis thus enabling an overall assessment of the population ES.

CHAPTER 2

POWER AND META-ANALYSES

Introduction

The power of a statistical test is the probability that it will yield statistically significant results (i.e., the probability of detecting an effect given that there actually is a real effect). Statistical power analysis exploits the relationship among the four variables involved in statistical inference: statistical power, population ES, alpha level, and the sample size. The population ES is the magnitude of the real effect to be detected. The alpha level specifies the maximum risk of mistakenly rejecting the null hypothesis, also known as the Type I error rate. Conventionally, alpha is usually set at $p = .05$. The sample size is the size of the sample taken from the population to which one wishes to generalise results. Given that the assumptions of the statistical model being used hold, the value of any one of these variables may be determined as a function of the other three.

Jacob Cohen is widely credited with introducing statistical power analysis to the behavioural sciences. To simplify the procedure of statistical power analysis, Cohen defined three levels of population ESs: small, medium, and large (Cohen, 1988). Since Cohen's (1962) first statistical power survey, many similar surveys have shown that small-to-medium ESs are the norm in behavioural science and that the statistical power to detect effects of these magnitudes is low (Rossi, 1990; Sedlmeier & Gigerenzer, 1989). The implication is that researchers generally have had a poor chance of detecting small-to-medium effects, even when the effects are real.

Within the framework of statistical significance testing, statistical power estimation is important for several reasons (Cohen, 1988; Rossi, 1990; Stevens, 1980). First, the estimation of statistical power is essential when interpreting nonsignificant effects. When statistical power is low, a null result is ambiguous because it may mean that either there is no effect, or there is an effect but the study was not sensitive enough to detect it (Fagley, 1985). This problem is more serious when the null hypothesis is the research hypothesis, and, thus, is confirmed when the results are nonsignificant. Null results should only be interpreted when statistical power is sufficient. Secondly, if researchers are going to invest time, effort, and resources into a study then they should give themselves the best chance of reaching the correct conclusion. Cohen (1988) suggests a minimum power level of .80 (i.e., an 80% chance of detecting an effect should one exist). Thirdly, and most importantly, investigators researching in an area where ESs are known to be, or suspected to be, small must ensure that statistical power is as high as possible. Only then do they have a reasonable chance of detecting an effect, or of replicating some important result when using statistical significance testing (Whittington & Podd, 1996).

However, the use of statistical significance testing has been debated since the 1950's (Eysenck, 1960). In every decade since then, commentators have argued that the reliance on statistical significance testing in psychology and related disciplines has had a detrimental effect on the cumulation of knowledge (e.g., Cohen, 1994; Cronbach, 1975; Eysenck, 1960; Salsburg, 1985; but see Chow, 1988, for an alternative view). Likewise, the authors of several major statistical texts used by psychologists have highlighted the problems associated with statistical significance testing (e.g., Hayes, 1973; Keppel, 1991). Moreover, Schmidt (1996) showed very clearly that we should abandon the significance test and instead use point estimates of ES and confidence intervals around these point estimates for the analysis of data from individual studies. Recently, Zakzanis (1998) argued that readers of neuropsychological reports would be better served if the words *significance* and *nonsignificance* were not used to describe research findings. With regard to analysing the data from multiple studies, there is now

substantial support for the use of meta-analysis (e.g., Rosenthal, 1991, 1995; Schmidt, 1992, 1996; Zakzanis, 1998). Within the field of psychology the use of meta-analysis and statistical power analysis is becoming increasingly common, and the reporting of ESs is advocated by a number of journals, such as those published by American Psychological Association (1994).

One danger of relying on statistical significance testing is that researchers often take a statistically significant *p* value as evidence for an important effect (Borenstein, Rothstein, & Cohen, 1997). Conversely, a nonsignificant finding is often taken to imply that there is no effect at all. However, the *p* value is a function of two factors, ES and sampling error (Carver, 1993). Carver noted that when sampling error is very small (large *n*), even differences of a trivial size can be statistically significant. Conversely, when sampling error is large (small *n*), large differences may not be statistically significant. Furthermore, Schafer (1993) pointed out that "the event of nonsignificance suggests only that the data are not sufficient to estimate a parameter. In other words, the range of likely parameter estimates includes zero. This does not mean the data estimate the parameter to be zero!" (p. 384-385). By reporting the magnitude of the effect and the associated confidence interval, a clearer picture of the findings is provided.

With respect to the analysis of data from multiple studies, meta-analysis was developed independently by Glass (1976) and Schmidt and Hunter (1977), with important contributions from other researchers (Hedges & Olkin, 1985; Rosenthal & Rubin, 1982). Meta-analysis is a perspective that uses objective techniques for research synthesis. It was designed to be a more precise and quantitative method for integrating study findings than the traditional narrative review (Hunter & Schmidt, 1990). The advantage of a meta-analysis over a narrative review is most clearly seen when trying to integrate conflicting findings across large numbers of studies. In these situations, the statistical techniques developed by Hunter and Schmidt allow a meta-analyst to estimate how much variance in a set of ESs can be attributed to sampling error. Large amounts of variance not accounted for by sampling error indicate possible moderators, whereas small amounts of variance

can usually be ignored (Hunter & Schmidt, 1990). Finally, it should be noted that a meta-analysis will not always provide a conclusive answer to a research question, but may highlight productive directions for future research (Eagly & Wood, 1994).

Purpose of the Power and Meta-Analyses

The present investigation aimed to clarify the degree of impairment in PD recognition memory by using statistical power analysis[4] and meta-analysis to provide a comprehensive view of the literature on the topic. The meta-analysis was designed to isolate possible sources of variability among the studies involving recognition memory and to provide a quantitative estimate of the magnitude of any memory deficit. Given the problems associated with significance testing, no attempt is made in the meta-analyses to calculate the statistical significance of the observed effect.

The power analysis was conducted first in order to quantify the level of statistical power in previous research. It was designed to avoid some of the problems associated with past analyses. Specifically, power surveys in the behavioural sciences have suffered from at least three limitations. First, the sample period generally has been limited to just one year of one journal (Rossi, 1990). Second, the specific definitions of small, medium, and large ESs are, by Cohen's (1988) own admission, arbitrary. Therefore, the power surveys will only be as representative as the ESs are of the actual effects in research (Rossi, 1990). However, it can be noted in passing that Cohen's estimates have now been supported by several recent ES surveys (Rossi, 1990; but see Murray & Dosser, 1987, for survey limitations). Third, nearly all power surveys have covered broad research domains (Rossi, 1990) and therefore tell us little about specific areas of research. By narrowing the research domain and expanding the time-frame and

[4] Power computations were conducted by Melanie Kan as part of a small Honours project.

number of journals, the present power analysis provides a more accurate assessment of statistical power.

Method

Literature Search

The literature search was restricted to journal articles published between the years 1978 and 1997 inclusive. The main reason for only using studies published in this 20-year period was to avoid those that may have used participants who had disorders other than idiopathic PD (e.g., ShyDrager syndrome, and progressive supranuclear palsy). Such disorders were once classified as PD but are now recognised as separate disease entities (Hietanen & Teravainen, 1986). Unpublished studies were not used in the present power and meta-analyses. Meta-analysts usually include unpublished material to reduce publication bias (i.e., bias due to selectively publishing only significant results). However, Collins and Miller (1994) argue that publication bias is unlikely in some cases; their argument applies here. First, the literature on memory deficits in PD contains many inconsistent findings and several published studies have failed to find a memory impairment. Second, in many studies, the main focus was not recognition memory; so it is less likely that a nonsignificant recognition effect would exert a systematic bias on published research. Therefore, given what appeared to be a small risk of publication bias, it was deemed too costly to search and gather all unpublished work.

Several methods were used to search the published literature. First, computer-based information searches on PsycLit and MedLine were conducted. Key words used in the computer searches included *Parkinson's disease*, paired with the descriptors *cognition*, *cognitive deficit*, *cognitive impairment*, *memory deficit*, and *memory impairment*. Second, a manual literature search was conducted on the journals considered most likely to publish studies on cognitive deficits in PD (see

Appendix A for a list of these journals). These journals were searched issue by issue for the years 1993–1995. In addition, the references of review articles were searched for further articles. From the large pool of references generated, only those that met a number of inclusion criteria were used (see Appendix B for a full listing of the references included in the power and meta-analyses[5]).

General Inclusion Criteria

All studies selected for inclusion had to be published in a peer-reviewed, English-language journal with a publication date between 1978 and 1997. To be included in the power analysis, a study had to compare the performance of people with PD to healthy controls on a task explicitly described as measuring memory. To be included in the meta-analysis, the memory task had to measure recognition memory (see recognition task inclusion criteria below). In addition, a study had to report sufficient information to permit ES estimation as described below.

Power Analysis

Exclusion Criteria

Preliminary calculations showed that the average number of statistical tests in each article for which power could be calculated numbered about 10. A power analysis for small, medium, and large ESs would have amounted to some 2600 power calculations. Therefore, a set of criteria was developed to limit the number of calculations.

First, all statistical tests were excluded other than *t* and *F* tests. These statistics were by far the most commonly used tests for evaluating the most important hypotheses. Procedures for calculating power for *t* and *F* tests are well established, unlike for some other statistics (e.g., nonparametric tests). Second,

[5] Note that all references used in the power and meta-analyses appear in the main reference section regardless of whether or not they are cited in the text.

tests that were peripheral to the main hypotheses were excluded (e.g., initial tests to determine participant suitability, or participant demographics). Finally, a few studies that failed to provide information such as sample size and degrees of freedom, necessary to calculate power, were also excluded.

A total of 46 from an initial set of 88 studies remained for the power analysis. Where alpha-adjusted procedures were used to control for multiple significance tests, power was calculated as if no adjustment had been made. In such cases, power calculations will overestimate the true power by a small amount (Sedlmeier & Gigerenzer, 1989).

Determining Effect Sizes

Statistical power for each test statistic was calculated for small, medium, and large ESs as defined by Cohen (1988). There is evidence that Cohen's estimates for these population ESs are reasonably accurate (Rossi, 1990). Previous power analyses (Cohen, 1962; Rossi, 1990; Sedlmeier & Gigerenzer, 1989; Whittington & Podd, 1996) have tended to use Cohen's population estimates rather than those based on the sample data for each investigation for several reasons. First, ES estimates based on sample data can be perturbed by error inherent in all measures of behaviour. Such error tends to result in underestimates of the true population ES. Second, several of the studies in the present power analysis had small sample sizes. In general, the smaller the sample the less reliable the ES estimate. Third, published research often fails to report enough information to calculate a sample ES. A fourth and final reason for following the conventional approach to estimating population ESs was that data from the present study lend support to the estimates of population *d* values given by Cohen (1988). The ordered *d* values from the entire data set were divided into three equally sized groups yielding average *d* values of .18 (small), .52 (medium), and 1.09 (large). These values are in good agreement with Cohen's (1988) estimated population *d* values of .20, .50, and .80, respectively.

Procedure

Statistical power was calculated using the GPOWER computer program developed by Erdfelder, Faul, and Buchner (1996). Power can be calculated for a range of univariate test statistics, including *t* and *F*. The program calculates power using Cohen's (1988) small, medium, and large ESs as the default. However, any ES value can be entered. For each power calculation, the appropriate values for alpha, ES, and *N* were entered and power was then calculated by the program for 1380 test statistics. Where a study used a particular statistic more than once (almost always the case), the mean power level was calculated separately for small, medium, and large ESs for that study.

Meta-Analysis

Exclusion Criteria

A total of 39 journal papers were identified; 17 were not used in the meta-analysis because they did not use a nonneurologically impaired control group (Helkala, Laulumaa, Soininen, & Riekkinen, 1989; Pillon, Deweer, Agid, & Dubois, 1993), or did not report sufficient information (Allain et al., 1995; Direnfeld et al., 1984; Jacobs et al., 1995; Reid et al., 1989; Tweedy et al., 1982), or used a possibly biased measure of recognition (Bondi & Kaszniak, 1991; Buytenhuijs et al., 1994; Daum et al., 1995; Hartikainen, Helkala, Soininen, & Riekkinen, 1993; Pillon et al., 1996), or used a recognition test that did not meet the criteria for inclusion in the present analysis (Dubois et al., 1987; Huberman et al., 1994; Levin et al., 1989; Sagar et al., 1988; Tsai, Lu, Hua, Lo, & Lo, 1994). Specifically, tasks that involved batch testing were excluded from the analysis. Therefore, only data from tasks testing item information using yes–no or forced-choice tests with accuracy as the dependent variable were included.

Variables coded from each study

Coding was carried out by the present author. Reliability of coding was checked by having a research assistant code a randomly selected subset of 33% of the sample. Discrepancies in coding were measured with Kappa's coefficient and Pearson's correlation for categorical data and non-categorical data, respectively (Bryman & Cramer, 1994). Overall, the coding was very reliable (mean r = .89, SD = .19). Where low reliabilities were found, after discussion between coders, the particular variable was re-coded by the author. Studies were not weighted on the basis of their research quality.

General information coded from each study included: year of publication and geographical location of sample (North America, United Kingdom, Europe, Australia, New Zealand, or other).

Participant characteristics for each study's PD and control samples were coded as follows: gender (expressed as the percentage of females); handedness (expressed as the percentage of right-handers); and, average age of participants.

The following characteristics were coded for the PD sample only: Intelligence measure and mean score; premorbid intelligence measure and mean score; depression measure and mean score; affective state of sample (depressed, nondepressed, or unselected); medication status (medicated, withdrawn from medication, never medicated, or unselected); laterality of disease symptoms (unilateral, bilateral, or unselected); age of onset of symptoms (early, less than 45 years old; late, greater than 60 years old; middle, between 45 and 60 years old; or, unselected); physical symptoms measure and mean score; number of participants at each Hoehn and Yahr stage (Hoehn & Yahr, 1967); stage of disease (early, late, or unselected); disease duration (years); dementia measure and mean score; and cognitive status of sample (demented, nondemented, or unselected). In addition, a note was made of whether the PD group and control group were matched (yes/no), and if so, the variables used to match.

The following task characteristics were also coded: Name of task; recognition performance measure (e.g., percent correct, *d'*, discriminability); task type (yes-no, *n*-alternative); modality (verbal, nonverbal[6]); delay (delay, no delay); recall (recall before recognition, recognition only); and administration method (computer, manual).

Finally, all statistical data needed to conduct the meta-analysis were recorded (e.g., sample sizes, mean scores and standard deviations, inferential statistics, degrees of freedom, significance levels).

Appendix C contains selected sample descriptors (Table C1), and task information and ES data (Table C2) for each study included in the analysis.

Procedure

The basic approach was modelled on the meta-analytic techniques of Hunter & Schmidt (1990). Most computations were made with the help of a spreadsheet or Schwarzer's (1989) Meta-Analysis Programs (Version 5.3). Schwarzer's software was not used for all computations because modifications made by Hunter and Schmidt to some of their formulae have not yet been incorporated into the program (R. Schwarzer, personal communication, 1998).

Calculation of effect sizes. The standardised mean difference, *d* (Cohen, 1988) was used as the estimate of ES (see Appendix D for the formulae used in the computation of *d*). The *d* statistic can be defined as the difference between the group means divided by the within-group standard deviation. Where an ES could not be calculated directly from means and standard deviations, *d* was transformed from *t*-values or *F*-values where, $d = t\,[(n_e + n_c) / (n_e n_c)]^{1/2}$, and $t = \sqrt{F}$. For *F* values

[6] Stimulus material that can not be readily labelled.

≤ 1, a *d* value of zero was assumed. Where only the significance level was reported, this was converted to *r* and then transformed to *d*, where $d = (1/(pq)^{1/2})\ r / (1-r^2)^{1/2}$ and *p* and *q* are the proportions of persons in the two groups. With regard to calculating *r* from significance levels, when significance or nonsignificance was reported without further information, the normal (and conservative) approach was taken. That is, when the results were reported as significant, $p = .05$ was assumed and when nonsignificance was reported, $p = .50$ was assumed.

Multiple effect sizes from single studies. There is no generally agreed method for dealing with multiple ESs from a single study. Including all separate ESs regardless of their interdependence will actually lead to an underestimate of the degree of homogeneity across studies, but will not produce any systematic effect on the mean ES (Hunter & Schmidt, 1990). Where ESs come from different study characteristics that act as real and substantial moderators, these ESs can be entered into the meta-analysis as independent outcome values (Hunter & Schmidt, 1990). However, others argue that ESs should be independent with each study represented only once in an analysis (e.g., Bangert-Drowns, 1986). In the present analysis there was little basis for determining, *a priori*, whether a variable was a real moderator or not. Therefore, initially, a weighted average ES was calculated for each study and entered into the overall meta-analysis. (Note that a weighted average will be an underestimate of the value that would have been obtained from an overall composite variable if one could be formed; Rosenthal & Rubin, 1986.) There was one exception to this rule: Where a study reported separate results for two or more independent PD groups (e.g., demented and nondemented participants), a single ES was computed for each group individually. Separate control groups were not required as long as sufficient information was available to calculate ESs for each PD group.

Two studies used some of the same participants to obtain recognition data from different stimulus modalities (Sagar et al., 1988, and Sullivan & Sagar, 1989). For present purposes the two reports were treated as one. Overall, the procedure used in the present analysis produced more ES estimates than the number of

papers from which the data were taken. In total, the 22 published reports yielded a total of 32 independent PD groups. Therefore, in the overall meta-analysis there were 32 independent ESs.

Summary analysis of effect sizes. For the overall analysis, the weighted mean and variance were computed based on formulae reported by Hunter & Schmidt (1990): mean $d = \sum w_i d_i / \sum w_i$ and observed variance of $d = \sum w_i [d_i - D]^2 / \sum w_i$, where D is the weighted average of d, and w_i is the sample size of each group. A weighted average ES was used since large N studies have less sampling error than small N studies and therefore deserve more weight.

The population (residual) variance, s^2_{res}, was then computed by subtracting the sampling error variance, s^2_e, from the observed variance, s^2_r, where the sampling error variance was computed using the formula, $s^2_e = [(N - 1) / (N - 3)] [(4 / N) (1 + D^2 / 8)]$, N being the average sample size across all groups.

The estimate of the population variance served as the multiplier in the formula for the 95% confidence interval (CI): $d - 1.96 (s_{res}) < ES < d + 1.96 (s_{res})$. Hedges and Olkin (1985, p. 80) show that d has a small sample bias. By defining a new estimator, d^*, they remove this bias: $d^* = d / a$, where $a = 1 + .75 / (N - 3)$. Hedges and Olkin use the symbol "d" for their approximately unbiased estimator. To avoid confusion, d^* will be used to denote this estimator in the present study. The correction can be applied either study by study or, as in the present analysis, after the meta-analysis has been done using the average sample size (Hunter & Schmidt, 1990).

To reliably interpret the estimated population ES, the underlying data set should be homogeneous. Hunter and Schmidt (1990) suggest that a data set can be considered homogeneous when most of the observed variance is explained by artifacts (e.g., sampling error and error of measurement in the dependent and independent variables). Artifacts such as measurement error will produce both a systematic reduction in the mean ES and a systematic increase in the variance of

ESs (Hunter & Schmidt, 1990). It is possible to correct a meta-analysis for the effect of artifacts, provided sufficient information exists about the artifact. However, substantial evidence suggests that sampling error is responsible for as much as 85% of this artifactual variance (Stoffelmayr, Dillavou, & Hunter, 1983). Therefore, the decision was made to correct only for sampling error. Furthermore, it would be difficult to correct for other artifacts since little information is available in PD research with which to make the correction.

Once the effect of sampling error is removed from the data, the question remains: To what extent can any residual variance be explained as due to artifacts not corrected for? Hunter and Schmidt (1990) suggest that if 75% of the observed variance is accounted for by artifacts, then the data set can be considered homogeneous. In the present study, Hunter and Schmidt's rule of thumb was used, but the actual percentage of variance explained by sampling error is also reported to allow readers to judge for themselves.

Some authors (e.g., Rosenthal & Rubin, 1982) advocate the use of a statistical test based on the chi-square distribution to help determine whether there is heterogeneity in a data set. However, this method suffers from all the usual problems associated with statistical significance testing. Most problematic is the power of the test. In data sets containing a large number of studies, the test will have high power and almost any residual variance will produce a significant result. Conversely, in situations of low power, even large amounts of residual variance may not be detected leading to the conclusion that the data set is homogeneous when in fact it is not.

After the effect of artifacts has been removed from a data set, heterogeneity may remain as a result of moderating variables. To examine the influence of possible moderators in the present analysis, the data set was subdivided as a function of the potential moderator and each subset was subjected to further analysis. At this level, a single study may contribute ESs to each of the subsets provided that a separate ES can be calculated as a function of the moderator. Two requirements

had to be met in order for a variable to be classified as a moderator. First, the population ES had to vary between subsets. Second, the residual variance in each subset had to be smaller than that seen in the data set as a whole (Hunter & Schmidt, 1990). It is legitimate to search for moderators even after the conclusion has been made that all of the ES variability is due to sampling error (Rosenthal, 1995). This approach was taken in the present study when further moderators were suspected.

Hunter and Schmidt's (1990) procedure using the least squares method for estimating the mean and variance of a distribution of ESs is based on the assumption that the data do not contain outliers. They suggest that even a single outlier can greatly distort the observed standard deviation, and to a lesser extent the mean. Therefore, outliers were removed from each analysis by excluding values greater than three standard deviations from the mean. Subsequently, stem and leaf graphs were used to detect other possible outliers.

Results

Power Analysis

Table 1 summarises the results of the statistical power analysis, based on 48 studies and 1360 power calculations. The median power values are in good agreement with those obtained by other researchers (Cohen, 1962; Sedlmeier & Gigerenzer, 1989) in the area of abnormal psychology. For small ESs, none of the 48 studies examined in the present study had sufficient average power for *t* and *F* tests. Cohen suggests a minimum power value of 80%, but even for a medium ES only 15 of the studies reached this level. Assuming (unrealistically) a large ES for recognition memory deficits in PD, 12 (25%) of the studies had an average power value of less than 80%.

Table 1 *Summary Statistics for the Statistical Power Analysis*

	Effect size[a]		
	Small	Medium	Large
Mean	20	63	85
Median	16	67	94
Standard deviation	12	24	19
Minimum	6	15	32
Maximum	55	100	100
95% confidence interval	17–23	56–70	80–90

Note. The values represent power expressed as a percentage.
[a]Cohen's (1988) conventions for effect size (e.g., small: $d = .20$, medium: $d = .50$, large: $d = .80$).

In summary, most studies of memory deficits in PD have adequate power to detect only large ESs; the following results of the meta-analysis strongly suggest that large ESs occur only in certain subsets of participants in this field of research.

Meta-analysis

General characteristics. Data from a total of 32 independent PD groups taken from 22 studies were used in the analysis. At the study level, 77% of the studies were published between 1988 and 1997, with the remainder published between 1978 and 1987. At the sample level, 47% (15) of the PD samples were taken from North America and 44% (14) from the UK. The remaining were sourced from Europe (9%). The vast majority of PD samples (94%) were matched with a healthy control group, most commonly on age and gender. The remaining samples (6%) were compared to an unmatched control group.

PD group characteristics. At the sample level, 84% of the PD groups contained less than 25 participants, while only 3% exceeded 50. The mean sample size was 18.53 ($SD = 14.83$). From the reports that provided sufficient information, most groups (70%) contained 40% or fewer females. In most groups (81%), the average age of participants was greater than 60 years. Across all groups, the mean age was 63:41 years ($SD = 4.68$). A majority of the PD samples (75%) contained predominantly nondemented participants, while 16% were demented,

and 9% were either unselected or insufficient information was given to determine cognitive status.

In all cases where cognitive status was reported, it was confirmed with the use of a quantitative measure (e.g., MMSE). With regard to depression, 28% of the studies contained predominantly nondepressed participants while no study contained predominantly depressed participants. However, in most studies (72%) participants were either unselected or this information was not reported. Sixty-nine per cent of groups contained participants on antiparkinsonian medication, while 22% contained participants who had never received medication, the remainder being either unselected or this information was not reported. Just under half of groups (48%) had a mean disease duration of between 5-10 years, 24% had less than five years, and 15% had a mean duration over 10 years. However, this information was not reported for 12% of the groups used in the present analysis. Furthermore, for many groups (34%), information about disease stage was not reported. However, in those reports that did provide this information, 33% contained participants in the early stages of Parkinson's, while 24% were made up of participants in the later stages of PD. Nonreporting of participant characteristics, or using unselected groups, were factors in many studies. Such characteristics included laterality of symptoms, age of onset of symptoms, IQ, and estimated premorbid IQ.

Control group characteristics. The mean sample size of the 23 independent control groups was 24.61 (*SD* = 13.91). The mean age of controls was 61.91 years (*SD* = 7.37). Most (67%) control groups consisted of 40 to 60% males.

Results of the meta-analysis. Table 2 presents the results of the meta-analysis that addressed the issue of whether recognition memory is impaired in PD. The overall analysis indicated that substantial heterogeneity existed in the data set. Therefore, the data were subdivided according to the cognitive status of the participants; that is, demented, nondemented, and unselected samples. (It was expected that cognitive status would be a moderator since memory impairment is

Table 2 *Cumulated Effect Size (ES) Estimates and Residual Variation as a Function of Cognitive Status After Accounting for Sampling Error*

Moderator	*k*	*N*	*d**	S^2e	S^2res	95% CI	% Var
Overall	32	1401	.32	.10	.12	-.35–1.01	44
Demented[a]	4	130	1.30	.16	.00	1.30–1.30	100
Nondemented[b]	23	1039	.16	.09	.00	.16– .16	100
Unselected	3	157	.52	.08	.05	.11– .95	64

Note. *k* = number of ESs; *N* = total sample size; *d** = Hedges and Olkin's (1985) unbiased ES statistic; S^2e = variance due to sampling error; S^2res = residual variance; 95% CI = 95% confidence interval for *d**; % Var = percentage of variance attributable to sampling error. [a]One study (Heindel, Salmon, Shults, Walicke, & Butters, 1989) produced an ES that was considered an outlier and so was excluded from the analysis. [b]One study (Sahakian et al., 1988) produced an ES that was considered an outlier and so was excluded from the analysis.

a principle component of dementia. Studies that used demented participants were included in the meta-analysis so as to quantify the recognition deficit for all participants.) The analysis was then repeated for each subset.

In the demented subanalysis, the 95% confidence intervals for individual ESs overlap substantially (see Figure 2). In addition, the ES for every study was greater than .80. ESs of this magnitude can be described as large effects (Cohen, 1988). As illustrated in Table 2, sampling error accounted for 100% of the observed variance, and so the mean weighted ES (*d** = 1.30) can be considered representative of the entire set. This indicates that the recognition memory performance of the controls exceeded the performance of demented PD participants by over one standard deviation.

In the nondemented subanalysis, the confidence intervals also overlap (see Figure 3) and the analysis indicates homogeneity (see Table 2). The estimated population ES was .16. An ES of this magnitude indicates a small recognition deficit in nondemented PD participants.[7] As could be expected from the separate analyses of demented and nondemented cases, the third subset of unselected participants was not homogenous, although the ES was of moderate magnitude. Taken together, these results show that the estimated population ES varied

[7] This is likely to be an underestimate of the population ES as several papers (Breen, 1993; Cooper et al., 1993; Heindel et al., 1989; Taylor, Saint-Cyr, & Lange, 1987) reported only that their results were nonsignificant. Therefore, an ES of zero was entered into the meta-analysis, whereas the actual ES may have been greater than zero.

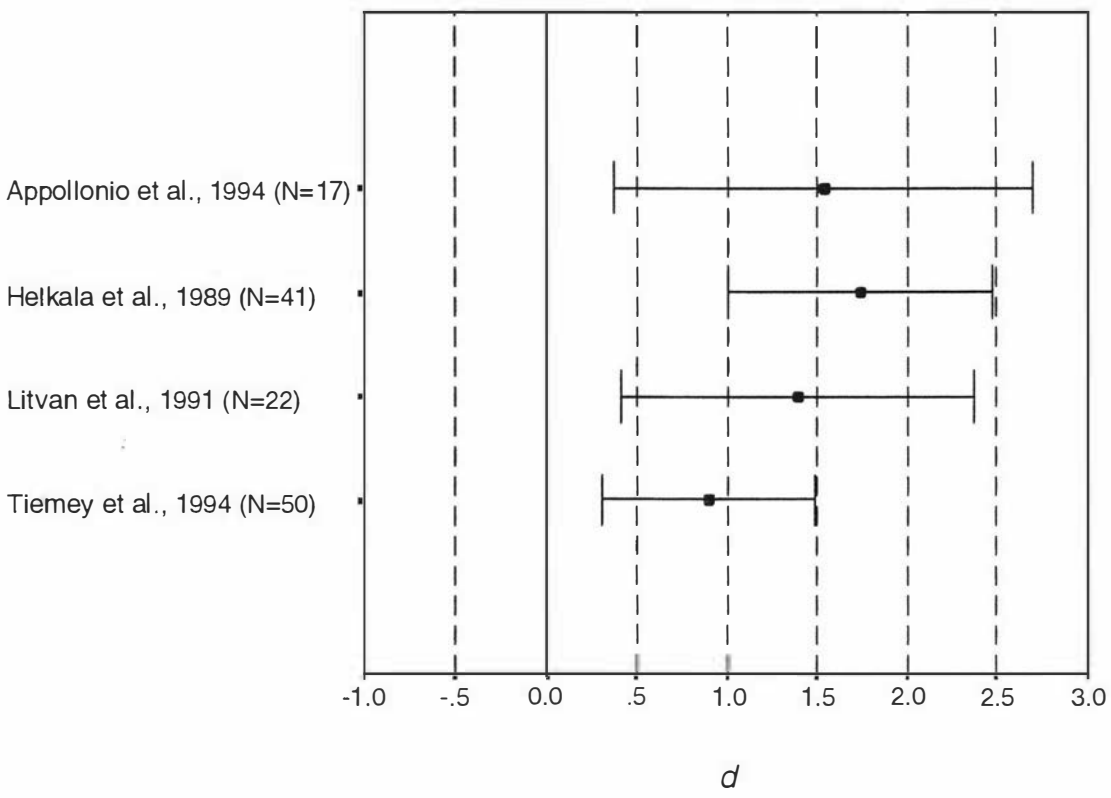


Figure 2. Meta-analysis plot of the effect sizes for each study entered into the demented subgroup analysis. The mean *d* for each study is shown by the points, and the horizontal lines show the 95% confidence interval.

between subsets and the residual variance in each subset was smaller than that seen in the overall analysis. Therefore, there is sufficient evidence to suggest that cognitive status is a moderator (Hunter & Schmidt, 1990).

Further analysis of the nondemented subset indicated that just 35% of the ESs were greater than .20, 17% of them were greater than .50, and only one exceeded .80. Moreover, 30% of the PD groups consisted of patients who were newly diagnosed and had not yet received medication, whereas the remainder were receiving antiparkinsonian medication. These *de novo* participants had a much smaller mean disease duration (M = 1.96 years, SD = 0.54) than the medicated participants (M = 8.55 years, SD = 2.15), d = 3.58, 2.20–4.97. In order to assess whether *de novo* and medicated PD participants differed with regard to recognition memory, the meta-analysis was repeated on these groups separately.

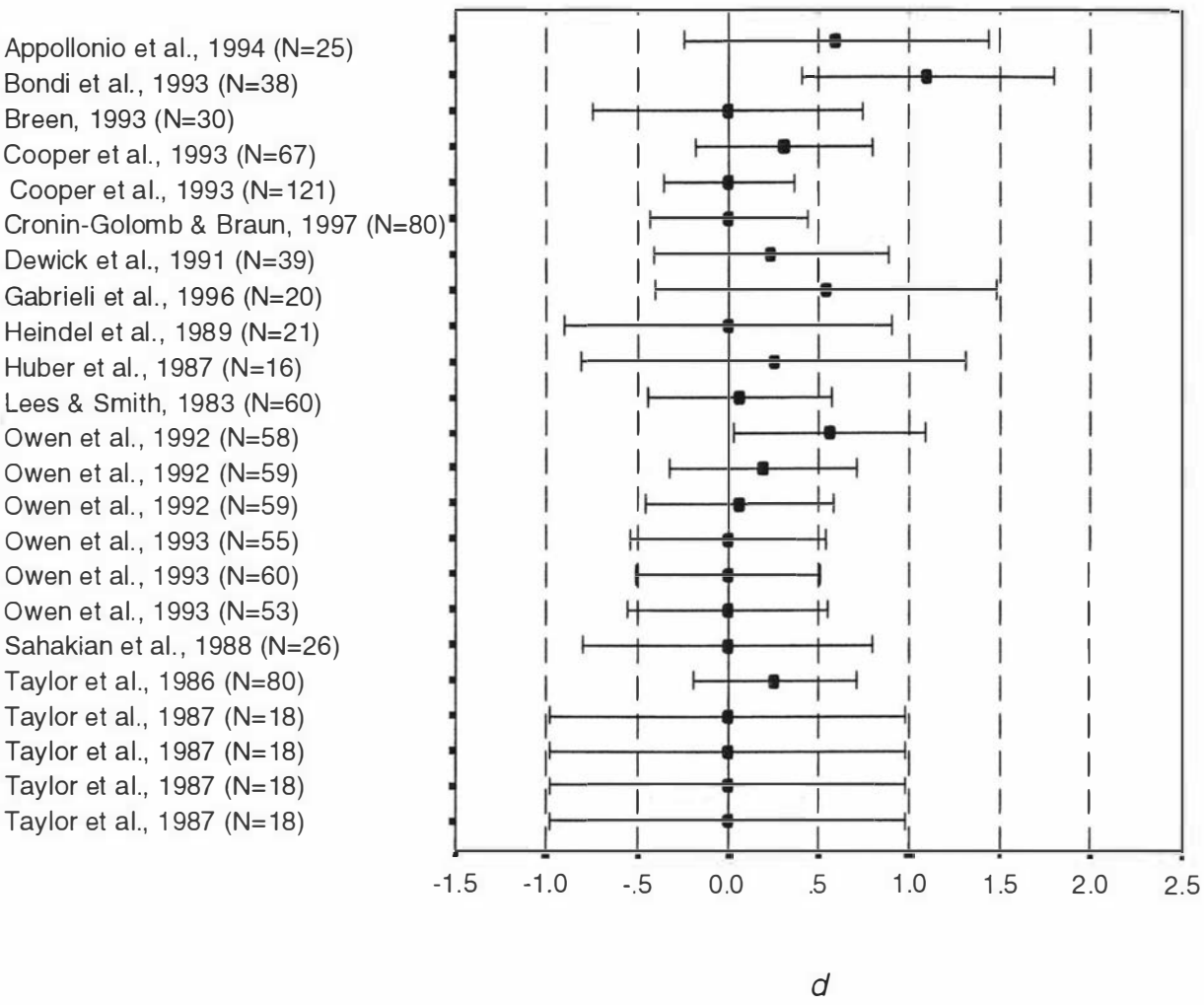


Figure 3. Meta-analysis plot of the effect sizes for each study entered into the nondemented subgroup analysis. The mean *d* for each study is shown by the points, and the horizontal lines show the 95% confidence interval. Four studies have multiple entries because they used independent groups of PD participants.

As can be seen in Table 3, the ES varied between subsets and the variance in each subset was explained by sampling error suggesting that medication status acts as a moderator of recognition memory deficits in PD. Medicated PD participants had poorer performance relative to controls, whereas *de novo* participants did not. This was indicated by an ES of .23 for the medicated subset, a population value that the associated 95% CI suggests can be viewed with high confidence. There was little evidence of a difference in recognition memory performance between *de novo* PD participants and controls (d^* = .05).

Table 3 *Cumulated Effect Size (ES) Estimates and Residual Variation as a Function of Medication Status After Accounting for Sampling Error*

Moderator	*k*	*N*	*d**	S^2e	S^2res	95% CI	% Var
Nondemented							
Medicated[a]	16	675	.23	.10	.00	.23–.23	100
De novo	7	364	.05	.08	.00	.05–.05	100

Note. *k* = number of ESs; *N* = total sample size; *d** = Hedges and Olkin's (1985) unbiased ES statistic; S^2e = variance due to sampling error; S^2res = residual variance; 95% CI = 95% confidence interval for *d**; % Var = percentage of variance attributable to sampling error; *De novo* = never received antiparkinsonian medication.
[a]One study (Sahakian et al., 1988) produced an ES that was considered an outlier and so was excluded from the analysis.

Task characteristics within cognitive status subsets. Within the demented subset, all of the studies used a verbal recognition task. In half of the studies, the task involved recall before the recognition phase. In three out of the four studies, the task had a forced delay between the presentation phase and the recognition phase. Also, in three of the studies, a yes–no task was used.

Within the nondemented subset, the majority (70%) of the studies used a verbal task and a majority (61%) also used tasks involving only recognition (i.e., no recall component). In 39% of the studies the task had a forced delay, while in a further 43% there was no delay. The remaining studies used both delay and non-delay tasks, but the results were averaged across this factor for the purpose of the meta-analysis. Finally, in the majority (65%) of the studies a *m*-AFC task was used, while in 31% a yes–no task was used, and in one study the results were averaged across this factor.

Discussion

Power Analysis

The statistical power analysis of studies on memory deficits in PD confirmed that for at least small and medium ESs, power was inadequate. Maximum power for a small ES was just 55%, and only 15 of 48 studies had power ≥ 80% for a medium ES. The median power values in the present study (16, 67, and 94%) are consistent with power analyses conducted in the area of abnormal psychology (Cohen, 1962: 17, 46, and 89%; Sedlmeier & Gigerenzer, 1989: 14, 44, and 90%).

The general picture emerging from a number of power analyses in a wide range of subdisciplines in psychology is that many studies simply do not have the power to detect the effects they may predict, especially when ESs are small. Under such circumstances, it is impossible to decide whether an acceptance of the null hypothesis is due to there truly being a population ES of zero, or due to a too high Type II error probability (due to low power).

If it is generally true that studies producing small ESs are underpowered, why not remedy the situation by increasing power? It appears that the reasons are largely of a practical nature. Power is completely determined by the ES, *N*, and the alpha (*p*) level. Increases in any one or more of these will increase power (Cohen, 1988). Researchers (and journal editors) seem reluctant to adjust the alpha level upwards, despite the fact that most well-known statistical handbooks (e.g., Keppel, 1991) advocate setting the alpha level only after careful consideration of the expected ES and the costs and values associated with drawing the wrong conclusion from a set of results. In regard to the ES, researchers probably try to the best of their ability to eliminate experimental noise that may have the effect of causing the sample ES to underestimate the true population ES. Thus, there is probably little room to manoeuvre here, in terms of increasing power. The

remaining variable is *N*, the number of participants. However, for small to moderate ESs, sample sizes for most common statistical tests have to number in the hundreds to produce around 80% power. The cost of such a large *N* is usually prohibitive.

These difficulties in obtaining sufficient power, especially for small-to-medium ESs, make it highly likely that studies of recognition deficits in PD will be underpowered. That is, because the probability of a Type II statistical error is high, it is unlikely that small memory deficits will produce significant results. As the power analysis shows, power levels have generally been inadequate to detect small-to-medium ESs.

In summary, the statistical power in many investigations of memory deficits in PD has been too low. There are strong reasons for believing that researchers would find it difficult in practice to increase power to a satisfactory level, even if they wanted to. Schmidt (1996) suggested that trying to improve power in individual studies may be too difficult. A better approach may be to increase statistical power by combining the results from several studies in a meta-analysis.

Meta-analysis

The current meta-analytic findings indicate that cognitive status moderates the recognition memory deficit in PD. In particular, the analysis of 4 studies using demented PD participants revealed a large recognition deficit (d^* = 1.30). An additional analysis involving those studies using nondemented PD participants indicated that medication was also a moderating factor. The meta-analysis of 16 studies using medicated PD participants suggested a small recognition deficit (d^* = .23), but the analysis of seven studies using *de novo* participants indicated little evidence of a deficit (d^* = .05).

Many previous reviews and some primary research examining memory functioning in PD have made little distinction between demented and nondemented participants with regard to recognition memory. The present analysis makes it clear that global cognitive status needs to be taken into consideration, especially at the primary research level. Nevertheless, the concept of PD dementia is controversial (see chapter 1), and is outside the scope of the present study. Therefore, the remainder of the discussion will focus on the analysis involving nondemented participants.

The meta-analysis provides evidence for recognition memory deficits in medicated patients, but not in patients who had never received medication. It is therefore possible that antiparkinsonian medication contributes to the observed impairment in these participants. However, this seems unlikely because the available evidence suggests that dopaminergic therapy has no effect, or may actually improve memory functioning (Cooper et al., 1993; El-Awar et al., 1987; Growdon et al., 1990; Growdon et al., 1998; Lange et al., 1992; Mohr et al., 1989).

Other differences between medicated and *de novo* patients provide a more plausible explanation for the difference in recognition memory functioning. For instance, the *de novo* patients generally were at an earlier stage in the progression of the disease than the medicated participants. As described in chapter 1, the relationship between motor disability and cognitive decline is not straightforward (Starkstein & Robinson, 1991). However, current evidence indicates that late stage participants have greater recognition memory deficits than those participants in the early stages (Lees & Smith, 1983; Owen et al., 1992; Sahakian et al., 1988).

Depression has been related to the severity of cognitive impairment in PD (Starkstein et al., 1989). This is potentially problematic for the present analysis because nearly three-quarters of the studies sampled do not appear to have excluded participants with depression or controlled for differences in the level of depression. Nevertheless, there is some evidence that depression may have little

impact on recognition memory in PD (Owen et al., 1993) or in healthy elderly people (Boone et al., 1995).

Taken together, the existing research suggests that medication and affective symptomology are unlikely to account for the clear distinction between *de novo* and medicated PD participants seen in the present meta-analysis. Although, it should not be inferred that both memory and motor deficits share a common pathology, the dissociation in memory dysfunction is more likely related to motor disability. While tremor and rigidity result from a dysfunction of dopaminergic mechanisms, there is little correlation between these signs and cognitive performance (Pillon et al., 1989). However, cognitive dysfunction is more likely to be related to motor disability that is unresponsive to levodopa therapy (e.g., gait disorder), in the sense that these two impairments may share a common nondopaminergic neurochemical pathology (Pillon et al., 1989; Sagar et al., 1995). Moreover, recognition memory deficits may be subserved by "by nondopaminergic mechanisms such as alterations of cholinergic, noradrenergic, or serotoninergic projections to the neocortex and hippocampus" (Lange, Paul, Robbins, & Marsden, 1993, p. 477). Support for this line of reasoning comes from research indicating that there may be two subgroups of PD. Zetusky et al. (1985) suggested that one subgroup presents with predominant tremor, earlier age at onset, and relatively normal functional and cognitive status. The other presents with predominant postural instability and gait disorder, later age at onset, greater functional and cognitive impairment, dysarthria, dysphagia, and more rapid progression.

In summary, nondemented, medicated PD participants appear to be impaired on tests of recognition memory, but participants early in the course of the disease who have never received medication appear to have normal recognition. It is unlikely that medication alone can account for the difference. Rather, it seems that nondopaminergic CNS abnormalities underlie both the recognition memory deficit and motor disability not alleviated by levodopa therapy. Such a view is consistent with the finding that recognition deficits are greater in PD patients with

more severe motor disability. Finally, PD is a complex disease and the consensus of opinion suggests that memory disturbances may be the result of degeneration in multiple systems (Brown & Marsden, 1990; Hammond-Tooke & Pollock, 1992).

Conclusion

The power analysis revealed that, in general, studies of memory deficits in PD (as in many other research areas) have not had sufficient statistical power to detect small-to-medium ESs. Yet there is no doubt, for both practical and theoretical reasons, that it is important to demonstrate the presence of these smaller ESs. The least controversial way of increasing power is to use larger participant numbers, but for small ESs, very large (and often unattainable) numbers are required. A very acceptable alternative is to pool the results of several small studies in a meta-analysis. This was done in the present study and the data provide compelling evidence that, as might be expected, demented PD participants suffer from a large recognition memory deficit relative to controls. Nondemented Parkinsonians have a much smaller deficit while *de novos* appear to have normal recognition memory. There is now a need for researchers to consider the mechanisms responsible for both impaired recall and impaired recognition memory in PD.

It seems unlikely (and perhaps even undesirable) that most individual studies will ever be able to achieve satisfactory statistical power levels for analysing small- to medium-sized effects. Without adequate power, statistical significance tests can not be correctly interpreted. Moreover, problems inherent in statistical significance testing limit its usefulness (Schmidt, 1996). Thus, with respect to interpreting the results of individual studies, point estimates of ES and their associated confidence intervals provide a good alternative. With respect to interpreting the results of multiple studies, meta-analyses will be required to help researchers reach general conclusions. These meta-analyses can only be as good as the data they are based on, most of which is obtained from individual research papers.

Therefore, researchers (and journal editors) must ensure that published results contain sufficient information to allow them to be included in meta-analytic reviews.

Finally, the initial literature review (chapter 1) highlighted a number of factors that may be associated with recognition memory functioning in PD (task difficulty, gender, age at onset, disease stage, and depression). While the results of the meta-analysis indicate that disease stage may be important, the other factors could not be examined due to a lack of primary research. The review also highlighted the concept of prospective memory and the fact that it has not been studied in PD participants. To directly address these issues, a study was conducted with nondemented PD participants. A recognition memory task was specifically designed to look at task difficulty, and two tasks were used to examine prospective memory. The results were analysed in terms of gender, disease stage, age at onset, and depression.

CHAPTER 3

RECOGNITION, RECALL, AND PROSPECTIVE MEMORY DEFICITS

Introduction

Previous research suggests an association between task demands and memory impairment in PD, such that more difficult tasks produce greater deficits (e.g., Appollonio et al., 1994; Weingartner et al., 1984). Research has also demonstrated that PD participants have greater deficits on recall tasks relative to recognition memory tasks (e.g., Breen, 1993). In fact, the predominant opinion is that recognition is relatively normal in people with PD (e.g., Brown & Marsden, 1990). The apparent dissociation between free recall and recognition memory has been explained in terms of task demands, recall being the more demanding (e.g., Knight et al., 1988; Taylor et al., 1988). In addition, the dissociation has been used as evidence that the deficit lies with retrieval strategies rather than encoding (e.g., Beatty, 1992; Mahurin et al., 1993). However, procedural differences between recall and recognition tasks have made it hard to directly compare data from each task. To avoid this problem, a task could be designed with more than one level of difficulty. In this way, the effect of task demands could be studied without the type of task confounding the results. To date, no study has done this using PD participants.

Furthermore, although males and females generally differ in memory performance (Kramer et al., 1988), few studies have explicitly examined whether gender moderates memory dysfunction in PD. In addition, cognitive deficits may vary as a function of PD-related factors such as age at symptom onset, disease stage, and

depression (Beatty, 1992; Levin et al., 1992a; Starkstein et al., 1989). However, little is known about how these factors affect recognition memory or interact with task difficulty.

Prospective memory is memory for activities to be performed in the future (Einstein & McDaniel, 1990). It appears that no research to date has examined whether people with PD have poor prospective memory, despite the prevalence in everyday life of activities that require this type of memory.

Given that our knowledge about several important aspects of memory functioning in PD is limited, or based on indirect evidence, a major goal of the present study was to examine these issues directly. To look at the effect of task difficulty, a task was specifically designed with two levels of difficulty: hard and easy. The task used a 2AFC recognition memory paradigm with nonverbal stimuli (drawings of abstract objects). Recognition was also assessed with a verbal (common words) analogue of the nonverbal task, with the exception that there was no manipulation of task difficulty. Prospective memory was measured with two different event-based tasks: one involving a question and the other involving an object. The data were analysed separately for the men and for the women.

There are very few longitudinal studies in the PD literature (e.g., Piccirilli, D'Alessandro, Finali, & Piccinin, 1994); thus, little is known about the rate at which memory declines. In the present study, an assessment of short-term decline was made by testing each participant on two occasions, at least six months apart. The second testing session also serves as a conceptual replication (Hunter & Schmidt, 1990), and therefore, consistency of ESs across the two sessions will provide support for the existence of any effect. A second (but equally important) aim of the present study was to examine memory as a function of disease stage, depression, and age at symptom onset. To do this, the entire PD sample was divided into subgroups based on these variables and separate analyses were conducted for each subgroup.

Method

This section contains three main subsections. A materials and procedures subsection is given first to provide the reader with the rationale and procedure for using each task. The statistical analyses subsection contains a rationale for the statistical methods, as well as specific information about each main analysis. Finally, the participants subsection provides a detailed account of the demographic and neuropsychological characteristics of the sample, analysed as a function of gender.

Materials and Procedures

Most of the tasks were computer-administered. Computerised testing provides consistency and uniformity of stimulus presentation, and automates scoring which reduces the effects of experimental bias (Youngjohn, Larrabee, & Crook, 1992). A computer may also alter the dynamics of the testing situation in a way that provides a nonthreatening challenge to the individual, thereby encouraging participation (Baker, Letz, & Fidler, 1985). Furthermore, pilot work for the present study showed that, while many of the elderly participants were unfamiliar with computers, they were generally interested and enthusiastic about taking part in the study. The recognition memory tasks were specifically designed for computer use. A NEC Versa 2000 Laptop computer was used in conjunction with the enhanced Experimental RunTime System (ERTSVIPL) software (version 3.11; Beringer, 1995) to control stimulus presentation, visual feedback, and the experimental contingencies. The ERTSVIPL software uses a VESA-based Intelligent Preload technique for displaying 256-colour graphics on SVGA cards. Stimuli were presented in yellow on a blue background, using a NEC Versa 2000C active-matrix 24 cm colour screen controlled by a SVGA graphics card. The ERTSVIPL software operates with millisecond accuracy to control response registration and event timing. Participants' responses were registered using two

metallic keys measuring 40 mm x 70 mm, mounted 80 mm apart in the center of a response console measuring 300 mm x 264 mm. An external keyboard was connected to the laptop computer for use by the experimenter.

Other tasks (e.g., the SFT) were adapted for computer-administration, but were still scored manually. The tasks were categorised into three groups: (a) neurological assessment tasks, (b) neuropsychological tasks, and (c) experimental tasks. The rationale for inclusion of the tasks, validity information, and the procedures used to administer the tasks are presented in the following sections. But first, the general procedure is described.

General Procedure

The assessment procedures were conducted in each participant's own home. In most cases the equipment was set up on a dinning-room table and the positioning of the equipment was adjusted to suit the participant. Generally, the computer screen was centered 0.5 m away at eye-level when the participant was seated.

The administration of tasks was arranged so as to minimise fatigue and to avoid interference from tasks of similar nature. Each testing session was begun by asking participants to read both an information sheet (see Appendix E) and a brief overview of the tasks (see Appendix F). The information sheet gave a complete and accurate description of the goals, procedures, risks and benefits associated with the study. All procedures were approved by the Massey University Human Ethics Committee, the Manawatu-Whanganui Ethics Committee, the Hawkes Bay Ethics Committee, and the Wellington Ethics Committee. All participants were asked to complete a general consent form (see Appendix G) and informed of their right to withdraw from the study at any time. In addition, the PD participants were asked to complete a video consent form (see Appendix H).

Each participant was then taken through a series of automated tasks using both standardised verbal instructions and instructions presented on the computer screen. Practice trials were given where necessary to allow participants to become familiar with the test procedures. If a participant required clarification of the task instructions, or appeared initially to misunderstand them, standardised verbal prompts were provided. Participants completed each task (used to assess a specific function) in the order shown in Table 4. Note that for all participants the order of tasks was the same. The measure of cognitive status (dementia) was only used at the end of Time 2.

The interval between Time 1 and Time 2 was approximately six months. More specifically, the male PD group had a mean intersession interval of 188.92 days (SD = 7.70) and this was very similar to that of the male HC group (M = 188.52, SD = 9.13), $F < 1$. The female PD group's interval (M = 189.63, SD = 6.03) was

Table 4 *Order of Experimental Tasks and Associated Functions*

Task	Function
Prospective Memory Question Task	Event-based prospective memory
Visual Analogue Scale - Begin	Anxiety at the beginning of the session
Structured Interview	Demographic and disease data
Modified Activities of Daily Living Scale	Functional disability
Video Assessment	Hoehn and Yahr disease stage
Prospective Memory Object Task	Event-based prospective memory
Verbal Recognition Memory Task	Verbal recognition memory
Supermarket Fluency Test	Semantic verbal fluency
Nonverbal Recognition Memory Task	Nonverbal recognition memory
FAS Fluency Test	Letter fluency
Kendrick Object Learning Task	Picture recall
National Adult Reading Test	Premorbid IQ (WAIS-R); Premorbid letter fluency
Geriatric Depression Scale	Affective status
Mackworth Clock Test	Vigilance
Visual Analogue Scale - End	Anxiety at the end of the session
Short Orientation-Memory-Concentration Test	Cognitive status

also similar to that of the female HC group (M = 192.72, SD = 12.41), $F < 1$.

Every effort was made to keep the time of day constant between Time 1 and Time 2 for each participant. Practical constraints meant that this was not always possible. In the group of male PD participants, 92% were seen at the same time of the day compared to 83% of the male HC participants, χ^2 (1, N = 48) = .97, p = .33, d = .29 (-.29–.87). For the female participants, 69% of the PD participants in comparison to 61% of the HC participants were seen at the same time of the day, χ^2 (1, N = 34) = .22, p = .64, d = .16 (-.53–.86).

Neurological Assessment

Structured Interview

Rationale. The structured interview was used to record the following data from all participants: date of birth, gender, marital status, years of education, handedness, problems with colour vision and eye-sight, occupation, health, and medication. In addition, the following data were collected from the PD participants: year of diagnosis, person who made diagnosis, cause, first symptom, date of first symptom, laterality at onset, laterality of current symptoms, antiparkinsonian medication, medication effectiveness, and state of balance.

Procedure. Each participant was lead through the questions and their responses were entered directly into the computer.

Categories of Diagnosis

Rationale. Calne et al.'s (1992) categories of diagnosis were used to support the diagnosis of idiopathic PD. Calne et al. proposed three categories for idiopathic parkinsonism (IP):

1. Clinically possible IP. The presence of any one of the salient features: tremor, rigidity, or bradykinesia. Impairment of postural reflexes is not included because it is too nonspecific. The tremor must be of recent onset, but may be postural or resting.

2. Clinically probable IP. A combination of any two of the cardinal features: resting tremor, rigidity, bradykinesia, or impaired postural reflexes. Alternatively, asymmetrical resting tremor, asymmetrical rigidity, or asymmetrical bradyphrenia are sufficient.

3. Clinically definite IP. Any combination of three of the features: resting tremor, rigidity, bradykinesia, or impairment of postural reflexes. Alternatively sufficient are two of these features, with one of the first three displaying asymmetry. (p. 126)

Calne et al. (1992) did not include pharmacological criteria in their categories of diagnosis because dopaminergic therapy can produce a positive response in the parkinsonism plus syndromes. However, they did suggest that failure to respond to levodopa may be grounds for exclusion.

Procedure. Each participant was categorised according to Calne et al.'s (1992) criteria, based on observations made during the experimental session.

Hoehn and Yahr Disability Rating Scale

Rationale. The Hoehn and Yahr Disability Rating Scale (H&Y; Hoehn & Yahr, 1967) was used to rate the disease stage of each participant (see Appendix I for a copy of the scale). The H&Y was designed as a method to delineate disease severity in patients with PD, not as a linear description of progression (Hoehn, 1998). Although the H&Y provides a rather coarse assessment of disease stage, it has the advantage of being relatively easy to employ, and is widely utilised in the

PD research literature. The H&Y was developed before the side-effects of long-term drug therapy were known, and provides no assessment of motor fluctuations or adverse reactions to therapy. However, it is relatively insensitive to these fluctuations (Diamond & Markham, 1983). Nevertheless, in the present study an attempt was made to avoid problems caused by drug-induced side-effects by testing participants at their time of optimal therapeutic response. The H&Y has been shown to have "substantial" interobserver reliability (Geminiani et al., 1991).

Procedure. Each participant was asked two questions about their disease symptoms and then asked to perform three motor tasks that were recorded by video camera (see Appendix J). At a later date, the data from the questions and the video assessment were used to rate each participant according to H&Y. This was done both at Time 1 and at Time 2.

Activities of Daily Living Scale

Rationale. The modified Activities of Daily Living (mADL) scale (van Hilten et al., 1994) was used to provide a self-report measure of functional disability. The mADL is a 8-item version of the Activities of Daily Living section from the Unified Parkinson's Disease Rating Scale. Construction of the short version was made using a sample of 111 PD participants. Principal components analysis was used to extract 8 items (from the original 13 items) that more accurately target disability rather than the presence and severity of physical symptoms. The items are scored from 0 to 4 on a Likert-type scale. A score of 0 represents normal functioning and a score of 4 represents severe disability (see Appendix K). The mADL was found to have good internal consistency (Cronbach's alpha = .85) and concurrent validity (correlation of .66 with the H&Y stages) (van Hilten et al., 1994).

Procedure. Participants were asked to rate themselves on each item of the mADL. Ratings were then entered directly into the computer. The total score was used to provide a global disability rating.

Neuropsychological Tasks

Geriatric Depression Scale

Rationale. A computerised version of the Geriatric Depression Scale (GDS) was used to screen for depression. The GDS was specifically designed for rating depression in the aged (Yesavage et al., 1983). It is particularly well suited for use with Parkinsonians because it contains only one somatic item that may relate directly to the patient's physical disability, focusing instead on psychological aspects of depression. Moreover, in a sample of PD participants, the GDS was found to be a more sensitive measure of depression than the Beck Depression Inventory (Youngjohn, Beck, Jogerst, & Caine, 1992). The GDS has a reported Cronbach's alpha of .94 and a test-retest reliability of .85 (over one week), and is highly associated with conceptually related scales (Yesavage et al., 1983). Yesavage et al. (1983) also reported a sensitivity rate of 0.84 and a specificity of 0.95. Furthermore, using either current Research Diagnostic Criteria or the then current DSM-III diagnosis as the criterion, the GDS (with a cut-off score of 10) accurately identified 77% of a geriatric medical outpatient sample (Norris, Gallagher, Wilson, & Winograd, 1987). Sensitivity was 0.89 and specificity was 0.73. The GDS uses a self-report, yes-no format and contains 30 items. A third of the items indicate depression when answered negatively whereas the remaining items indicate depression when answered positively. Appendix L contains a copy of the GDS, including the scoring information.

Procedure. In the present study, each item was presented one at a time in the center of the computer screen. Participants indicated their answer to each question by pressing the left response key for "yes" and the right response key for "no." A cut-off score of 10 was used to detect those participants possibly suffering from depression.

National Adult Reading Test - Premorbid IQ

Rationale. The National Adult Reading Test (NART) was used to assess premorbid IQ. It was used primarily to check that the PD and HC groups did not differ in terms of premorbid IQ. The NART is a measure of a person's ability to pronounce 50 "irregular" words (see Appendix M). Each "irregular" word can only be read correctly if the participant knows of the word and recognises its written form (Nelson & O'Connell, 1978). Considerable evidence suggests that word-reading ability is preserved in people with mild cognitive impairment and that it serves as a useful predictor of premorbid intelligence (Christensen, Hadzi-Pavlovic, & Jacomb, 1991; Nelson & O'Connell, 1978; Ryan & Paolo, 1992). Nelson and O'Connell standardised the NART on 120 participants aged 20-70 years. Data from this sample were used to form regression equations that can be used to predict WAIS Verbal, Performance, and Full-scale IQ from NART errors. Subsequently, Ryan & Paolo built new regression equations for the prediction of WAIS-R IQ. The new equations and their SEMs are presented below:

Estimated VIQ = 132.3893 + (NART errors) (-1.164) *SE* = 7.70
Estimated PIQ = 123.0684 + (NART errors) (-0.823) *SE* = 12.08
Estimated FSIQ = 131.3845 + (NART errors) (-1.124) *SE* = 8.83.

Procedure. Each participant was presented with each word one at a time on the computer screen and asked to say each one aloud without time constraint. All responses were recorded onto audio tape. Later, the audio tapes were reviewed and the number of pronunciation errors were recorded for each participant. Correct pronunciations were based on the Heinemann New Zealand Dictionary (Orsman, 1989) and The New Zealand Contemporary Dictionary (Foreman, 1972).

National Adult Reading Test - Premorbid Verbal Fluency

Rationale. Premorbid verbal fluency (an estimate of FAS fluency) was calculated from a NART-based equation (Crawford, Moore, & Cameron, 1992). The regression equation and its *SE* are:

$$\text{Premorbid FAS} = 57.5 - (0.76 \times \text{NART errors}) \qquad SE = 9.09.$$

Crawford et al. built the regression equation using data from a healthy sample of 142 participants aged 16-88 years. The equation was validated using a sample of 38 neurologically impaired participants.

Orientation-Memory-Concentration Test

Rationale. The Orientation-Memory-Concentration Test (OMCT; Katzman et al., 1983) was used to provide a dementia rating. The OMCT is a 6-item version of the Information-Memory-Concentration (IMC) test created by Blessed, Tomlinson, & Roth (1968). Katzman et al. constructed the short version using a sample of 322 participants and then validated it using three additional samples. The OMCT was found to have good test-retest reliability (over three weeks) and be highly predictive of scores on the IMC[8]. Furthermore, scores on the short version were shown to correlate with plaque counts in the temporal, parietal, and frontal cortex of 38 autopsied participants (r = .54) (Katzman et al., 1983).

The OMCT includes three orientation questions (year, month, and time), a name and address memory phrase, counting from 20 to 1, and saying the months backwards (see Appendix N). The OMCT is particularly good for use in people with PD as there is no copying task that may be affected by the physical symptoms of PD independently of dementia.

[8] The IMC has excellent test-retest reliability and is highly correlated with the Mini-Mental State Examination and Dementia Rating Scale (Salmon, Thal, Butters, & Heindel, 1990).

Procedure. Each participant's scores were entered directly into the computer; one point for each incorrect response. The computer was programmed to produce a total weighted error score by multiplying each item by a weighting factor (Katzman et al., 1983). As suggested by Katzman et al., the data from three participants were excluded from the statistical analyses because their total weighted error score was greater than 10.

Verbal Fluency Tasks

Rationale. Past research studying the effects of PD on verbal fluency has indicated that letter (phonemic) and semantic (category) fluency may be differentially effected in PD (Azuma et al., 1997). Furthermore, fluency has been described as a multifactorial task in which the number of correct words, perseverations, and intrusions may each capture different aspects of performance deficits (Troyer, Moscovitch, & Winocur, 1997). Additionally, multiple brain regions appear to be involved in the performance of fluency tasks. Evidence suggests that letter fluency is associated with frontal lobe impairment whereas semantic fluency is related to temporal deficits (Troyer et al., 1997). Therefore, both letter and semantic fluency were assessed in the present study.

Procedure. Letter fluency was examined with the Controlled Oral Word Association Test, commonly called FAS (Lezak, 1995). Participants were asked to generate as many words as possible beginning with the letters F, A, and S. The duration of each letter trial was 60 s. Participants were asked not to use proper names, or the same word with different suffixes. Semantic fluency was assessed with the Supermarket Fluency Test (SFT). Participants were required to name as many examples of supermarket items as they could in 60 s. All participant responses were audiotaped and later transcribed for scoring purposes.

The data were scored by assigning one point for each correct exemplar. In addition, perseverations (the number of times a word was repeated) and intrusion

errors (addition of items not related to category) were recorded. In the FAS, proper names and words with the same suffix were not counted, nor were words that began with the same phonetic sound but not the same letter (e.g., *photo* for words beginning with F). In the SFT, words that were not considered supermarket items were not counted.

Visual Analogue Scale

Rationale. A computerised, self-report version of the Visual Analogue Scale (VAS) was used to measure anxiety both at the beginning and at the end of each testing session (see Appendix O). The VAS has been shown to correlate well with questionnaire measurement of anxiety (Henderson, 1988, cited in Cockburn & Smith, 1994). McCormack, Horne, and Sheather (1988) reviewed validation studies of the VAS as a measure of anxiety. Their review indicated reasonable convergent validity with the STAI-State Anxiety Scale in a sample of medical students and in a sample of psychiatric patients. Furthermore, the VAS has several advantages (e.g., it can be administered repeatedly and it is easy to administer and score). However, McCormack et al. (1988) raised concerns over the choice of scoring interval and whether scores can be treated as interval data. In the present study these concerns were alleviated by conducting both parametric and nonparametric analyses.

Procedure. Following the recommendations of McCormack et al. (1988), the VAS was presented in a horizontal format with no defining intermediate points. However, to facilitate data collection the VAS was presented on the computer screen. Participants were instructed to move a small vertical bar on a 10 cm horizontal line to indicate how anxious they felt at that point in time. The left-most end indicated "not at all anxious," whereas the right-most end indicated "very anxious." The position of the bar was adjusted by using the left and right arrow keys on the keyboard. Participant responses were scored to the nearest millimetre, producing a 100-point scale.

Mackworth Clock Test

Rationale. A computerised version of the Mackworth Clock (MC) test was used to measure vigilance (sustained attention). First devised by Mackworth (1948), the original test involved a clock-type device that had a large pointer that moved on in steps in front of a white surface devoid of scale markings or reference points. The task is ideally suited for use in the present study as it was designed to keep motor, memory, and problem-solving abilities to a minimum (Sheer & Schrock, 1986). A deficit in vigilance should be reflected in a deficit in perceptual sensitivity, rather than a change in response criterion (Sheer & Schrock, 1986).

Procedure. The computerised version consisted of a circle of 60 points presented on a blue background. The points were dark yellow in colour, and the circle they formed was 9 cm in diameter. During the task each point consecutively turned white appearing to "light up," moving in a counter-clockwise direction. Each point turned white for 300 ms with an interstimulus interval of 2200 ms. During the 6.5 min test period, at irregular intervals, a single point would not turn white. Instead, the point following it would turn white creating the appearance of a "double jump." These double jumps, 15 in number, were the targets. The screen was positioned approximately 0.5 m in front of the participant, but this distance was reduced for those experiencing difficulty focusing on the points.

At the beginning of the instruction period, participants were told that they would be working on a "watch-keeping" task that was demanding and would require their full attention. On-screen instructions then informed participants of what they were about to see and how to respond. Responses were made by pressing the left response key with the preferred hand. Participants were then given a practice period with 3 to 4 double jumps. During the practice period, participants were prompted, whenever necessary, to make sure that they understood the task.

Following practice, participants were told that the actual task would take 6.5 min. They were then reminded that the task was demanding and would require their full

attention. They were also asked to "respond as fast as possible, but even if late or if they have any doubt, they should still respond." It was pointed out that false presses would not affect their score. Participants were then given an opportunity to ask questions before being left to complete the task without interruption.

The position of the 15 double jumps was determined randomly, with the constraint that one double jump occurred within every set of 10 single jumps. Immediately following the test, participants were asked several questions concerning the strategy they had used to perform the task. The questions (based on the categories used by Giambra, Quilter, Phillips, & Hiscock, 1988) were as follows: (a) "Did you count the single jumps and try and work out a pattern to help you detect the double jumps?", (b) "Did you count the single jumps without trying to work out a pattern?", (c) "Did you watch the clock and move your eyes ahead to the next point after each single jump?" The participants' answers were used to rate their strategy as being based on either internal or external resources.

Performance on the MC was assessed with the nonparametric measures of sensitivity, A', and response bias, B'' (MacMillan & Creelman, 1991). These measures are suitable in situations where it is impossible to conduct the large number of signal and noise trials necessary to compute SDT statistics (McNicol, 1972). The measure, A', is an estimate of the proportion of the total area which lies beneath the Receiver-operating Characteristic, or ROC, curve (McNicol, 1972). The area under the ROC curve can be estimated with only a single pair of hit (the conditional probability of responding yes to a target) and false alarm (the conditional probability of responding yes to a nontarget) rates, calculated by:

$$A' = .5 + (H - F)(1 + H - F) / [4H(1 - F)] \quad \text{if } H > F, \text{ and}$$

$$A' = .5 + (F - H)(1 + F - H) / [4F(1 - H)] \quad \text{if } H < F,$$

where H = hit rate and F = false alarm rate (MacMillan & Creelman, 1991). Similarly, response bias was calculated by:

$$B'' = [H(1 - H) - F(1 - F)] / [H(1 - H) + F(1 - F)] \quad \text{if } H \geq F, \text{ and}$$

$$B'' = [F(1 - F) - H(1 - H)] / [H(1 - H) + F(1 - F)] \quad \text{if } H < F.$$

Experimental Tasks

Verbal Recognition Memory Task

Rationale. The Verbal Recognition Memory Task (VRMT) was created specifically for the present study and provides a measure of verbal recognition relatively free from response bias. A pool of 80 words was created for the VRMT from words generated by Lesch and Pollatsek (1993) and Francis and Kucera (1982). Half of the words served as targets in the memory test, and the remaining half were used as distractors. All stimuli were nouns of 3-6 letters ($M = 4.35$, $SD = .74$) which ranged in frequency from 1 times per million words to 201 times per million in Kucera and Francis' (1982) word count, with a median frequency of 13.5 times per million. Table P1 (Appendix P) contains a list of all words used in the VRMT, including the frequency and length of each word.

Procedure. The recognition task consisted of two phases: study and test. Initially, participants were shown a single study list of 40 target words. Each stimulus was presented one at a time on a computer screen at a 3-s rate, and participants were required to read each word silently. Stimuli were typed using Helvetica 60 pt. lower case lettering. Participants were given as much time as necessary to read standardised instructions displayed on the screen before beginning the study phase. Participants were instructed to press the space bar to begin the study phase.

The second (test) phase consisted of a 2AFC recognition test. In this phase, target words were paired with distractors of equal word length. There was little difference in word frequency between distractors ($M = 35.38$, $SD = 39.93$) and

their respective target words (M = 28.63, SD = 44.49), t (38) = .71. The order that the stimuli were presented in both the study phase and the test phase was randomised for each participant.

Participants began the test phase by pressing the right response key. A trial consisted of the following sequence. A clear screen was presented for 1 s, then two alternative answers appeared side by side in the centre of the screen. The stimulus display remained on the screen for a maximum of 6 s, or until a response was made. Participants were required to press the left response key if they thought the word on the left was a target word, or the right response key if they thought the word on the right was the target word. They were instructed to guess if neither word appeared familiar or if both appeared familiar, with the proviso that they try not to favour one response key over the other. Instructions displayed on the screen emphasised accuracy over speed. The left-right locations of the targets were randomised with the restraint that 50% occurred on each side.

After a response was made, the screen was cleared for 500 ms and then instructions on the screen asked the participant how confident they were that their response was correct. Four levels of confidence (very confident, confident, somewhat confident, or not at all confident) were displayed on the screen, and participants were instructed to say their response aloud. Ratings were entered directly into the computer via the keyboard. The confidence rating interval was a maximum of 6 s, or terminated by a response. This procedure was repeated with all 40 target words with a 500-ms intertrial interval. The complete sequence of events is presented in Figure 4, and the temporal sequence of events for one trial (in the test phase) is presented in Figure 5.

Accuracy on the VRMT was expressed as the percentage correct (PC). In addition, the frequency of confidence ratings was expressed as the proportion of the total trials. The recognition task began with a practice study list, consisting of four target numbers followed by a recognition test using the four target numbers paired with four unrelated numbers (distractors).

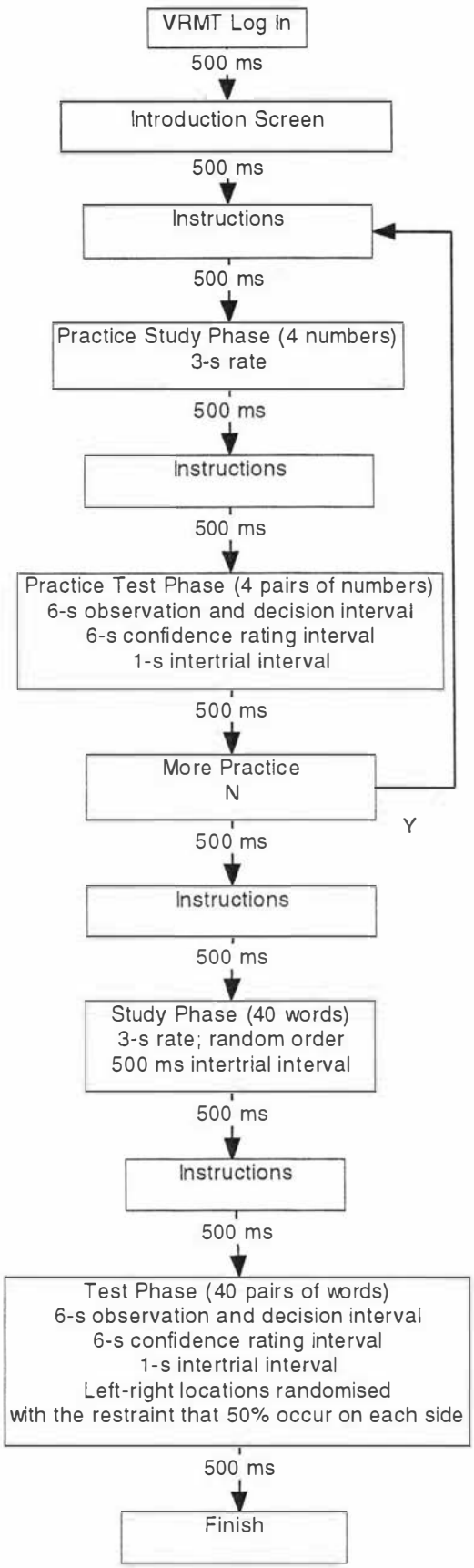


Figure 4. Flow chart of the Verbal Recognition Memory Task (VRMT) procedure.

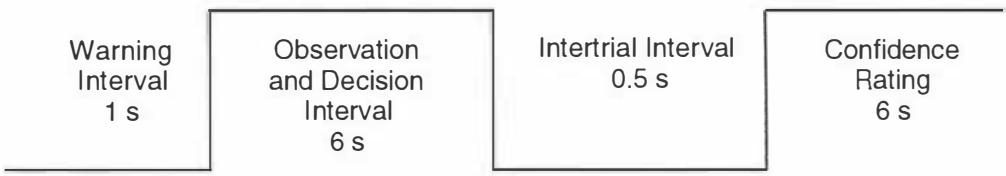


Figure 5. Temporal sequence of events for one trial in the test phase of the Verbal Recognition Memory Task.

Nonverbal Recognition Memory Task

Rationale. The Nonverbal Recognition Memory Task (NRMT) was created specifically for the present investigation and provides a measure of nonverbal recognition memory relatively free from response bias. A pool of 80 abstract drawings was created for the task by the author and from the complex figures used in the Embedded Figures Test[9] (Witkin, Oltman, Raskin, & Karp, 1971); half of the drawings served as targets in the NRMT, and the remaining half were used as distractors. To create two levels of difficulty within the one task, half of the targets were paired with similar distractors whereas the other half were paired with less similar ("dissimilar") distractors (see Appendix Q for examples of targets and similar/dissimilar distractors). Similarity was based on form and colour. Specifically, dissimilar distractors differed in terms of structural features and colour, whereas similar distractors differed only in terms of form. Past research has demonstrated that recognition memory is poorer for targets paired with structurally similar distractors than for targets paired with structurally dissimilar distractors (Bower & Glass, 1976; Wickelgren, 1977). Furthermore, Bower and Glass demonstrated that this same result occurs both for perceptual and for verbal encoding of the stimulus material. Nevertheless, in the present study an effort was made to select stimuli that were difficult to encode verbally. Although no attempt was made to explicitly examine the modality of encoding, post-test feedback indicated that most participants found it difficult to assign verbal labels

[9] Modification and reproduction of stimuli was by special permission of the Publisher, Consulting Psychologists Press, Palo Alto, CA 94303 from Embedded Figures Test by Herman A. Witkin. Copyright 1971 by Consulting Psychologists Press, Inc. All rights reserved. Further reproduction is prohibited without the Publisher's written consent.

to the stimuli. Therefore, it is likely that they used perceptual encoding. Stimulus order was randomised for each participant in both phases of the task.

A pilot study was conducted to check the procedure and make sure that the two levels of difficulty produced scores as predicted. Three people with PD (two males and one female) and five normal elderly (one male and four females) participated after providing informed consent. Percentage of correct responses on the NRMT were computed for the similar target-distractor pairs (hard level of difficulty) and for dissimilar pairs (easy level of difficulty). As can be seen in Table 5, mean PC was lower for the hard level of difficulty relative to the easy level for both groups. The calculated ESs (*d*) can be described as moderate in size for the PD group and large for the HC group.

Table 5 *Means and SDs for Percentage Correct on the Nonverbal Recognition Memory Task for the Parkinson's Disease (PD) and Healthy Control (HC) Participants who Participated in the Pilot Study*

	Hard			Easy				
	M	*SD*	*Range*	*M*	*SD*	*Range*	*d*[a]	95% CI[b]
PD[c]	55.00	15.00	40–70	63.33	10.41	55–75	.65	-1.47 –2.76
HC[d]	51.00	11.05	35–65	65.00	10.00	55–75	1.33	-.22 –2.88

Note. Hard = hard level of difficulty; Easy = easy level of difficulty.
[a]Cohen's (1988) effect size statistic, *d*. [b]95% confidence interval for *d*. [c]*n* = 3. [d]*n* = 5.

Procedure. Administration of the NRMT followed exactly the same procedure as that described above for the VRMT. The only exception was that no practice trials were given because the NRMT always followed the VRMT and pilot work indicated that participants did not need to practice again. In addition to PC, the frequency of confidence ratings were computed for each level of difficulty.

Kendrick Object Learning Task

Rationale. The Kendrick Object Learning Task (KOLT) was originally designed to be administered together with the Kendrick Digit Copying Test (KDCT) to screen for dementia in the elderly (Kendrick, 1985). The KOLT is a test of recall of everyday objects whereas the KDCT is a test of speed performance. Kendrick stated that to screen for dementia "it is essential that the two tests be used in conjunction with each other and if possible with a retesting six weeks later" (p. 8). An alternative form of the KOLT is available for the purpose of retesting. The two forms have been shown to correlate highly (r = .91). Both test-retest reliability (over 24 hours) and internal consistency were also shown to be high (r's = .92).

In the present study, the data from the KOLT were not used to screen for dementia, although originally included for this purpose, because of the PD participants' limitation in performing tests that contain a motor component (i.e., the KDCT). Furthermore, it was not possible to retest after six weeks. However, rather than discard the data from the KOLT, they were used as a measure of recall. Kendrick (1985) suggested that it could be used for this purpose, and it seems particularly suitable in the present study because it contains two equivalent forms and is relatively easy and quick to administer. Furthermore, raw scores could be converted to age-scaled quotients, thus avoiding the potential confounding effect of age (Kendrick, 1985).

Procedure. Administration and scoring of the KOLT was conducted in the manner outlined by Kendrick (1985). Briefly, four large cards with pictures of common objects were shown to participants. Each card has 10, 15, 20, or 25 objects and the time allowed for inspection of each card is based upon a 3-s viewing time per object. After viewing each card, participants were asked to recall as many objects as possible, in any order. Scoring was done by giving one point for each correct response or acceptable alternative. This raw score was then converted to an age-scaled quotient using tables provided by Kendrick.

Prospective Memory Tasks

Rationale. Prospective memory tasks require the participant to execute a planned action at some point in the future and can be event-based or time-based (Einstein & McDaniel, 1990). Despite the pervasiveness of this type of memory in daily activities, no studies have examined it for deficits in PD sufferers. Therefore, two prospective memory tasks were used in the present study. The tasks were adapted from Huppert and Beardsall (1993) who demonstrated that these tasks are sensitive to the mildest form of cognitive impairment in dementia.

Procedure. The measures of prospective memory were: Prospective Memory for a Question (PMQT), and Prospective Memory for an Object (PMOT). Both tasks were event-based; the first involved a short time period, whereas the second involved a longer time period.

The PMQT involved remembering to ask two questions when an alarm sounded. Instructions were presented on the computer screen and participants were given as much time as they needed to read them. At Time 1, participants were told that an alarm would sound in 5 min, at which point they should stop what they are doing and ask the experimenter two questions concerning their next testing session. The two questions were, "When will the next session be?" and "What will the procedure involve?" At Time 2 the same procedure was used except the two questions were related to the results of the research. That is, "When will a summary of the results be sent to me?" and "What will the summary include?" If the participant failed to spontaneously ask either question within 15 s, they were prompted with the question, "What were you supposed to do when the alarm sounded?" Participants were scored one point for each question recalled and additional points for each response that was made without a prompt. The maximum score was four.

The PMOT involved remembering to ask for an object when the testing session concluded. At approximately one-third of the way through the testing session, participants were asked for a small personal belonging. It was explained that their

task was to remember to ask for it back when the testing session finished, and to also remember where the object was put during the session. During the remainder of the session the object was stored in one of the equipment boxes. Participants were reassured that their belonging would be returned if they did not remember to ask for it back. At the end of the testing session, it was made clear that the session had finished and the equipment was packed away. If the participant did not ask for their belonging during this time, they were prompted with, "Was there something you were going to ask for now that we have finished?" Participants were scored one point for remembering to ask for the object, and another point if they recalled the location of the hidden object. Additional points were given for each response that was made without a prompt. The maximum score was 4.

Statistical Analyses Procedure

The data were analysed using the Statistical Package for the Social Sciences (SPSS; SPSS Inc., 1995). The main analyses were conducted using one-way analysis of variance (ANOVA) with planned comparisons. The specific planned comparisons, or contrasts, used in the present study are outlined below. However, it should be noted that although the SPSS procedure produces an omnibus *F* test, only the results of the contrasts were interpreted. Both the omnibus *F* test and the contrasts are based on the same statistical theory, and there is no reason why the overall *F* test should have priority (Lindman, 1992). In other analyses, differences between data sets were evaluated with multivariate analysis of variance (MANOVA), univariate mixed ANOVA, repeated measures ANOVA, and/or *t*-tests for two independent samples (two-tailed, unless otherwise stated).

Although the data were analysed with traditional tests of significance, the results were interpreted by examining point estimates of ES and the 95% confidence interval around these point estimates. Some researchers (e.g., Carver, 1993; Schmidt, 1996) have argued that we should abandon the statistical significance test in favour of point estimates of ES when interpreting the data in individual studies (see chapter 2). Tests of significance provide no indication of the

importance of an effect and are problematic under many circumstances, for example, in situations of low statistical power[10]. In addition, statistical significance tests should only be used if the assumption of randomisation (random sampling and/or assignment) is met (Shaver, 1993). Shaver argued that "without randomness, the result of the test of statistical significance is meaningless or, at best, its relevance to a statement of probability is indeterminate" (p. 299). Poole (1987) went so far as to state that tests of significance are "worse than useless" (p. 196), even "misleading." Furthermore, significant tests do not hold the overall error rate to the desired level. However, point estimates of ES and confidence intervals, when interpreted correctly, do not suffer from the problems associated with significance testing (Borenstein et al., 1997; Schmidt, 1996). Confidence intervals have been used in the past as tests of significance by looking for null values within the confidence interval (Poole, 1987). For example, a 95% confidence interval that does not contain zero indicates a statistically significant effect at $p < .05$. However, this practice suffers from exactly the same problems as any other statistical significance testing procedure, and therefore, is not used here.

In the present study, where an ES for the between-groups comparison was equal to, or greater than, Cohen's (1988) definition of a small effect, it was considered potentially important. Cohen argued that scientifically important effects may be of this modest order of magnitude. Indeed, in the present investigation, the meta-analysis strongly suggests that recognition memory deficits may be quite small; but this does not mean they are unimportant. Furthermore, observed ESs are usually attenuated by irrelevant sources of variance and so are likely to be smaller than what they would be if measurements could be made without error (Cohen, 1988; Hunter & Schmidt, 1990). For example, Kim and Hunter (1993) found that after correction for measurement error and dichotomization of

[10] In the present study, small effect sizes were predicted for the recognition memory tasks based on the results of the meta-analysis. Given the number of participants used, there would have been just 15% power to detect a small recognition memory deficit using a two-tailed t-test with an alpha level of .05. Under these conditions the sample size would have had to be increased to nearly 800 to provide satisfactory power. A sample of this size would have been impossible. Thus, the low level of statistical power available in the present study is good reason to interpret the results with respect to the magnitude of the effect rather than statistical significance.

variables, the mean ES for the relationship between attitudes and behaviour increased from .47 to .79. For each within-groups comparison in the present study, the correlation between Time 1 and Time 2 (test-retest reliability coefficient) was used to correct for attenuation due to between-subjects variation (see Hunter & Schmidt, 1990). The magnitude of the effect was quantified in terms of Cohen's ES statistic, *d*. Cohen offered the following conventional definitions of ES: small: $d = .20$, medium: $d = .50$, large: $d = .80$. Cohen made it clear that although he suggested conventions for ES, researchers should interpret the meaning of an ES in light of the context in which it is embedded. Here, a small effect was taken as $\geq .20$ and $< .50$; a medium effect as $\geq .50$ and $< .80$; a large effect as $\geq .80$.

For each *F* test, estimates of ES were calculated from sample means or η^2 (eta squared: used by SPSS for Windows) provided that $F > 1$. (See Appendix D for the formulas used in the computation of *d*.) Traditional *p* values are still reported alongside the ES statistics to allow for a comparison with the usual method of interpreting the results. Where the results of an analysis indicated $F \leq 1$, this was reported in the text without an ES estimate. However, all ESs that could be calculated from sample statistics were reported, regardless of the associated *F* value in ES summary tables.

Prior to the statistical analyses, the data set was checked for missing values. Few cases with missing values were found with the exception of the MC test (24% missing) and the NART measured at Time 2 (over 50% missing). Missing data on these tasks were mainly due to technical problems and rarely due to participant-related factors. Because complete values of the NART measured at Time 1 were available, data from Time 2 were dropped from the analysis. The MC data were still analysed but cases with missing values were dropped. The few missing values from other tasks were replaced with the appropriate group mean (Tabachnick & Fidell, 1989).

The statistical assumptions of the tests used in the present study were checked prior to the main analyses. SPSS for Windows was used to validate the

assumptions where appropriate. In particular, Z scores were calculated for each dependent variable (DV) at each level of the independent variable (IV), to check for univariate outliers. Where a participant's score was more than three standard deviations from the mean, it was checked for accuracy. Where the score was accurate, it was considered valid and left in the analysis. The *F* max test was used to assess the assumption of homogeneity of variance (Coakes & Steed, 1996). An *F* max value was calculated by dividing the largest variance by the smallest variance; values of less than three were taken to indicate that the assumption was met.

Recognition memory data. The percentage of correct responses at each level of difficulty of the NRMT at Time 1 and at Time 2 were analysed separately with planned contrasts. The first contrast was used to compare the male PD group with the male HC group. The second compared the female PD group with the female HC group. To limit the number of contrasts, no direct comparisons were made between the PD subgroups. Repeated measures ANOVA was used to confirm that the manipulation of task difficulty was effective (i.e., that each group had lower mean PC scores at the more difficult level of the task relative to the easier level). The VRMT data were analysed in the same way the NRMT data were analysed. To examine whether memory declined over the 6-month period, a comparison was made between the Time 1 and the Time 2 ESs. A substantially larger ES at the second session was taken to indicate memory decline in the PD group.

Recognition memory confidence ratings data. The confidence data were scored as the number of times each of the four confidence ratings were used, expressed as a proportion of the total trials. Preliminary analysis indicated that the confidence data contained many missing values. Missing data are problematic in this case because they may distort the proportion of confidence ratings. Therefore, the data were first analysed by comparing those participants with missing data to those participants without missing data. Also, to assess whether the number of participants with missing data was the same between PD and HC groups, the χ^2 test was used. Males and females were assessed separately. The preliminary analysis also indicated that the data did not meet the assumptions necessary for

parametric analysis. Therefore, the nonparametric Mann-Whitney U test was used. Comparisons were made between PD and HC groups as a function of gender and time.

Recall and prospective memory data. The recall data (KOLT age-scaled quotients) and the prospective memory data (PMQT and PMOT) were analysed in the same way the VRMT data were analysed.

Disease stage data. To examine the potential moderating effect of disease stage, the participants were classified into two groups based on a cut-off of H&Y II, and planned comparisons were conducted. The groups were comprised of participants in the early stages (early PD) (H&Y rating of I or II) and those in the more advanced stages (advanced PD) (H&Y rating of III or IV). Because participants were rated on the H&Y at each session, separate analyses were conducted for Time 1 and for Time 2. The contrasts were made between the early PD group and the HC group, and between the advanced PD group and the HC group.

Depression data. To examine the potential moderating effect of depression, the participants were divided into three groups based on a cut-off score of 10 on the GDS, and planned comparisons were conducted. The groups consisted of PD participants who met the criteria for possible depression, PD participants who did not show signs of depression, and HC participants who did not show signs of depression. Those participants in the control group who showed signs of depression were not used in this analysis. Because different numbers of participants met the criteria for depression at each session, separate analyses were conducted. Planned contrasts were made between the depressed PD group and the nondepressed HC group, and between the nondepressed PD group and the nondepressed HC group.

Age at onset data. To find out whether the age at symptom onset moderated any memory impairment found in PD, the participants were classified into two groups based on a cut-off of 60 years of age at onset, and planned comparisons were conducted. There were 18 participants (7 females, 11 males) with symptom onset

before 60 years of age (early-onset PD), and 23 participants (9 females, 14 males) with onset at or after 60 years of age (late-onset PD). Each subgroup was matched with an equal number of controls on the basis of gender, age, education, and premorbid IQ. Planned contrasts were used to compare each PD subgroup with their respective control group.

Correlation Analysis. To further evaluate the relationship between memory and the potential moderators described above, correlations were calculated between the memory task scores and H&Y stage, age at onset, and GDS score. Correlations were determined with the Pearson product-moment correlation coefficient, *r*. Again, Cohen's (1988) conventions were used to interpret the size of the correlation coefficients: small: $r = .10$, medium: $r = .30$, large: $r = .50$.

In addition, correlations were computed between Time 1 and Time 2 for each of the memory tasks as a function of group and gender, so as to provide an indication of test-retest reliability.

Principal components analysis of memory tasks. Because the number of DVs used to measure memory was relatively large, it was deemed appropriate to conduct an empirical summary of the data set. Principal components analysis (PCA) is a statistical technique that can be applied to a set of variables in order to produce a smaller set of orthogonal components (Tabachnick & Fidell, 1989). This approach "produces a more reliable assessment of the conceptual constructs and increases power both though increased reliability and through preservation of degrees of freedom" (Bondi et al., 1993, p. 100). In the current analysis, PCA was used with Varimax rotation. The following variables for each session separately were entered into the analysis: VRMT, NRMT, KOLTQ, PMQT, and PMOT. Principal components analysis has an additional benefit in that it can be used to produce factor scores for each participant, and these scores are usually more reliable than scores on individual observed variables (Tabachnick & Fidell, 1989). Differences between data sets were examined with MANOVA and ANOVA. All assumptions and practical considerations underlying the application of PCA, MANOVA, and ANOVA were confirmed prior to the analyses.

Two different sets of analyses were conducted. First, MANOVA was used to look at whether gender moderated differences between the PD and HC groups. Each component served as the DV, and group (PD, HC) and gender (males, females) served as the IVs. Further interpretation of any multivariate effects was made with univariate tests. Second, MANOVA was used to look at whether disease stage moderated differences between each PD subgroup and the HC group. Each component served as the DV, and group (advanced PD, HC, or early PD, HC) served as the IV. Again, further interpretation of any multivariate effects was made with univariate tests. Because disease stage was rated at each session, there was the potential for a participant to be classified as advanced PD at one time but early PD at the other time. Therefore, for the purpose of this analysis, participants classified as advanced at either time formed the advanced PD group, while the remaining participants formed the early PD group.

Participants

Fifty eight volunteers with idiopathic PD were recruited for the present study. Volunteers were sourced through five Parkinsonism Societies in the lower North Island of New Zealand. Prior to taking part in the study, each participant was screened by telephone to ensure that they met the following criteria: (a) no other medical conditions with known central nervous system complications; (b) no hereditary diseases (e.g., Wilson's disease, Huntington's disease); (c) never suffered from alcohol abuse; (d) never suffered from a neurological impairment (e.g., Alzheimer's disease, stroke, brain tumour, head trauma with loss of consciousness greater than 1hr); (e) never had neurosurgery; and (f) no premorbid history of significant emotional disturbance before the onset of PD. (See Appendix R for a copy of the screening questionnaire.) Those participants that reported a diagnosis of clinical depression subsequent to the onset of PD were not excluded because the role of depression was to be examined in the research programme. In addition, participants who were colour blind, had poor eye-sight (unless corrected artificially), or had difficulty following the task instructions were excluded so as to prevent these factors confounding

performance on the memory tasks. Finally, participants who reported a poor response to antiparkinsonian medication were excluded on the grounds that they may not have idiopathic PD (Bakheit, 1995; Calne et al., 1992). Seven (12%) participants were excluded on the basis of the above criteria. A further seven participants were excluded from the eventual experimental sample because they could not take part in the second test session. Three more (5%) were also excluded because they showed signs of dementia.

Sixteen female and 25 male PD Participants remained in the experimental sample. This corresponds to a male-to-female ratio of 1:1.56, similar to that reported by others in the PD literature (Diamond, Markham, Hoehn, McDowell, & Muenter, 1990). The healthy control (HC) group consisted of 41 participants (23 males, 18 females), spouses of the PD participants and other volunteers from the community selected to match as closely as possible the PD group in terms of age and education (see Table 6). Exclusion criteria for the controls were the same as for the PD participants.

Planned comparisons were conducted to examine whether there were important demographic and psychometric differences between the PD and HC groups. As can be seen in Table 6, there was little evidence of important differences in mean age between the male PD and the male HC groups ($F < 1$), or between the female PD and female HC groups, $F(1, 78) = 1.95$, $p = .17$, $d = .48$ (-.22–1.19). The analysis also indicated that the PD and the HC groups were equivalent in terms of years of education ($Fs < 1$), and in males, the two groups were equivalent with regard to premorbid IQ ($F < 1$). The female PD group had a mean IQ 5 points lower than that of the female HC group, $F(1, 78) = 1.81$, $p = .18$, $d = .54$ (-.16–1.25). Finally, there was no evidence that the ratio of males and females in the PD group was different to that of the HC group, χ^2 (1, $N = 82$) = .20, $p = .65$, $d = .10$ (-.34–.54).

Table 6 *Means and SDs for the Demographic and Neuropsychological Characteristics of the Parkinson's Disease (PD) Group and the Healthy Control (HC) Group as a Function of Gender and Time*

	Male PD		Male HC		Female PD		Female HC	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (years)	70.48	5.93	69.13	5.43	67.94	6.16	70.67	5.18
Education (years)	3.92	3.00	4.22	3.27	4.06	2.82	4.50	2.75
Premorbid IQ	109.62	11.19	108.76	9.88	109.00	9.47	114.00	6.91
OMCT	4.40	3.42	2.43	2.09	3.19	3.17	2.33	2.77
GDS								
Time 1	10.72	5.86	5.35	4.37	8.63	5.41	4.17	3.62
Time 2	10.64	5.22	4.48	3.81	9.69	5.88	3.78	4.52
SFT								
Time 1	22.20	6.05	26.70	7.12	26.94	6.38	32.61	6.62
Time 2	23.21	6.38	28.54	4.44	28.38	6.03	32.97	5.26
FAS								
Time 1	36.98	13.73	40.06	11.97	38.38	12.94	45.44	12.76
Time 2	37.92	14.18	43.00	11.48	37.27	11.89	46.82	14.19
VAS - Beg								
Time 1	24.48	23.60	20.91	17.15	24.50	24.88	12.22	14.58
Time 2	27.68	22.20	19.78	19.30	24.13	21.68	13.33	16.33
VAS - End								
Time 1	22.94	24.41	19.91	22.73	14.00	19.49	11.44	19.97
Time 2	19.84	20.76	14.30	16.41	21.25	25.71	13.50	21.23

Note. Premorbid IQ = Premorbid full-scale IQ (WAIS-R) estimated from National Adult Reading Test; OMCT = Orientation-Memory-Concentration Test; GDS = Geriatric Depression Scale; SFT = Supermarket Fluency Test; FAS = FAS Fluency Test; VAS - Beg = Visual Analogue Scale - Anxiety at the beginning; VAS - End = Visual Analogue Scale - Anxiety at the end.

Dementia Rating. Although participants with OMCT scores that indicated dementia were excluded from the study, PD participants generally had higher (poorer) scores relative to the HC participants. As can be seen in Table 6, the male PD group had a poorer mean dementia rating relative to the male HC group, $F(1, 78) = 5.67$, $p = .02$, $d = .69$ (.10–1.29), but there was little difference between the two female groups ($F < 1$).

The distributions of OMCT scores for the two male groups are shown in Figure 6. It appears that considerably more male PD participants scored above 8 relative to

the male HC participants, but relatively less male PD participants had scores of 2. Figure 7 shows that the female PD distribution was similar to that of the female HC group, except that a greater number of controls had scores of two.

Depression Rating. From Table 6, it can be seen that at both Time 1 and Time 2 the male PD group had markedly higher mean GDS scores than their respective control group, $F(1, 78) = 14.16$, $p < .001$, $d = 1.03$ (.42–1.65), and $F(1, 78) = 19.28$, $p < .001$, $d = 1.34$ (.70–1.98). The distributions of these scores, at Time 1, for both the male PD group and the male HC group are shown in Figure 8. Forty-four percent (11) of the male PD group scored above the cut-off (GDS = 10), and, therefore, were classified as depressed. Conversely, only 13% (3) of the male HC group met the criteria for depression. At Time 2, 64% (16) of the male PD group were classified as depressed, whereas, only 9% (2) of the male HC group met the criteria for depression (see Figure 9).

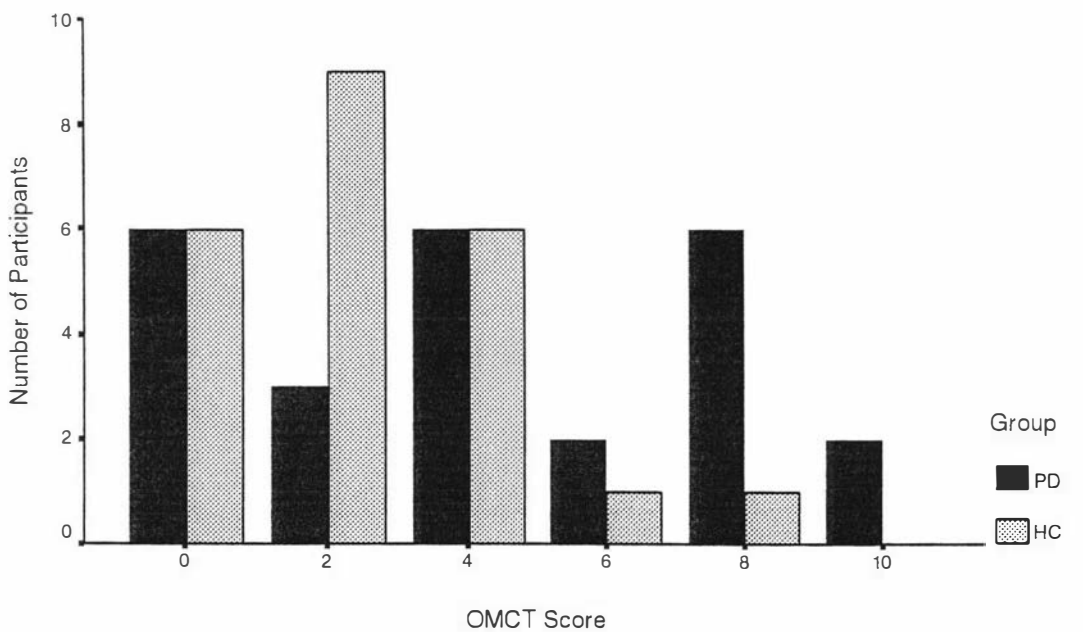


Figure 6. Orientation-Memory-Concentration Test (OMCT) scores for the 25 male PD participants and the 23 male HC participants. Higher scores indicate poorer cognitive status.

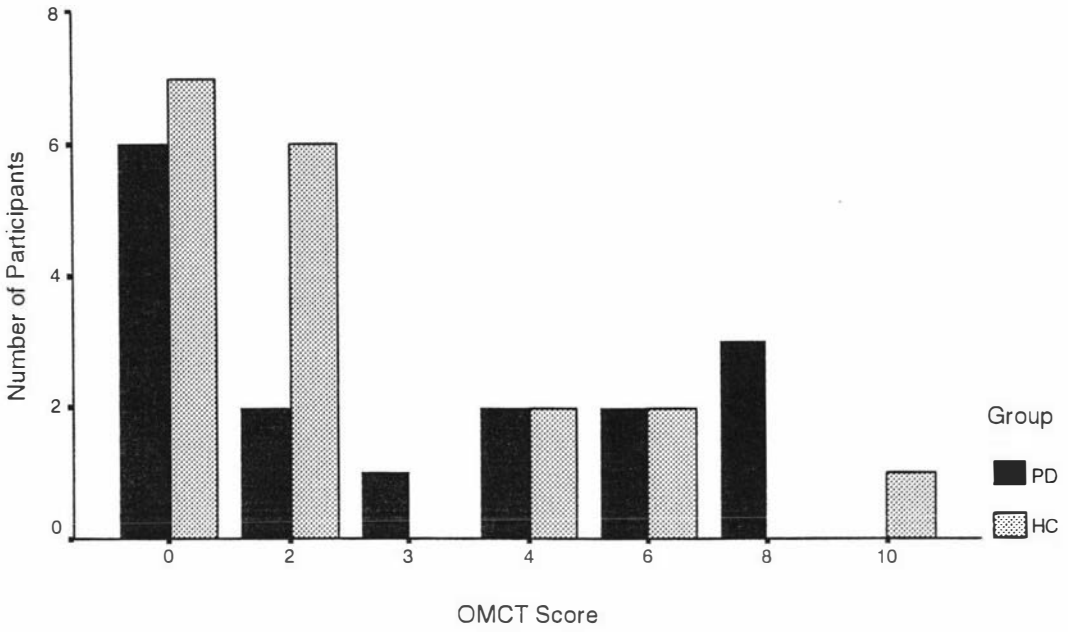


Figure 7. Orientation-Memory-Concentration Test (OMCT) scores for the 16 female PD participants and the 18 female HC participants. Higher scores indicate poorer cognitive status.

At both Time 1 and Time 2 the female PD group had markedly higher mean GDS scores than their respective control group (see Table 6), $F(1, 78) = 6.90$, $p = .01$,

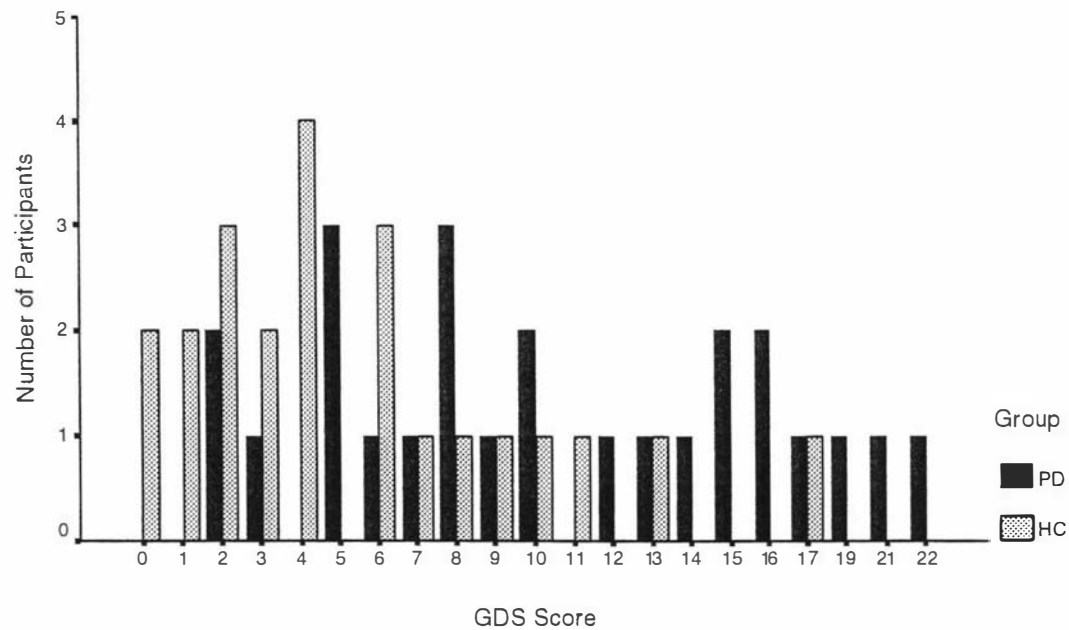


Figure 8. Geriatric Depression Scale (GDS) scores for the 25 male PD participants and the 23 male HC participants at Time 1. Higher scores indicate greater depression.

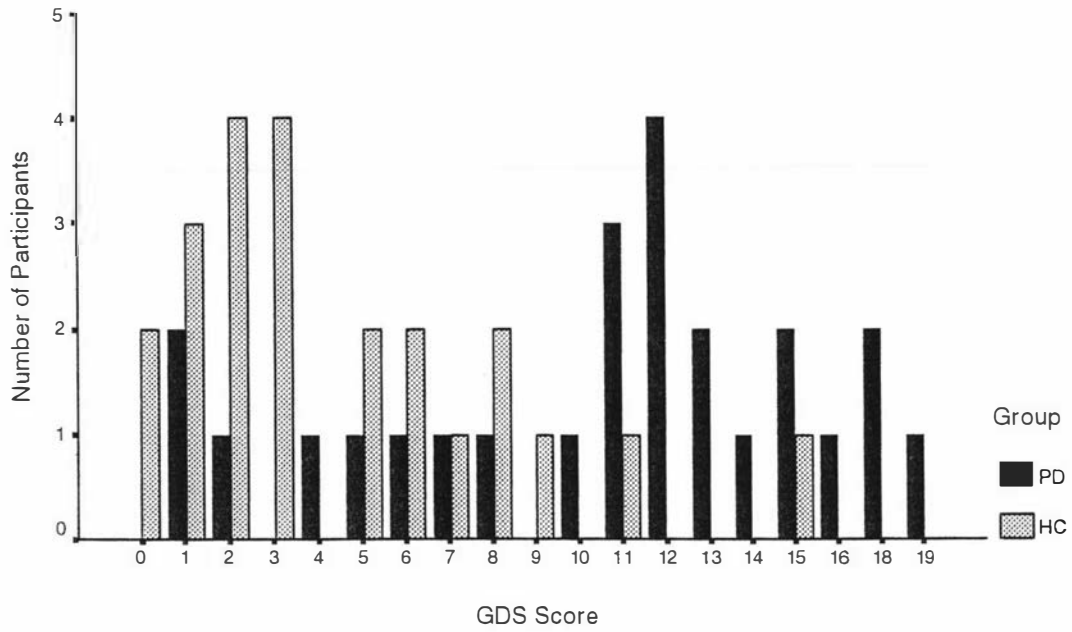


Figure 9. Geriatric Depression Scale (GDS) scores for the 25 male PD participants and the 23 male HC participants at Time 2. Higher scores indicate greater depression.

d = .98 (.25–1.71), and $F(1, 78) = 12.54$, $p = .001$, *d* = 1.14 (.39–1.88). At Time 1, 44% (7) of the female PD group met the criteria for depression, whereas none of the female HC group were classified as depressed (see Figure 10). At Time 2, 44% (7) of the female PD participants met the criteria for depression, but only 11% (2) of the female HC participants were classified as depressed (see Figure 11).

Semantic Fluency. Table 6 indicates that the male PD group had poorer semantic fluency relative to the male HC group, both at Time 1, $F(1, 78) = 5.64$, $p = .02$, *d* = .68 (.09–1.28), and at Time 2, $F(1, 78) = 10.95$, $p = .001$, *d* = .96 (.35–1.57). Similarly, the female PD group had poorer performance relative to the female HC group, both at Time 1, $F(1, 78) = 6.35$, $p = .01$, *d* = .87 (.15–1.60), and at Time 2, $F(1, 78) = 5.75$, $p = .02$, *d* = .82 (.09–1.54).

Letter Fluency. The results of the premorbid letter fluency test provided no evidence of a difference between the male PD group and the male HC group (*M* = 42.72, *SD* = 7.64, and *M* = 42.26, *SD* = 6.63, respectively), $F < 1$. However, the female PD group had somewhat poorer premorbid letter fluency (*M* = 42.31, *SD* =

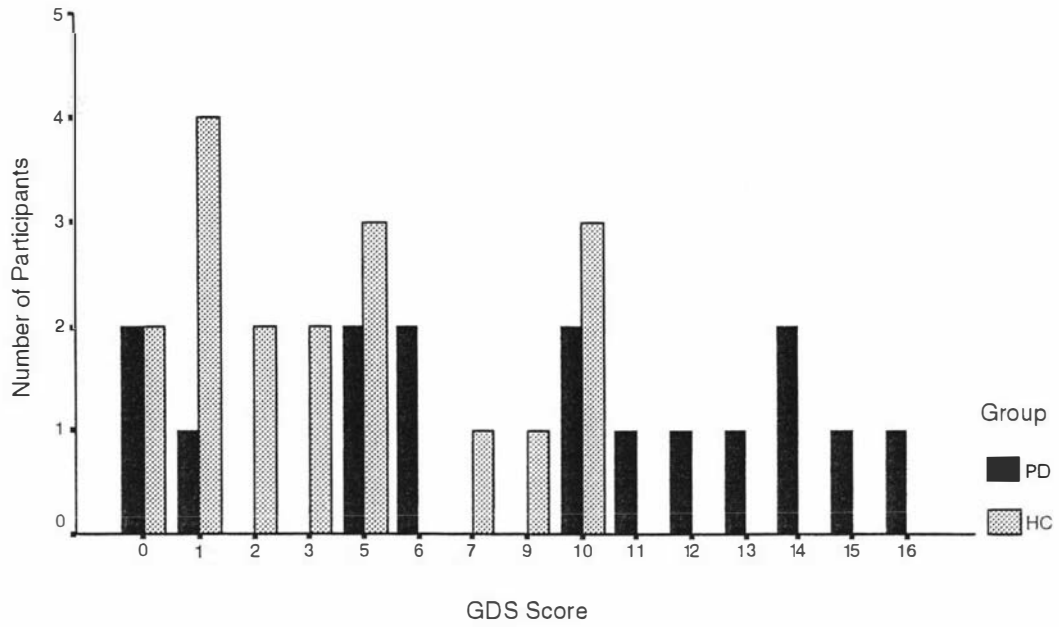


Figure 10. Geriatric Depression Scale (GDS) scores for the 16 female PD participants and the 18 female HC participants at Time 1. Higher scores indicate greater depression.

6.38) relative to the female HC group (M = 45.44, SD = 4.66), $F(1, 78) = 1.93$, p = .17, d = .57 (-.14–1.27).

In terms of current letter fluency (see Table 6), the male PD group appeared to

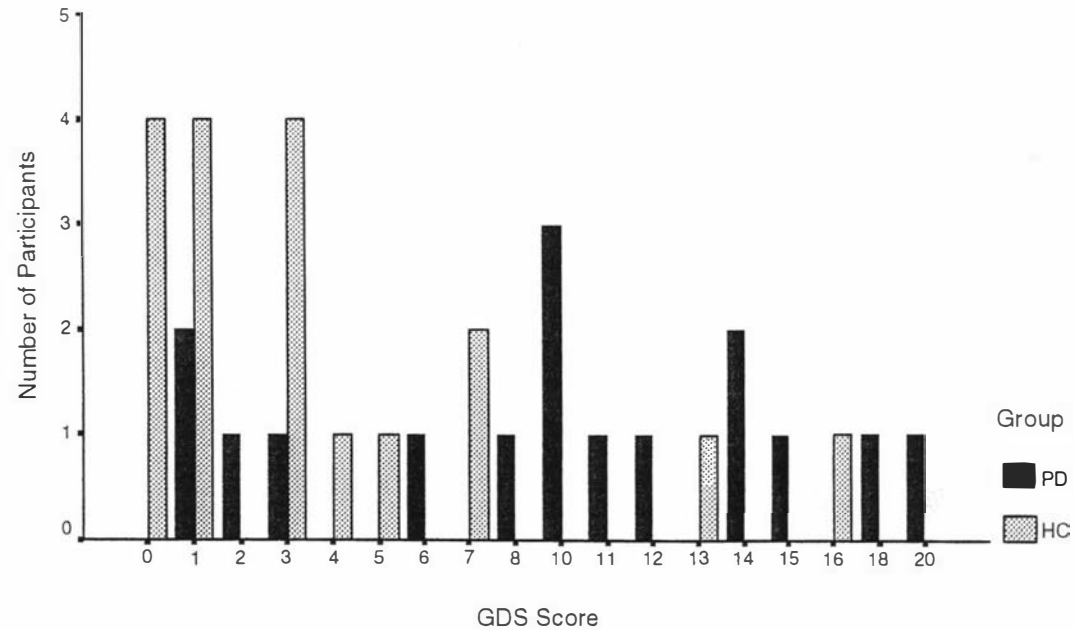


Figure 11. Geriatric Depression Scale (GDS) scores for the 16 female PD participants and the 18 female HC participants at Time 2. Higher scores indicate greater depression.

have normal letter fluency relative to their respective controls at Time 1 ($F < 1$), and slightly poorer letter fluency at Time 2, $F(1, 78) = 1.82$, $p = .18$, $d = .39$ (-.19–.98). The female PD group had poorer performance than the female HC group, both at Time 1, $F(1, 78) = 2.54$, $p = .12$, $d = .55$ (-.16–1.26), and at Time 2, $F(1, 78) = 4.54$, $p = .02$, $d = .73$ (.01–1.44).

Anxiety. As can be seen in Table 6, there was little evidence of a difference in anxiety between the male PD group and the male HC group, at the beginning of Time 1 ($F < 1$), or at the beginning of Time 2, $F(1, 78) = 1.84$, $p = .18$, $d = .38$ (-.20–1.27). In contrast, the female PD group had higher anxiety than the female HC group, both at Time 1, $F(1, 78) = 3.04$, $p = .09$, $d = .61$ (-.10–1.32), and at Time 2, $F(1, 78) = 2.43$, $p = .12$, $d = .57$ (-.14–1.27). At the end of the testing session, anxiety levels were not elevated in the male PD group at either time (Fs< 1). Similarly, the female PD group had similar levels of anxiety to those of their respective control group, both at Time 1 ($F < 1$), and at Time 2, $F(1, 78) = 1.17$, $p = .28$, $d = .33$ (-.37–1.03).

Finally, although there were some differences in VAS anxiety scores, the level of reported anxiety was generally very low. For instance, no participant scored above 83 on the 100-point scale, and the 75th percentile was ≤ 50 for all groups.

Vigilance. Overall, performance was generally high; no participant scored below .70, and the 75th percentile was ≥ .98 for all groups. Even so, the male PD group had markedly poorer sensitivity than the male HC group (see Table 7), both at Time 1, $F(1, 36) = 15.27$, $p < .001$, $d = 1.10$ (.39–1.80), and at Time 2, $F(1, 42) = 10.52$, $p = .002$, $d = 1.20$ (.54–1.86). The female PD group had similar sensitivity to that of the female HC group, at Time 1 ($F < 1$), but poorer sensitivity at Time 2, $F(1, 30) = 7.35$, $p = .01$, $d = 1.00$ (.24–1.76).

With respect to response bias, one-sample t-tests indicated that at Time 1 the male PD group produced positive bias, $t(16) = 2.15$, $p = .05$, $d = .53$, (-.50–1.56), whereas the male HC group produced negative bias, $t(20) = -3.10$, $p = .01$, $d =$

Table 7 *Means and SDs for the Mackworth Clock for the Parkinson's Disease (PD) Group and Healthy Control (HC) Group as a Function of Gender and Time*

	Male PD		Male HC		Female PD		Female HC	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
A'								
Time 1[a]	.95	.05	.99	.02	.97	.04	.98	.02
Time 2[b]	.93	.09	.99	.01	.94	.07	.99	.02
B"								
Time 1[a]	.39	.74	-.55	.82	-.26	.97	-.09	.96
Time 2[b]	-.12	.74	-.58	.79	-.01	.78	-.35	.92

Note. A' = Nonparametric measure of sensitivity; *B"* = Nonparametric measure of response bias. [a]*n* = 17 PD Males, 21 HC Males, 13 PD Females, 10 HC Females. [b]*n* = 22 PD Males, 22 HC Males, 15 PD Females, 17 HC Females.

.67, (-.26–1.60); a between-groups comparison revealed a large difference between the male PD and HC groups, $F(1, 36) = 13.51$, $p < .001$, $d = 1.20$ (.49–1.91). At Time 2, the male PD group had no response bias, $t(21) = -.75$, while the male HC group had a larger negative bias $t(21) = -3.48$, $p = .002$, $d = .73$, (-.18–1.64); a between-groups comparison indicated a medium effect, $F(1, 42) = 4.11$, $p = .05$, $d = .63$ (.01–1.25). At Time 1, both the female PD and the female HC group had little response bias, $t(12) = -.97$, and $t(9) = -.30$; a between-groups comparison showed no practical difference between the groups ($F < 1$). At Time 2, the female PD group had no response bias, $t(14) = -.05$, whereas the female HC group produced a small negative bias, $t(16) = -1.56$, $p = .14$, $d = .38$, (-.65–1.41); a between-groups comparison revealed a small effect, $F(1, 30) = 1.23$, $p = .28$, $d = .34$ (-.38–1.06).

To summarise, vigilance was generally high, although the PD participants generally had poorer sensitivity than the controls. There was no consistent evidence of response bias in the PD participants, but the male PD group responded somewhat differently to that of their respective controls.

The strategies used to complete the MC task were classified as being either internal or external. At Time 1, more PD males (94%) used external strategies than HC males (71%), χ^2 (1, N = 39) = 3.49, p = .06, d = .63 (-.03–1.29). However, at Time 2 there was little difference between the numbers of PD and HC males using external strategies (81% and 82%, respectively), χ^2 (1, N = 43) = 0.01, p = .94, d = .02 (-.59–.63). For the female participants, more of the PD sample than the HC sample used external strategies at Time 1 (92% and 80%, respectively), χ^2 (1, N = 23) = 0.75, p = .38, d = .37 (-.49–1.24). Similarly, at Time 2 a greater number of PD females (87%) used external strategies than HC females (71%), χ^2 (1, N = 32) = 1.21, p = .27, d = .40 (-.33–1.12).

Clearly, most participants relied on external strategies to complete the vigilance task, and any difference between the PD and control groups was generally small.

Parkinson's Disease Diagnosis and Disease Stage. The diagnosis of PD had been made by each participant's own general practitioner and confirmed by a neurologist in 93% of cases (based on self-report data). Support for the diagnosis of PD was made with the help of Calne et al.'s (1992) categories of diagnosis. In the male PD group, 72% were categorised as being clinically definite, 24% as being clinically probable, and only one participant was classified as clinically possible. In the female PD group, 75% were classified as being clinically definite, while the remaining participants were classified as clinically probable.

Disease stage was rated with the H&Y (Hoehn & Yahr, 1967). The number of participants in each of the five stages of the rating scale at Time 1 (Figure 12) and at Time 2 (Figure 13) shows that most participants were rated at stage II or III, indicating these PD participants had mild-to-moderate physical symptoms.

Table 8 presents a summary of the functional severity (mADL), disease duration (i.e., years since the onset of symptoms), and age at onset of symptoms data.

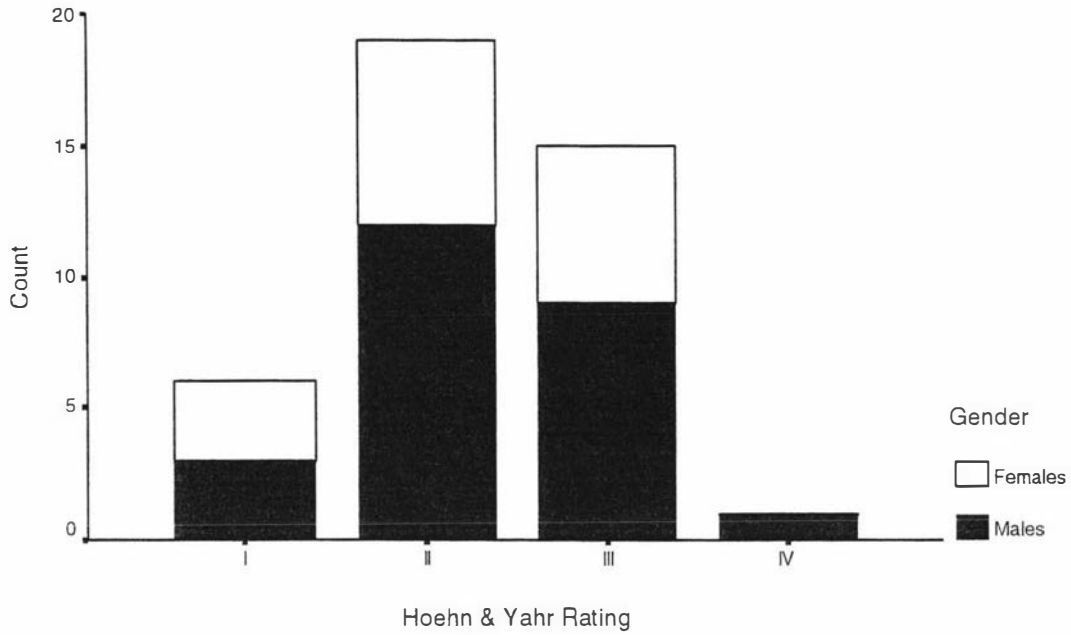


Figure 12. Number of participants at each stage of the Hoehn and Yahr Disease Rating Scale as a function of gender at Time 1. The higher the rating, the more severe the symptoms.

Medication at Time 1. The medication profile of the male and female PD participants is shown in Table 9. Out of the 25 male PD participants, 92% were

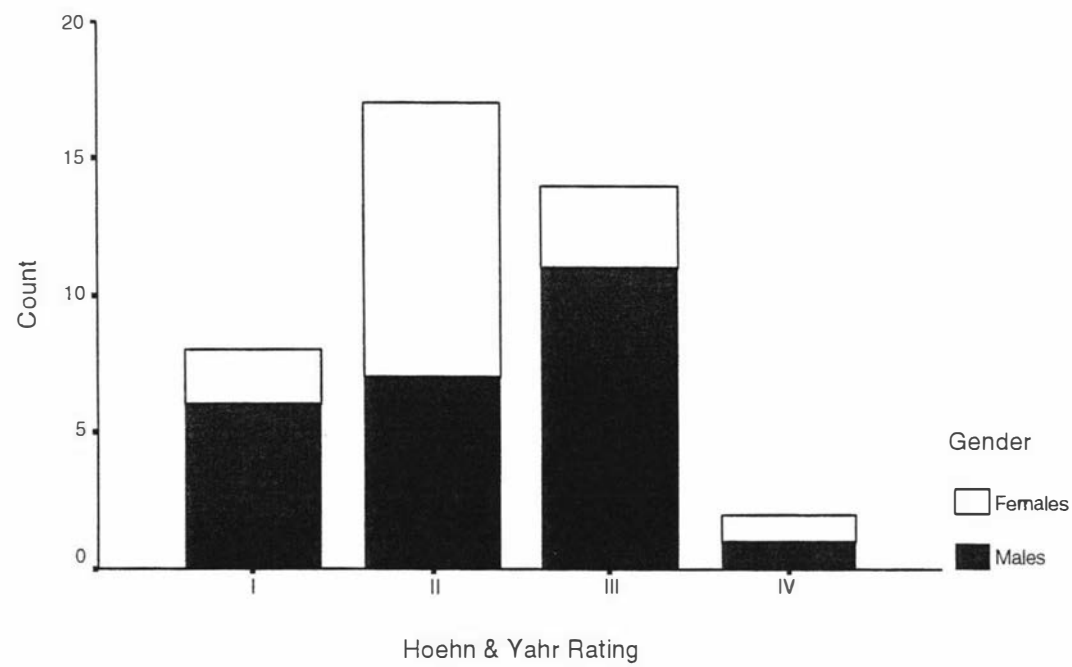


Figure 13. Number of participants at each stage of the Hoehn and Yahr Disease Rating Scale as a function of gender at Time 2. The higher the rating, the more severe the symptoms.

Table 8 *Means and SDs for the Parkinson's Disease (PD) Group Characteristics as a Function of Gender*

	Male PD			Female PD		
	M	*SD*	*Range*	*M*	*SD*	*Range*
mADL						
Time 1[a]	11.40	3.59	3–18	8.00	4.10	0–16
Time 2[b]	10.92	4.15	2–20	8.56	4.38	0–16
Disease duration[c]	10.16	5.91	1–25	7.22	5.03	1–21
Age at onset[c]	60.32	7.81	39–74	60.72	10.06	38–77

Note. mADL = modified Activities of Daily Living Scale.
[a]*n* = 20 males, 14 females. [b]*n* = 24 males, 16 females. [c]Years.

taking antiparkinsonian medications. Of those taking medication, all were using levodopa with a decarboxylase inhibitor, 30% were taking a dopamine agonist, 26% were taking a monoamine oxidase-B inhibitor, and one was on an anticholinergic. Two participants were also receiving tricyclic antidepressants. Fifty-seven percent of the males were taking more than one type of antiparkinsonian medication.

Of the 16 female PD participants, 94% were taking antiparkinsonian medication. In this medicated subgroup, 93% were taking levodopa plus decarboxylase inhibitor, 13% were taking a dopamine agonist, 13% were taking a monoamine oxidase-B inhibitor, and 33% were taking an anticholinergic. Two participants were also using tricyclic antidepressants. Sixty percent were receiving more than one type of antiparkinsonian medication.

As closely as possible, participants were tested at a time of optimal therapeutic effect. One female participant had temporarily stopped taking all antiparkinsonian medication, and two males were early in the course of PD and had not yet started taking antiparkinsonian medication.

Medication at Time 2. Of the 25 male PD participants, all but one were taking antiparkinsonian medication (Table 9). Of those using medication, all were taking levodopa plus decarboxylase inhibitor, 26% were taking a dopamine agonist, 46%

Table 9 *Medication Profile of Parkinson's Disease (PD) Participants as a Function of Gender and Time*

Medication Profile[a]	Male PD		Female PD	
	n	(%)	*n*	(%)
Antiparkinsonian medication				
Time 1				
L-Dopa	10	(40)	6	(38)
L-Dopa + Dopamine Agonist	5	(20)	1	(6)
L-Dopa + MAOI	5	(20)	1	(6)
L-Dopa + Dopamine Agonist + MAOI	1	(4)	0	(0)
L-Dopa + Dopamine Agonist + Anticholinergic	1	(4)	0	(0)
L-Dopa + MAOI + Anticholinergic	0	(0)	1	(6)
L-Dopa + Anticholinergic	0	(0)	3	(19)
L-Dopa + Tricyclic	1	(4)	2	(13)
Dopamine Agonist + Anticholinergic	0	(0)	1	(6)
Withdrawn from L-Dopa	0	(0)	1	(6)
De novo	1	(4)	0	(0)
De novo + Tricyclic	1	(4)	0	(0)
Time 2				
L-Dopa	8	(33)	6	(38)
L-Dopa + Dopamine Agonist	3	(13)	1	(6)
L-Dopa + MAOI	6	(25)	5	(31)
L-Dopa + Dopamine Agonist + MAOI	3	(13)	0	(0)
L-Dopa + Dopamine Agonist + Anticholinergic	1	(4)	1	(6)
L-Dopa + MAOI + Anticholinergic	1	(4)	0	(0)
L-Dopa + Anticholinergic	0	(0)	2	(13)
L-Dopa + MAOI + Anticholinergic + Tricyclic	1	(4)	0	(0)
L-Dopa + Dopamine Agonist + Tricyclic	0	(0)	1	(6)
De novo + Tricyclic	1	(4)	0	(0)
Effectiveness of antiparkinsonian medication				
Time 1				
Good	13	(57)	9	(60)
Fluctuating	10	(43)	6	(40)
Time 2[b]				
Good	16	(67)	9	(56)
Fluctuating	7	(29)	7	(44)
Currently on Antidepressants				
Time 1				
Yes	4	(16)	2	(12)
No	21	(84)	14	(88)
Time 2				
Yes	2	(8)	1	(6)
No	22	(92)	15	(94)

Note. L-Dopa = levodopa/decarboxylase inhibitor; MAOI = Monoamine oxidase-B inhibitor; Tricyclic = Tricyclic antidepressant.
[a]Based on participant's self-report. [b]One participant had missing data.

were taking a monoamine oxidase-B inhibitor, and 13% were on an anticholinergic. Two participants were also receiving tricyclic antidepressants.

Sixty-three percent were taking more than one type of antiparkinsonian medication. All 16 female PD participants were taking antiparkinsonian medication: 38% were using levodopa plus decarboxylase inhibitor, 19% were taking a dopamine agonist, 31% were taking a monoamine oxidase-B inhibitor, and 19% were taking an anticholinergic. One participant was also taking a tricyclic antidepressant. Sixty-three percent were taking more than one type of antiparkinsonian medication. Appendix S contains a list of most of the antiparkinsonian medications commercially available in New Zealand.

To determine whether medication with anticholinergic properties adversely affected performance on the memory tasks, participants receiving anticholinergic drugs were compared with the remaining participants. These supplementary analyses revealed no detrimental affects of anticholinergic medication on any of the memory tasks.

Of the male PD participants taking antiparkinsonian medication at Time 1, 57% rated its effectiveness as good, and 44% rated it as fluctuating (see Table 9). At Time 2, 67% rated it as good, whereas 29% rated it as fluctuating. In the female group at Time 1, 60% rated the medication effectiveness as good, and 40% as fluctuating. At Time 2, 56% rated it as good, and 44% rated it as fluctuating.

For those PD participants taking a levodopa preparation, the daily dose for males and females separately is shown in Table 10.

Table 10 *Mean, SD, and Range of the Daily Levodopa Dose for the Parkinson's Disease (PD) Participants as a Function of Gender and Time*

	Male PD			Female PD		
	M	*SD*	Range	*M*	*SD*	Range
Time 1[a]	477.17	233.33	100-1100	330.36	166.10	125-700
Time 2[b]	526.63	223.55	300-1000	308.59	135.07	100-600

Note. Values represent mean daily levodopa dose in milligrams.
[a] n = 23 males, 14 females. [b] n = 23 males, 16 females.

In summary, the demographic variables described above indicate that the PD participants were generally well matched to the controls. However, the neuropsychological analyses show the PD participants had poor generalised cognitive status and verbal fluency, although no participant met the criteria for dementia. In addition, the PD participants generally had higher levels of depression.

Results

Nonverbal Recognition Memory as a Function of Gender, Task Difficulty, and Time

Planned comparisons. The percentage of correct responses on the NRMT for each gender subgroup were analysed using planned comparisons. The within-groups comparisons revealed that the manipulation of the task difficulty factor was effective for the male PD group, both at Time 1, $F(1, 24) = 14.27$, $p = .001$, $d = 1.41$ (.72–2.11), and at Time 2, $F(1, 24) = 6.16$, $p = .02$, $d = 1.02$ (.35–1.70) (see Table 11). This was also true for the male HC group, both at Time 1, $F(1, 22) = 16.89$, $p < .001$, $d = 1.43$ (.75–2.11), and at Time 2, $F(1, 22) = 13.91$, $p = .001$, $d = 2.24$ (1.11–3.37). Interestingly, manipulation of the task difficulty factor was *not* effective for either female group at Time 1 (Fs < 1), and only weak for the HC females at Time 2, $F(1, 17) = 1.16$, $p = .30$, $d = .22$ (-.10–.55). However, it was effective for the PD females at Time 2, $F(1, 15) = 8.43$, $p = .01$, $d = 1.33$ (.46–2.20).

The between-groups comparisons revealed that the male PD group had impaired nonverbal recognition memory relative to the male HC group, both on the hard level, $F(1, 78) = 7.68$, $p = .01$, $d = .79$ (.19–1.39), and on the easy level, $F(1, 78) = 3.79$, $p = .06$, $d = .61$ (.02–1.20), at Time 1. However, at Time 2 there was no evidence to suggest that the male PD group were impaired on the hard level of the task ($F < 1$). However, the deficit remained on the easy level, $F(1, 78) = 2.13$, $p = .15$, $d = .45$ (-.14–1.03).

Table 11 *Summary of Results From the Nonverbal Recognition Memory Task as a Function of Gender, Difficulty Level, and Time for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Males[a]			
Hard level						
Time 1	58.58	10.32	40 – 83	66.30	9.02	50 – 90
Time 2	62.25	14.45	35 – 82	64.57	10.72	50 – 85
Easy level						
Time 1	68.04	11.78	42 – 89	74.26	8.15	60 – 95
Time 2	69.79	11.29	47 – 90	74.43	9.20	55 – 95
			Females[b]			
Hard level						
Time 1	66.06	9.27	53 – 85	69.28	9.75	53 – 90
Time 2	63.25	11.25	47 – 85	67.78	12.76	40 – 85
Easy level						
Time 1	67.56	14.34	31 – 85	71.06	9.87	55 – 90
Time 2	71.75	9.56	58 – 90	70.28	13.61	45 –100

Note. Values are scores on the NRMT expressed as the percent correct.
[a]*n* = 25 PD, 23 HC. [b]*n* = 16 PD, 18 HC.

Conversely, there was little evidence to suggest that the female PD group was impaired relative to the female HC group at Time 1, on either the hard or the easy level of the NRMT (*F*s < 1). At Time 2, there was some evidence of a small deficit on the hard level, $F(1, 78) = 1.11$, $p = .30$, $d = .38$ (-.32–1.07), but not on the easy level ($F < 1$).

Verbal Recognition Memory as a Function of Gender and Time

Planned comparisons. The percentage of correct responses on the VRMT were analysed using planned contrasts. The data in Table 12 suggest that the male PD group had impaired verbal recognition relative to the male HC group, both at Time 1, $F(1, 78) = 1.18$, $p = .28$, $d = .33$ (-.25–.91), and at Time 2, $F(1, 78) = 3.24$, $p = .08$, $d = .48$ (-.11–1.06). Similarly, the female PD group was impaired relative to the female HC group, both at Time 1, $F(1, 78) = 6.29$, $p = .01$, $d = .81$ (.09–1.53), and at Time 2, $F(1, 78) = 3.11$, $p = .09$, $d = .71$ (.00–1.43).

Table 12 *Summary of Results From the Verbal Recognition Memory Task as a Function of Gender and Time for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Males[a]			
Time 1	80.86	7.14	64 – 98	83.59	9.32	67 – 98
Time 2	82.60	8.96	67 –100	87.00	9.52	70 –100
			Females[b]			
Time 1	77.41	9.68	59 – 89	84.89	8.89	70 –100
Time 2	79.34	7.35	65 – 90	84.47	7.11	75 –100

Note. Values are scores on the VRMT expressed as the percent correct.
[a]*n* = 25 PD, 23 HC. [b]*n* = 16 PD, 18 HC.

Recognition Memory Confidence Ratings

Missing data analysis. A preliminary analysis of the data indicated substantial missing data. The Mann-Whitney U Test was used to check for differences in confidence ratings between participants with and without missing data. The results indicated important differences between these two groups with ESs ranging from .02 to .71 for the VRMT confidence ratings, and .05 to 1.17 for the NRMT.

The χ^2 test was used to check that the percentage of participants with missing data was the same within each group. For the VRMT at Time 1, the results showed there were less male PD participants (56%) with missing data compared to male HC participants (87%), χ^2 (1, N = 48) = 5.56, p = .02, d = .72 (.13–1.32). Similarly, at Time 2 there were less male PD participants (44%) compared to their respective controls (74%), χ^2 (1, N = 48) = 4.41, p = .04, d = .64 (.04–1.23). The situation was somewhat different for female participants. At Time 1 there were more female PD participants (81%) with missing data relative to female HC participants (61%), χ^2 (1, N = 34) = 1.66, p = .20, d = .45 (-.25–1.16). Similarly, at

Time 2 there were more female PD participants (69%) with missing data compared to their respective controls (50%), χ^2 (1, N = 34) = 1.23, p = .27, d = .39 (-.31–1.09).

For the hard level of the NRMT at Time 1, there were less male PD participants (32%) with missing data relative to male HC participants (57%), χ^2 (1, N = 48) = 2.93, p = .09, d = .58 (-.08–1.10). At Time 2, there was little difference between groups (36% and 35% for the male PD and HC participants, respectively), χ^2 (1, N = 48) = .008, p = .93, d = .03 (-.55–.60). For the female participants there was evidence that at Time 1 there were more in the PD group (50%) with missing data compared to the HC group (22%), χ^2 (1, N = 34) = 2.86, p = .09, d = .61 (-.10–1.32). However, at Time 2 there was no evidence of a difference between groups (38% and 44% for the female PD and HC participants, respectively), χ^2 (1, N = 34) = .17, p = .68, d = .14 (-.55–.84).

For the easy level of the NRMT at Time 1, there was only a small difference between the male PD participants and their respective controls (44% and 35%, respectively), χ^2 (1, N = 48) = .43, p = .51, d = .19 (-.39–.77). At Time 2 there were more male PD participants (48%) with missing data compared to male HC participants (22%), χ^2 (1, N = 48) = 3.61, p = .06, d = .57 (-.02–1.16). For the female participants at Time 1 there were slightly more in the PD group (38%) with missing data compared to the HC group (28%), χ^2 (1, N = 34) = .37, p = .55, d = .21 (-.49–.90). However, at Time 2 there was little difference between groups (25% and 28% for the female PD and HC participants, respectively), χ^2 (1, N = 34) = .03, p = .86, d = .06 (-.63–.76).

In summary, participants with missing data performed differently to those without missing data on both the verbal and the nonverbal recognition tasks. Moreover, the numbers of participants with missing data varied as a function of group and task.

Confidence rating analysis. Because the missing data analysis indicated subgroup differences in the confidence rating data set, no further analysis or discussion will be made. Nevertheless, the mean proportion of trials for which each confidence rating was selected as a function of gender, time, and task can be found in Appendix T (Table T1).

Recall as a Function of Gender and Time

Planned comparisons. Recall was assessed with the KOLT - age-scaled quotients. From Table 13 it is clear that the male PD group had impaired recall, both at Time 1, $F(1, 78) = 3.91$, $p = .05$, $d = .55$ (-.04–1.14), and at Time 2, $F(1, 78) = 10.86$, $p = .001$, $d = .94$ (.33–1.55). Similarly, the female PD group had impaired recall, both at Time 1, $F(1, 78) = 5.24$, $p = .03$, $d = .82$ (.10–1.55), and at Time 2, $F(1, 78) = 10.83$, $p = .002$, $d = 1.15$ (.41–1.90).

Table 13 *Summary of Results From the Kendrick Object Learning Task as a Function of Gender and Time for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Males[a]			
Time 1	93.72	14.37	61 – 123	101.00	11.65	77 – 126
Time 2	91.24	14.89	62 – 134	102.80	8.61	89 – 119
			Females[b]			
Time 1	97.81	11.18	85 – 123	107.83	12.94	89 – 133
Time 2	94.13	13.36	73 – 131	107.83	10.39	87 – 126

Note. Values are Kendrick Object Learning Task - mean age-scaled quotients.
[a] $n = 25$ PD, 23 HC. [b] $n = 16$ PD, 18 HC.

Prospective Memory as a Function of Gender and Time

Remembering a Question

Planned comparisons. There was insufficient evidence to suggest that the male PD group had impaired prospective memory for a question relative to the male HC group at Time 1 ($F < 1$) (Table 14). However, at Time 2 it appears that the male PD group were impaired relative to the male HC group, $F(1, 78) = 7.88$, $p = .006$, $d = .84$ (.24–1.44). Similarly, the female PD group had normal performance at Time 1 ($F < 1$), but impaired performance at Time 2, $F(1, 78) = 5.65$, $p = .02$, $d = .78$ (.06–1.50).

Remembering an Object

Planned comparisons. Table 14 indicates that the male PD group had impaired

Table 14 *Summary of Results From the Prospective Memory Tasks as a Function of Gender and Time for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD[a]			HC[b]		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Males			
PMQT						
Time 1	2.52	1.42	0–4	2.87	1.46	0–4
Time 2	2.08	1.55	0–4	3.16	0.92	1–4
PMOT						
Time 1	3.00	0.96	2–4	3.61	0.78	2–4
Time 2	3.60	0.82	2–4	3.19	0.98	2–4
			Females			
PMQT						
Time 1	3.13	1.09	1–4	3.17	1.34	0–4
Time 2	2.13	1.50	0–4	3.22	1.31	0–4
PMOT						
Time 1	3.44	1.03	1–4	3.44	0.92	2–4
Time 2	3.19	0.98	2–4	3.44	0.92	2–4

Note. PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a] n = 25 Males, 16 Females. [b] n = 23 Males, 18 Females.

prospective memory for an object relative to the male HC group at Time 1, $F(1, 78) = 5.26$, $p = .03$, $d = .69$ (.10–1.29). However, at Time 2 there was no evidence of a deficit in the male PD group. There was also no evidence that the female PD group were impaired either at Time 1 or at Time 2 (Fs < 1).

Summary: Group Differences as a Function of Gender

The ESs computed for both the male and the female participants showed that both PD subgroups suffered from consistent deficits in nonverbal recognition (at the more difficult level), verbal recognition, and recall (Table 15). Compared to the PD females, the male PD group had a larger and more consistent impairment with respect to nonverbal recognition (at the easier level) and prospective memory for a question. Conversely, the female group had larger deficits in verbal recognition

Table 15 *Summary of Effect Sizes From the Memory Tasks as a Function of Time for the Comparisons Involving the Male and Female Parkinson's Disease (PD) and their Respective Healthy Control Groups*

	Male PD		Female PD	
	d	95% CI	*d*	95% CI
NRMT - hard level				
Time 1	.79	.19–1.39	.34	-.36–1.04
Time 2	.18	-.40– .76	.38	-.32–1.07
NRMT - easy level				
Time 1	.61	.02–1.20	.29	-.41– .98
Time 2	.45	-.14–1.03	-.12	-.82– .57
VRMT				
Time 1	.33	-.25– .91	.81	.09–1.53
Time 2	.48	-.11–1.06	.71	.00–1.43
KOLTQ				
Time 1	.55	-.04–1.14	.82	.10–1.55
Time 2	.94	.33–1.55	1.15	.41–1.90
PMQT				
Time 1	.24	-.34– .82	.03	-.66– .73
Time 2	.84	.24–1.44	.78	.06–1.50
PMOT				
Time 1	.69	.10–1.29	.00	-.69– .69
Time 2	-.46	-1.04– .13	.24	-.43– .96

Note. d = ES statistic (Cohen, 1988); 95% CI = 95% confidence interval for d; NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.

memory and recall.

Memory decline, as indicated by a substantially larger ES at the second session, can be seen in both subgroups, but only for recall and prospective memory for a question (Table 15). Nevertheless, considerable overlap of the confidence intervals for these two tasks suggests that sampling variability may be responsible for the observed differences in ES.

Effects of Disease Stage

To examine the potential moderating effect of disease stage, planned comparisons were made between the early-stage PD group and the HC group, and between the advanced-stage PD group and the HC group. Separate analyses were made at each time.

Stage Effects at Time 1

At Time 1 there were 25 participants in the early PD group, and 16 in the advanced PD group based on a cut-off of H&Y II. Table 16 shows that the early PD and the late PD subgroups did not differ from the HC group with respect to gender, χ^2 (1, N = 66) = .10, p = .76, d = .08 (-.41–.57), and χ^2 (1, N = 57) = .19, p = .66, d = .13 (-.40–.66), respectively. The subgroups were also matched to the controls in terms of age, education, and premorbid IQ (Fs < 1). Compared to the HC group, both the early PD group and the advanced PD group had higher dementia ratings, F(1, 79) = 1.78, p = .19, d = .34 (-.15–.84), and F(1, 79) = 8.21, p = .01, d = .94 (.38–1.50), respectively, and the magnitude of the effect was greater in the advanced group. There was also considerable evidence that both the early PD and the advanced PD group had higher levels of depression than the HC group, F(1, 79) = 12.76, p = .001, d = .94 (.42–1.45). and F(1, 79) = 16.78, p <

Table 16 *Characteristics of the Early-Stage Parkinson's Disease (Early PD) Group, the Advanced-Stage Parkinson's Disease (Advanced PD) Group, and the Healthy Control (HC) Group at Time 1*

	Early PD[a]		Advanced PD[b]		HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age (years)	68.84	6.02	70.50	6.22	69.80	5.31
Education (years)	3.92	2.98	4.06	2.84	4.34	3.02
Premorbid IQ	109.17	9.38	109.82	12.20	110.15	8.94
OMCT	3.36	3.45	4.81	3.04	2.39	2.38
GDS	9.32	5.84	10.81	5.58	4.83	4.05
Disease duration	8.94	6.16	9.13	5.10	-	-
Age at onset	59.90	8.27	61.38	9.38	-	-
Levodopa	430.68	193.14	408.33	261.34	-	-
mADL	8.71	4.05	12.08	3.40	-	-

Note. Premorbid IQ = Premorbid full-scale IQ (WAIS-R) estimated from National Adult Reading Test; OMCT = Orientation-Memory-Concentration Test; GDS = Geriatric Depression Scale; Levodopa = mean daily levodopa dose in milligrams; mADL = modified Activities of Daily Living Scale.
[a]n = 15 Males, 10 Females. [b]n = 10 Males, 6 Females. [c]n = 23 Males, 18 Females.

.001, d = 1.32 (.74–1.91), respectively. As with the dementia rating, the ES was greater in the advanced group compared to the early group.

For each PD subgroup, the mean disease duration, the mean age at onset, the mean daily levodopa dose, and the mean mADL are shown in Table 16. With respect to disease stage, the majority of the early PD participants were rated as H&Y stage II (I: n = 6; II: n = 19), whereas the majority of advanced PD participants were rated as stage III (III: n = 15; IV: n = 1).

Table 17 presents a summary of the results from the memory tasks. The within-groups comparisons revealed that the manipulation of NRMT difficulty factor was effective for both the early PD group and the HC group, $F(1, 24) = 22.73$, $p < .001$, d = 2.66 (1.92–3.39), and $F(1, 40) = 9.98$, $p = .003$, d = 1.00 (.61–1.39), respectively. However, this was not the case for the advanced PD group ($F < 1$).

The between-groups comparisons indicated that, compared to the HC group, both the early PD group and the advanced PD group had impaired nonverbal recognition at the more difficult level of the task, $F(1, 79) = 4.97$, $p = .03$, d = .65 (.15–1.16), and $F(1, 79) = 5.37$, $p = .02$, d = .62 (.08–1.17), respectively. The

Table 17 *Mean PC and SD for the Verbal and Nonverbal Recognition Memory Tasks for the Early-Stage Parkinson's Disease (Early PD) Group, the Advanced-Stage Parkinson's Disease (Advanced PD) Group, and the Healthy Control (HC) Group at Time 1*

	Early PD[a]		Advanced PD[b]		HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
NRMT - hard	61.96	7.39	60.79	14.31	67.61	9.35
NRMT - easy	71.88	9.00	61.57	15.13	72.85	8.98
VRMT	80.34	8.02	78.22	8.79	84.16	9.04
KOLTQ	96.28	15.15	93.18	9.76	104.00	12.55
PMQT	3.08	1.19	2.25	1.39	3.00	1.40
PMOT	3.24	0.93	3.06	1.12	3.54	0.84

Note. NRMT-hard = Nonverbal Recognition Memory Task - hard level of difficulty; NRMT-easy = Nonverbal Recognition Memory Task - easy level of difficulty; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task. [a]$n = 25$. [b]$n = 16$. [c]$n = 41$.

pattern of results was somewhat different at the easier level of the NRMT. The advanced PD group had impaired nonverbal recognition, $F(1, 79) = 13.47$, $p < .001$, $d = 1.03$ (.46–1.59), but the early PD group was not impaired ($F < 1$).

Both the early PD group and the advanced PD group had impaired verbal recognition relative to the HC group, $F(1, 79) = 3.00$, $p = .09$, $d = .44$ (-.06–.94), and $F(1, 79) = 5.37$, $p = .02$, $d = .66$ (.12–1.21), respectively.

Both the early PD group and the advanced PD group had impaired recall relative to the HC group, $F(1, 79) = 5.52$, $p = .02$, $d = .57$ (.07–1.07), and $F(1, 79) = 7.13$, $p = .01$, $d = .86$ (.31–1.41), respectively.

The advanced PD group showed signs of impaired prospective memory for a question compared to the HC group, $F(1, 79) = 3.63$, $p = .06$, $d = .54$ (.00–1.08). A similar deficit was not evident in the early PD group ($F < 1$). With respect to prospective memory for an object, both the early PD group and the advanced PD group were impaired, $F(1, 79) = 1.59$, $p = .21$, $d = .34$ (-.15–.83), and $F(1, 79) = 3.02$, $p = .09$, $d = .51$ (-.03–1.05), respectively, with the advanced group again showing the greater deficit.

Stage Effects at Time 2

As for Time 1, a cut-off of H&Y II was used to divide the PD sample into two subgroups; comparisons were made between each subgroup and the HC group. At Time 2, 84% (21) of the early PD group could still be classified as having mild PD, whereas 75% (12) of the advanced PD group remained classified as having more advanced symptoms.

As shown in Table 18, the early PD group did not differ from the HC group with respect to gender, χ^2 (1, N = 66) = .11, p = .76, d = .08 (-.41-.57), but the advanced PD group had proportionally less female participants, χ^2 (1, N = 57) = 1.74, p = .19, d = .39 (-.14-.93). The subgroups were matched to the controls in terms of education and premorbid IQ (Fs < 1). In terms of age, both the early PD group and the advanced PD group had a similar mean age to that of the HC group, F(1, 79) = 1.20, p = .28, d = .28 (-.21–.77), and F < 1, respectively. Compared to the HC group, both the early PD and the advanced PD group had higher dementia scores, F(1, 79) = 5.24, p = .03, d = .58 (.08–1.08), and F(1,

Table 18 *Characteristics of the Early-Stage Parkinson's Disease (Early PD) Group, the Advanced-Stage Parkinson's Disease (Advanced PD) Group, and the Healthy Control Group (HC) at Time 2*

	Early PD[a]		Advanced PD[b]		HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age (years)	68.24	6.08	71.44	5.72	69.80	5.31
Education (years)	3.88	2.91	4.13	2.96	4.34	3.02
Premorbid IQ	108.99	8.52	110.10	13.16	110.88	8.94
OMCT	4.08	3.67	3.69	2.82	2.39	2.38
GDS	9.44	5.59	11.56	5.09	4.17	4.10
Disease duration	8.26	5.93	10.19	5.30	-	-
Age at onset	59.98	8.19	61.25	9.51	-	-
Levodopa	391.67	162.59	510.00	278.32	-	-
mADL	8.48	4.19	12.47	3.44	-	-

Note. Premorbid IQ = Premorbid full-scale IQ (WAIS-R) estimated from National Adult Reading Test; OMCT = Orientation-Memory-Concentration Test; GDS = Geriatric Depression Scale; Levodopa = mean daily levodopa dose in milligrams; mADL = modified Activities of Daily Living Scale.
[a]n = 13 Males, 12 Females. [b]n = 12 Males, 4 Females. [c]n = 23 Males, 18 Females.

79) = 2.29, p = .13, d = .52 (-.02–1.06), respectively. Likewise, both the early PD group and the advanced PD group had substantially higher levels of depression than the HC group, $F(1, 79)$ = 18.82, $p < .001$, d = 1.12 (.59–1.64), and $F(1, 79)$ = 27.45, $p < .001$, d = 1.68 (1.07–2.30), respectively. Once again, as might be expected, the level of depression was a little higher for the advanced group in comparison to the early group.

For the two PD subgroups, the mean disease duration, age at onset, daily levodopa dose, and mADL are shown in Table 18. With respect to disease stage, a majority of the early PD participants had an H&Y rating of II (I: n = 8; II: n = 17), whereas a majority of the advanced PD participants were stage III (III: n = 14; IV: n = 2).

Table 19 presents a summary of the results from the memory tasks. The within-groups comparisons confirmed that the NRMT difficulty factor was effective for the early PD, the advanced PD group, and the HC group, $F(1, 24)$ = 11.77, p = .002, d = 1.14 (.66–1.63), $F(1, 15)$ = 2.79, p = .12, d = 1.25 (.55–1.96), and $F(1, 40)$ = 12.57, p = .001, d = .85 (.50–1.20), respectively.

Table 19 *Mean PC and SD for the Verbal and Nonverbal Recognition Memory Tasks for the Early-Stage Parkinson's Disease (Early PD) Group, the Advanced-Stage Parkinson's Disease (Advanced PD) Group, and the Healthy Control (HC) Group at Time 2*

	Early PD[a]		Advanced PD[b]		HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
NRMT - hard	63.56	13.43	61.20	13.01	65.98	11.62
NRMT - easy	72.44	10.16	67.61	10.84	72.61	11.39
VRMT	82.00	8.11	80.28	9.07	85.89	8.54
KOLTQ	92.96	16.74	91.44	9.42	105.00	9.65
PMQT	2.40	1.44	1.63	1.54	3.19	1.09
PMOT	3.60	0.82	3.19	0.98	3.30	0.95

Note. NRMT-hard = Nonverbal Recognition Memory Task - hard level of difficulty; NRMT-easy = Nonverbal Recognition Memory Task - easy level of difficulty; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotients; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a]n = 25. [b]n = 16. [c]n = 41.

The between-groups comparisons showed that the advanced PD group had impaired nonverbal recognition at the more difficult level of the NRMT, $F(1, 79) = 1.69$, $p = .20$, $d = .40$ (-.14–.93), while the early PD group did not ($F < 1$). Likewise, at the easier level the advanced PD group were impaired, $F(1, 79) = 2.41$, $p = .13$, $d = .44$ (-.09–.98), but there was no evidence of a deficit in the early PD group ($F < 1$).

Compared to the HC group, both the early PD group and the advanced PD group showed signs of impaired verbal recognition memory, $F(1, 79) = 3.24$, $p = .08$, $d = .46$ (-.03–.96), and $F(1, 79) = 4.99$, $p = .03$, $d = .65$ (.10–1.19), respectively.

Both the early PD group and the advanced PD group had impaired recall relative to the HC group, $F(1, 79) = 15.10$, $p < .001$, $d = .94$ (.43–1.46), and $F(1, 79) = 14.20$, $p < .001$, $d = 1.41$ (.82–2.01), respectively.

Relative to the HC group, both the early PD and the advanced PD group had inferior prospective memory for a question, $F(1, 79) = 5.67$, $p = .02$, $d = .64$ (.13–1.14), and $F(1, 79) = 16.68$, $p < .001$, $d = 1.27$ (.69–1.85), respectively. Prospective memory for an object was not impaired in the advanced PD group ($F < 1$), or in the early PD group (the mean memory score in the early group was higher than in the HC group).

In summary, the ESs computed for both the early and the advanced participants indicated that both PD subgroups suffered from consistent recognition and recall memory impairment, whereas only the advanced group had consistent prospective memory impairment (Table 20). Furthermore, the advanced PD group generally had larger memory deficits than the early group. Table 20 also shows memory decline, as indicated by a substantially larger ES at the second session, in both subgroups, but only for recall and prospective memory for a question.

Table 20 *Summary of Effect Sizes From the Memory Tasks as a Function of Time for the Comparisons Involving the Early-Stage Parkinson's Disease (Early PD) Group, the Advanced-Stage Parkinson's Disease (Advanced PD) Group, and the Healthy Control Group*

	Early PD[a]		Advanced PD[b]	
	d	95% CI	*d*	95% CI
NRMT - hard level				
Time 1	.65	.15–1.16	.62	.08–1.17
Time 2	.20	-.30– .69	.40	-.14– .93
NRMT - easy level				
Time 1	.11	-.38– .60	1.03	.46–1.59
Time 2	.02	-.47– .51	.44	-.09– .98
VRMT				
Time 1	.44	-.06– .94	.66	.12–1.21
Time 2	.46	-.03– .96	.65	.10–1.19
KOLTQ				
Time 1	.57	.07–1.07	.86	.31–1.41
Time 2	.94	.43–1.46	1.41	.82–2.01
PMQT				
Time 1	-.06	-.55– .43	.54	.00–1.08
Time 2	.64	.13–1.14	1.27	.69–1.85
PMOT				
Time 1	.34	-.15– .83	.51	-.03–1.05
Time 2	-.33	.83– .16	.11	-.41– .64

Note. *d* = ES statistic (Cohen, 1988); 95% CI = 95% confidence interval for *d*; NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a]*n* = 25 PD, 41 HC. [b]*n* = 16 PD, 41 HC.

Effects of Depression

To establish whether depression moderated performance on the memory tasks, planned comparisons were made between the depressed PD group and the nondepressed HC group, and between the nondepressed PD group and the nondepressed controls. Separate analyses were made for each time.

Mood Effects at Time 1

At Time 1 there were 18 depressed PD participants, 23 nondepressed PD participants, and 38 nondepressed controls based on a cut-off of 10 on the GDS.

Table 21 shows that the depressed and nondepressed PD subgroups did not differ from the nondepressed controls in terms of age (*F*s < 1) or gender, χ^2 (1, *N* = 56) = .36, *p* = .55, *d* = .17 (-.36–.71), and χ^2 (1, *N* = 61) = .39, *p* = .53, *d* = .17 (-.35–.68). However, the depressed PD participants had, on average, less years of education than the controls, *F*(1, 76) = 3.37, *p* = .07, *d* = .56 (.01–1.10). There was no evidence of a difference between the nondepressed PD group and the controls in terms of education (*F* < 1). There was some evidence that the depressed PD group had lower mean premorbid IQ relative to the controls, *F*(1, 76) = 2.00, *p* = .16, *d* = .40 (-.14–.94). However, there was no difference in IQ between the nondepressed PD group and the controls (*F* < 1). Both the depressed PD group and the nondepressed PD group had higher dementia ratings compared to the controls, *F*(1, 76) = 5.48, *p* = .02, *d* = .71 (.16–1.26), and *F*(1, 76) = 2.18, *p* = .14, *d* = .41 (-.10–.93), respectively. As expected, the depressed PD group had considerably higher levels of depression than the controls, *F*(1, 76) = 151.19, *p* < .001, *d* = 3.54 (2.68–4.39). Finally, although the nondepressed PD participants did not meet the criteria for depression, they did have higher levels of depression than the controls, *F*(1, 76) = 3.49, *p* = .07, *d* = .49 (-.03–1.00).

Table 21 *Characteristics of the Depressed Parkinson's Disease (Depressed PD) Group, the Nondepressed Parkinson's Disease (Nondepressed PD) Group, and the Nondepressed Healthy Controls (Nondepressed HC) at Time 1*

	Depressed PD[a]		Nondepressed PD[b]		Nondepressed HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age (years)	69.67	6.46	69.35	5.90	70.05	5.36
Education (years)	2.94	2.21	4.78	3.15	4.42	2.82
Premorbid IQ	107.53	11.94	110.91	9.07	111.33	8.18
OMCT	4.39	3.42	3.57	3.30	2.42	2.42
GDS	15.28	2.99	5.70	3.21	4.13	3.22
Disease duration	10.28	6.22	8.02	5.19	-	-
Age at onset	59.39	10.83	61.33	6.57	-	-
Levodopa	450.00	205.19	400.00	233.42	-	-
mADL	12.25	3.44	8.00	3.66	-	-

Note. Premorbid IQ = Premorbid full-scale IQ (WAIS-R) estimated from National Adult Reading Test; OMCT = Orientation-Memory-Concentration Test; GDS = Geriatric Depression Scale; Levodopa = mean daily levodopa dose in milligrams; mADL = modified Activities of Daily Living Scale.
[a]*n* = 11 Males, 7 Females. [b]*n* = 14 Males, 9 Females. [c]*n* = 20 Males, 18 Females.

For the PD subgroups, mean disease duration, age at onset, daily levodopa dose, and mADL are shown in Table 21. With respect to disease stage, all of the depressed PD group were either stage II or III (II: $n = 10$; III: $n = 8$), whereas the nondepressed PD group ranged from stage I to IV (I: $n = 6$; II: $n = 9$; III: $n = 7$; IV: $n = 1$).

The results from the memory tasks are shown in Table 22. The within-groups comparisons between the two difficulty levels of the NRMT revealed that the manipulation of this factor was effective for the depressed PD group, the nondepressed PD group, and the controls, $F(1, 17) = 1.51$, $p = .24$, $d = .57$ (-.16–1.29), $F(1, 22) = 13.19$, $p = .001$, $d = 1.02$ (.50–1.55), and $F(1, 37) = 7.37$, $p = .01$, $d = .88$ (.37–1.39), respectively.

Both the depressed and the nondepressed PD groups had impaired nonverbal recognition at the more difficult level of the task relative to the controls, $F(1, 76) = 5.54$, $p = .02$, $d = .68$ (.13–1.23), and $F(1, 76) = 5.40$, $p = .02$, $d = .63$ (.11–1.16), respectively. In contrast, at the easier level the depressed PD group had impaired nonverbal recognition, $F(1, 76) = 5.22$, $p = .03$, $d = .66$ (.11–1.21), but the

Table 22 *Mean PC and SD for the Verbal and Nonverbal Recognition Memory Tasks for the Depressed Parkinson's Disease (Depressed PD) Group, the Nondepressed Parkinson's Disease (Nondepressed PD) Group, and the Nondepressed Healthy Controls (Nondepressed HC) at Time 1*

	Depressed PD[a]		Nondepressed PD[b]		Nondepressed HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
NRMT - hard	61.17	11.06	61.76	10.25	67.92	9.38
NRMT - easy	65.39	14.10	69.78	11.37	72.63	9.17
VRMT	76.47	5.87	81.89	9.21	83.84	9.03
KOLTQ	92.50	15.27	97.52	11.23	104.68	11.85
PMQT	2.44	1.46	3.00	1.17	3.00	1.43
PMOT	3.00	1.08	3.30	0.93	3.58	0.83

Note. NRMT-hard = Nonverbal Recognition Memory Task - hard level of difficulty; NRMT-easy = Nonverbal Recognition Memory Task - easy level of difficulty; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotients; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a] $n = 18$. [b] $n = 23$. [c] $n = 38$.

nondepressed PD group did not ($F < 1$).

The depressed PD group had impaired verbal recognition memory relative to the controls, $F(1, 76) = 9.22$, $p = .003$, $d = .90$ (.34–1.46). There was no evidence, however, of a deficit in the nondepressed PD group ($F < 1$).

Both the depressed PD group and the nondepressed PD group had impaired recall when compared to the controls, $F(1, 76) = 11.55$, $p = .001$, $d = .94$ (.37–1.50), and $F(1, 76) = 4.68$, $p = .03$, $d = .62$ (.09–1.14), respectively.

The depressed PD group showed signs of impaired prospective memory for a question, $F(1, 76) = 2.01$, $p = .16$, $d = .39$ (-.15–.93), but a similar deficit was not evident in the nondepressed PD group ($F < 1$). Prospective memory for an object was impaired both in the depressed PD group and in the nondepressed PD group, $F(1, 76) = 4.85$, $p = .03$, $d = .63$ (.09–1.18), and $F(1, 76) = 1.28$, $p = .26$, $d = .32$ (-.19–.84), respectively.

Mood Effects at Time 2

As at Time 1, a cut-off score of 10 on the GDS was used to divide the total sample into three groups; comparisons were made between the PD subgroups and the nondepressed controls. Out of the total group of PD participants, 39% (16) still showed signs of depression at Time 2. A further 17% (7), who were not depressed at Time 1, met the criteria for depression at Time 2. Only 5% (2) of PD participants, who were depressed at Time 1, were not classified as depressed at Time 2. In the HC group, 2% (1) were still depressed, but a further 7% (3), who were not depressed at Time 1, met the criteria at Time 2.

As can be seen in Table 23, the depressed and nondepressed PD subgroups had approximately the same ratio of males-to-females as the controls, χ^2 (1, $N = 60$) = .98, $p = .32$, $d = .27$ (-.25–.78), and χ^2 (1, $N = 55$) = .22, $p = .64$, $d = .14$ (-.40–

Table 23 *Characteristics of the Depressed Parkinson's Disease (Depressed PD) Group, the Nondepressed Parkinson's Disease (Nondepressed PD) Group, and the Nondepressed Healthy Controls (Nondepressed HC) at Time 2*

	Depressed PD[a]		Nondepressed PD[b]		Nondepressed HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age (years)	70.43	6.66	68.28	5.17	69.46	4.89
Education (years)	3.61	2.74	4.44	3.05	4.32	3.00
Premorbid IQ	110.47	11.59	108.09	8.87	110.09	8.98
OMCT	4.70	3.34	2.94	3.15	2.54	2.43
GDS	14.17	2.81	5.28	3.51	3.14	2.64
Disease duration	10.46	6.52	7.17	3.89	-	-
Age at onset	59.98	10.48	61.11	5.70	-	-
Levodopa	483.70	217.27	370.31	210.05	-	-
mADL	12.09	3.64	7.12	3.57	-	-

Note. Premorbid IQ = Premorbid full-scale IQ (WAIS-R) estimated from National Adult Reading Test; OMCT = Orientation-Memory-Concentration Test; GDS = Geriatric Depression Scale; Levodopa = mean daily levodopa dose in milligrams; mADL = modified Activities of Daily Living Scale.
[a]*n* = 16 Males, 7 Females. [b]*n* = 9 Males, 9 Females. [c]*n* = 21 Males, 16 Females.

.67), respectively. Neither of the PD subgroups differed from the nondepressed controls in terms of age, education, or premorbid IQ (*F*s < 1). The depressed PD group had a higher mean OMCT dementia rating compared to the controls, $F(1, 75) = 7.88$, $p = .006$, $d = .77$ (.23–1.30). But, the nondepressed PD group did not differ from the controls ($F < 1$). As expected, the depressed PD group had a higher level of depression than the controls, $F(1, 75) = 204.69$, $p < .001$, $d = 4.08$ (3.18–4.99). Finally, as at Time 1, the nondepressed PD group also had a higher level of depression than the controls, $F(1, 75) = 6.58$, $p = .01$, $d = .73$ (.17–1.28).

For both the PD subgroups, mean disease duration, age at disease onset, daily levodopa dose, and mADL are shown in Table 23. In terms of disease stage, the depressed PD group ranged from stage I to IV (I: $n = 1$; II: $n = 10$; III: $n = 10$; IV: $n = 2$), whereas the nondepressed PD group ranged from stage I to III (I: $n = 7$; II: $n = 7$; III: $n = 4$).

Table 24 summarises the results from the memory tasks. The within-groups comparisons between the two difficulty levels of the NRMT indicated that this factor was effective for the depressed PD group, the nondepressed PD group, and the controls, $F(1, 22) = 7.13$, $p = .01$, $d = 1.17$ (.42–1.91), $F(1, 17) = 6.16$, $p = .02$,

Table 24 *Mean PC and SD for the Verbal and Nonverbal Recognition Memory Tasks for the Depressed Parkinson's Disease (Depressed PD) Group, the Nondepressed Parkinson's Disease (Nondepressed PD) Group, and the Nondepressed Healthy Controls (Nondepressed HC) at Time 2*

	Depressed PD[a]		Nondepressed PD[b]		Nondepressed HC[c]	
	M	*SD*	*M*	*SD*	*M*	*SD*
NRMT - hard	62.62	12.79	62.67	13.97	66.49	11.78
NRMT - easy	70.82	12.31	70.22	8.16	72.59	11.83
VRMT	81.02	7.90	81.72	9.27	85.96	8.15
KOLTQ	91.39	14.06	93.61	14.72	105.76	9.33
PMQT	1.91	1.50	2.34	1.53	3.37	0.94
PMOT	3.22	1.00	3.72	0.67	3.39	0.92

Note. NRMT-hard = Nonverbal Recognition Memory Task - hard level of difficulty; NRMT-easy = Nonverbal Recognition Memory Task - easy level of difficulty; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task. [a]$n = 23$. [b]$n = 18$. [c]$n = 37$.

$d = 1.02$ (.32–1.72), and $F(1, 36) = 10.09$, $p = .003$, $d = .73$ (.35–1.10), respectively.

Both the depressed PD group and the nondepressed PD group had a small impairment in nonverbal recognition at the more difficult level of the task, $F(1, 75) = 1.33$, $p = .25$, $d = .32$ (-.20–.84), and $F(1, 75) = 1.11$, $p = .30$, $d = .30$ (-.24–.85), respectively. However, at the easier level there was no evidence of a difference between either of the PD subgroups and the controls (Fs < 1).

Compared to the controls, both the depressed PD group and the nondepressed PD group had verbal recognition deficits, $F(1, 75) = 4.96$, $p = .03$, $d = .61$ (.09–1.14), and $F(1, 75) = 3.12$, $p = .08$, $d = .50$ (-.05–1.04), respectively.

Similarly, the depressed and the nondepressed PD groups had impaired recall, $F(1, 75) = 19.66$, $p < .001$, $d = 1.27$ (.70–1.83), and $F(1, 75) = 12.00$, $p = .001$, $d = 1.07$ (.50–1.65), respectively.

Both the depressed PD group and the nondepressed PD group had inferior prospective memory for a question when compared to the controls, $F(1, 75) =$

18.44, $p < .001$, $d = 1.23$ (.67–1.79), and $F(1, 75) = 7.87$, $p = .006$, $d = .89$ (.32–1.45), respectively. Prospective memory for an object was not impaired in the depressed PD group ($F < 1$) or the nondepressed PD group. In fact, the latter group had better scores, on average, compared to the controls, $F(1, 75) = 1.70$, $p = .20$, $d = -.39$ (-.93–.15).

In summary, the ESs computed for both the depressed and the nondepressed participants indicate that both PD subgroups suffered from deficits in recognition, recall, and prospective memory for a question (Table 25). Moreover, the depressed PD group generally had larger memory deficits than the nondepressed PD group. With respect to memory decline, both PD subgroups showed decline in recall and prospective memory for a question, but not recognition or prospective memory for an object.

Table 25 *Summary of Effect Sizes From the Memory Tasks as a Function of Time for the Comparisons Involving the Depressed and Nondepressed Parkinson's Disease (PD) Groups and the Nondepressed Healthy Controls*

	Depressed PD[a]		Nondepressed PD[b]	
	d	95% CI	*d*	95% CI
NRMT - hard level				
Time 1	.68	.13–1.23	.63	.11–1.16
Time 2	.32	-.20– .84	.30	-.24– .85
NRMT - easy level				
Time 1	.66	.11–1.21	.28	-.23– .80
Time 2	.15	-.37– .63	.22	-.32– .76
VRMT				
Time 1	.90	.34–1.46	.21	-.30– .73
Time 2	.61	.09–1.14	.50	-.05–1.04
KOLTQ				
Time 1	.94	.37–1.50	.62	.09–1.14
Time 2	1.27	.70–1.83	1.07	.50–1.65
PMQT				
Time 1	.39	-.15– .93	.00	-.51– .51
Time 2	1.23	.67–1.79	.89	.32–1.45
PMOT				
Time 1	.63	.09–1.18	.32	-.19– .84
Time 2	.18	-.34– .69	-.39	-.93– .15

Note. d = ES statistic (Cohen, 1988); 95% CI = 95% confidence interval for d; NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a] n = 18 PD, 38 HC (Time 1); 23 PD, 37 HC (Time 2). [b] n = 23 PD, 38 HC (Time 1); 18 PD, 37 HC (Time 2).

Effects of Age at Onset

To examine the potential moderating effect of age at symptom onset, planned comparisons were made between the early-onset PD group and the early-onset HC group, and between the late-onset PD group and the late-onset HC group.

The early- and late-onset PD groups did not differ from their respective controls with regard to gender (39% female in each group), age, education, or premorbid IQ (*F*s < 1) (see Table 26). Both the early- and late-onset PD groups had higher mean dementia ratings than their respective controls, $F(1, 34) = 2.91$, $p = .10$, $d = .57$ (-.12–1.26), and $F(1, 44) = 3.67$, $p = .06$, $d = .57$ (-.04–1.17), respectively. Finally, the level of depression was greater in the early-onset PD group compared to their respective controls, both at Time 1, $F(1, 34) = 9.41$, $p = .004$, $d = 1.02$ (.31–1.74), and at Time 2, $F(1, 34) = 16.02$, $p < .001$, $d = 1.33$ (.59–2.08).

Table 26 *Summary of Demographic and Neuropsychological Characteristics as a Function of Time and Age at Onset for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Early-onset[a]			
Age (years)	65.50	5.70	58 – 79	66.28	3.85	60 – 74
Education (years)	4.56	2.99	0 – 10	4.61	3.74	0 – 10
Premorbid IQ	109.03	11.14	91 –126	111.28	9.43	89 –124
OMCT	3.78	3.28	0 – 10	2.22	2.05	0 – 6
GDS						
Time 1	10.44	6.55	0 – 22	4.61	4.70	0 – 17
Time 2	10.44	6.34	1 – 20	3.50	3.75	0 – 15
			Late-onset[b]			
Age (years)	72.61	4.34	64 – 79	72.22	4.80	65 – 82
Education (years)	3.52	2.79	0 – 10	3.43	2.63	0 – 9
Premorbid IQ	109.74	10.08	92 –128	109.20	9.22	89 –124
OMCT	4.04	3.44	0 – 10	2.43	2.09	0 – 8
GDS						
Time 1	9.48	5.07	0 – 21	5.26	4.05	0 – 13
Time 2	10.13	4.76	1 – 18	4.52	4.39	0 – 16

Note. Premorbid IQ = Premorbid full-scale IQ (WAIS-R) estimated from National Adult Reading Test; OMCT = Orientation-Memory-Concentration Test; GDS = Geriatric Depression Scale.
[a]*n* = 11 males and 7 females in each group. [b]*n* = 14 males and 9 females in each group.

Similarly, the level of depression was greater in the late-onset PD group compared to their controls, both at Time 1, $F(1, 44) = 9.72$, $p = .003$, $d = .92$ (.30–1.54), and at Time 2, $F(1, 44) = 17.26$, $p < .001$, $d = 1.23$ (.58–1.87).

Mean disease duration, age at symptom onset, daily levodopa dose, and mADL for the two PD subgroups are shown in Table 27. In terms of disease stage, the early-onset PD group ranged from stage I to IV, both at Time 1 (I: $n = 3$; II: $n = 8$; III: $n = 6$; IV: $n = 1$) and at Time 2 (I: $n = 3$; II: $n = 9$; III: $n = 5$; IV: $n = 1$). The late-onset PD group ranged from stage I to III, both at Time 1 (I: $n = 3$; II: $n = 11$; III: $n = 9$) and at Time 2 (I: $n = 5$; II: $n = 8$; III: $n = 9$; IV: $n = 1$).

The within-groups comparisons showed the manipulation of the NRMT difficulty factor to be effective at Time 1 for the early-onset PD and HC groups, and the late-onset PD and HC groups, $F(1, 17) = 19.77$, $p < .001$, $d = 1.35$ (.78–1.93), $F(1, 17) = 5.55$, $p = .03$, $d = 1.01$ (.44–1.59), $F(1, 22) = 1.13$, $p = .30$, $d = .36$ (-.08–.81), and $F(1, 22) = 1.43$, $p = .24$, $d = .81$ (.25–1.37), respectively. Similarly, at Time 2 the difficulty factor was effective for the early-onset PD and HC groups, and the late-onset PD and HC groups, $F(1, 17) = 5.98$, $p = .03$, $d = 1.09$ (.50–1.68), $F(1, 17) = 7.13$, $p = .02$, $d = 1.33$ (.69–1.97), $F(1, 22) = 7.41$, $p = .01$, $d = 1.15$ (.62–1.69), and $F(1, 22) = 9.67$, $p = .01$, $d = .98$ (.50–1.45), respectively.

Table 27 *Summary of Disease Characteristics for the Early- and Late-Onset Parkinson's Disease Groups*

	Early-onset				Late-onset			
	n	*M*	*SD*	*Range*	*n*	*M*	*SD*	*Range*
Age at onset	18	52.97	6.44	38 – 59	23	66.35	4.62	61 – 77
Disease duration	18	12.53	5.97	4 – 25	23	6.26	3.68	1 – 14
Levodopa								
Time 1	16	439.84	234.75	100–1000	21	407.74	213.07	125–1100
Time 2	18	429.17	233.89	100– 825	21	444.05	210.90	188–1000
mADL								
Time 1	17	10.00	3.64	3 –16	17	10.00	4.65	0 –18
Time 2	18	10.78	4.51	2 –20	22	9.32	4.20	0 –17

Note. Levodopa = mean daily levodopa dose in milligrams; mADL = modified Activities of Daily Living Scale.

Table 28 indicates that the early-onset PD group had inferior nonverbal recognition at the more difficult level of the task compared to their respective controls, both at Time 1, $F(1, 34) = 6.09$, $p = .02$, $d = .82$ (.12–1.52), and at Time 2, $F(1, 34) = 1.59$, $p = .22$, $d = .42$ (-.26–1.10). For the late-onset PD group the evidence was inconsistent, with a small deficit at Time 1, $F(1, 44) = 1.94$, $p = .17$, $d = .41$ (-.19–1.01), but not at Time 2, ($F < 1$).

With respect to the easier level of the NRMT, the early-onset PD group had a small deficit in nonverbal recognition, both at Time 1, $F(1, 34) = 1.45$, $p = .24$, $d = .40$ (-.28–1.08), and at Time 2, $F(1, 34) = 1.63$, $p = .21$, $d = .43$ (-.25–1.11). Again, the evidence was inconsistent for the late-onset PD group, with a small deficit at Time 1, $F(1, 44) = 1.40$, $p = .24$, $d = .35$ (-.25–.94), but not at Time 2, ($F < 1$).

The early-onset PD group had poorer verbal recognition compared to their respective controls, both at Time 1 and at Time 2, $F(1, 34) = 3.03$, $p = .09$, $d = .58$ (-.11–1.27), and $F(1, 34) = 3.73$, $p = .06$, $d = .64$ (-.05–1.33), respectively (see

Table 28 *Summary of Results From the Nonverbal Recognition Memory Task as a Function of Difficulty Level, Time, and Age at Onset for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Early-onset[a]			
Hard level						
Time 1	58.89	10.28	40 – 70	67.44	10.52	53 – 90
Time 2	62.50	16.05	35 – 85	68.28	10.97	53 – 85
Easy level						
Time 1	69.61	11.13	50 – 85	73.94	10.42	55 – 95
Time 2	71.56	10.79	55 – 90	75.89	9.52	65 –100
			Late-onset[b]			
Hard level						
Time 1	63.55	10.40	50 – 85	67.61	9.34	50 – 90
Time 2	62.75	10.75	40 – 82	63.26	11.84	40 – 85
Easy level						
Time 1	66.48	13.84	31 – 89	70.39	7.76	55 – 82
Time 2	69.77	10.56	47 – 90	70.65	10.13	45 – 95

Note. Values are scores on the NRMT expressed as percent correct.
[a] $n = 11$ males and 8 females in each group. [b] $n = 14$ males and 9 females in each group.

Table 29). Similarly the late-onset PD group had impaired verbal recognition, both at Time 1, $F(1, 44) = 4.85$, $p = .03$, $d = .65$ (.04–1.26), and at Time 2, $F(1, 44) = 3.38$, $p = .07$, $d = .54$ (-.06–1.14).

The early-onset PD group had impaired recall, both at Time 1 and at Time 2, $F(1, 34) = 3.21$, $p = .08$, $d = .60$ (-.09–1.28), and $F(1, 34) = 7.40$, $p = .01$, $d = .91$(.20–1.61), respectively (see Table 30). Similarly, the late-onset PD group were impaired, both at Time 1 and at Time 2, $F(1, 44) = 4.85$, $p = .03$, $d = .65$ (.04–1.26), and $F(1, 44) = 20.54$, $p < .001$, $d = 1.34$ (.68–1.99), respectively.

Table 30 shows that the early-onset PD group did not have impaired prospective memory for a question at Time 1 ($F < 1$), but at Time 2, there was strong evidence to suggest a deficit, $F(1, 34) = 10.12$, $p = .003$, $d = 1.06$ (.34–1.78). In addition, the late-onset PD group had impaired prospective memory for a question, both at Time 1, $F(1, 44) = 1.73$, $p = .20$, $d = .39$ (-.21–.99), and at Time 2, $F(1, 44) = 9.84$, $p = .003$, $d = .92$ (.30–1.54).

Turning to the prospective memory for an object task, there was no evidence that the early-onset PD group performed differently from their control group, either at Time 1 ($F < 1$) or at Time 2 (early-onset PD mean score was greater than that of

Table 29 *Summary of Results From the Verbal Recognition Memory Task as a Function of Age at Onset and Time for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Early-onset[a]			
Time 1	80.31	7.96	63 – 98	85.31	9.24	70 –100
Time 2	82.28	8.25	67 –100	87.81	8.91	74 –100
			Late-onset[b]			
Time 1	78.89	8.66	59 – 89	84.35	8.13	70 –100
Time 2	80.59	8.67	65 –100	85.13	8.08	70 – 98

Note. Values are scores on the VRMT expressed as the percent correct.
[a] $n = 11$ males and 7 females in each group. [b] $n = 14$ males and 9 females in each group.

Table 30 *Summary of Results From the Prospective Memory and Kendrick Object Learning Tasks as a Function of Age at Onset and Time for the Parkinson's Disease (PD) and Healthy Control (HC) Groups*

	PD			HC		
	M	*SD*	*Range*	*M*	*SD*	*Range*
			Early-onset[a]			
KOLTQ						
Time 1	93.56	14.87	61 –123	102.11	13.74	77 –127
Time 2	93.06	17.21	62 –134	105.56	9.17	89 –123
PMQT						
Time 1	3.33	0.97	1 – 4	3.22	1.22	0 – 4
Time 2	2.45	1.46	0 – 4	3.63	0.59	2 – 4
PMOT						
Time 1	3.22	0.94	2 – 4	3.28	0.96	2 – 4
Time 2	3.67	0.77	2 – 4	3.22	1.00	2 – 4
			Late-onset[b]			
KOLTQ						
Time 1	96.70	11.94	77 –123	104.74	12.81	87 –133
Time 2	91.83	11.73	73 –131	106.57	10.28	87 –126
PMQT						
Time 1	2.30	1.40	0 – 4	2.87	1.52	0 – 4
Time 2	1.83	1.53	0 – 4	3.06	1.11	1 – 4
PMOT						
Time 1	3.13	1.06	1 – 4	3.52	0.85	1 – 4
Time 2	3.26	0.96	2 – 4	3.39	0.94	2 – 4

Note. KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a] n = 11 males and 7 females in each group. [b] n = 14 males and 9 females in each group.

their control's mean score). Conversely, the late-onset PD group had a deficit in prospective memory for an object at Time 1, $F(1, 44) = 1.92$, $p = .17$, $d = .41$ (-.19–1.00), but not at Time 2 ($F < 1$).

In summary, the magnitude of the ESs computed for both the early-onset and the late-onset participants indicated that both PD subgroups suffered from a variety of memory disturbances (see Table 31). The early-onset PD group had consistently greater deficits in nonverbal recognition than the late-onset group, but the opposite occurred for recall and prospective memory for an object. In addition, deficits in verbal recognition and prospective memory for a question were smaller in the early-onset group at Time 1, but larger at Time 2.

Table 31 *Summary of Effect Sizes From the Memory Tasks as a Function of Time for the Comparisons Involving the Early- and Late-Onset Parkinson's Disease (PD) Groups and Their Respective Healthy Control Groups*

	Early-Onset PD[a]		Late-Onset PD[b]	
	d	95% CI	*d*	95% CI
NRMT - hard level				
Time 1	.82	.12–1.52	.41	-.19–1.01
Time 2	.42	-.26–1.10	.05	-.55– .64
NRMT - easy level				
Time 1	.40	-.28–1.08	.35	-.25– .94
Time 2	.43	-.25–1.11	.09	-.51– .68
VRMT				
Time 1	.58	-.11–1.27	.65	.04–1.26
Time 2	.64	-.05–1.33	.54	-.06–1.14
KOLTQ				
Time 1	.60	-.09–1.28	.65	.04–1.26
Time 2	.91	.20–1.61	1.34	.68–1.99
PMQT				
Time 1	-.10	-.77– .56	.39	-.21– .99
Time 2	1.06	.34–1.78	.92	.30–1.54
PMOT				
Time 1	.06	-.61– .74	.41	-.19–1.00
Time 2	-.50	-1.19– .18	.14	-.46– .73

Note. *d* = ES statistic (Cohen, 1988); 95% CI = 95% confidence interval for *d*; NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a]*n* = 18 in each group. [b]*n* = 23 in each group.

Memory decline, as indicated by a substantially larger ES at the second session, was evident in both subgroups, but only for recall and prospective memory for a question (Table 31).

Correlation Analysis

Correlation coefficients were calculated between the memory scores and H&Y stage, GDS scores, and age at onset. These correlations provided an indication of the strength of the relationship between these potential moderators and memory. A preliminary analysis indicated that correlations were similar across difficulty level (for NRMT) and across Time. Therefore, the correlations reported here are computed from task scores averaged across these factors. In addition,

correlations were computed between Time 1 and Time 2 for all memory measures to provide an indication of test-retest reliability.

Disease stage as measured by the H&Y was negatively associated with performance on all memory tasks (Table 32). The size of the correlation coefficient was weak for recall memory, but could be described as small-to-medium for the other measures (Cohen, 1988). Depression was negatively associated with all of the memory tasks, and the size of effect was small-to-medium. Finally, age at onset was also negatively correlated with the memory tasks, with the exception of recall. Again, the magnitude of these effects can be described as small-to-medium.

Table 33 presents the test-retest reliability coefficients and their associated 95% confidence intervals for each memory task. In terms of performance on the recall and recognition tasks, the correlations between Time 1 and Time 2 were medium-to-large (Cohen, 1988), with the exception of the VRMT for the female PD group (r = .26), and the KOLTQ for the male HC group (r = .25). In contrast, for every group the prospective memory task coefficients indicated that both tasks had relatively poor external reliability.

Table 32 *Zero-Order Correlations Between Memory Scores and Hoehn and Yahr (H&Y) Stage, Age at Onset, and Geriatric Depression Scale (GDS) Score (Averaged Across Time)*

	H&Y		GDS		Age at Onset	
	r	95% CI[a]	*r*	95% CI[a]	*r*	95% CI[a]
NRMT[b]	-.32 *	-.62–-.02	-.16	-.49– .18	-.11	-.51– .29
VRMT	-.22	-.50– .06	-.29	-.55–-.03	-.09	-.38– .19
KOLTQ	-.05	-.33– .24	-.26	-.59– .06	.16	-.10– .42
PMQT	-.34 *	-.60–-.08	-.20	-.47– .07	-.32 *	-.57–-.06
PMOT	-.35 *	-.61–-.09	-.36 *	-.57–-.16	-.17	-.47– .14

Note. NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a] 95% confidence interval of r (N = 41). [b]Total score across difficulty level.
*p < .05.

Table 33 *Test-Retest Reliability Coefficients for Each Group as a Function of Gender*

	PD		HC	
	r	95% CI[a]	*r*	95% CI[a]
		Males		
NRMT[b]	.55 **	.30–.79	.47 *	.17–.78
VRMT	.65 **	.41–.89	.58 **	.23–.93
KOLTQ	.71 **	.52–.90	.25	-.09–.59
PMQT	.26	-.08–.60	.07	-.42–.56
PMOT	-.11	-.51–.30	.28	-.12–.68
		Females		
NRMT[b]	.51 *	.12–.91	.51 *	.12–.90
VRMT	.26	-.27–.78	.54 *	.26–.82
KOLTQ	.76 **	.51–.99	.72 **	.53–.91
PMQT	.08	-.04–.55	-.09	-.37–.20
PMOT	.31	-.13–.74	-.11	-.39–.18

Note. NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMQT = Prospective Memory Question Task; PMOT = Prospective Memory Object Task.
[a] 95% confidence interval of *r* (*N* = 41). [b]Total score across difficulty level.
*$p < .05$. **$p < .01$.

Principal Components Analysis

Five memory tasks (NRMT, VRMT, KOLT, PMQT, PMOT) were used to measure different aspects of memory on two separate occasions. To provide an empirical summary of this relatively large data set, principal components analysis with varimax rotation was used to produce a smaller set of components for further data analysis.

Loadings of variables on components, communalities, and percent of variance are shown in Table 34. Variables are grouped by size of loading to facilitate interpretation. Loadings under .45 (20% of variance) were replaced by zeros. Four components (with eigenvalues > 1) were extracted from the data set. The first component, labelled as recognition, was derived from the two recognition memory tasks (VRMT and NRMT) measured at Time 1 and Time 2. The second component, labelled as recall, was derived from the KOLTQ scores measured at

Table 34 *Component Loadings*[a]*, Communalities (h^2), and Percent of Variance Derived From Principal Components Analysis with Varimax Rotation for all Participants on the Memory Tasks*

Task	Component 1	2	3	4	h^2
NRMT (Time 2)	*.85*	.00	.00	.00	.73
NRMT (Time 1)	*.74*	.00	.00	.00	.67
VRMT (Time 2)	*.63*	.00	.00	.00	.63
VRMT (Time 1)	*.50*	.00	.00	.00	.51
KOLTQ (Time 1)	.00	*.87*	.00	.00	.78
KOLTQ (Time 2)	.00	*.85*	.00	.00	.80
PMOT (Time 2)	.00	.00	.00	*.80*	.68
PMQT (Time 2)	.00	.00	.00	*.71*	.65
PMQT (Time 1)	.00	.00	*.80*	.00	.52
PMOT (Time 1)	.00	.00	*.62*	.00	.74
Percent of variance	20.51	19.54	11.34	15.77	
Proposed identity of component	Recognition	Recall	Initial Prospective Memory	Retest Prospective Memory	

Note. NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task; KOLTQ = Kendrick Object Learning Task - age-scaled quotient; PMOT = Prospective Memory Object Task; PMQT = Prospective Memory Question Task.
[a]Loadings under .45 (20% of variance) are replaced by zeros. [b]Italicised loadings are the highest within a column and were used in the interpretation of the identity of the factor.

Time 1 and Time 2. The third component, labelled as initial prospective memory, was derived from the two prospective memory tasks (PMQT and PMOT) measured at Time 1. The fourth component, labelled as retest prospective memory, was derived from the two prospective memory tasks (PMQT and PMOT) measured at Time 2.

Gender by group MANOVA. The four components were subjected to a 2 x 2 (Gender x Group) MANOVA. The results indicated a large main effect for group, $F(4, 75) = 6.31$, $p < .001$, $d = 1.16$ (.69–1.63), and a two-way (Group x Gender) interaction, $F(4, 75) = 1.03$, $p = .40$, $d = .47$ (.02–.91), but no main effect for gender ($F < 1$). Univariate tests revealed that for the recognition component, there was a small main effect for group, $F(1, 78) = 3.30$, $p = .07$, $d = .41$ (-.03–.85), but no two-way (Group x Gender) interaction ($F < 1$). Inspection of the marginal

means indicated that the PD group had impaired recognition memory relative to the HC group (M = -.21, SD = 1.02, and M = .20, SD = 1.00, respectively).

For the recall component, there was a large main effect for group, $F(1, 78)$ = 17.73, $p < .001$, d = .95 (.49–1.42), and a possible two-way interaction, $F(2, 78)$ = 1.41, p = .25, d = .38 (-.06–.82). However, from Figure 14, one can see that the size of the deficit was similar for both males and females, thus, only the main effect was interpreted.

For the initial prospective memory component there was no main effect for group ($F < 1$), but a possible two-way interaction, $F(1, 78)$ = 1.05, p = .35, d = .33 (-.11–.77). Figure 14 suggests that the male PD group had poorer prospective memory at Time 1 compared to their respective control group, $t(46)$ = 1.59, p = .12, d = .46 (-.13–1.04). However, there was no evidence of the same deficit in the female PD participants, $t(32)$ = .25.

Finally, for the retest prospective memory component, the analysis indicated a possible small main effect for group, $F(1, 78)$ = 1.69, p = .20, d = .29 (-.15–.74),

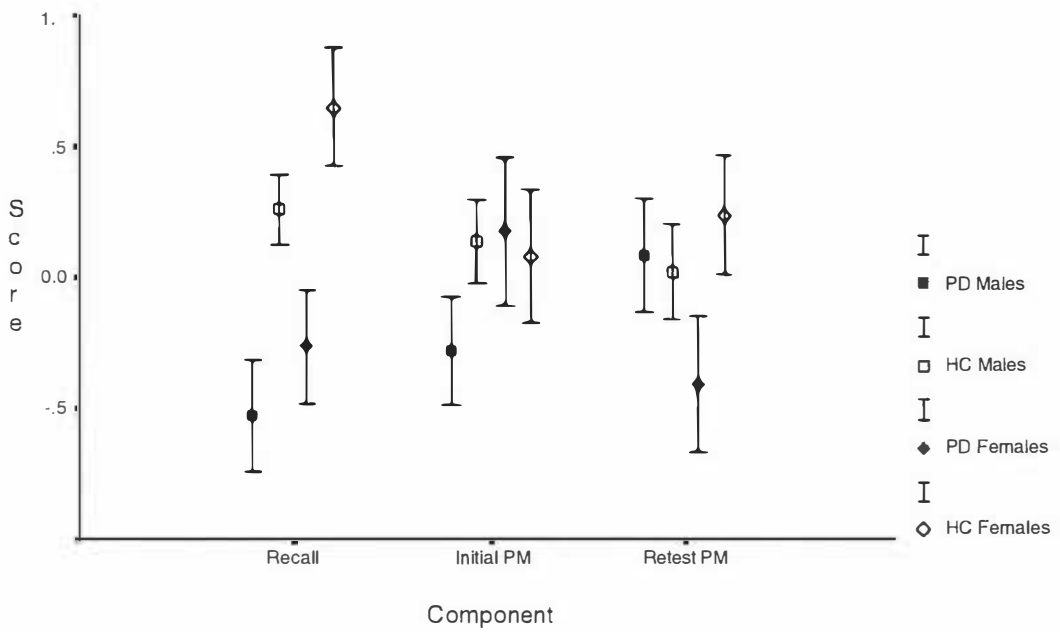


Figure 14. Mean component scores and *SE* (vertical bars) for the recall, initial prospective memory (PM), and retest PM components as a function of the Parkinson's disease (PD) and healthy control (HC) gender subgroups.

and a small two-way interaction, $F(1, 78) = 2.52$, $p = .12$, $d = .36$ (-.08–.80). As can be seen in Figure 14, there was no evidence that the male PD group had impaired prospective memory at Time 2 relative to the male HC group, $t(46) = .22$, but there was evidence of a deficit in the female PD group, $t(32) = 1.88$, $p = .07$, $d = .64$ (-.07–1.36).

Disease Stage MANOVA. For each PD disease stage subgroup, the four components were subjected to MANOVA, and where appropriate, ANOVA. For both the early PD group and the advanced PD group, the MANOVA indicated a main effect, $F(4, 57) = 2.64$, $p = .04$, $d = .91$ (.38–1.44), and $F(4, 56) = 9.12$, $p < .001$, $d = 1.72$ (1.12–2.31), respectively.

The univariate tests indicated that for the early PD group there was a main effect for recall, $F(1, 60) = 9.69$, $p = .003$, $d = .85$ (.32–1.38), but not for recognition or prospective memory (Fs < 1). Conversely, the advanced PD group produced a main effect for recognition, $F(1, 59) = 6.23$, $p = .02$, $d = .69$ (.16–1.22), for initial prospective memory, $F(1, 59) = 1.84$, $p = .18$, $d = .37$ (-.14–.89), for retest prospective memory, $F(1, 59) = 2.84$, $p = .10$, $d = .47$ (-.05–.99), as well as for recall, $F(1, 59) = 18.06$, $p < .001$, $d = 1.18$ (.62–1.73).

Discussion

Several important findings can be taken from the results of the present study. First, recognition, recall, and prospective memory may be affected by PD, but factors such as gender, disease stage, depression, and age at onset could moderate the size of these memory deficits. The most important moderator appears to be disease stage, suggesting that the progression of motor deficits operates in parallel with memory impairment. Depression seems to have a similar relationship with memory impairment, but is also associated with the motor deficits, complicating the issue. Gender and age at onset seem to have task specific affects, but these may be spurious. The data also indicate that task

difficulty may be important, but gender and disease stage may interact with difficulty.

To facilitate the interpretation of the results, the discussion below follows the format used in the results section. Also, given the large number of statistical tests conducted in the present analysis, more importance was placed on those results that were consistent across time. Before beginning the main discussion, a description of the group characteristics as a function of gender is summarised and discussed.

Participant Description as a Function of Gender

The 25 male and 16 female PD participants were generally well matched to their respective controls in terms of age, education, premorbid IQ, and handedness. The PD participants were also relatively high functioning as assessed by the OMCT, and had mild-to-moderate physical symptoms. Moreover, none of the participants were rated as demented and none were assessed as H&Y stage V. For all but three participants the diagnosis of PD was confirmed by a neurologist, and an independent assessment indicated that all but one participant was rated as clinically definite or probable PD. The mean NART estimate of premorbid IQ for the PD group was similar to that expected for their age group (Ivnik et al., 1992). Both the male and the female PD groups had consistent deficits in semantic fluency. In terms of letter fluency, only the female PD group had a consistent deficit. However, this finding may be accounted for by the NART estimate of premorbid letter fluency which indicated a considerable preexisting difference between the PD females and their respective controls. The finding of impaired semantic fluency, but relatively normal letter fluency has been reported in a number of previous studies (e.g., Auriacombe et al., 1993; Fama et al., 1998). Nevertheless, this apparent dissociation may be due to the nature of the task categories, rather than damaged semantic memory and intact lexical access (Azuma et al., 1997).

Anxiety measured before and after the experimental session was generally low, and there was little evidence that the PD participants differed from the controls at the end of the session. Nevertheless, at the beginning of the session the PD females had greater levels of anxiety than their respective controls. This appears to be due to very low levels of reported anxiety in the female controls, rather than elevated anxiety in the female PD participants. With regard to vigilance, the sensitivity measure indicated that performance was very high; no participant scored below .70. Nevertheless, both the male and the female PD groups' performance was poorer than that of their control groups, while response bias was generally minimal. This suggests a deficit in sustained attention, rather than a change in response criterion (Sheer & Schrock, 1986). However, given that sensitivity was generally high, the poorer performance of the PD participants may reflect differences in the strategy used to complete the task. Previous research has shown that strategies based on internal resources to guide behaviour produce the best performance on the Mackworth clock (Giambra et al., 1988), and it is precisely these strategies which may be impaired in PD (Taylor & Saint-Cyr, 1992). However, the present data do not provide convincing evidence that fewer PD participants used internal strategies compared to the controls. In fact, the vast majority in all groups did not use internal strategies. Because few, if any, studies have examined the Mackworth clock strategies in PD, clarification of this matter awaits further research.

Most PD participants were stabilised on a levodopa preparation, although males took higher doses, on average, than females. Of those who were taking antiparkinsonian medication at the time of testing, the majority rated its effectiveness as being good. Also, although participants were excluded from the analysis if they had ever suffered from major depression before the onset of PD, depression at the time of testing was elevated relative to the controls. Furthermore, scores on the GDS were comparable to previous PD studies (e.g., Dewick et al., 1991; Owen et al., 1992, 1993). Using the cut-off suggested by Yesavage et al. (1983), 44% of the male participants met the criteria for possible depression at Time 1 and 65% met the criteria at Time 2. For the female

participants, 44% of the group met the criteria for depression both at Time 1 and at Time 2. Because a sufficient number of participants met the criteria for depression, it was legitimate to divide the total PD sample into two groups to assess the impact of depression on recognition. This was important since few studies, one of which was Owen et al. (1993), have examined this complicated issue.

Nonverbal Recognition Memory as a Function of Task Difficulty and Gender

A major aim of the present study was to examine the role that task difficulty plays in PD memory impairment. In the past, the finding of recall deficits, but intact recognition has been explained in terms of differences in task difficulty. Breen (1993) suggested that if a recognition task was made more difficult, PD participants would begin to show deficits. To examine this issue without confounding task type, a single recognition task (NRMT) was designed with two levels of difficulty. A second aim of the current study was to examine the potential moderating effect of gender on memory. Thus, separate analyses were conducted to assess the performance of male and female PD participants on all experimental tasks.

The data provide evidence that nondemented PD participants may suffer from nonverbal recognition deficits, especially when the task is difficult. Furthermore, the males generally had larger deficits than the females. However, a supplementary analysis indicated that the male and female PD participants differed with respect to a number of potentially important characteristics, complicating the results. Specifically, the male group had a higher mean depression score and more severe disability (H&Y and mADL). Both of these factors have previously been shown to contribute to cognitive dysfunction in PD (Owen et al., 1993; Starkstein et al., 1990); this issue will be taken up again in later sections.

The effectiveness of the manipulation of task difficulty also complicates the present results. At Time 1 the male PD group had impaired nonverbal recognition, and the size of the deficit was greater when the task was most difficult. However, at Time 2 performance was normal at the most difficult level, although performance at the easier level remained inferior to that of the controls. Conversely, the female PD group had normal recognition at Time 1, but impaired recognition at Time 2 with respect to the more difficult level of the task only.

The inconsistency in the recognition deficit may be explained by looking at the effectiveness of the task difficulty factor and the change in performance from Time 1 to Time 2. For the PD males the task difficulty factor was effective at Time 1, and there was evidence of impaired recognition, especially at the more difficult level. However, the task difficulty factor was ineffective at Time 2, performance considerably improving at the more difficult level, but to a lesser extent at the easier level. Meanwhile, the male control group's performance declined at the more difficult level, but remained stable at the easier level. This contributed to a situation where, at Time 2, the male PD group did not appear to have a recognition deficit at the more difficult level, but their performance remained inferior at the easier level. For the female PD group there was a weak manipulation of task difficulty at Time 1, and interestingly there was no evidence of a deficit in recognition. However, at Time 2 the task difficulty factor was effectively manipulated, and the female PD group showed signs of impaired recognition at the more difficult level, but not at the easier level.

The most immediate factor to consider is that there probably was insufficient separation in performance levels between the easy and more difficult level–such that random variation (probably accounting for control group Time 1 to Time 2 differences) could confound the outcome. A supplementary analysis collapsing across time revealed that for the PD males, the recognition deficit was slightly smaller at the more difficult level (d = .53, -.06–1.12) than at the easier level (d = .64, .05–1.23). For the PD females, the data showed a larger deficit at the more difficult level (d = .44, -.27–1.14) than at the easier level (d = .10, -.59–.79).

Therefore, support for the effortfulness/automatic distinction was seen only in female PD group.

In this respect, the data support Breen's (1993) suggestion that deficits would be seen if a recognition task was made more difficult. It is possible that for the PD males used here, both the harder and the easier levels were demanding enough to produce deficits. Had the difficulty factor been more effectively manipulated, the present results may have provided stronger support for the moderating effect of task difficulty. Brown and Marsden (1990) proposed that their theory of processing resources accounts for the results seen in a number of different domains, including memory. Applied to the present results, task demands (especially at the more difficult level) may have exceeded the PD participants' central processing resources producing impaired recognition relative to controls.

Task demands have also been shown to be an important factor in the recognition memory deficits seen in alcoholic patients with Korsakoff's disease (Albert et al., 1979), in healthy elderly (Dujardin et al., 1995), and indirectly in PD (e.g., Dewick et al., 1991; Owen et al., 1993). While the present results provide some support for the importance of task demands, an alternative (although not necessarily mutually exclusive) explanation for these findings can be found in the work of Parkin (1993).

Parkin (1993) originally directed his argument at the apparent dissociation between recall and recognition in age-related memory decline (i.e., recall deteriorates with age while recognition does not). He argued that recognition memory involves both explicit and implicit processes, whereas free recall is purely explicit. Thus, preserved implicit processes would facilitate recognition, but not recall. It should be noted that other researchers have made similar suggestions (e.g., Mandler, 1980; Moscovitch, Winocur, & McLachlan, 1986; Richardson-Klavehn & Bjork, 1988; Tulving, 1992). However, Squire (1992, 1994) argued that his own research (Haist, Shimamura, & Squire, 1992; Squire, Shimamura, & Graf, 1985) does not support the idea that recognition typically involves implicit

processes, but did concede that implicit memory might contribute to recognition performance under some conditions. Johnston, Hawley, and Elliott (1991) conducted an extensive research program to examine this issue. They suggested that above chance recognition accuracy can be achieved through perceptual (implicit) memory in situations of minimal explicit memory. "Thus, a reliable level of recognition accuracy does not necessarily indicate a reliable level of explicit memory" (Johnston et al., 1991, p. 221). Parkin claimed that when performing a recognition task, a deficit in explicit memory may be compensated for by increased reliance on implicit processes (i.e., perceptual fluency cues). He went on to argue that there is sufficient evidence to support the involvement of implicit retrieval processes in recognition. The fact that implicit memory declines little with age explains why there is little recognition deficit with age.

This same logic can be applied to PD. Perceptual implicit memory appears to be mediated by the right occipital cortex (Fleischman et al., 1995; Gabrieli et al., 1995), a region not damaged by PD (Lezak, 1995; Wooten, 1984). The implication is that people with PD may sometimes show intact recognition because they are able to use perceptual implicit cues to recognise previously presented information. The plausibility of this suggestion is supported by research demonstrating that perceptual memory is relatively normal in PD, and that recognition is less impaired than recall (e.g., Appollonio et al., 1994; Bondi & Kaszniak, 1991; Bondi et al., 1993).

Appollonio et al. (1994) studied 13 (6 males, 7 females) nondemented PD patients (mean age = 57.2 years, *SD* = 11.6) and 12 (7 males, 5 females) normal controls (mean age = 62.4 years, *SD* = 13.4). Perceptual memory scores, assessed with a picture fragment test, were computed in such a way as to exclude improvement in performance due to skill learning. Recognition memory (hit rate less false alarm rate) was assessed with a yes-no test of 20 items (10 items served as distractors), and recall was assessed with free recall of 90 words. A reanalysis of Appollonio et al.'s results using ESs indicates that the PD participants had normal perceptual learning relative to the controls (*d* = .06, -.75–.88). While the PD group

did show signs of impaired recognition (d = .60, -.24–1.43), the deficit was smaller than that for free recall (d =.91, .05–1.77). This finding supports the earlier work of Bondi and Kaszniak (1991) and Bondi et al. (1993), suggesting that intact implicit memory may account for the dissociation between recall and recognition in PD.

Parkin's (1993) proposal may also account for the detrimental effect of task difficulty on the recognition of PD participants in the present study. That is, task difficulty was created by increasing the perceptual similarity between targets and distractors, thus reducing the efficacy of perceptual fluency cues. Therefore, at the easier level of the task the PD participants may have compensated for their explicit memory deficit by using intact implicit processes, but this may have been more difficult to do at the harder level. Thus, smaller recognition deficits would be seen at the easier level relative to the harder level. Unfortunately, the present study was not designed to distinguish between the role of implicit memory and the automatic/effortful distinction.

In the discussion above, Parkin's (1993) proposal was applied to nonverbal recognition, but it can also be applied to tasks that use verbal stimuli. Therefore, the general view that recall is dysfunctional, whereas verbal recognition is relatively intact may also be accounted for in terms of implicit memory processes. To examine verbal recognition memory functioning in PD, a verbal analogue of the nonverbal task (without the task difficulty factor) was used, and the results from this task are the focus of the next section. Therefore, Parkin's implicit memory proposal will be revisited at the end of the next section.

Verbal Recognition Memory as a Function of Gender

To examine verbal recognition and the potential moderating effect of gender, performance on the VRMT was assessed separately for males and females. The data indicate a consistent verbal recognition deficit both in the PD males and in the PD females, and, furthermore, the deficit was greater in the latter group.

Interestingly, the female PD group generally had poorer verbal recognition than the male PD group. These findings contrast sharply with most previous research showing that females have superior verbal memory (Kramer et al., 1988; Lezak, 1995). However, poor premorbid letter fluency in the PD females may be responsible for the current finding. It is also possible that the apparent greater deficit in the PD females is related to the fact that they found the task more difficult, given that the size of a deficit may grow as task difficulty increases. Nevertheless, regardless of the difference between the male and the female participants, performance was generally high on the VRMT. In light of the results from the nonverbal task, had the verbal task been more difficult, still greater deficits may have been seen in the PD participants.

Taken together, the deficits in verbal and nonverbal recognition memory can be contrasted with the frequently cited findings of Lees and Smith (1983), Flowers et al. (1984), and Taylor et al. (1986). As described in chapter 1, Lees and Smith reported that they could find no evidence of impaired word or face recognition memory in PD. A reanalysis of their results using ESs supported their claim. What, then, could account for the discrepancy between the present findings and those of Lees and Smith? It is unlikely that a difference in task difficulty is the explanation, because mean PC for the verbal task used here was approximately the same as that reported by Lees and Smith. A more likely explanation is related to sample differences. That is, in contrast to those studied by Lees and Smith, the present participants were, on average, approximately 10 years older, had longer disease duration, and were more likely to be in the later stages of PD (a factor that may be associated with memory impairment; see chapter 1). In addition, none of Lees and Smith's participants had ever received antiparkinsonian medication, whereas the majority of the present sample were stabilised on a levodopa preparation. Nevertheless, it is unlikely that medication alone could account for all of the difference in performance, because there is little evidence that antiparkinsonian medication directly causes deficits in recognition.

In contrast to Lees and Smith (1983), Flowers et al. (1984) used a PD sample that was generally older, more severely affected by the disease, and medicated. Nevertheless, their sample was still somewhat younger than the present sample. Flowers et al. reported that their results showed that PD participants did not have impaired recognition. However, a reanalysis of their data in terms of ES suggested that deficits in delayed recognition may have existed, but their study did not have sufficient statistical power to detect these (see chapter 1). Furthermore, floor and ceiling effects may have obscured the true extent of the deficits (Sahakian et al., 1988). Re-couching the Flowers et al. data in ESs actually brings them into line with the results of the present study.

The PD sample used by Taylor et al. (1986) was similar to those used here with respect to disease duration, disease stage, and medication, although Taylor et al.'s participants were generally younger. As with Flowers et al. (1984), Taylor et al. claimed that recognition was intact in their sample, although this is not surprising given their poor statistical power. Using ESs, their data provide evidence of a small recognition deficit (see chapter 1). Therefore, their results are also consistent with those of the present study.

While the studies described above have been instrumental in establishing the view that recognition is relatively normal in PD, other investigations have reported recognition deficits in nondemented patients (e.g., Bondi et al., 1993; Cooper et al., 1993; O'Sullivan, 1998; Owen et al., 1993; Sahakian et al., 1988). However, these authors provide a range of explanations for these deficits. For instance, both Bondi et al. and Cooper et al. suggested that frontal system dysfunction accounted for the recognition deficit. In addition, Owen et al. suggested that inadequate exclusion of demented participants may be responsible for the observed deficits in the study of Sahakian et al. Nevertheless, the present meta-analysis based on 16 studies (chapter 2), indicated the possibility of a small recognition deficit in more advanced stage (medicated) PD participants. Of particular interest here is the fact that the studies listed above (that reported deficits) used PD samples that were similar to the present sample in terms of

disease stage, a factor that may be important, and this issue will be taken up again in a later section.

Returning to the idea that implicit memory processes may facilitate recognition memory (Parkin, 1993), the present results show that the deficits in recognition were generally of a smaller magnitude than the recall deficits (discussed in the next section). Thus, it is possible that in the PD participants, deficits in explicit memory processes were compensated for by intact implicit memory, facilitating performance on the recognition tasks. However, small-to-medium recognition deficits remained visible suggesting that, at least in the present sample, implicit cues were not sufficient to return recognition to normal. It should also be noted that the recall task used here was not specifically designed to be comparable to the recognition tasks and so direct comparisons between recall and recognition deficits must be made with caution. Furthermore, the implicit memory proposal may operate in concert with Brown and Marsden's (1990) theory of processing resources; the present results are consistent with the idea that PD memory functioning will become dysfunctional as processing resources are exceeded.

Recall as a Function of Gender

The KOLT was used to measure recall, and scores were converted to age-scaled quotients; an alternate form was used at Time 2. The results show that both the male PD group and the female PD group had medium-to-large recall deficits that were consistent across time. The size of the deficit was somewhat greater in the female PD group, even though the male participants had, on average, poorer recall than the female participants. The stimuli used in the task were pictures of common objects and as such were easily verbalised. Therefore, the impaired performance seen here probably reflects a deficit in verbal recall. The fact that the female participants had superior recall relative to male participants supports past research (see chapter 1). However, the present study is the first to report that female PD participants (relative to controls) may have more severe recall deficits

than PD males. Nevertheless, a general deficit in recall memory is well documented in nondemented PD participants (see Brown & Marsden, 1990; Dubios et al., 1991).

Prospective Memory as a Function of Gender

Prior to the present study, prospective memory in PD has not been investigated. This is surprising since prospective memory plays an important role in everyday activities (see chapter 1). Moreover, prospective memory is most likely subserved by the frontal lobes (McDaniel et al., 1999), and as such, may be at risk in PD. The present study used two event-based tasks to measure prospective memory. The first task (PMQT) involved remembering to ask a question at some future point in time. At the first session, the question concerned the participant's next appointment, whereas at the second session the question concerned the results of the study. These questions were designed to be equally meaningful to the participant, but different so as to reduce practice effects. The second task (PMOT) involved remembering to ask for a personal belonging. Each participant chose the object to be remembered, and no attempt was made to insure that the object differed between Time 1 and Time 2.

The present data indicate that prospective memory for a question was poorer in both PD males and PD females relative to their controls, although the size of the deficit was much greater at Time 2. In addition, the ES was similar for the males and the females. Prospective memory for an object was only impaired in the male PD group at Time 1.

Taken together, the prospective memory data suggest that gender does not moderate the size of the deficit, and that the task used to measure prospective memory appears to be an important factor in detecting a deficit in PD. Using Brown and Marsden's (1990) theory of processing resources, it may be argued that the PMQT placed greater demands on processing resources than the PMOT.

The relative level of performance on each task generally supports this proposition. Thus, a deficit in the PD group's performance was observed because the PMQT demands exceeded resources when compared to the controls.

Memory as a Function of Disease Stage

To examine the potential moderating effect of disease stage, planned comparisons were made between the early PD group and the HC group, and between the advanced PD group and the HC participants, both at Time 1 and at Time 2. A preliminary analysis confirmed that the early PD group and the advanced PD group were similar to the controls with regard to gender, age, education, and premorbid IQ. Nevertheless, both PD subgroups had higher dementia and depression ratings relative to the HC group.

The results indicate that, at Time 1, the advanced PD group was impaired both at the more difficult level and at the easier level of the NRMT. However, the within-groups comparison revealed a weak manipulation of the difficulty factor (i.e., performance was almost equivalent at the two levels of difficulty). Nevertheless, at the second session where the difficulty factor was effective, the deficit remained at both levels of difficulty. Conversely, in the early PD group the manipulation of the difficulty factor was effective at both sessions, and the only evidence of impairment was at the more difficult level. Furthermore, at the more difficult level there was no consistent difference between the two PD subgroups with respect to the size of the deficit.

Both PD subgroups had impaired verbal recognition (consistent across time), and the magnitude of the deficit was greater in the advanced PD group. Similarly, for recall, both PD subgroups were consistently impaired, and the size of the deficit was greater in the advanced PD group.

In the case of prospective memory, the advanced PD group had inferior memory for a question relative to the HC group, and this was consistent across time. The early PD group also had impaired memory for a question, but only at Time 2. On the other hand, prospective memory for an object was consistently impaired in the advanced PD group, but impaired only at Time 1 in the early group. Again, the size of the prospective memory deficits were consistently greater in the advanced group.

In sum, the data show that the more advanced PD group were consistently impaired on all of the memory tasks. Conversely, the early PD group appear to have consistent recall and recognition deficits (except at the easiest level of the NRMT), but inconsistent prospective memory impairment. In addition, the size of the deficit was always greater in the advanced PD group relative to the early group (except at the hardest level of the NRMT). Taken together, these results support and extend earlier findings suggesting that patients in the advanced stages of PD may have more severe memory deficits than those in the early stages (e.g., Growdon et al., 1990; Lees & Smith, 1983; Owen et al., 1992, 1993; Sahakian et al., 1988). The present study extended previous results by showing that prospective memory is also at risk and such deficits may progress in parallel with the physical symptoms characteristic of PD. However, as discussed in chapter 2, this relationship should not be taken as evidence that both memory and motor deficits share a common pathology. It seems likely that memory deficits are associated with the disruption of nondopaminergic pathways, such as the ascending noradrenergic and cholinergic projections to the frontal cortex (Pillon et al., 1989; Sagar et al., 1995). On the other hand, most motor deficits are mediated by damage to the dopaminergic system. So, disruption to both systems may be responsible for the pattern of results that emerged in the present study.

With respect to the effect of task difficulty, it appears that those participants in the more advanced stages of PD produced nonverbal recognition deficits irrespective of task demands. Conversely, the early PD participants produced deficits of an important magnitude only at the more difficult level.

Finally, although the early PD group clearly had less severe memory dysfunction than the advanced PD group, differences in terms of depression complicate the interpretation of these data. A supplementary analysis directly comparing the two subgroups in terms of GDS scores showed that the advanced group reported, on average, higher levels of depression (d = .39, -.24–1.03). The impact of depression is discussed further in the next section.

Memory as a Function of Depression

To examine the potential moderating effect of depression, separate planned comparisons were made at Time 1 and at Time 2. At Time 1, the depressed PD group was well matched to the nondepressed HC group with respect to gender and age, but had slightly lower education, premorbid IQ, and cognitive status scores. At Time 2, the two groups were matched on gender, age, education, and premorbid IQ. As expected, depression scores were much higher in the depressed PD group at both sessions. The nondepressed PD group were well matched with the nondepressed controls in terms of gender, age, education, and premorbid IQ at both times. Although the nondepressed PD participants did not meet the criteria for depression, they did have higher depression scores compared to the nondepressed controls. The cognitive status score was also slightly higher in the nondepressed PD group at Time 1, but equal at Time 2. In terms of the nonverbal recognition task, the within-groups comparisons confirmed that the manipulation of task difficulty was effective for all three groups.
The data show that both the depressed PD group and the nondepressed PD group had impaired nonverbal recognition relative to the nondepressed controls, with little distinction between the two subgroups. In terms of task difficulty, the more difficult level generally produced greater deficits than the easier level.

The depressed PD group also had consistent deficits in verbal recognition, recall, and prospective memory. In contrast, the nondepressed PD group only had

consistent problems with verbal recognition and recall, and, furthermore, these deficits were smaller than those seen in the depressed PD group.

In conclusion, the results show that depressed and nondepressed PD participants may have similar levels of nonverbal recognition impairment, but differ with respect to verbal recognition, recall, and prospective memory. Thus, depression may partially contribute to the memory impairment seen in PD. However, a supplementary analysis indicated that disease stage confounds the relationship between affective status and memory functioning. Specifically, averaging across time, 48% of the depressed PD group compared to 26% of the nondepressed PD group were in the advanced stages (H&Y stage III or IV). In addition, the depressed PD group reported more functional disability (mADL) than the nondepressed PD group, both at Time 1 and at Time 2 (d = 1.19, .44–1.95 and d = 1.38, .67–2.08, respectively). For all PD participants, averaged across time, there was a positive correlation between the GDS scores and both the H&Y stage, $r(40) = .40$, $p = .009$, 95% CI = .15–.65, and the mADL scores, $r(39) = .67$, $p < .001$, 95% CI = .52–.82. Therefore, it is possible that disease severity is responsible for the observed difference between the depression subgroups. This seems likely for a number of reasons. First, previous research has indicated that depression has little relationship with memory dysfunction in PD, whereas motor deficits are clearly related to memory dysfunction (Owen et al., 1993). Second, the present data show little difference between the depressive subgroups with respect to nonverbal recognition deficits, but some evidence that disease stage subgroups differ in this respect.

Memory as a Function of Age at Onset

Another variable having a possible moderating effect on memory is age at onset. Planned comparisons were made both at Time 1 and at Time 2 to investigate this possibility. Preliminary analyses confirmed that both the early-onset and the late-onset PD groups were well matched to their respective controls with regard to gender, age, education, and premorbid IQ. Finally, as might be expected, the

early-onset PD group were approximately 13 years younger than the late-onset PD group at the time of symptom onset.

Overall, the data provide little evidence that age at onset has a consistent moderating effect on the magnitude of the memory impairment in PD. However, the early-onset PD group did have consistently larger nonverbal recognition deficits relative to the late-onset PD group. In contrast, the opposite occurred for recall and prospective memory for an object. For verbal recognition and prospective memory for a question, the difference between subgroups varied as a function of time.

Past research has indicated that PD participants who are older at the onset of their motor symptoms may develop more widespread cognitive deficits than those with early-onset of symptoms (e.g., Biggins et al., 1992; Caparros-Lefebvre et al., 1995; Dubois et al., 1990; Horiguchi et al., 1991; Hietanen & Teravainen, 1988; Katzen et al., 1998; Mahieux et al., 1998; Reid et al., 1989). However, previous research has also shown that age at onset does not appear to moderate specific memory impairment, both early- and late-onset PD participants showing similar deficits relative to healthy controls (Dubois et al., 1990; Haeske-Dewick, 1996; Hietanen & Teravainen, 1988). The present results may be interpreted in one of two ways. First, any difference in the size of the deficit between the early- and late-onset PD participants may have been due to random fluctuations, and, therefore, age at onset is not a moderating factor. Second, age at onset does have a moderating effect on memory impairment, but it is task specific. That is, early-onset participants have greater problems with nonverbal recognition, whereas late-onset participants have more severe recall and prospective memory deficits.

A supplementary analysis indicated that age and disease duration may confound the comparison made between age at onset subgroups in terms of ESs. The early-onset PD group were approximately 7 years younger, on average, than the late-onset PD group, $F(1, 39) = 20.60$, $p < .001$, $d = 1.43$ (.72–2.13), but had a

longer mean disease duration, $F(1, 39) = 17.12$, $p < .001$, $d = 1.30$ (.61–1.99). Nonetheless, it is difficult to see how these factors could account for the pattern of results described above.

Correlation Analysis

To further evaluate the relationship between memory and the potential moderators, correlations were calculated between the memory task scores and H&Y stage, GDS scores, and age at onset. In terms of disease stage (as measured with the H&Y) the correlations indicate that as PD progresses memory generally deteriorates, supporting the findings reported above that show that the advanced PD group had more severe memory deficits. Past studies have seldom reported correlations between indices of disease stage and recognition memory. Those that have, have found only small correlations (e.g., Cooper et al., 1993; Dewick et al., 1991). Studies that have reported correlations between disease stage and recall have generally found small-to-medium sized associations (Cooper & Sagar, 1993a; Cooper et al., 1992; Mohr et al., 1989; Sullivan et al., 1993). With respect to depression, small-to-medium sized correlations were found in the present study, indicating that higher levels of depression are associated with poorer memory. Previous investigators have reported small correlations between these factors in nondemented PD (e.g., Cooper et al., 1993; Dewick et al., 1991; Sullivan et al., 1993), although the size of the association may depend on the measure used to assess depression. That is, larger correlations have been found using the GDS in comparison to the Beck Depression Inventory (Youngjohn et al., 1992). Finally, in regards to age at onset, with one exception, the correlations reported here indicate that older age at onset is weakly associated with poorer memory. The exception was a positive correlation between recall and age at onset. Comparison with past research is difficult because few studies, if any, have published correlations between age at onset and memory performance.

In the present study, the test-retest reliability coefficients indicated that the recall and recognition tasks generally had moderate external reliability, with two exceptions. That is, poor reliability was found for the recall task when used by the male controls, and the verbal recognition task when used by the PD females. A previous study reported that in a sample of healthy participants, computerised tests of recognition and recall had moderate reliability when a mean test-retest interval of 21 days was used (Youngjohn et al., 1992). With respect to prospective memory, both of the tasks used here had poor test-retest reliability. No other studies appear to have reported external reliability data, so it is difficult to say whether these small coefficients are unique to the present study or related to the nature of the prospective memory task.

Principal Component Analysis

Although there is some debate, it is thought that recognition, recall, and prospective memory are mediated by different memory processes (see chapter 1). If this was true, it should be possible to extract different components reflecting these processes from the set of memory variables used in the present study. To do this, principal components extraction with varimax rotation was used. From the 10 dependent variables, four components were extracted (recognition memory, recall, initial prospective memory, and retest prospective memory). The extraction of these different components provides some evidence that these memory types are distinct from each other. In addition, because these components are orthogonal, they were used as dependent variables in MANOVA and ANOVA, which facilitated interpretation of the group comparisons.

In the case of gender, a MANOVA was conducted with each component serving as a separate dependent variable in order to look for group differences. The results revealed a multivariate interaction between group and gender. Further univariate analyses involving the recognition memory component showed that recognition was poorer in the PD group relative to the controls, and there was no

interaction with gender. The magnitude of the recognition deficit approached that classified by Cohen (1988) as medium. With respect to the recall component, the data clearly showed that the PD participants suffered from a large recall deficit, but there was little interaction with gender. Finally, the data provided some evidence of a prospective memory deficit: for the PD males, a small effect at Time 1 only, and for the PD females, a medium effect at Time 2 only. Taken together these results suggest that there is little difference between males and females in terms of the size of the memory deficit in PD.

With respect to disease stage, the analysis revealed a multivariate effect both for the early PD group and for the advanced PD group, with the ES being greater in the latter group. Further analysis of each memory component indicated that in the early PD group there was a large deficit in recall, but not in other memory types. The advanced PD group had a very large recall deficit, and in contrast to the early group, a medium recognition deficit and a small prospective memory deficit. These results provide further support for the association between disease stage and memory impairment, and indicate that deficits both in recognition and in prospective memory may emerge as PD progresses. It seems very likely that the progressive deterioration in the PD brain over time underlies these increasingly large memory deficits.

Principal components analysis has only rarely been reported in the PD memory literature (e.g., Bondi et al., 1993). In the present case, the analysis appears to clarify some of the findings taken from the planned comparisons. First, recognition, recall, and prospective memory may be mediated by different memory processes, although all are affected in some individuals with PD. Second, the moderating affect of gender (revealed by the planned comparisons) is not supported by the PCA and may be largely spurious. Third, participants in the early stages of PD may have recall deficits, but recognition and prospective memory are still intact. In contrast, those in the later stages of PD may have considerable problems with all three types of memory.

CHAPTER 4

GENERAL DISCUSSION AND CONCLUSIONS

Individual studies examining memory functioning in PD have typically had insufficient statistical power to detect all but the largest deficits. Given the difficulty in obtaining the large PD samples needed to increase statistical power, especially for small-to-medium ESs, meta-analysis may be the most useful method for integrating the results of individual studies. The current meta-analytic findings (chapter 2) demonstrate that theoretically important recognition memory deficits do exist in some nondemented medicated PD patients, contrary opinion notwithstanding (Brown & Marsden, 1990). The results of the present meta-analysis also raised questions about whether disease stage, age at onset, depression, and task difficulty moderate the magnitude of the memory deficit. In addition, another issue that has received little attention in the PD literature is whether males and females have similar memory deficits.

The major finding from the memory study (chapter 3) is that recognition, recall, and prospective memory are all affected by PD, and that the progression of these deficits operates in parallel with the progression of motor symptoms. Furthermore, task demands interacted with this finding, such that nonverbal recognition deficits were only seen in early-stage PD participants when the task was made more difficult. Conversely, advanced-stage participants produced deficits irrespective of the difficulty level. The parallel relationship between memory impairment and motor deficits has been noted before (Owen et al., 1993), but the interaction with task difficulty appears to be a new finding.

These combined results are in line with Brown and Marden's (1990) theory of processing resources, assuming that participants in the advanced stages of PD have a reduced level of central processing resources relative to those in the

earlier stages. Thus, in contrast to the advanced-stage participants, the early-stage participants' processing resources were not exceeded at the easier level of the NRMT and so recognition appeared intact. However, both subgroups experienced depleted processing resources at the hardest level, and so a performance deficit was observed. The results are also consistent with the idea that implicit memory processes facilitate recognition memory (Parkin, 1993). Furthermore, the degree of facilitation may depend both on the difficulty of the task and on disease progression. Thus, in the present study, participants in the advanced stages of PD showed recognition deficits regardless of task difficulty.

There was some evidence that the male and female PD groups' memory deficits differed in magnitude, but these differences were task specific. That is, the males generally had larger deficits in nonverbal recognition and prospective memory for a question, whereas the reverse was true for verbal recognition and recall. However, a number of factors make it hard to interpret these results. First, the two PD groups differed with respect to disease stage, functional disability, and depression. Second, a cursory examination of the male and female control groups' data suggests that performance levels were not equivalent, and this varied as a function of the measure used. Furthermore, the results of the principal components analysis showed that males and females have similar levels of memory impairment, suggesting that differences detected in the planned comparisons may have been spurious.

Direct comparisons between the male and female PD groups indicated that the females generally had better nonverbal recognition memory. In addition, they had poorer verbal recognition, despite the fact that they had better semantic fluency. These differences in recognition are in contrast to previous research (e.g., Kramer et al., 1988) that has generally shown that males perform better on visual tasks, while females have better verbal performance. Unfortunately, it is not readily apparent why the opposite result should exist for the verbal task used here, although it could well be related to PD. For the nonverbal task the possibility exists that the stimuli were actually encoded verbally, not visually. But, this seems

very unlikely because post-test feedback indicated that most participants found it difficult to verbalize the amorphous stimulus shapes. The difficulty of creating a purely visual memory task has been the focus of recent research (e.g., Eadie & Shum, 1995; Shum, O'Gorman, & Eadie, 1999).

The Shum Visual Learning Test (SVLT; Shum et al., 1999) appears to successfully overcome the problems associated with past visual memory tests. The test follows a yes-no recognition memory paradigm, with Chinese characters as stimuli. These stimuli were shown to be difficult to verbalize in participants who were not familiar with the Chinese language (Eadie & Shum, 1995). Nevertheless, Shum et al. found no clear distinction between the performance of males and females on the SVLT. The authors suggested that the male advantage for visuospatial tasks found in past research may have been a consequence of males having greater familiarity with the stimuli. The unfamiliarity of the stimuli used in the SVLT may have precluded the male advantage found in the past (Shum et al., 1999). This explanation may also account for the lack of a male advantage in the NRMT, given that the stimuli were unfamiliar abstract drawings.

As described in chapter 1, research suggests that nondemented PD participants with intact executive functions may have visuospatial deficits (e.g., Cronin-Colomb & Braun, 1997; Mohr et al., 1990). Assuming, then, that the NRMT is a measure of visual memory, it is possible that the observed deficits were actually due to difficulties in visuospatial functioning. However, two lines of evidence suggest that this scenario is unlikely. First, the NRMT uses a 2AFC recognition memory format that should serve to minimize the visuospatial component. Second, recognition memory deficits were also found using the VRMT. There seems little reason to suspect that this test of common words would be affected by visuospatial deficits. Therefore, visuospatial deficits can not adequately account for the deficits observed in recognition memory.

The correlation analysis supports the relationship between advancing disease stage and declining memory functioning. Furthermore, the size of the correlations

involving recognition and recall were similar to those reported in the literature (e.g., Cooper et al., 1992, 1993; Cooper & Sagar, 1993a; Dewick et al., 1991; Mohr et al., 1989; Sullivan et al., 1993). In the case of depression, similar correlations to those for disease stage were found, and these were larger than what has previously been reported (e.g., Cooper et al., 1993; Dewick et al., 1991; Sullivan et al., 1993). The correlations between age at onset and the memory tasks were generally weak and tend to support the between-groups analysis that indicated that age at onset has no consistent relationship with memory dysfunction.

Little is known about the rate of memory decline in PD because of a lack of longitudinal studies. In the present study, the conceptual replication conducted after an approximate 6 month interval offered the chance to assess the rate of decline over the short-term. If memory declined in the PD group over and above what was expected due to normal aging, the ES at the second session would be greater than that observed at the first session. The present data show no decline in recognition and prospective memory for an object, for any of the PD subgroups. However, all subgroups had greater deficits in recall and prospective memory for a question at the second session. These findings cannot be explained in terms of changes in antiparkinsonian medication or functional disability because there was no consistent change in these variables across groups.

Therefore, over a 6 month period, nondemented PD participants may have a task-specific decline in memory performance relative to healthy controls. Moreover, the decline occurred irrespective of gender, disease stage, affect, and age at onset. The decline was also not a function of task difficulty because even when the nonverbal recognition task was made more difficult, the ESs were generally smaller at the second session.

Finally, although the recall task was not specifically designed to be directly comparable to the recognition tasks, it should be noted that the ES of the recall deficit was, in some instances, equal to or smaller than the size of the recognition

memory deficits. This contrasts with the generally held view that recall deficits are more severe than recognition deficits (see chapter 1). This issue needs to be examined further with two tasks equated in terms of difficulty.

Limitations and Future Directions

Like most research, the present investigation faced a number of methodological limitations which need to be taken into account when interpreting the findings. This section covers both real and possible limitations of the current study and directions for the future.

While meta-analysis is not without criticism, it has many advantages over narrative reviews (see Hunter & Schmidt, 1990; Copper & Hedges, 1994). However, any meta-analysis is limited by the available database. Unfortunately, many potential moderators of the observed recognition impairment could not be tested in the analysis reported in chapter 2. This was due in part to a lack of research findings, but also because the aims of some studies did not necessitate providing information required for the present meta-analysis. Potential moderators include task variables such as modality and delay, and participant variables such as age at symptom onset, symptom duration, type of motor symptoms, and depression (e.g., Beatty, 1992; Karayanidis, 1989; Levin et al., 1992; Sagar & Sullivan, 1988; Saint-Cyr & Taylor, 1993; Starkstein et al., 1989).

If any sense is to be made of these moderators, investigators need to consider publishing sufficient information to allow their findings to be used in future meta-analyses. Moreover, the scope of new meta-analyses should be increased to examine a wider range of cognitive and motor deficits, providing a more complete picture of cognitive decline. Only then can we piece together a coherent theory.

In the primary level study reported in chapter 3, the assessment of memory decline was made over a 6-month period. Although little is known about the rate of

memory decline in PD, it seems likely that an interval of 6 months is too short to properly test for a genuine PD decline. However, in this case, time constraints prevented the use of a longer interval. Conversely, an interval of 6 months is a relatively long time to assess test-retest reliability.[10] It is possible that factors, such as age, interact with task performance reducing observed reliability.

The diagnosis of PD was, in all but three cases, made by a neurologist. Nevertheless, the accurate diagnosis of PD is difficult (Calne et al., 1992). For example, Hughes, Ben-Shlomo, Daniel, and Lees (1992) confirmed at *post mortem* that only 76% of patients, who had previously been diagnosed by a neurologist or geriatrician, actually had PD. Therefore, it is possible that in the present study some participants did not have idiopathic PD. Unfortunately, there is no way to determine to what extent this would have affected the results. Nevertheless, a number of procedures were followed to limit the impact of this issue. First, exclusion criteria were used to limit the number of participants with a condition other than idiopathic PD from taking part. For instance, those who reported any other neurological impairment or a poor response to antiparkinsonian medication were excluded. In addition, an independent diagnostic assessment was made using Calne et al.'s (1992) categories of diagnosis. The results of this showed that all but one participant could be categorized as having clinically definite or clinically probable idiopathic PD.

Participants were also excluded from the analysis if they showed signs of dementia. The meta-analysis clearly showed that dementia moderates the recognition memory deficit found in past research. Therefore, the size of the recognition deficits reported in chapter 3 would be inflated if participants with dementia were included. However, for a number of reasons it is unlikely that this occurred. First, cognitive status was assessed with the aid of the OMCT, which has previously been shown to have good reliability and validity (Katzman et al., 1983). Second, the nature of the tasks employed make it unlikely that a participant with dementia could have completed the relatively demanding testing session.

[10] It should be noted that at least one psychological text (Coolican, 1999) uses 6 months as an example of a test-retest interval.

The present study focused on whether task difficulty moderated recognition memory deficits in PD. The success in answering this question depended largely on how well the task difficulty factor was manipulated. Pilot work indicated that this factor was effective at the difficulty levels used. However, in the main study, the difficulty factor was not consistently effective when the results were analysed as a function of gender. This problem complicated the interpretation of the results. For instance, for the female PD group, performance at the more difficult level was nearly equivalent to that observed at the easier level. Had manipulation of the difficulty factor been stronger, it is possible that greater deficits would have been seen at the more difficult level of the task. Another limitation concerning task difficulty is that this factor was only manipulated in the nonverbal recognition task. Thus, any conclusion about increasing task demands is restricted to the nonverbal modality. Interestingly, performance deficits were found in the PD participants using the verbal recognition task, even though this task was seen to be considerably easier than the nonverbal task. It is possible that a still greater deficit in verbal recognition would have been detected had the task been more difficult. However, whether this is true or not awaits further study.

The way in which the confidence ratings data were collected proved to be another weakness in the present study. It will be recalled that participants were asked to rate their confidence in each response during the recognition memory phase of both the nonverbal and the verbal recognition memory tasks. If a participant did not respond within 6 s the next trial began. Pilot testing indicated that a 6-s interval was sufficient for participants to make a response. However, in actuality, many participants failed to respond during this time creating missing data. An analysis of these data indicated that important differences in confidence ratings existed between those participants with missing data and those without missing data. Furthermore, the percentage of participants with missing data varied as a function of group and gender. Therefore, the confidence ratings could not be analysed further. Fortunately, the confidence data were not an integral part of the study, but may have provided interesting information. The lessons to be learned

are that (a) it is worthwhile to do extensive checking before beginning the main study, and (b) PD participants can be very slow in making decisions.

The analyses of depression, age at onset, and disease stage involved categorizing the PD participants into two groups based on these variables. The accuracy of these categories therefore depends on the reliability and validity of these data. Previous research has confirmed that the GDS has good reliability and validity, so it seems reasonable to suggest that the depression categories are sufficiently accurate for the purposes of the present study. However, the age at onset data were collected by asking each participant to date the onset of their first symptom. Although it was impossible to independently verify this information, in many cases a participant's spouse was able to confirm the date of onset. In addition, as a further attempt to improve the accuracy of these data, this question was asked at both sessions and the average was used in the analysis. The data indicate reasonably consistent answers; only 13% of participants produced a discrepancy of more than four years, and the correlation between Time 1 and Time 2 was high ($r = .82$). The present data compare well with those of a previous study examining the reliability of symptom onset assessment in PD (Richards, Marder, Cote, & Mayeux, 1994).

In the case of disease stage, participants were classified according to their H&Y rating. The H&Y has been used extensively in the PD literature and has the advantage of being relatively easy to apply. The H&Y has also been shown to have "substantial" interobserver reliability (Geminiani et al., 1991). However, in the present study no independent assessment of the severity of motor symptoms was made and so the data can not be verified. Taylor et al. (1986, 1987, 1990) used a measure of manual speed and dexterity (the Purdue Pegboard) to measure severity, and this approach could have been used here. However, regardless of the measure used, one difficulty in rating disease stage or severity concerns the fluctuations in motor symptoms commonly seen in PD (Olanow & Koller, 1998). The H&Y was developed before the side-effects of long-term drug therapy were known; the scale provides no assessment of motor fluctuations or

adverse reactions to therapy. However, one advantage of the H&Y for the present purposes was that it is relatively insensitive to "change that is less than major" (Diamond & Markham, 1983, p. 1098). Therefore, fluctuations in therapeutic response were less likely to result in misclassification of the PD participants than if a more sensitive scale had been used. Additionally, participants were tested at their time of optimal therapeutic response to minimize problems caused by drug induced side-effects.

In summary, although the present study had several limitations, there is no reason to suspect that invalid or unreliable classification of depression, age at onset, and disease stage was responsible for the results found. However, further studies are needed to support these findings, especially longitudinal analyses of memory decline in PD, beginning in the earliest stage of this disorder and ending in the last stage.

Another limitation of the present study concerned the tasks used to measure prospective memory. The test-retest analysis indicated poor external reliability, which may stem from the design of the tasks. For instance, the task involving a question used a different question at the second session, whereas the object task required the participant to choose the object to be remembered, and so there was no control over task consistency. These factors may have reduced the reliability of the tasks, and may explain some of the inconsistent results. Clearly, the design problems highlighted here need to be addressed before selecting a prospective memory task. However, prospective memory is important in the daily lives of people with PD, and further research on this interesting memory type is urgently required.

In retrospect, the planned comparison approach taken here to examine the influence of moderators may have been conducted with other statistical techniques, such as multiple regression. However, such an approach would be complicated by the time factor. That is, disease stage and depression were measured both at Time 1 and at Time 2, and so the groups differed slightly at

each time. Furthermore, differences between the two age at onset PD subgroups with respect to current age necessitated matching each group to a different subgroup of controls. Thus, separate contrasts made at each time were needed. In addition, the approach taken proved useful for other reasons: (a) it revealed inconsistencies between the results of the two testing sessions, highlighting the problems of possible random error; (b) it provided an assessment of short-term memory decline; and (c) it revealed the inadequacies of using statistical significant testing (see below).

Statistical Significance Testing

As described in chapter 2, the use of statistical significance testing has been criticized on a number of grounds for many years now. One problem with significance testing occurs when a statistically significant result is taken as evidence of an important effect, and nonsignificance is taken as evidence of no effect. However, statistical significance depends on the magnitude of the effect and sampling error (Carver, 1993). Therefore, in situations of low statistical power an important effect may be "nonsignificant", whereas in situations of high statistical power a trivial effect may be "significant." Moreover, a nonsignificant finding can only be interpreted when statistical power is adequate. Even then, statistical significance tells us nothing about the size of the effect. To illustrate these points, selected results from the meta-analysis (chapter 2) and recognition memory study (chapter 3) are presented below.

The meta-analysis revealed that demented PD participants suffer from a large (d = 1.30, 1.30–1.30) recognition memory deficit. Had significance testing been used, this would have been described as a statistically significant effect, $t(129)$ = 7.26, $p < .001$. As well, there was evidence that nondemented, medicated PD participants suffer from a small (d = .23, .23–.23) and statistically significant, $t(674) = 2.87$, $p = .002$, recognition memory deficit. In contrast, there was little evidence that nondemented, *de novo* PD participants had impaired recognition memory (d = .05, .05–.05). This effect was nonsignificant, $t(363) = .38$, $p = .35$.

However, in this case there was only 48% power to detect a small effect (d = .20; Cohen, 1988). Without the estimate of ES, the significance test result is ambiguous because it may mean that either (a) there is no memory deficit, or (b) there is a deficit but the test was just not sensitive enough to detect it (Fagley, 1985).

Table 35 presents a summary of the results taken from the disease stage analysis. The purpose of the table is to facilitate a comparison of the significance levels versus the ESs. If the data are interpreted solely on the basis of the traditional critical p value of 0.05, it would be concluded that early PD participants do not have consistent recognition memory deficits, whereas the advanced participants have a consistent deficit only in verbal recognition. However, if interpretation is based on ES and the associated confidence interval, it may be concluded that early participants have consistent deficits both in verbal recognition and in nonverbal recognition, but only when the task is more difficult. In addition, participants in the advanced stages of PD have consistent deficits both in verbal and in nonverbal recognition, regardless of task difficulty. One of the problems of relying solely on significance testing is very clear from the data in Table 35: Small to medium ESs frequently do not reach significance, not because

Table 35 *Significance Levels Versus Effect Sizes for the Recognition Memory Tasks as a Function of Time for the Comparisons Involving the Early-Stage Parkinson's Disease (Early PD), the Advanced-Stage Parkinson's Disease (Advanced PD), and the Healthy Control Groups*

	Early PD[a]			Advanced PD[b]		
	p	*d*	95% CI	*p*	*d*	95% CI
NRMT-H						
Time 1	.03	.65	.15–1.16	.02	.62	.08–1.17
Time 2	.44	.20	-.30– .69	.20	.40	-.14– .93
NRMT-E						
Time 1	.67	.11	-.38– .60	.001	1.03	.46–1.59
Time 2	.95	.02	-.47– .51	.13	.44	-.09– .98
VRMT						
Time 1	.09	.44	-.06– .94	.02	.66	.12–1.21
Time 2	.08	.46	-.03– .96	.03	.65	.10–1.19

Note. p = Alpha level; d = ES statistic (Cohen, 1988); 95% CI = 95% confidence interval for d; NRMT = Nonverbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task.
[a]n = 25 PD, 41 HC. [b]n = 16 PD, 41 HC.

the ES is zero but because the power to detect these ESs is not sufficient.

In both the meta-analysis and the recognition memory study, the use of ES estimates and confidence intervals provide a more accurate picture of the memory impairment in PD than the use of significance testing alone. As illustrated above, different conclusions would have been drawn had significance testing been used exclusively. In the primary study, the only consistent statistically significant results would have involved: recall in PD males and in PD females, verbal recognition memory and recall in the advanced-stage PD group, recall in the early-stage PD group, verbal recognition memory and recall in the depressed PD group, recall in the nondepressed PD group, and recall in the late-onset PD group. Given the size of the sample, statistical power would have been poor, thus all other nonsignificant findings would have been ambiguous. What sample size would be needed to increase power to the conventional 80% level, assuming a small ES? In order to detect small effects at the 0.05 level of significance, a total sample size of 788 participants would be needed! Clearly, the resources required to achieve such a large sample size are beyond most researchers. The use of ES estimates and confidence intervals is a good alternative in individual studies. Unfortunately, most previous research has drawn conclusions about effects based on significance tests alone. Therefore, making comparisons between the present and previous studies would have been difficult without first calculating the ESs.

Thesis Conclusion

In general, the results presented in chapters 2 and 3 support the view that recognition memory deficits do occur in PD. The power analysis demonstrated that past research in this area has not had sufficient statistical power to detect small, but theoretically important deficits. The meta-analysis revealed the potential moderating effect of symptom severity; advanced stage (medicated) nondemented PD participants produced small recognition memory deficits,

whereas early stage (*de novo*) participants do not. The results of the recognition memory study support this view, and highlight the need to consider task difficulty.

Thus, the simple theory that parkinsonians have little or no difficulty encoding new information (intact recognition system) while having problems with information retrieval (faulty recall system; Ruberg & Agid, 1988), needs modification. In fact, recent research suggests that different aspects of both retrieval and encoding may be at risk in PD (Faglioni et al., 1997). Thus, a general theory of cognitive dysfunction in PD must take into account not only that a recognition deficit does occur in this neurodegenerative disorder, but that this deficit may grow with disease severity and task difficulty.

The present results can be explained by Brown and Marsden's (1990) theory of processing resources. With respect to the advanced stage PD participants, it is possible that the tasks used here, regardless of difficulty, exceeded these participants' central processing resources. On the other hand, the early stage participants generally showed less impairment, presumably because their initial level of processing resources was greater than that of the advanced participants. Moreover, the early stage participants showed some evidence of a nonverbal recognition deficit on the more difficult level of the task, but not on the easier level. It is possible that in these participants only the more difficult level exceeded their central processing resources. These results also provide some support for Weingartner et al.'s (1984) effortfulness/automatic distinction, and Breen's (1993) view that the dissociation between recall and recognition memory may simply be an artifact of task difficulty.

Other factors may also be involved in the apparent dissociation between recall and recognition. For instance, performance on recognition memory tasks may be facilitated by perceptual implicit processes that remain intact in PD. It is possible that as task difficulty increases PD participants are less able to take advantage of these perceptual fluency cues. It is equally possible that the ability to use these perceptual cues is partly determined by disease stage. Hence, participants in the

advanced stages of PD produce recognition memory deficits regardless of task difficulty.

Finally, the research presented here should serve to highlight the danger of relying on the statistical significance of individual findings as a method for assessing the existence of cognitive deficits in PD. Small-sample studies have well-known methodological problems, especially with respect to sampling error, which may lead to error in the statistical significance test. Fortunately, when assessing an individual study an alternative exists; that is, point estimates of ES and confidence intervals. The results of many isolated studies, usually lacking sufficient statistical power, can then be integrated using meta-analysis. Such an approach provides a more reliable estimate of the population effect and the degree of certainty associated with this estimate.

References

Agniel, A., Celsis, P., Viallard, G., Montastruc, J. L., Rascol, O., Demonet, J. F., Marc-Vergnes, J. P., & Rascol, A. (1991). Cognition and cerebral blood flow in lateralised Parkinsonism: Lack of functional lateral asymmetries. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 783-786.

Albert, M. S., Butters, N., & Levin, J. (1979). Temporal gradients in the retrograde amnesia of patients with alcoholic Korsakoff's disease. *Archives of Neurology*, *36*, 211-216.

Ali, A. S., & Morley, J. E. (1999). *Treatment of Parkinson's disease*. (On-line). Available: http://www.cyberounds.com/conferences/geriatrics/conferences/current/conference.html

Allain, H., Lieury, A., Quemener, V., Thomas, V., Reymann, J., & Gandon, J. (1995). Procedural memory in PD. *Dementia*, *6*, 174-178.

American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behaviour*, *22*, 261-295.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*, 97-123.

Anderson, J. R., & Bower, G. H. (1974). A propositional theory of recognition memory. *Memory and Cognition*, *2*, 406-412.

Appollonio, I., Grafman, J., Clark, K., Nichelli, P., Zeffiro, T., & Hallett, M. (1994). Implicit and explicit memory in patients with Parkinson's disease with and without dementia. *Archives of Neurology*, *51*, 359-367.

Atkinson, R. C., & Juola, J. F. (1974). Factors influencing the speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583-612). New York: Academic Press.

Auriacombe, S., Grossman, M., Carvell, S., Gollomp, S., Stern, M. B., & Hurtig, H. I. (1993). Verbal fluency deficits in Parkinson's disease. *Neuropsychology*, *7*, 182-192.

Azuma, T., Bayles, K. A., Cruz, R. F., Tomoeda, C. K., Wood, J. A., McGeagh, A., & Montgomery, E. B. (1997). Comparing the difficulty of letter, semantic, and name fluency tasks for normal elderly and patients with Parkinson's disease. *Neuropsychology*, *11*, 488-497.

Baddeley, A. D. (1992a). Is working memory working? *Quarterly Journal of Experimental Psychology*, *44A*, 1-31.

Baddeley, A. D. (1992b). Working memory. *Science*, *255*, 556-559.

Baddeley, A. D. (1994). Working memory: The interface between memory and cognition. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems 1994* (pp. 351-367). Cambridge, MA: MIT Press.

Baker, E. L., Letz, R., & Fidler, A. (1985). A computer-administered neurobehavioral evaluation system for occupational and environmental epidemiology. *Journal of Occupational Medicine*, *27*, 206-212.

Bakheit, A. M. O. (1995). Early diagnosis of Parkinson's disease. *Postgraduate Medicine Journal*, *71*, 151-152.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin*, *99*, 388-399.

Beatty, W. W. (1992). Memory disturbances in Parkinson's disease. In S. J. Huber & J. L. Cummings (Eds.), *Parkinson's Disease: Neurobehavioral Aspects* (pp. 49-58). Oxford: Oxford University Press.

Berent, S., & Giordani, B. (1998). Functional neuroimaging correlates of memory dysfunction in neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 128-139). Cambridge: Cambridge University Press.

Beringer, J. (1995). *Experimental run time system*, Version 3.11 [Computer program manual]. Frankfurt: BeriSoft Cooperation.

Bernbach, H. A. (1967). Decision processes in memory. *Psychological Review*, *74*, 462-480.

Bieliauskas, L. A., Klawans, H. L., & Glantz, R. H. (1986). Depression and cognitive changes in Parkinson's disease: A review. In M. D. Yahr & K. J. Bergmann (Eds.), *Advances in Neurology: Vol. 45* (pp. 437-438). New York: Raven.

Biggins, C. A., Boyd, J. L., Harrop, F. M., Madeley, P., Mindham, R. H. S., Randall, J. I., & Spokes, E. G. S. (1992). A controlled, longitudinal study of dementia in Parkinson's disease. *Journal of Neurology, Neurosurgery, & Psychiatry, 55*, 566-571.

Bleecker, M. L., Bolla-Wilson, K., Agnew, J., & Meyers, D. A. (1988). Age-related sex differences in verbal memory. *Journal of Clinical Psychology, 44*, 403-411.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry, 114*, 797-811.

Blonder, L. X., Gur, R. E., Gur, R. C., Saykin, A. J., & Hurtig, H. I. (1989). Neuropsychological functioning in hemiparkinsonism. *Brain and Cognition, 9*, 244-257.

Bondi, M. W., & Kaszniak, A. W. (1991). Implicit and explicit memory in Alzheimer's disease and Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 13*, 339-358.

Bondi, M. W., Kaszniak, A. W., Bayles, K. A., & Vance, K. T. (1993). Contributions of frontal system dysfunction to memory and perceptual abilities in Parkinson's disease. *Neuropsychology, 7*, 89-102.

Boone, K. B., Lesser, I. M., Miller, B. L., Wohl, 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.

Borenstein, M., Rothstein, H., & Cohen, J. (1997). *Power and precision: A computer program for statistical power analysis and confidence intervals* [computer program manual]. Teaneck, NJ: Biostat.

Bower, G. H., & Glass, A. L. (1976). Structural units and the redintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 456-466.

Boyd, J. L., Cruickshank, C. A., Kenn, C. W., Madeley, P., Mindham, R. H. S., Oswald, A. G., Smith, R. J., & Spokes, E. G. S. (1991). Cognitive impairment

and dementia in Parkinson's disease: A controlled study. *Psychological Medicine*, *21*, 911-921.

Bradshaw, J. L., Waterfall, M. L., Phillips, J. G., Iansek, R., Mattingley, J. B., & Bradshaw, J. A. (1993). Re-orientation of attention in Parkinson's disease: An extension to the vibrotactile modality. *Neuropsychologia*, *31*, 51-66.

Brainerd, C. J., Howe, M. L., & Desrochers, A. (1982). The general theory of two stage learning: A mathematical review with illustrations from memory development. *Psychological Bulletin*, *91*, 634-665.

Breen, E. K. (1993). Recall and recognition memory in Parkinson's disease. *Cortex*, *29*, 91-102.

Brown, J. (1976). An analysis of recognition and recall and of problems in their comparison. In J. Brown (Ed.), *Recall and recognition* (pp. 1-36). London: John Wiley & Sons.

Brown, R. G., & Marsden, C. D. (1988a). Internal versus external cues and the control of attention in Parkinson's disease. *Brain*, *111*, 323-345.

Brown, R. G., & Marsden, C. D. (1988b). Subcortical dementia: The neuropsychological evidence. *Neuroscience*, *25*, 363-387.

Brown, R. G., & Marsden, C. D. (1990). Cognitive function in Parkinson's disease: From description to theory. *Trends in Neuroscience*, *13*, 21-29.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305-327.

Bryman, A., & Cramer, D. (1994). *Quantitative data analysis for social scientists* (Rev. ed.). London: Routledge.

Burt, D. A., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin*, *117*, 285-305.

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

Buytenhuijs, E. L., Berger, H. J. C., van Spaendonck, K. P. M., Horstink, M. W. I. M., Borm, G. F., & Cools, A. R. (1994). Memory and learning strategies in patients with Parkinson's disease. *Neuropsychologia*, *32*, 335-342.

Calev, A. (1984). Recall and recognition in chronic nondemented schizophrenics: Use of matched tasks. *Journal of Abnormal Psychology*, *93*, 172-177.

Calne, D. B., Snow, B. J., & Lee, C. (1992). Criteria for diagnosing Parkinson's disease. *Annals of Neurology*, *32* (Suppl.), 125-127.

Caparros-Lefebvre, D., Pecheux, N., Petit, V., Duhamel, A., & Petit, H. (1995). Which factors predict cognitive decline in Parkinson's disease? *Journal of Neurology, Neurosurgery, & Psychiatry*, *58*, 51-55.

Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education*, *61*, 287-292.

Chow, S. L. (1988). Signficance test or effect size? *Psychological Bulletin*, *103*, 105-110.

Christensen, H., Hadzi-Pavlovic, D., & Jacomb, P. (1991). The psychometric differentiation of dementia from normal aging: A meta-analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 147-155.

Coakes, S. L., & Steed, L. G. (1996). *SPSS for Windows: analysis without anguish*. Milton, Australia: Jacaranda Wiley.

Cockburn, J., & Smith, P. T. (1994). Anxiety and errors of prospective memory among elderly people. *British Journal of Psychology*, *85*, 273-282.

Cohen, J. (1962*).* The statistical power of abnormal-social psychological research: a review. *Journal of Abnormal and Social Psychology*, *65*, 145-153.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997-1003.

Cohen, N., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*, 207-209.

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, *116*, 457-475.

Coolican, H. (1999). *Research methods and statistics in psychology* (3rd ed.). Oxon, United Kingdom: Hodder and Stoughton.

Cooper, H. & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.

Cooper, J. A., & Sagar, H. J. (1993a). Encoding deficits in untreated Parkinson's disease. *Cortex*, *29*, 261-265.

Cooper, J. A., & Sagar, H. J. (1993b). Incidental and intentional recall in Parkinson's disease: an account based on diminished attentional resources. *Journal of Clinical and Experimental Neuropsychology*, *15*, 713-731.

Cooper, J. A., Sagar, H. J., Doherty, S. M., Jordan, N., Tidswell, P., & Sullivan, E. V. (1992). Different effects of dopaminergic and anticholinergic therapies on cognitive and motor function in Parkinson's disease. *Brain*, *115*, 1701-1725.

Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S., & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain*, *114*, 2095-2122.

Cooper, J. A., Sagar, H. J., & Sullivan, E. V. (1993). Short-term memory and temporal ordering in early Parkinson's disease: effects of disease chronicity and medication. *Neuropsychologia*, *31*, 933-949.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51-110). Hillsdale, NJ: Lawrence Erlbaum Associates.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 474-479.

Crawford, J. R., Moore, J. W., & Cameron, I. M. (1992). Verbal fluency: A NART-based equation for the estimation of premorbid performance. *British Journal of Clinical Psychology*, *31*, 327-329.

Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement*, *10*, 3-31.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*, 116-127.

Cronin-Golomb, A., & Braun, A. E. (1997). Visuospatial dysfunction and problem solving in Parkinson's disease. *Neuropsychology*, *11*, 44-52.

Cummings, J. L., & Benson, D. F. (1988). Psychological dysfunction accompanying subcortical dementias. *Annual Review of Medicine*, *39*, 53-61.

Daum, I., Schugens, M. M., Spieker, S., Poser, U., Schonle, P. W., & Birbaumer, N. (1995). Memory and skill acquistion in Parkinson's disease and frontal lobe dysfunction. *Cortex*, *31*, 413-432.

De Keyser, J., Herregodts, P., & Ebinger, G. (1990). The mesoneocortical dopamine neuron system. *Neurology*, *40*, 1660-1662.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test manual*. New York: The Psychological Corporation.

DeLong, M. R., & Wichmann, T. (1993). Basal ganglia-thalamocortical circuits in parkinsonian signs. *Clinical Neuroscience*, *1*, 18-26.

Dewick, H. C., Hanley, J. R., Davies, A. D. M., Playfer, J., & Turnbull, C. (1991). Perception and memory for faces in Parkinson's disease. *Neuropsychologia*, *29*, 785-802.

Diamond, S. G., & Markham, C. H. (1983). Evaluating the evaluations: Or how to weigh the scales of parkinsonian disability. *Neurology*, *33*, 1098-1099.

Diamond, S. G., Markham, C. H., Hoehn, M. M., McDowell, F. H., & Muenter, M. D. (1990). An examination of male-female differences in progression and mortality of Parkinson's diease. *Neurology*, *40*, 763-766.

Direnfeld, L. K., Albert, M. L., Volicer, L., Langlais, P. J., Marquis, J., & Kaplan, E. (1984). Parkinson's disease: The possible relationship of laterality to dementia and neurochemical findings. *Archives of Neurology*, *41*, 935-941.

Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology*, *41*, 209-222.

Downes, J. J., Sharp, H. M., Costall, B. M., Sagar, H. J., & Howe, J. (1993). Alternating fluency in Parkinson's disease. An evaluation of the attentional control theory of cognitive impairment. *Brain*, *116*, 887-902.

Dubois, B., Boller, F., Pillon, B., & Agid, Y. (1991). Cognitive deficits in Parkinson's disease. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 5, pp. 195-240). Amsterdam: Elsevier.

Dubois, B., Danze, F., Pillon, B., Cusimano, G., Lhermitte, F., & Agid, Y. (1987). Cholingergic-dependent cognitive deficits in Parkinson's disease. *Annals of Neurology*, *22*, 26-30.

Dubois, B., Pillon, B., Legault, F., Agid, Y., & Lhermitte F. (1988). Slowing of cognitive processing in progressive supranuclear palsy. *Archchives of Neurology*, *45*, 1194-1198.

Dubois, B., Pillon, B., Sternic, N., Lhermitte, F., Agid, Y. (1990). Age-induced cognitive disturbances in Parkinson's diease. *Neurology*, *40*, 38-41.

Dujardin, K., Bourriez, J. L., & Guieu, J. D. (1995). Event-related desynchronization (ERD) patterns during memory processes: Effects of aging and task difficulty. *Electroencephalography and Clinical Neurophysiology*, *96*, 169-182.

Eadie, K., & Shum, D. (1995). Assessment of visual memory: A comparison of Chinese characters and geometric figures as stimulus materials. *Journal of Clinical and Experimental Neuropsychology*, *17*, 731-739.

Eagly, A. H., & Wood, W. (1994). Using research synthesis to plan future research. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 485-500). New York: Russell Sage Foundation.

Eichenbaum, H. (1997). Declarative memory: Insights from cognitive neurobiology. *Annual Review of Psychology*, *48*, 547-572.

Einstein, G. O., Holland, L. J., McDaniel, M. A., & Guynn, M. J. (1992). Age-related deficits in prospective memory: The influence of task complexity. *Psychology and Aging*, *7*, 471-478.

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 717-726.

Einstein, G. O., McDaniel, M. A., Richardson, S. L., Guynn, M. J., & Cunfer, A. R. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 996-1007.

El-Awar, M., Becker, J. T., Hammond, K. M., Nebes, R. D., & Boller, F. (1987). Learning deficit in Parkinson's disease: Comparison with Alzheimer's disease and normal aging. *Archives of Neurology*, *44*, 180-184.

Ellis, H. D. (1981). Theoretical aspects of face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces.* London: Academic Press.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: a general power analysis program. *Behaviour Research Methods, Instruments, & Computers, 28*, 1-11.

Eysenck, H. J. (1960). The concept of statistical signficance and the controversey about one-tailed tests. *Psychological Review, 67*, 269-271.

Fagley, N. S. (1985). Applied statistical power analysis and the interpretation of nonsignificant results by research consumers. *Journal of Counseling Psychology, 32*, 391-396.

Faglioni, P., Botti, C., Scarpa, M., Ferrari, V., & Saetti, M. C. (1997). Learning and forgetting processes in Parkinson's disease: A model-based approach to disentangling storage, retention and retrieval contributions. *Neuropsychologia, 35*, 767-779.

Fama, R., Sullivan, E. V., Shear, P. K., Cahn-Weiner, D. A., Yesavage, J. A., Tinklenberg, J. R., Pfefferbaum, A. (1998). Fluency performance patterns in Alzheimer's disease and Parkinson's disease. *The Clinical Neuropsychologist, 12*, 487-499.

Farina, E., Cappa, S. F., Polimeni, M., Magni, E., Canesi, M., Zecchinelli, A., Scarlato, G., Mariani, C. (1994). Frontal dysfunction in early Parkinson's disease. *Acta Neurologica Scandinavica, 90*, 34-38.

Ferraro, F. R., Balota, D. A., & Connor, L. T. (1993). Implicit memory and formation of new associations in nondemented Parkinson's disease individuals and individuals with senile dementia of the Alzheimer type: A serial reaction time (SRT) investigation. *Brain and Cognition, 21*, 163-180.

Filoteo, J. V., Delis, D. C., Demadura, T. L., Salmon, D. P., Roman, M. J., & Shults, C. W. (1994). Abnormally rapid disengagement of covert attention in global and local stimulus levels may underlie the visuoperceptual impairment in Parkinson's patients. *Neuropsychology, 8*, 210-217.

Filoteo, J. V., Delis, D. C., Demadura, T. L., Salmon, D. P., Roman, M. J., & Shults, C. W. (1997). An examination of the nature of attentional deficits in

patients with Parkinson's disease: Evendence from a spatial orienting task. *Journal of the International Neuropsychological Society*, *3*, 337-347.

Fink, J. W., & Randolph, C. (1998). Semantic memory in neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 197-209). Cambridge: Cambridge University Press.

Fleischman, D. A., Gabrieli, J. D. E., Reminger, S., Rinaldi, J., Morrell, F., & Wilson, R. (1995). Conceptual priming in perceptual identification for patients with Alzheimer's disease and a patient with a right occipital lobectomy. *Neuropsychology*, *9*, 187-197.

Flowers, K. A., Pearce, I., & Pearce, J. M. S. (1984). Recognition memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 1174-1181.

Foreman, J. B. (Ed.).(1972). *The New Zealand contemporary dictionary* (4th ed.). Christchurch, New Zealand: Whitcome and Tombs.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*, 87-115.

Gabrieli, J. D. E., Fleischman, D. A., Keane, M. M., Reminger, S. L., & Morrell, F. (1995). Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science*, *6*, 76-82.

Gabrieli, J. D. E., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence for the role of a frontostriatal system in working strategic memory. *Neuropsychology*, *10*, 322-332.

Geminiani, G., Cesana, B. M., Tamma, F., Contri, P., Pacchetti, C., Carella, F., Piolti, R., Martignoni, E., Giovannini, P., Girotti, F., & Caraceni, T. (1991). Interobserver reliability between neurologists in training of Parkinson's disease rating scales. *Movement Disorders*, *6*, 330-335.

Geschwind, N. (1979). Specializations in the human brain. *Scientific American*, *241*, 158-168.

Giambra, L. M., Quilter, R. E., Phillips, P. B., & Hiscock, B. S. (1988). Performance on a sustained attention task as a function of strategy: A cross-sectional investigation using the Mackworth clock-test. *Bulletin of the Psychonomic Society*, *26*, 333-335.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1-67.

Glass, G. V. (1976). Primary, secondary, and meta-analysis research. *Educational Researcher*, *5*, 3-8.

Glass, A. L., Holyoak, K. J., & Santa, J. L. (1979). *Cognition*. Reading, MA: Addison-Wesley.

Goetz, C. G., Tanner, C. M., Stebbins, G. T., & Buchman, A. S. (1988). Risk factors for progression in Parkinson's disease. *Neurology*, *38*, 1841-1844.

Graybiel, A. M. (1993). Functions of the nigrostriatal system. *Clinical Neuroscience*, *1*, 12-17.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*, New York: Wiley.

Grossman, M., Carvell, S., Peltzer, L., Stern, M. B., Gollomp, S., & Hurtig, H. I. (1993). Visual construction impairments in Parkinson's disease. *Neuropsychology*, *7*, 536-547.

Grossman, M., Carvell, S., Stern, M. B., Gollomp, S.,& Hurtig, H.I. (1992). Sentence comprehension in Parkinson's disease: the role of attention and memory. *Brain and Language*, *42*, 437-384.

Growdon, J. H., & Corkin, S. (1986). Cognitive impairments in Parkinson's disease. In M. D. Yahr & K. J. Bergmann (Eds.), *Advances in neurology*, *Vol. 45* (pp. 383-392). New York: Raven.

Growdon, J. H., Corkin, S., & Rosen, T. J. (1990). Distinctive aspects of cognitive dysfunction in Parkinson's disease. In M. D. Streifler, A. D. Korczyn, E. Melamed, & M. B. H. Youdim (Eds.), *Advances in neurology*, *Vol. 53: Parkinson's disease: Anatomy*, *pathology*, *and therapy* (pp. 365-376). New York: Raven.

Growdon, J. H., Kieburtz, K., McDermott, M. P., Panisset, M., Friedman, J. H., & the Parkinson Study Group. (1998). Levodopa impoves motor function

without impairing cognition in mild nondemented Parkinson's disease patients. *Neurology*, *50*, 1327-1331.

Gurd, J. M., & Ward, C. D. (1989). Retrieval from semantic and letter initial categories in patients with Parkinson's disease. *Neuropsychologia*, *27*, 743-746.

Haeske-Dewick, H. C. (1996). Are perception and memory for faces influenced by a specific age at onset factor in Parkinson's disease? *Neuropsychologia*, *34*, 315-320.

Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 691-702.

Hammond-Tooke, G. D., & Pollock, M. (1992). Depression, dementia, and Parkinson's disease. In A. B. Joseph & R. R. Young (Eds.), *Movement disorders in neuropsychiatry*. Cambridge, MA: Blackwell.

Hanley, J. R., Dewick, H. C., Davies, A. D., Playfer, J., & Turnball, C. (1990). Verbal fluency in Parkinson's disease. *Neuropsychologia*, *28*, 737-741.

Harrington, D. L., Haaland, K. Y., Yeo, R. A., & Marder, E. (1990). Procedural memory in Parkinson's disease: Impaired motor but not visuoperceptual learning. *Journal of Clinical and Experimental Neuropsychology*, *12*, 323-339.

Hartikainen, P., Helkala, E. L., Soininen, H., & Riekkinen, P. (1993). Cognitive and memory deficits in untreated Parkinson's disease and amyotrophic lateral sclerosis patients: A comparative study. *Journal of Neural Transmission*, *6*, 127-137.

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

Hayes, W. L. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart, and Winston.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Heindel, W. C., Salmon, D. P., & Butters, N. (1993). Cognitive approaches to the memory disorders of demented patients. In P. B. Sutker & H. E. Adams

(Eds.), *Comprehensive handbook of psychopathology* (2nd ed.) (pp. 735-761). New York: Plenum.

Heindel, W. C., Salmon, D. P., Shults, C. W., Walicke, P. A., & Butters, N. (1989). Neuropsychological evidence for multiple implicit memory systems: A comparison of Alzheimer's, Huntington's, and Parkinson's disease patients. *The Journal of Neuroscience*, *9*, 582- 587.

Helkala, E. L., Laulumaa, V., Soininen, H., & Riekkinen, P. J. (1989). Different error pattern of episodic and semantic memory in Alzheimer's disease and Parkinson's disease with dementia. *Neuropsychologia*, *27*, 1241- 1248.

Henderson, V. W., & Buckwalter, J. G. (1994). Cognitive deficits of men and women with Alzheimer's disease. *Neurology*, *44*, 90-96.

Herlitz, A., Nilsson, L., & Backman, L. (1997). Gender differences in episodic memory. *Memory and Cognition*, *25*, 801-811.

Hietanen, M., & Teravainen, H. (1986). Cognitive performance in early Parkinson's disease. *Acta Neurologica Scandinavia*, *73*, 151-159.

Hietanen, M., & Teravainen, H. (1988). The effect of age of disease onset on neuropsychological performance in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *51*, 244-249.

Hintzman, D. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528-551.

Hoehn, M. M. (1998). Commentary. *Neurology*, *50*, 318.

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology*, *17*, 427-442.

Hopfensperger, K., & Koller, W. C. (1991). Recognizing early Parkinson's disease. *Postgraduate Medicine*, *90*, 49-59.

Horiguchi, J., Nishimatsu, O., Inami, Y., Sukegawa, T., & Shoda, T. (1991). A clinical study on intellectual impairment in parkinsonian patients during long-term treatment. *The Japanese Journal of Psychiatry & Neurology*, *45*, 13-18.

Huber, S. J., & Bornstein, R. A. (1992). Neuropsychological evaluation of Parkinson's disease. In S. J. Huber & J. L. Cummings (Eds.), *Parkinson's disease: Neurobehavioral aspects* (pp. 32-45). Oxford: Oxford University Press.

Huber, S. J., Freidenberg, D. L., Shuttleworth, E. C., Paulson, G. W., & Christy, J. A. (1989). Neuropsychological impairments associated with severity of Parkinson's disease. *Journal of Neuropsychiatry*, *1*, 154-158.

Huber, S. J., Freidenberg, D. L., Shuttleworth, E. C., Paulson, G. W., & Clapp, L. E. (1989). Neuropsychological similarities in lateralised Parkinsonism. *Cortex*, *25*, 461-470.

Huber, S. J., Shulman, H. G., Paulson, G. W., & Shuttleworth, E. C. (1987). Fluctuations in plasma dopamine level impair memory in Parkinson's disease. *Neurology*, *37*, 1371-1375.

Huber, S. J., Shuttleworth, E. C., & Paulson, G. W. (1986). Dementia in Parkinson's disease. *Archives of Neurology*, *43*, 987-990.

Huberman, M., Moscovitch, M., & Freedman, M. (1994). Comparison of patients with Alzheimer's and Parkinson's disease on different explicit and implicit tests of memory. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *7*, 185-193.

Hughes, A. J., Ben-Shlomo, Y., Daniel, S. E., & Lees, A. J. (1992). What features improve the accuracy of clinical diagnosis in Parkinson's disease: A clinicopathologic study. *Neurology*, *42*, 1142-1146.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging*, *5*, 356-368.

Hultsch, D. F., Hertzog, C., Small, B. J., McDonald-Miszczak, L., & Dixon, R. A. (1992). Short-term longitudinal change in cognitive performance in later life. *Psychology and Aging*, *7*, 571-584.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage Publications.

Huppert, F. A., & Beardsall, L. (1993). Prospective memory impairment as an early indicator of dementia. *Journal of Clinical and Experimental Neuropsychology*, *15*, 805-821.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Peterson, R. C., Kokmen, E., & Kurland, L. T. (1992). Mayo's older Americans normative studies: WAIS-R norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6* (Suppl.), 1-30.

Jacobs, D. M., & Schofield, P. (1998). Biological and psychosocial risk factors for dementia and memory loss. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 243-254). Cambridge: Cambridge University Press.

Jacobs, D. M., Marder, K., Cote, L. J., Sano, M., Stern, Y., & Mayeux, R. (1995). Neuropsychological characteristics of preclinical dementia in Parkinson's disease. *Neurology, 45*, 1691-1696.

Jankovic, J. (1993). Theories on the etiology and pathogenesis of Parkinson's disease. *Neurology, 43* (Suppl. 1), 21-23.

Johnson, M. K., & Hasher, L. (1987). Human learning and memory. *Annual Review of Psychology, 38*, 631-668.

Johnson, S. C., Pinkston, J. B., Bigler, E. D., & Blatter, D. D. (1996). Corpus callosum morphology in normal controls and traumatic brain injury: Sex differences, mechanisms of injury, and neuropsychological correlates. *Neuropsychology, 10*, 408-415.

Johnston, W. A., Hawley, K. J., & Elliott, J. M. G. (1991). Contribution of perceptual fluency to recognition judgements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 210-223.

Kalat, J. W. (1998). *Biological psychology* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Karayanidis, F. (1989). Parkinson's disease: A conceptualization of neuropsychological deficits within an information-processing framework. *Biological Psychology, 29*, 149-179.

Katzen, H. L., Levin, B. E., & Llabre, M. (1998). Age of disease onset influences cognition in Parkinson's disease. *Journal of the International Neuropsychological Society, 4*, 285-290.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a short orientation-memory-concentration test of cognitive impairment. *American Journal of Psychiatry, 140*, 734-739.

Kendrick, D. C. (1985). *Kendrick cognitive tests for the elderly*. Windsor, United Kingdom: NFER Nelson.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Kim, M., & Hunter, J. E. (1993). Attitude-behavior relations: A meta-analysis of attitudinal relevance and topic. *Journal of Communication*, *43*, 101-142.

Kintsch, W. (1967). Memory and decision aspects of recognition learning. *Psychological Review*, *74*, 496-504.

Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory* (pp. 307-373). New York: Academic Press.

Knight, R. G. (1992). *The neuropsychology of degenerative brain diseases.* Hillsdale, NJ: Lawrence Erlbaum.

Knight, R. G. (1998). Prospective memory in aging and neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 172-183). Cambridge: Cambridge University Press.

Knight, R. G., Godfrey, H. P. D., & Shelton, E. J. (1988). The psychological deficits associated with Parkinson's disease. *Clinical Psychology Review*, *8*, 391-410.

Koller, W. C. (1992a). How accurately can Parkinson's disease be diagnosed? *Neurology*, *42* (Suppl. 1), 6-16.

Koller, W. C. (1992b). When does Parkinson's disease begin? *Neurology*, *42* (Suppl. 4), 27-31.

Koller, W. C. (1993). Differential diagnosis of Parkinson's disease. *Neurology*, *43* (Suppl. 1), 11.

Koller, W. C., Silver, D. E., & Lieberman, A. (1994). An algorithm for the management of Parkinson's disease. *Neurology*, *44* (Suppl. 10), 1-52.

Kramer, J. H., Delis, D. C., & Daniel, M. (1988). Sex differences in verbal learning. *Journal of Clinical Psychology*, *44*, 907-915.

Kurtzke, J. F., & Murphy, F. M. (1990). The changing patterns of death rates in parkinsonism. *Neurology*, *40*, 42-49.

Kvavilashvili, L. (1987). Remembering intention as a distinct form of memory. *British Journal of Psychology*, *78*, 507-518.

Lange, K. W., Paul, G. M., Robbins, T. W., & Marsden, C. D. (1993). L-DOPA and frontal cognitive function in Parkinson's disease. In H. Narabayashi, T. Nagatsu, N. Yanagisawa, & Y. Mizuno (Eds.), *Advances in neurology: Vol. 60*, (pp. 475-478). New York: Raven.

Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A. M., & Paul, G. M. (1992). L-Dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology*, *107*, 394-404.

Larsen, J. P., Dupoint, E., & Tandberg, E. (1994). Clinical diagnosis of Parkinson's diease: Proposal of diagnostic subgroups classified at different levels of confidence. *Acta Neurologica Scandinavia*, *89*, 242-251.

Latash, M. L., & Anson, J. G. (1996). What are "normal movements" in atypical populations? *Behavioral and Brain Sciences*, *19*, 55-106.

Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. *Brain*, *106*, 257-270.

Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 285-294.

Levin, B. E., & Katzen, H. L. (1995). Early cognitive changes and nondementing behavioral abnormalities in Parkinson's disease. In W. J. Weiner and A. E. Lang (Ed.), *Advances in neurology: Vol. 65. Behavioral neurology of movement disorders* (pp. 85-95). New York: Raven.

Levin, B. E., Llabre, M. M., Reisman, S., Weiner, W. J., & Brown, M. C. (1991). A retrospective analysis of the effects of anticholinergic medication on memory performance in Parkinson's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *3*, 412-416.

Levin, B. E., Llabre, M. M., & Weiner, W. J. (1989). Cognitive impairments associated with early Parkinson's disease. *Neurology*, *39*, 557-561.

Levin, B. E., Tomer, R., & Rey, G. J. (1992a). Clinical correlates of cognitive impairment in Parkinson's disease. In S. J. Huber & J. L. Cummings (Eds.), *Parkinson's disease: Neurobehavioral aspects* (pp. 97-106). Oxford: Oxford University Press.

Levin, B. E., Tomer, R., & Rey, G. J. (1992b). Cognitive impairments in Parkinson's disease. *Neurologic Clinics*, *10*, 471-485.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). Oxford: Oxford University Press.

Lieberman, A. (1998). Managing the neuropsychiatric symptoms of Parkinson's disease. *Neurology*, *50* (Suppl. 6), 33-38.

Lieberman, A., Dziatolowski, M., Kupersmith, M., Serby, M., Goodgolc, A., Korein, J., & Goldstein, M. (1979). Dementia in Parkinson Disease. *Annals of Neurology*, *6*, 355-359.

Lindman, H., R. (1992). *Analysis of variance in experimental design*. New York: Springer-Verlag.

Litvan, I., Mohr, E., Williams, J., Gomez, C., & Chase, T. N. (1991). Differential memory and executive functions in demented patients with Parkinson's and Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 25-29.

Lyons, K. E., Hubble, J. P., Troster, A. I., Pahwa, R., & Koller, W. C. (1998). Gender differences in Parkinson's disease. *Clinical Neuropharmacology*, *21*, 118-121. (From PsycLit, 1998)

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*, 6-21.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Maddox, W. T., Filotea, J. V., Delis, D. C., & Salmon, D. P. (1996). Visual selective attention deficits in patients with Parkinson's disease: A quatitative model-based approach. *Neuropsychology*, *10*, 197-218.

Mahieux, F., Fenelon, G., Flahault, A., Manifacier, M., Michelet, D., & Boller, F. (1998). Neuropsychological prediction of dementia in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, *64*, 178-183.

Mahurin, R. K., Feher, E. P., Nance, M. L., Levy, J. K., & Pirozzolo, F. J. (1993). Cognition in Parkinson's disease and related disorders. In R. W. Parks, R. F. Zec, & R. S. Wilson (Eds.), *Neuropsychology of Alzheimer's disease and other dementias* (pp. 308-349). Oxford: Oxford University Press.

Mandler, G. (1980). Recognizing: The judgement of previous occurrence. *Psychological Review*, *87*, 252-271.

Markham, C. H., & Diamond, S. G. (1993). Clinical overview of Parkinson's disease. *Clinical Neuroscience*, *1*, 5-11.

Markowitsch, H. J. (1998). The biology of neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 140-153). Cambridge: Cambridge University Press.

Marsden, C. D. (1990). Parkinson's disease. *The Lancet*, *335*, 948-952.

Marttila, R. J., & Rinne, U. K. (1981). Epidemiology of Parkinson's disease: An overview. *Journal of Neural Transmission*, *51*, 135-148.

Massman, P. J., Delis, D. C., Butters, N., Levin, B. E., & Salmon, D. P. (1990). Are all subcortical dementias alike? Verbal learning and memory in Parkinson's and Huntington's disease patients. *Journal of Clinical and Experimental Neuropsychology*, *12*, 729-724.

Mayes, A. (1997). [Review of *Prospective memory: Theory and applications*]. *Neuropsychologia*, *35*, 1423.

Mayes, A. R. (1998). Memory in neurodegenerative disease: what has been learned about the organization of memory? In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 226-239). Cambridge: Cambridge University Press.

Mayeux, R. P. (1994). Behavioral and cognitive dysfunction. In A. M. Cohen & W. J. Weiner. *Comprehensive neurologic rehabilitation: Vol. 8. The comprehensive management of Parkinson's disease* (pp. 119-130). New York: Demos.

Maylor, E. A. (1990). Age and prospective memory. *The Quarterly Journal of Experimental Psychology*, *42A*, 471-493.

Maylor, E. A. (1993). Aging and forgetting in prospective and retrospective memory. *Psychology and Aging*, *8*, 420-428.

McCormack, H. M., Horne, D. J. L., & Sheather, S. (1988). Clinical applications of visual analogue scales: A critical review. *Psychological Medicine*, *18*, 1007-1019.

McDaniel, M. A., Glisky, E. L., Rubin, S. R., Guynn, M. J., & Routhieaux, B. C. (1999). Prospective memory: A neuropsychological study. *Neuropsychology, 13*, 103-110.

McNicol, D. (1972). *A primer of signal detection theory*. London: George Allen & Unwin.

Mohr, E., Fabbrini, G., Williams, J., Schlegel, J., Cox, C., Fedio, P., & Chase, T. N. (1989). Dopamine and memory function in Parkinson's disease. *Movement Disorders, 4*, 113-120.

Mohr, E., Juncos, J., Cox, C., Litvan, I., Fedio, P., & Chase, T. N. (1990). Selective deficits in cognition and memory in high-functioning Parkinsonian patients. *Journal of Neurology, Neurosurgery, and Psychiatry, 53*, 603-606.

Mohr, E., Mendis, T., & Grimes, J. D. (1995). Late cognitive changes in Parkinson's disease with an emphasis on dementia. In W. J. Weiner and A. E. Lang (Eds.), *Advances in neurology: Vol. 65. Behavioral neurology of movement disorders* (pp. 97-113). New York: Raven.

Morris, P. E. (1992). Prospective memory: Remembering to do things. In M. Gruneberg & P. Morris (Eds.), *Aspects of memory: Vol. 1. The practical aspects* (2nd ed.) (pp. 196-222). London: Routledge.

Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 51*, 757-766.

Mortimer, J. A., Pirozzolo, F. J., Hansch, E. C., & Webster, D. D. (1982). Relationship of motor symptoms to intellectual deficits in Parkinson's disease. *Neurology, 32*, 133-137.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems 1994* (pp. 269-310). Cambridge, MA: MIT Press.

Moscovitch, M., Winocur, G., & McLachlan, D. (1986). Memory as assessed by recognition and reading time in normal and memory-impaired people with Alzheimer's disease and other neurological disorders. *Journal of Experimental Psychology: General, 115*, 331-347.

Murdock, B. B., Jr. (1974). *Human memory: Theory and data*. Maryland: Erlbaum.

Murdock, B. B., Jr. (1982). Recognition memory. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 1-26). New York: Academic Press.

Murray, L. W., & Dosser, D. A., Jr. (1987). How significant is a significant difference? Problems with the measurement of magnitude of effect. *Journal of Counseling Psychology*, *34*, 68-72.

Nelson, H. E., & O'Connell, A. (1978). Dementia: The estimation of premorbid intelligence levels using the new adult reading test. *Cortex*, *14*, 234-244.

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

Ogden, J. A., Growden, J. H., & Corkin, S. (1990). Deficits on visuospatial tests involving forward planning in high-functioning Parkinsonians. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *3*, 125-139.

Olanow, C. W., Hauser, R. A., Gauger, L., Malapira, T., Koller, W., Hubble, J., Bushenbark, K., Lilienfeld, D., & Esterlitz, J. (1995). The effect of deprenyl and levodopa on the progression of Parkinson's disease. *Annals of Neurology*, *38*, 771-777.

Olanow, C. W., & Koller, W. C. (1998). An algorithm (decision tree) for the management of Parkinson's disease: Treatment guidelines. *Neurology*, *50* (Suppl. 3), 1-57.

Orsman, H. W. (Ed.). (1989). *Heinemann New Zealand dictionary* (2nd ed.). Auckland, New Zealand: Heinman Reed.

Oscar-Berman, M., & Bardenhagen, F. (1998). Nonhuman primate models of memory dysfunction in neurodegenerative disease: Contributions from comparative neuropsychology. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 3-20). Cambridge: Cambridge University Press.

O'Sullivan, M. (1998). *Memory deficits in Parkinson's disease.* Unpublished Master's thesis, Massey University, Palmerston North, New Zealand.

Owen, A. M., Beksinska, M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Sahakian, B. J., & Robbins, T. W. (1993). Visuo-spatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia*, *31*, 627-644.

Owen, A. M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Lange, K. W., & Robbins, T. W. (1992). Fronto-striatal cognitive deficits at different stages of Parkinson's disease. *Brain*, *115*, 1727-1751.

Owen, A. M., Sahakian, B. J., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1995). Dopamine-dependent frontostriatal planning deficits in early Parkinson's disease. *Neuropsychology*, *9*, 126-140.

Owen, A. M., Sahakian, B. J., & Robbins, T. W. (1998). The role of executive deficits in memory disorders in neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 157-171). Cambridge: Cambridge University Press.

Oyebode, J. R., Barker, W. A., Blessed, G., Dick, D. J., & Britton P. G. (1986). Cognitive functioning in Parkinson's disease: In relation to prevalence of dementia and psychiatric diagnosis. *British Journal of Psychiatry*, *149*, 720-725.

Parkin, A. J. (1993). Implicit memory across the lifespan. In P. Graf & M. E. J. Masson (Eds.), *Implicit memory: New directions in cognition, development, & Neuropsychology* (pp. 191-206). Hillsdale, NJ: Lawrence Erlbaum.

Parkin, A. J., & Lawrence, A. (1994). A dissociation in the relation between memory tasks and frontal lobe tests in the normal elderly. *Neuropsychologia*, *32*, 1523-1532.

Parkinson, J. (1955). An essay on the shaking palsy. In M. Critchley, W. M. McMenemey, F. M. R. Walshe, & J. G. Greenfield (Eds.), *James Parkinson (1755-1824). A bicentenary volume of dealing with Parkinson's disease*. London: MacMillan. (Reprinted from *An essay on the shaking palsy*, 1817, London: Sherwood, Neeley, and Jones).

Perry, E. K., Curtis, M., Dick, D. J., Candy, J. M., Atack, J. R., Bloxham, C. A., Blessed, G., Fairbairn, A., Tomlinson, B. E., & Perry, R. H. (1985). Cholinergic

correlates of cognitive impairment in Parkinson's disease: Comparisons with Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 48*, 413-421.

Piccirilli, M., D'Alessandro, P., Finali, G., & Piccinin, G. L. (1994). Neuropsychological follow-up of parkinsonian patients with and without cognitive impairment. *Dementia, 5*, 17-22.

Pillon, B., Deweer, B., Agid, Y., & Dubois, B. (1993). Explicit memory in Alzheimer's, Huntington's, and Parkinson's disease. *Archives of Neurology, 50*, 374-379.

Pillon, B., Dubois, B., Cusimano, G., Bonnet, A. M., Lhermitte, F., & Agid, Y. (1989). Does cognitive impairment in Parkinson's disease result from non-dopaminergic lesions? *Journal of Neurology, Neurosurgery, and Psychiatry, 52*, 201-206.

Pillon, B., Dubois, B., Lhermitte, F., & Agid, Y. (1986). Heterogeneity of cognitive impairment in progressive supranuclear palsy, Parkinson's disease, and Alzheimer's disease. *Neurology, 36*, 1179-1185.

Pillon, B., Ertle, S., Deweer, B., Sarazin, M., Agid, Y., & Dubois, B. (1996). Memory for spatial location is affected in Parkinson's disease. *Neuropsychologia, 34*, 77-85.

Poewa, W. H. (1994). Clinical aspects of motor fluctuations in Parkinson's disease. *Neurology, 44* (Suppl. 6), 6-9.

Poole, C. (1987). Beyond the confidence interval. *American Journal of Public Health, 77*, 195-199.

Randolph, C., Braun, A. R., Goldberg, T. E., & Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: dissociation of storage and retrieval failures. *Neuropsychology, 7*, 82-88.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, 285-308.

Reid, W. G. J., Broe, G. A., Hely, M. A., Morris, J. G. L., Willimson, P. M., O'Sullivan, D. J., Rail, D., Gence, S., & Moss, N. G. (1989). The neuropsychology

of de novo patients with idiopathic Parkinson's disease: The effects of age of onset. *International Journal of Neuroscience*, *48*, 205-217.

Richards, M., Marder, K., Cote, L., & Mayeux, R. (1994). Reliability of symptom onset assessment in Parkinson's disease. *Movement Disorders*, *9*, 340-342.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology*, *39*, 475-543.

Rinne, J. O. (1991). Nigral degeneration in Parkinson's disease in relation to clinical features. *Acta Neurologica Scandinavia*, *84* (Suppl. 136), 87-90.

Robbins, T. W., James, M., Owen, A. M., Lange, K. W., Lees, A. J., Leigh, P. N., Marsden, C. D., Quinn, N. P., & Summers, B. A. (1994). Cognitive deficits in progressive supranuclear palsy, Parkinson's disease, and multiple system atrophy in tests sensitive to frontal lobe dysfunction. *Journal of Neurology, Neurosurgery, and Psychiatry*, *57*, 79-88.

Rosenthal, R. (1991). Meta-analysis: A review. *Psychosomatic Medicine*, *53*, 247-271.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Copper & L. V. Hedges (Eds.)., *The handbook of research synthesis* (pp. 231-244). New York: Russell Sage Foundation.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*, 183-192.

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin*, *92*, 500-504.

Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin*, *99*, 400-406.

Rosselli, M., & Ardila, A. (1991). Effects of age, education, and gender on the Rey-Osterrith Complex Figure. *The Clinical Neuropsychologist*, *5*, 370-376.

Rossi, J. S. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology*, *58*, 646-656.

Ruberg, M., & Agid, Y. (1988). Dementia in Parkinson's disease. In L. Iversen, S. D. Iversen, & S. H. Snyder (Eds.), *Handbook of psychopharmacology: Vol. 4. Psychopharmacology of the aging nervous system.* New York: Plenum.

Ryan, J. J., & Paolo, A. M. (1992). A screening procedure for estimating premorbid intelligence in the elderly. *The Clinical Neuropsychologist*, *1*, 53-62.

Sagar, H. J., Atchison, L., Doherty, S. M., Ball, M. C., & Cooper, J. A. (1995). The contribution of non-dopaminergic pathology to cognitive, motor and affective disability in Parkinson's disease. In I. Hanin et al. (Eds.), *Alzheimer's and Parkinson's diseases* (pp. 69-76). New York: Plenum.

Sagar, H. J., Cohen, N. J., Sullivan, E. V., Corkin, S., & Growdon, J. H. (1988). Remote memory function in Alzheimer's disease and Parkinson's disease. *Brain*, *111*, 185-206.

Sagar, H. J., & Sullivan, E. V. (1988). Patterns of cognitive impairment in dementia. In C. Kennard (Ed.), *Recent advances in clinical neurology: Number 5.* (pp. 47-86). Edinburgh: Churchill Livingstone.

Sagar, H. J., Sullivan, E. V., Cooper, J. A., & Jordan, N. (1991). Normal release from proactive interference in untreated patients with Parkinson's disease. *Neuropsychologia*, *29*, 1033-1044.

Sagar, H. J., Sullivan, E. V., Gabriel, J. D. E., Corkin, S., & Growdon, J. H. (1988). Temporal ordering and short-term memory deficits in Parkinson's disease. *Brain*, *111*, 525-539.

Sage, J. I., & Mark, M. H. (1994). Basic mechanisms of motor fluctuations. *Neurology*, 44 (Suppl. 6), 10-14.

Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain*, *111*, 695-718.

Saint-Cyr, J. A., & Taylor, A. E. (1993). Cognitive dysfunction in Parkinson's disease. In J. S. Schneider & M. Gupta (Eds.), *Current concepts in Parkinson's disease research.* Kirkland, WA: Hogrefe and Huber.

Saint-Cyr, J. A., Taylor, A. E., & Lang, A. E. (1988). Procedual learning and neostriatal dysfunction in man. *Brain*, *111*, 941-959.

Salmon, D. P., Lineweaver, T. T., & Heindel, W. C. (1998). Nondeclarative memory in neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 210-225). Cambridge: Cambridge University Press.

Salmon, D. P., Thal, L. J., Butter, N., & Heindel, W. C. (1990). Longitudinal evaluation of dementia of the Alzheimer type: A comparison of 3 standardized mental status examinations. *Neurology*, *40*, 1225-1230.

Salsburg, D. S. (1985). The religion of statistics as practiced in medical journals. *The American Statistician*, *39*, 220-223.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory Systems 1994* (pp. 1-38). Cambridge, MA: MIT Press.

Schafer, W. D. (1993). Interpreting statistical significance and nonsignificance. *Journal of Experimental Education*, *61*, 383-387.

Schmidt, F. L. (1992). What do data really mean?: Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, *47*, 1173-1181.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods*, *1*, 115-129.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, *62*, 529-540.

Schneider, J. S. (1993). Parkinson's disease research: a brief introduction. In J. S. Schneider & M. Gupta (Eds.), *Current concepts in Parkinson's disease research* (pp. 3-8). Kirkland, WA: Hogrefe and Huber.

Scholz, O. B., & Sastry, M. (1985). Memory characteristics in Parkinson's disease. *International Journal of Neuroscience*, *27*, 229-234.

Schwab, R. (1960). Progression and prognosis in Parkinson's disease. *Journal of Nervous and Mental Disease*, *130*, 556-566.

Schwarzer, R. (1989). *Meta-analysis programs.* Durham, NC: Duke University.

Sedlmeier, P. & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin*, *105*, 309-316.

Sharpe, M. H. (1992). Auditory attention in early Parkinson's disease: An impairment in focused attention. *Neuropsychologia*, *30*, 101-106.

Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Education*, *61*, 293-316.

Sheer, D. E., & Schrock, B. (1986). Attention. In H. J. Hannay (Ed.), *Experimental techniques in human neuropsychology* (pp. 95-137). New York: Oxford University Press.

Shoben, E. J. (1992). Semantic memory. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 581-585). New York: Macmillan.

Shum, D. H. K., O'Gorman, J. G., & Eadie, K. (1999). Normative data for a new memory test: The Shum Visual Learning Test. *The Clinical Neuropsychologist*, *13*, 1-15.

Singh, S. N., & Cole, C. A. (1985). Forced-choice recognition tests: A critical review. *Journal of Advertising*, *14*, 52-58.

SPSS Inc. (1995). *SPSS for windows*. Chicago: Author.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195-231.

Squire, L. R. (1994). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems 1994* (pp. 203-231). Cambridge, MA: MIT Press.

Squire, L. R., Shimamura, A. P., & Graf, P. (1985). Independence of recognition memory and priming effects: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 37-44.

Stacy, M., & Jankovic, J. (1992). Differential diagnosis of Parkinson's disease and the parkinsonism plus syndromes. *Neurologic Clinics*, *10*, 341-359.

Starkstein, S. E., Bolduc, P. L., Mayberg, H. S., Presziosi, T. J., & Robinson, R. G. (1990). Cognitive impairments and depression in Parkinson's disease: A follow up study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *53*, 597-602.

Starkstein, S. E., Preziosi, T. J., Berthier, M. L., Bolduc, P. L., Mayberg, H. S., & Robinson, R. G. (1989). Depression and cognitive impairment in Parkinson's disease. *Brain*, *112*, 1141-1153.

Starkstein, S. E., & Robinson, R. G. (1991). Dementia of depression in Parkinson's disease and stroke. *The Journal of Nervous and Mental Disease*, *179*, 593-601.

Stevens, J. P. (1980). Power of the multivariate analysis of variance tests. *Psychological Bulletin*, *88*, 728-737.

Stoffelmayr, B. E., Dillavou, D., & Hunter, J. E. (1983). Premorbid functioning and recidivism in schizophrenia: A cumulative analysis. *Journal of Consulting and Clinical Psychology*, *51*, 338-352.

Sullivan, E. V., & Sagar, H. J. (1989). Nonverbal recognition and recency discrimination deficits in Parkinson's disease and Alzheimer's disease. *Brain*, *112*, 1503-1517.

Sullivan, E. V., Sagar, H. J., Cooper, J. A., & Jordan, N. (1993). Verbal and nonverbal short-term memory impairment in untreated Parkinson's disease. *Neuropsychology*, *7*, 396-405.

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper and Row.

Taylor, A. E, & Saint-Cyr, J. A. (1992). Executive function. In S. J. Huber & J. L. Cummings (Eds.), *Parkinson's Disease: Neurobehavioral Aspects* (pp. 74-85). Oxford: Oxford University Press.

Taylor, A. E, Saint-Cyr, J. A., & Lang, A. E. (1985). Dementia prevalence in Parkinson's disease. *Lancet*, *1*, 1037.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1986). Frontal lobe dysfunction in Parkinson's disease. *Brain*, *109*, 845- 883.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1987). Parkinson's disease. *Brain*, *110*, 35-51.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1988). Idiopathic Parkinson's disease: Revised concepts of cognitive and effective status. *Canadian Journal of Neurological Science*, *15*, 106-113.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1990). Memory and learning in early Parkinson's disease: evidence for a "frontal lobe syndrome". *Brain and Cognition*, *13*, 211-232.

Testa, J. A., Brumback, R. A., Baik, T., Leech, R. W., & Cannon, T. C. (1998). Neuropathology and memory dysfunction in neurodegenerative disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive, and clinical perspectives* (pp. 36-86). Cambridge: Cambridge University Press.

Tierney, M. C., Nores, A., Snow, W. G., Fisher, R. H., Zorzitto, M. L., & Reid, D. W. (1994). Use of the Rey auditory verbal learning test in differentiating normal aging from Alzheimer's and Parkinson's dementia. *Psychological Assessment*, *6*, 129-134.

Tombaugh, T. N., Grandmaison, L. J., & Schmidt, J. P. (1995). Prospective memory: Relationship to age and retrospective memory in the learning and memory battery (LAMB). *The Clinical Neuropsychologist*, *9*, 135-142.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*, 138-146.

Tsai, C. H., Lu, C. S., Hua, M. S., Lo, W. L., & Lo, S. K. (1994). Cognitive dysfunction in early onset parkinsonism. *Acta Neurologica Scandinavica*, *89*, 9-14.

Tulving, E. (1992). Episodic memory. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 161-163). New York: Macmillan.

Tweedy, J. R., Langer, K. G., & McDowell, F. H. (1982). The effect of semantic relations on the memory deficit associated with Parkinson's disease. *Journal of Clinical Neuropsychology*, *4*, 235-247.

Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science*, *270*, 769-775.

van Hilten, J. J., van der Zwan, A. D., Zwinderman, A. H., & Roos, R. A. C. (1994). Rating impairment and disability in Parkinson's disease rating scale. *Movement Disorders*, *9*, 84-88.

van Spaendonck, K. P. M., Berger, H. J. C., Horstink, M. W. I., Buytenhuijs, E. L., & Cools, A. R. (1993). Impaired cognitive shifting in Parkinsonian patients on anticholinergic therapy. *Neuropsychologia*, *31*, 407-411.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*, 254-256.

Waterfall, M. L., & Crowe, S. F. (1995). Meta-analytic comparison of the components of visual cognition in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, *17*, 759-772.

Weingartner, H., Burns, S., Diebel, R., & LeWitt, P. A. (1984). Cognitive impairment in Parkinson's disease: Distinguishing between effort-demanding and automatic cognitive processes. *Psychiatry Research*, *11*, 223-235.

Whittington, C. J., & Podd, J. V. (1996). Human performance and physiology: A statistical power analysis of ELF electromagnetic field research. *Bioelectromagnetics*, *17*, 274-278.

Whittington, C. J., Podd, J., & Kan, M. M. (in press). Recognition memory impairment in Parkinson's disease: Power and meta-analyses. *Neuropsychology*.

Wickelgren, W. A. (1977). *Learning and memory*. Englewood Cliffs, NJ: Prentice-Hall.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the Embedded Figures Test*. CA: Consulting Psychologists Press.

Woodruff-Pak, D. S. (1997). *The neuropsychology of aging*. Malden, MA: Blackwell.

Wooten, G. F. (1984). Parkinsonism. In A. L. Pearlman & R. C. Collins (Eds.), *Neurological pathophysiology* (3rd ed.) (pp. 365-377). New York: Oxford University Press.

Wright, M. J., Burns, R. J., Geffen, G. M., & Geffen, L. B. (1990). Covert orientation of visual attention in Parkinson's disease: An impairment in the maintenance of attention. *Neuropsychologia*, *28*, 151-159.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37-49.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1341-1354.

Youdim, M. B. H., & Riederer, P. (1997). Understanding Parkinson's disease. *Scientific American*, *276*, 38-45.

Youngjohn, J. A., Beck, J., Jogerst, G., & Caine, C. (1992). Neuropsychological impairment, depression, and Parkinson's disease. *Neuropsychology*, *6*, 149-158.

Youngjohn, J. A., Larrabee, G. J., Crook, T. H. (1992). Test-retest reliability of computerized, everyday memory measures and traditional memory tests. *The Clinical Neuropsychologist*, *6*, 274-286.

Zakzanis, K. K. (1998). Brain is related to behavior ($p < .05$). *Journal of Clinical and Experimental Neuropsychology*, *20*, 419-427.

Zelinski, E. M., Gilewski, M. J., & Schaie, K. W. (1993). Individual differences in cross-sectional and 3-year longitudinal memory performance across the adult life span. *Psychology and Aging*, *8*, 176-186.

Zelinski, E. M., & Stewart, S. T. (1998). Individual differences in 16-year memory changes. *Psychology and Aging*, *13*, 622-630.

Zetusky, W. J., Jankovic, J., & Pirozzolo, F. J. (1985). The heterogeneity of Parkinson's disease: Clinical and prognostic implications. *Neurology*, *34*, 522-526.

APPENDIX A
JOURNAL TITLES

Journals searched issue by issue for studies on cognitive deficits in Parkinson's disease

Acta Neurologica Scandinavica; Advances in Neurology; Annals of Neurology; Annals of the New York Academy of Sciences; Annual Review of Medicine; Archives of Clinical Neuropsychology; Archives of Neurology; Brain and Cognition; Brain and Language; Brain Research; British Journal of Psychiatry; Clinical Neuropsychologist; Clinical Psychology Review; Cognitive Neuropsychology; Cortex; International Journal of Neuroscience; Journal of Clinical and Experimental Neuropsychology; Journal of Nervous and Mental Disease; Journal of Neurology, Neurosurgery, and Psychiatry; Neurology; Neuropsychiatry, Neuropsychology, and Behavioral Neurology; Neuropsychobiology; Neuropsychologia; Neuropsychology; Psychological Assessment; Psychological Medicine; Psychological Reports; and, *Trends in Neurosciences.*

APPENDIX B
LIST OF REFERENCES USED IN THE POWER AND META-ANALYSES

References marked with a single asterisk indicate studies included only in the meta-analysis. References marked with a double asterisk indicate studies included only in the power analysis. A triple asterisk indicates studies that were included in both the meta-analysis and the power analysis.

**Agniel, A., Celsis, P., Viallard, G., Montastruc, J. L., Rascol, O., Demonet, J. F., Marc-Vergnes, J. P., & Rascol, A. (1991). Cognition and cerebral blood flow in lateralised Parkinsonism: Lack of functional lateral asymmetries. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 783-786.

***Appollonio, I., Grafman, J., Clark, K., Nichelli, P., Zeffiro, T., & Hallet, M. (1994). Implicit and explicit memory in patients with Parkinson's disease with and without dementia. *Archives of Neurology*, *51*, 359- 367.

**Auriacombe, S., Grossman, M., Carvell, S., Gollomp, S., Stern, M. B., & Hurtig, H. I. (1993). Verbal fluency deficits in Parkinson's disease. *Neuropsychology*, *7*, 182-192.

**Blonder, L. X., Gur, R. E., Gur, R. C., Saykin, A. J., & Hurtig, H. I. (1989). Neuropsychological functioning in hemiparkinsonism. *Brain and Cognition*, *9*, 244-257.

**Bondi, M. W., & Kaszniak, A. W. (1991). Implicit and explicit memory in Alzheimer's disease and Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, *13*, 339-358.

***Bondi, M. W., Kasniak, A. W., Bayles, K. A., & Vance, K. T. (1993). Contributions of frontal system dysfunction to memory and perceptual abilities in Parkinson's disease. *Neuropsychology*, *7*, 89- 102.

**Boyd, J. L., Cruikshank, C. A., Kenn, C. W., Madeley, P., Mindham, R. H. S., Oswald, A. G., Smith, R. J., & Spokes, R. G. S. (1991). Cognitive impairment and dementia in Parkinson's disease: a controlled study. *Psychological Medicine*, *21*, 911-921.

***Breen, E. K. (1993). Recall and recognition memory in Parkinson's disease. *Cortex*, *29*, 91-102.

**Cooper, J. A., & Sagar, H. J. (1993a). Encoding deficits in untreated Parkinson's disease. *Cortex*, *29*, 261-265.

**Cooper, J. A., & Sagar, H. J. (1993b). Incidental and intentional recall in Parkinson's disease: an account based on diminished attentional resources. *Journal of Clinical and Experimental Neuropsychology*, *15*, 713-731.

***Cooper, J. A., Sagar, H. J., & Sullivan, E. V. (1993). Short-term memory and temporal ordering in early Parkinson's disease: effects of disease chronicity and medication. *Neuropsychologia*, *31*, 933-949.

**Cooper, J. A., Sagar, H. J., Doherty, S. M., Jordan, N., Tidswell, P., & Sullivan, E. V. (1992). Different effects of dopaminergic and anticholinergic therapies on cognitive and motor function in Parkinson's disease. *Brain*, *115*, 1701-1725.

**Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S., & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain*, *114*, 2095-2122.

*Cronin-Golomb, A., & Braun, A. E. (1997). Visuospatial dysfunction and problem solving in Parkinson's disease. *Neuropsychology*, *11*, 44-52.

*Dewick, H. C., Hanley, J. R., Davies, A. D. M., Playfer, J., & Turnbull, C. (1991). Perception and memory for faces in Parkinson's disease. *Neuropsychologia*, *29*, 785-802.

**Dubois, B., Pillon, B., Legault, F., Agid, Y., & Lhermitte F. (1988). Slowing of cognitive processing in progressive supranuclear palsy. *Archchives of Neurology*, *45*, 1194-1198.

**El-Awar, M., Becker, J. T., Hammond, K. M., Nebes, R. D., & Boller, F. (1987). Learning deficit in Parkinson's disease: Comparison with Alzheimer's disease and normal aging. *Archives of Neurology*, *44*, 180-184.

*Flowers, K. A., Pearce, I., & Pearce, J. M. S. (1984). Recognition memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 1174-1181.

*Gabrieli, J. D. E., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence for the role of a frontostriatal system in working strategic memory. *Neuropsychology*, *10*, 322-332.

**Grossman, M., Carvell, S., Peltzer, L., Stern, M. B., Gollomp, S., & Hurtig, H. I. (1993). Visual construction impairments in Parkinson's disease. *Neuropsychology*, *7*, 536-547.

**Grossman, M., Carvell, S., Stern, M. B., Gollomp, S.,& Hurtig, H.I. (1992). Sentence comprehension in Parkinson's disease: the role of attention and memory. *Brain and Language*, *42*, 437-384.

**Gurd, J. M., & Ward, C. D. (1989). Retrieval from semantic and letter initial categories in patients with Parkinson's disease. *Neuropsychologia*, *27*, 743-746.

**Harrington, D. L., Haaland, K. Y., Yeo, R. A., & Marder, E. (1990). Procedural memory in Parkinson's disease: impaired motor but not visuoperceptual learning. *Journal of Clinical and Experimental Neuropsychology*, *12*, 323-339.

***Heindel, W. C., Salmon, D. P., Shults, C. W., Walicke, P. A., & Butters, N. (1989). Neuropsychological evidence for multiple implicit memory systems: a comparison of Alzheimer's, Huntington's, and Parkinson's disease patients. *The Journal of Neuroscience*, *9*, 582- 587.

**Helkala, E. L., Laulumaa, V., Soininen, H., & Riekkinen, P. J. (1989). Different error pattern of episodic and semantic memory in Alzheimer's disease and Parkinson's disease with dementia. *Neuropsychologia*, *27*, 1241- 1248.

**Hietanen, M., & Teravainen, H. (1986). Cognitive performance in early Parkinson's disease. *Acta Neurologica Scandinavica*, *73*, 151-159.

**Hietanen, M., & Teravainen, H. (1988). The effect of age of disease onset on neuropsychological performance in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *51*, 244-249.

**Huber, S. J., Freidenberg, D. L., Shuttleworth, E. C., Paulson, G. W., & Clapp, L. E. (1989). Neuropsychological similarities in lateralised Parkinsonism. *Cortex*, *25*, 461-470.

***Huber, S. J., Shulman, H. G., Paulson, G. W., & Shuttleworth, E. C. (1987). Fluctuations in plasma dopamine level impair memory in Parkinson's disease. *Neurology*, *37*, 1371-1375.

**Huber, S. J., Shuttleworth, E. C., & Paulson, G. W. (1986). Dementia in Parkinson's disease. *Archives of Neurology*, *43*, 987-990.

**Huberman, M., Moscovitch, M., & Freedman, M. (1994). Comparison of patients with Alzheimer's and Parkinson's disease on different explicit and implicit tests of memory. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 7*, 185-193.

**Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A. M., & Paul, G. M. (1992). L-Dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology, 107*, 394-404.

*Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. *Brain, 106*, 257-270.

**Levin, B. E., Llabre, M. M., & Weiner, W. J. (1989). Cognitive impairment associated with early Parkinson's disease. *Neurology, 39*, 557- 561.

*Litvan, I., Mohr, E., Williams, J., Gomez, C., & Chase, T. N. (1991). Differential memory and executive functions in demented patients with Parkinson's and Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 54*, 25-29.

***Massman, P. J., Delis, D. C., Butters, N., Levin, B. E., & Salmon, D. P. (1990). Are all subcortical dementias alike? Verbal learning and memory in Parkinson's and Huntington's disease patients. *Journal of Clinical and Experimental Neuropsychology, 12*, 729-724.

**Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 51*, 757-766.

**Ogden, J. A., Growden, J. H., & Corkin, S. (1990). Deficits on visuospatial tests involving forward planning in high-functioning Parkinsonians. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 3*, 125-139.

*Owen, A. M., Beksinska, M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Sahakian, B. J., & Robbins, T. W. (1993). Visuo-spatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia, 31*, 627-644.

***Owen, A. M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Lange, K. W., & Robbins, T. W. (1992). Fronto-striatal cognitive deficits at different stages of Parkinson's disease. *Brain, 115*, 1727-1751.

**Owen, A. M., Sahakian, B. J., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1995). Dopamine-dependent frontostriatal planning deficits in early Parkinson's disease. *Neuropsychology, 9*, 126-140.

**Oyebode, J. R., Barker, W. A., Blessed, G., Dick, D. J., & Britton P. G. (1986). Cognitive functioning in Parkinson's disease: In relation to prevalence of dementia and psychiatric diagnosis. *British Journal of Psychiatry, 149*, 720-725.

**Pillon, B., Dubois, B., Lhermitte, F., & Agid, Y. (1986). Heterogeneity of cognitive impairment in progressive supranuclear palsy, Parkinson's disease, and Alzheimer's disease. *Neurology, 36*, 1179-1185.

**Randolph, C., Braun, A. R., Goldberg, T. E., & Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: dissociation of storage and retrieval failures. *Neuropsychology, 7*, 82-88.

**Robbins, T. W., James, M., Owen, A. M., Lange, K. W., Lees, A. J., Leigh, P. N., Marsden, C. D., Quinn, N. P., & Summers, B. A. (1994). Cognitive deficits in progressive supranuclear palsy, Parkinson's disease, and multiple

system atrophy in tests sensitive to frontal lobe dysfunction. *Journal of Neurology, Neurosurgery, and Psychiatry, 57*, 79-88.

**Sagar, H. J., Sullivan, E. V., Cooper, J. A., & Jordan, N. (1991). Normal release from proactive interference in untreated patients with Parkinson's disease. *Neuropsychologia, 29*, 1033-1044.

***Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain, 111*, 695-718.

*Sullivan, E. V., & Sagar, H. J. (1989). Nonverbal recognition and recency discrimination deficits in Parkinson's disease and Alzheimer's disease. *Brain, 112*, 1503-1517.

**Sullivan, E. V., Sagar, H. J., Cooper, J. A., & Jordan, N. (1993). Verbal and nonverbal short-term memory impairment in untreated Parkinson's disease. *Neuropsychology, 7*, 396-405.

***Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1986). Frontal lobe dysfunction in Parkinson's disease. *Brain, 109*, 845- 883.

*Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1987). Parkinson's disease. *Brain, 110*, 35-51.

**Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1990). Memory and learning in early Parkinson's disease: evidence for a "frontal lobe syndrome". *Brain and Cognition, 13*, 211-232.

*Tierney, M. C., Nores, A., Snow, W. G., Fisher, R. H., Zorzitto, M. L., & Reid, D. W. (1994). Use of the Rey auditory verbal learning test in differentiating normal aging from Alzheimer's and Parkinson's dementia. *Psychological Assessment, 6*, 129-134.

APPENDIX C
DESCRIPTION OF STUDIES
USED IN THE META-ANALYSIS

Table C1 *Description of Studies Included in Meta-analysis: Parkinson's Disease (PD) and Control Samples*

Study	Sample	Description	Age (Years)	M/F	H&Y Stage	Disease Duration (Years)
Appollonio et al., 1994	Control	Spouses of PD sample or paid participants matched for age, gender and educational level	62.4	5/7	N/A	N/A
	NPD	Idiopathic PD, none met the MDRS criteria for dementia, but had BDI criteria for elevated depression; mostly medicated	57.2	6/7	NR	6.25
	DPD	Idiopathic PD, all met the MDRS criteria for dementia, none met the BDI criteria for depression; mostly medicated	62.4	5/0	NR	5.75
Bondi et al., 1993	Control	No information given other than matched for age, estimated premorbid intelligence and education level	69.3	7/12	N/A	N/A
	NPD	Idiopathic PD, none met the modified Hachinski criteria for dementia, or the HDRS for depression; all medicated	67.3	13/7	I = 5 II = 9 III = 5 IV = 0 V = 0	8.00 (median)
Breen, 1993	Control	Spouses of PD sample matched approximately for educational and socio-economic level	64.5	6/9	N/A	N/A
	NPD	None met self-report criteria for dementia or depression; all medicated	64.8	9/6	NR	8.80
Cooper et al., 1993	Control	Spouses of PD sample and volunteers matched for age and gender	59.1	24/21	N/A	N/A
	NPD De novo	Newly diagnosed PD, none met DSM-III criteria for dementia, but some met BDS criteria for dementia, and BDI criteria for depression; *de novo*	58.6	39/37	NR	1.40
	NPD Medicated	Chronically medicated PD, none met DSM-III criteria for dementia, but some met BDS criteria for dementia, and BDI criteria for depression; all medicated	62.0	17/5	NR	7.60
Cronin-Golomb & Braun, 1997	Control	Spouses and other volunteers matched for age and education level	66.9	29/10	N/A	N/A

(table continues)

Table C1 *(continued)*

Study	Sample	Description	Age (Years)	M/F	H&Y Stage	Disease Duration (Years)
	NPD	Idiopathic PD, none met BDS criteria for dementia, or BDI/DSM criteria for depression; all medicated	67.4	17/33	I = 9 II = 29 III = 7 IV = 5 V = 0	7.2
Dewick et al., 1991	Control	No information given other than matched for education level, age, MMSE, verbal ability, and premorbid intelligence	70.9	9/12	N/A	N/A
	NPD	Idiopathic PD, none met MMSE criteria for dementia, but some met GDS criteria for depression; all but one medicated	72.9	10/9	I = 2 II = 6 III = 8 IV = 2 V = 1	4.60
Flowers et al., 1984	Control	Spouses of the PD sample and patients with peripheral nerve or spinal cord afflictions and local volunteers matched for age, occupation and background	61.1	24/20	N/A	N/A
	PD	All medicated	62.5	28/16	NR	1-33
Gabrieli et al., 1996	Control	Hospital volunteers, community volunteers, and staff members matched for age, education, and MMSE scores	55.5	2/8	N/A	N/A
	NPD	None met MMSE criteria for dementia; *de novo*	60.1	6/4	I = 0 II = 9 III = 1 IV = 0 V = 0	2.9
Heindel et al., 1989	Control	Spouses of the PD sample or paid participants obtained through newspaper advertisements	71.3	2/10	N/A	N/A
	NPD	Idiopathic PD, none met DRS criteria for dementia; all medicated	62.7	7/2	NR	11.00
	DPD	Idiopathic PD, all met DRS criteria for dementia; all medicated	72.4	8/0	NR	6.40
Helkala et al., 1989	Control	Relatives of staff members and other patients matched for age	68.0	5/18	N/A	N/A
	DPD	Idiopathic PD, all met the CDR criteria for dementia	70.0	11/7	I = 0 II = 2 III = 2 IV = 14 V = 0	NR
Huber et al., 1987	Control	Mainly spouses of the PD sample	NR	NR	N/A	N/A
	NPD	Idiopathic PD, none met the criteria for dementia; all medicated	NR	NR	NR	NR

(table continues)

Table C1 *(continued)*

Study	Sample	Description	Age (Years)	M/F	H&Y Stage	Disease Duration (Years)
Lees & Smith, 1983	Control	Hospital inpatients mainly waiting for elective surgery for prolapsed intervertebral discs or carpal tunnel syndrome	53.6	17/13	N/A	N/A
	NPD	Idiopathic PD, none met the Hachinski Ischaemia score criteria for dementia, or the Zung scale criteria for depression; *de novo*	58.7	19/11	I = 12 II = 13 III = 5 IV = 0 V = 0	2.4
Litvan et al., 1991	Control	No information given other than matched for age, and education level	66.0	6/5	N/A	N/A
	DPD	Idiopathic PD, all met DSM-IIIR criteria for dementia; all but one medicated	69.0	10/1	NR	13
Massman et al., 1990	Control	Recruited from the community with comparable age and education level	53.2	13/6	N/A	N/A
	PD	Idiopathic PD; all medicated	57.5	10/9	I = 5 II = 4 III = 5 IV = 5 V = 0	6.3
Owen et al., 1992	Control	No information given other than matched for age and premorbid verbal IQ	60.8	NR	N/A	N/A
	NPD De novo	Idiopathic PD; *de novo*	55.7	NR	I = 3 II = 10 III = 2 IV = 0 V = 0	1.50
	NPD Medicated Mild	Idiopathic PD, none met the MMSE or KOLT criteria for dementia, but some met the GDS criteria for depression; all medicated	58.9	NR	I = 3 II = 12 III = 0 IV = 0 V = 0	7.10
	NPD Medicated Severe	Idiopathic PD, none met the MMSE or KOLT criteria for dementia, but some met the GDS criteria for depression; all medicated	65.9	NR	I = 0 II = 0 III = 8 IV = 6 V = 0	10.20
Owen et al., 1993	Control	No information given other than matched for age and premorbid verbal IQ	65.6	21/21	N/A	N/A
	NPD De novo	Idiopathic PD; *de novo*	61.4	11/7	I = 9 II = 7 III = 2 IV = 0 V = 0	1.75
	NPD Medicated Mild	Idiopathic PD, none met the MMSE or KOLT criteria for dementia, but some met the GDS criteria for depression; all medicated	63.3	7/4	I = 2 II = 9 III = 0 IV = 0 V = 0	9.73

(table continues)

Table C1 *(continued)*

Study	Sample	Description	Age (Years)	M/F	H&Y Stage	Disease Duration (Years)
	NPD Medicated Severe	Idiopathic PD, none met the MMSE or KOLT criteria for dementia, but some met the GDS criteria for depression; all medicated	66.5	8/5	I = 0 II = 0 III = 7 IV = 6 V = 0	10.60
Sagar et al., 1988	Control	No information given other than matched for age and years of education	63.1	7/8	N/A	N/A
	PD	Only three met the DSM-III criteria for dementia; all were medicated	62.5	13/2	NR	6.00
Sahakian et al., 1988	Control	Volunteers drawn from the community matched for age and verbal IQ	61.9	NR	N/A	N/A
	NPD De novo	Idiopathic PD, none were diagnosed as demented; *de novo*	61.3	8/5	I = 7 II = 4 III = 2 IV = 0 V = 0	1.7
	Control	Volunteers drawn from the community matched for age and verbal IQ	65.1	NR	N/A	N/A
	NPD Medicated	Idiopathic PD, none were diagnosed as demented; all medicated	64.2	8/6	I = 2 II = 3 III = 9 IV = 0 V = 0	7.6
Sullivan & Sagar, 1989	Controls	All controls had also participated in an earlier study (Sagar et al., 1988).	63.1	7/8	N/A	N/A
	PD	All of the PD sample had also participated in an earlier study (Sagar et al., 1988). Only one from that study could not participate in this study	64.1	12/2	I = 1 II-IV =13	5.9
Taylor et al., 1986	Control	Relatives of the PD sample or drawn through the Canadian PD Foundation matched for age, education, and verbal IQ	60.8	21/19	N/A	N/A
	NPD	Idiopathic PD, none were demented; all except 12 were medicated	60.5	25/15	I = 10 II = 13 III = 14 IV = 3 V = 0	6.62
Taylor et al., 1987	Control	Volunteers matched for age, education, and premorbid IQ	62.4	NR	N/A	N/A
	NPD De novo	Idiopathic PD, none met clinical or neuropsychological criteria for dementia; *de novo*	61.1	NR	I = 7 II = 2 III = 0 IV = 0 V = 0	2.1

(table continues)

Table C1 *(continued)*

Study	Sample	Description	Age (Years)	M/F	H&Y Stage	Disease Duration (Years)
	NPD Good response to treatment	Idiopathic PD, none met clinical or neuropsychological criteria for dementia; all medicated	62.0	NR	I = 7 II = 2 III = 0 IV = 0 V = 0	9.3
	NPD Fluctuations and abnormal involuntary movements	Idiopathic PD, none met clinical or neuropsychological criteria for dementia; all medicated	61.1	NR	I = 7 II = 2 III = 0 IV = 0 V = 0	13.8
	NPD Poor response to treatment	Idiopathic PD, none met clinical or neuropsychological criteria for dementia; all medicated	61.2	NR	I = 7 II = 2 III = 0 IV = 0 V = 0	8.0
Tierney et al., 1994	Control	No information given, but age is comparable to the PD sample	75.4	16/22	N/A	N/A
	DPD	All met the NINCDS-AD-RDA Workgroup and the DSM-III criteria for dementia	72.5	11/1	NR	NR

Note. NPD = nondemented Parkinson's disease; DPD = demented Parkinson's disease; PD = unselected Parkinson's disease; De novo = never received antiparkinsonian medication; Medicated = currently receiving antiparkinsonian medication; Mild = mild disease severity; Severe = severe disease severity; M/F = number of male/female participants; H&Y Stage = Hoehn & Yahr staging of disease severity (Roman numerals indicate stage, Arabic numbers indicate the number of participants) ; N/A = not appropriate; NR = data not reported; Exclusion criteria = criteria used by each study to exclude participants from the sample.

Table C2 *Description of Studies Included in Meta-analysis: Recognition Task and Effect Size*

Study	PD Sample	Task	Type	Meas	Mod	*N*	*d*
Appollonio et al., 1994	NPD	Word Fragment Completion Test (recognition)	20 yes–no (10 distractors)	H - FA	V	25	.60
	DPD	Word Fragment Completion Test (recognition)	20 yes–no (10 distractors)	H - FA	V	17	1.54
Bondi et al., 1993	NPD	Verbal Continuous-Recognition-Memory Task	48, 2AFC	PC	V	38	1.10
Breen, 1993	NPD	Word list learning	120, yes–no (60 distractors)	H - FA	V	30	.00
Cooper et al., 1993	NPD De novo	Verbal Temporal Ordering (content recognition: averaged across all stimulus-test intervals)	5, 2AFC	PC	V	121	.00
	NPD Med	Verbal Temporal Ordering (content recognition: averaged across all stimulus-test intervals)	5, 2AFC	PC	V	67	.31
Cronin-Golomb & Braun, 1997	NPD	Randt Memory Test (story subtest-immediate recognition)	7, 4AFC	PC	V	77	.00
		Randt Memory Test (story subtest-delayed recognition)	7, 4AFC	PC	V	82	.00
Dewick et al., 1991	NPD	Warrington Recognition Memory Test (words)	50, 2AFC	PC	V	40	.22
		Words & Numbers Verbal Recognition Memory Test (immediate recognition)	24, yes–no (12 distractors)	*d'*	V	39	.30
		Words & Numbers Verbal Recognition Memory Test (delayed recognition)	24, yes–no (12 distractors)	*d'*	V	39	.17
Flowers et al., 1984	PD	Object test (immediate recognition)	16, 5AFC	PC	V	98	.17
		Object test (delayed recognition)	16, 5AFC	PC	V	98	.24
		Histogram test (immediate recognition)	8, 2AFC	PC	NV	88	.07

(table continues)

Table C2 *(continued)*

Study	PD Sample	Task	Type	Meas	Mod	*N*	*d*
		Histogram test (delayed recognition)	8, 2AFC	PC	NV	88	.21
		Words & numbers test (immediate recognition)	12 yes–no (6 distractors)	*d'*	V	99	.33
		Words & numbers test (delayed recognition)	12 yes–no (6 distractors)	*d'*	V	99	.27
		Abstract art test (immediate recognition)	10 yes–no (5 distractors)	*d"*	NV	69	.16
		Abstract art test (delayed recognition)	10 yes–no (5 distractors)	*d'*	NV	69	.38
Gabrieli et al., 1996	NPD	Word-list task	24, 2AFC (12 distractors)	PC	V	20	.54
Heindel et al., 1989	NPD	Word-list learning	10, 2AFC	PC	V	21	.00
	DPD	Word-list learning	10, 2AFC	PC	V	20	.00
Helkala et al., 1989	DPD	Shopping-list word learning	20, yes–no (10 distractors)	R	V	41	1.74
Huber et al., 1987	NPD	Word-list learning	24, 3-AFC	PC	V	16	.25
Lees & Smith, 1983	NPD	Warrington recognition memory test	50, 2AFC	PC	V	60	.06
Litvan et al., 1991	DPD	RAVLT	n, 4AFC	PC	V	22	1.40
Massman et al., 1990	PD	California Verbal Learning Test-recognition	44, yes–no (28 distractors)	D	V	38	1.06
Owen et al., 1992	NPD De novo	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	59	.06
	NPD Med Mild	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	59	.19
	NPD Med Severe	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	58	.56
Owen et al., 1993	NPD De novo	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	60	.00

(table continues)

Table C2 *(continued)*

Study	PD Sample	Task	Type	Meas	Mod	*N*	*d*
	NPD Med Mild	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	53	.00
	NPD Med Severe	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	55	.00
Sagar et al., 1988	PD	Verbal Temporal Ordering Test	45, 2AFC	PC	V	30	.67
Sahakian et al., 1988	NPD De novo	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	26	.00
	NPD Med	Cambridge Neuropsychological Test Automated Battery-pattern recognition	24, 2AFC	PC	NV	28	1.44
Sullivan & Sagar, 1989	PD Same sample as Sagar et al., 1988	Nonverbal Temporal Ordering Test	45, 2AFC	PC	NV	29	.86
Taylor et al., 1986	NPD	RAVLT	30, yes-no (15 distractors)	H - FA	V	80	.26
Taylor et al., 1987	NPD De novo	RAVLT	30, yes-no (15 distractors)	H - FA	V	18	0
	NPD[a]	RAVLT	30, yes-no (15 distractors)	H - FA	V	18	0
	NPD[b]	RAVLT	30, yes-no (15 distractors)	H - FA	V	18	0
	NPD[c]	RAVLT	30, yes-no (15 distractors)	H - FA	V	18	0
Tierney et al., 1994	DPD	RAVLT	30, yes-no (15 distractors)	D	V	50	.90

Note. PD Sample = Parkinson's disease sample; NPD = nondemented Parkinson's disease; DPD = demented Parkinson's disease; PD = unselected Parkinson's disease; De novo = never received antiparkinsonian medication; Med = currently receiving antiparkinsonian medication; Mild = mild disease severity; Severe = severe disease severity; Task = name of recognition memory task; RAVLT = Rey Auditory Verbal Learning Test; Type = type of recognition task and number of stimuli; AFC = m-alternative, forced-choice procedure; Measure = measure of recognition performance; PC = percent correct; *d'* = *d* - prime; R = Recognition score, calculated from the formula [Hits - ((Hits + False Alarms)/2)] + 8; D = Discriminability index, calculated from the formula [(true positives + true negatives) / 30] x 100; Mod = stimuli were classified as either verbal (V) stimuli (those that could be easily labelled) including visual stimuli (e.g., pictures of common objects), or as nonverbal (NV) stimuli (e.g., abstract drawings); *N* = total number of Parkinson's disease and control participants; *d* = Cohen's (1988) effect size statistic.
[a]Good response to treatment. [b]Fluctuations and abnormal involuntary movements. [c]Poor response to treatment.

APPENDIX D
EFFECT SIZE FORMULAE

Formulae were obtained from Hunter and Schmidt (1990) and Schwarzer (1989). For the between-groups planned comparisons, *d* was defined as the difference between the group means divided by the within-group standard deviation:

$$d = (Y_e - Y_c) / S_W,$$

where S_W is the within-group standard deviation, defined by taking the square root of the within-group variance, V_W, and

$$V_W = [(n_e - 1)\ V_e + (n_c - 1)\ V_c\ ] / [(n_e - 1) + (n_c - 1)].$$

V_e is the variance for the experimental group, V_c the variance for the control group, and n_e and n_c are the sample sizes for the experimental group and control group, respectively.

For the within-groups planned comparisons, *d* was defined as the difference between the group means divided by the true score standard deviation:

$$d = (Y_e - Y_c) / \sqrt{[r\ V_W]},$$

where *r* is the correlation between repeated measures and V_W is the within-group variance.

Where multi-factorial ANOVA was used, *d* was computed for all main effects and interactions using eta-squared (η^2). The transformation formula is:

$$d = 2\sqrt{\eta^2 / (1 - \eta^2)},$$

where η^2 is defined by:

$$\eta^2 = [F(\text{df effect})] / [F(\text{df effect}) + \text{df error}]$$

The *d* statistic can also be estimated from the Pearson chi-square statistic (χ^2) after first transforming χ^2 to the phi coefficient (ϕ):

$$\phi = \sqrt{\frac{\chi^2}{N}}.$$

For a 2 x 2 contingency table, the phi coefficient is equal to the Pearson correlation coefficient and can be used in the formula:

$$d = [(1/\sqrt{pq})\,\phi] / [\sqrt{1-\phi^2}],$$

where *p* and *q* are the proportion of persons in the two groups.

The *d* statistic can also be estimated from *r*, using the maximum likelihood formula:

$$d = 2r / \sqrt{(1-r^2)}.$$

The *d* statistic can also be estimated from Mann-Whitney's U after first transforming to *r*:

$$r = 1 - 2U / (n_1 n_2).$$

The 95% confidence interval for the population value of the *d* statistic is given to a close approximation by:

$$d - 1.96\,SE < \delta < d + 1.96\,SE,$$

where *SE* is the sampling error standard deviation or *standard error*. For the between-groups case, *SE* was estimated by:

$$SE = \sqrt{[(N-1)/(N-3)][(4/N)(1 + d^2/8)]}.$$

For the within-groups case, *SE* was estimated by:

$$SE = \sqrt{[2\ (1 - r)\ /\ N] + [d^2\ (1 + r^2)\ /\ \{4\ (N - 1)\}]}\ ,$$

where *r* is the correlation between repeated measures.

APPENDIX E

PARTICIPANT INFORMATION SHEET

Information for Participants

We are conducting an investigation that will help us find out if Parkinson's disease causes memory problems, and if so, whether these problems progress over a period of time. We need both individuals with and without Parkinson's disease to aid us in our study. The researchers for this study are Dr. John Podd and Craig Whittington (both from the Department of Psychology, Massey University). Dr. Podd can be contacted at work by phoning xxx-xxxx. His home phone number is xxx-xxxx. Craig Whittington's work phone number is xxx-xxxx and home phone number is xxx-xxxx.

Craig Whittington will visit you on two occasions separated by approximately six months. During each visit you will be asked questions about your health and medical history, and will be asked to do some simple memory tasks. Most of the tasks will be presented on a computer and you will give your answers by pressing specially designed buttons or speaking aloud. It does not matter if you have never used a computer before, or even seen one. We use it simply to display what we want you to remember. The computer also automatically records your answers. Plenty of practice is given and you will find the tasks quite interesting.

The time involved during each visit is about 1 - 1 1/2 hours. Participation is on a strictly voluntary basis without monetary compensation. The questions about your health and medical history will take about 15 minutes and the memory tasks will take about 45 minutes, including practice and rest intervals. The extra 30 minutes is to introduce you to the study and for you to be able to ask any questions. Of course, you can stop the study at any time if you feel too tired to continue, or if any of the tasks bother you. If you do find the procedure distressing and wish to talk to someone about this, you can phone Dr. John Podd on xxx-xxxx.

If you agree to take part in our study, you have the following rights:

* Refuse to answer any particular question we might ask.

* Be given time to consider and discuss participation with others if desired.

* Withdraw from the study at any time without adverse affects to any further health care.

* Ask questions as they occur to you at any time during your participation.

* Provide information on the understanding that it is completely confidential to the researchers. All information is collected anonymously, and it will not be possible to identify you in any reports prepared from the study.

* Be given access to your own personal data, and a copy of it if required.

* Be given access to a summary of the findings from the study when it is concluded.

* Contact the Manawatu-Whanganui Ethics Committee and discuss any ethical concerns you have about the study; phone (xx) xxx-xxxx or write to PO Box xxxx, Palmerston North.

Does Parkinson's disease cause memory problems? And if so, do they progress over the short-term? These are interesting questions for us. We hope you will be willing to help us find some answers by agreeing to participate in our study. In return for your participation, we will be very willing to tell you as much as we can about memory problems and why they interest us so much.

Dr. John Podd (Department of Psychology, Massey University)
Phone: xxx-xxxx

Craig Whittington (Department of Psychology, Massey University)
Phone: xxx-xxxx

APPENDIX F
DESCRIPTION OF MEASURES

Description of Measures Used in the Study

Consent forms

An ethical requirement.

Two Prospective Memory Tasks

These measure your memory for doing something in the future.

Visual Analogue Scale

This measures how nervous you feel about taking part in this study.

Structured Interview

Used to collect important information about you. All information will be kept confidential.

Two Recognition Memory Tasks

These measure your recognition memory for words and abstract drawings.

Three Recall Tasks

These measure your recall of words and pictures.

National Adult Reading Task

This measures your ability to pronounce uncommon words.

Mood Scale

This measures how you have been feeling over the past week.

Mackworth Clock

This measures your watch-keeping ability.

Short Orientation-Memory-Concentration Test

This is a brief measure of your orientation, memory and concentration.

Video Assessment (Parkinson's Participants Only)

Used to assess how Parkinson's effects you physically.

APPENDIX G

PARTICIPANT CONSENT FORM

Consent Form

Project Title: COGNITIVE DEFICITS IN PARKINSON'S DISEASE

Principal Investigator(s): Dr. John Podd and Mr Craig Whittington

Participant's Name: ..... ID .....

Name of Institution: Massey University

English	I wish to have an interpreter	Yes / No
Maori	E hiahia ana koe ki tetahi tangata hei korero Maori kia koe	Ae / Kao

1. Craig Whittington has explained to me the reasons for this study and the procedures involved in it.

2. I have read the participant information sheet, and my questions have been answered to my satisfaction. I understand that I am able to ask further questions at any time during the study.

3. I understand the compensation arrangements that are available to me.

4. I understand that I am free to withdraw from the study at any time, and that such withdrawal will not adversely affect my further health care.

5. I have been assured that my results will remain confidential and that my identity will not be revealed in any written or verbal reports about the study.

6. I understand that the study will be discontinued if it appears that it could cause me harm or if I do not follow the required procedures.

7. I understand that if I have any ethical concerns regarding the study, that I may contact the Manawatu-Whanganui Ethics Committee on 0-6 xxx-xxxx.

I agree / do not agree to take part in this study.

Signature..... (participant) ..... (date)

..... (parent/guardian) ..... (date)

..... (witness)..... (date)

..... (name of witness, please print)

Investigator's statement:

I have discussed with ..... (participant's name) the aims of and procedures involved in this study.

Signature..... (investigator) ..... (date)

Craig Whittington..... (name of investigator)

APPENDIX H

VIDEO ASSESSMENT CONSENT FORM

Consent Form for Video Assessment

Project Title: COGNITIVE DEFICITS IN PARKINSON'S DISEASE
Principal Investigators: Dr. John Podd and Mr Craig Whittington
Participant's Name: .....ID:.....
Name of Institution: Massey University

In order to assess how Parkinson's disease has effected you physically, it will be useful to video you performing a number of actions. For the video assessment to be done accurately, it may be necessary for a professional such as a neurologist to view the tape. The tape will be kept strictly confidential and only viewed by the investigator and persons qualified to assess the severity of Parkinson's disease. At the conclusion of the study, the tape will be erased. Please note, the video assessment is optional and if you decline to undergo the video assessment you may still take part in the rest of the study.

I understand the above and **agree / do not agree** to undergo the video assessment.

Signature.....(participant) .....(date)
Signature.....(witness) .....(date)
.....(name of witness, please print)
Signature.....(investigator) .....(date)
Craig Whittington.....(name of investigator)

APPENDIX I

HOEHN AND YAHR DISABILITY RATING SCALE

Stage I Unilateral involvement only, usually with minimal or no functional impairment.

Stage II Bilateral or midline involvement, without impairment of balance.

Stage III First sign of impaired righting reflexes. This is evident by unsteadiness as the patient turns or is demonstrated when he [or she] is pushed from standing equilibrium with the feet together and eyes closed. Functionally the patient is somewhat restricted in his [or her] activities but may have some work potential depending upon the type of employment. Patients are physically capable of leading independent lives, and their disability is mild to moderate.

Stage IV Fully developed, severely disabling disease; the patient is still able to walk and stand unassisted but is markedly incapacitated.

Stage V Confinement to bed or wheelchair unless aided.

(Hoehn & Yahr, 1967, p. 433)

APPENDIX J

VIDEO ASSESSMENT

INSTRUCTIONS: The next task is used so I can assess how Parkinson's has affected you physically. I will ask you a number of questions and then get you to perform a number of actions and record you doing these with the video camera.

Questions:

1) Are your disease symptoms restricted to one side of your body?

2) Is your balance as good as it used to be?

Motor Tasks:

Finger tapping:

Please tap the index finger of your left hand for 30 seconds while your hand is rested flat upon the table, palm down. Please repeat with your right hand.

Rising from chair:

Please sit in a hard, upright chair. Now attempt to rise with your arms folded across your chest. If you can't do this use your arms to help you stand up.

Walking and turning:

Please walk at a natural pace for about 4 metres, turn on the spot, and return to the starting point.

APPENDIX K

MODIFIED ACTIVITIES OF DAILY LIVING SCALE

Speech:-

0-normal
1-mildly affected, no difficulty being understood
2-moderately affected, may be asked to repeat
3-severely affected, frequently asked to repeat
4-unintelligible most of time

Swallowing:-

0-normal
1-rare choking
2-occasional choking
3-requires soft food
4-requires NG tube or G-tube

Handwriting:-

0-normal
1-slightly small or slow
2-all words small but legible
3-severely affected, not all words legible
4-majority illegible

Cutting Food/Handling Utensils:-

0-normal
1-somewhat slow and clumsy but no help needed
2-can cut most foods, some help needed
3-food must be cut, but can feed self
4-needs to be fed

Dressing:-

0-normal
1-somewhat slow, no help needed
2-occasional help with buttons or arms in sleeves
3-considerable help required but can do some things alone
4-helpless

Hygiene:-

0-normal
1-somewhat slow but no help needed
2-needs help with shower or bath or very slow in hygienic care
3-requires assistance for washing, brushing teeth, going to bathroom
4-helpless

Turning in Bed/Adjusting Bed Clothes:-

0-normal
1-somewhat slow no help needed
2-can turn alone or adjust sheets but with great difficulty
3-can initiate but not turn or adjust alone
4-helpless

Walking:-

0-normal
1-mild difficulty, may drag legs or decrease arm swing
2-moderate difficulty requires no assistance
3-severe disturbance requires assistance
4-cannot walk at all even with assistance

(van Hilten et al., 1994)

APPENDIX L
GERIATRIC DEPRESSION SCALE

INSTRUCTIONS: Choose the best answer for how you felt over the past week.

Press the left key to answer "Yes" or the right key to answer "No"

1. Are you basically satisfied with your life?

2. Have you dropped many of your activities and interests?

3. Do you feel that your life is empty?

4. Do you often get bored?

5. Are you hopeful about the future?

6. Are you bothered by thoughts you can't get out of your head?

7. Are you in good spirits most of the time?

8. Are you afraid that something bad is going to happen to you?

9. Do you feel happy most of the time?

1. Do you often feel helpless?

11. Do you often get restless and fidgety?

12. Do you prefer to stay at home, rather than going out and doing new things?

13. Do you frequently worry about the future?

14. Do you feel you have more problems with memory than most?

15. Do you think it is wonderful to be alive now?

16. Do you often feel downhearted and blue?

17. Do you feel pretty worthless the way you are now?

18. Do you worry a lot about the past?

19. Do you find life very exciting?

2. Is it hard for you to get started on new projects?

21. Do you feel full of energy?

22. Do you feel that your situation is hopeless?

23. Do you think that most people are better off than you are?

24. Do you frequently get upset over little things?

25. Do you frequently feel like crying?

26. Do you have trouble concentrating?

27. Do you enjoy getting up in the morning?

28. Do you prefer to avoid social gatherings?

29. Is it easy for you to make decisions?

3. Is your mind as clear as it used to be?

For the purpose of the present study, one point was scored for each of the following answers.

1. no	6. yes	11. yes	16. yes	21. no	26. yes
2. yes	7. no	12. yes	17. yes	22. yes	27. no
3. yes	8. yes	13. yes	18. yes	23. yes	28. yes
4. yes	9. no	14. yes	19. no	24. yes	29. no
5. no	10. yes	15. no	20. yes	25. yes	30. no

APPENDIX M
WORDS USED IN THE
NATIONAL ADULT READING TEST

Ache

Debt

Psalm

Depot

Chord

Bouquet

Deny

Capon

Heir

Aisle

Subtle

Nausea

Equivocal

Naive

Thyme

Courteous

Gaoled

Procreate

Quadruped

Catacomb

Superfluous

Radix

Assignate

Gist

Hiatus

Simile

Rarefy

Cellist

Zealot

Abstemious

Gouge

Placebo

Facade

Aver

Leviathan

Aeon

Detente

Gauche

Drachm

Idyll

Beatify

Banal

Sidereal

Puerperal

Topiary

Demesne

Campanile

Labile

Synocope

Prelate

APPENDIX N

ORIENTATION-MEMORY-CONCENTRATION TEST

Participant ID:- ###

"What year is it now?" - Max error score = 1; Weight = 4

1. Score:- #

"What month is it now?" - Max error score = 1; Weight = 3

2. Score:- #

Memory phase - "Repeat this phrase after me:

John Brown, 42 Market Street, Chicago."

"About what time is it?" - Max error score = 1; Weight = 3

3. Score:- #

"Count backwards from 20 to 1." - Max error = 2 (score 1 if correct, but only after self-correction) ; Weight = 2

20:-_ 19:-_ 18:-_ 17:-_ 16:-_ 15:-_ 14:-_ 13:-_ 12:-_ 11:-_

10:-_ 9:-_ 8:-_ 7:-_ 6:-_ 5:-_ 4:-_ 3:-_ 2:-_ 1:-_

4. Score:- #

"Say the months in reverse order." - Max error = 2 (score 1 if correct, but only after self-correction) ; Weight = 2

Dec:-_ Nov:-_ Oct:-_ Sep:-_ Aug:-_ Jul:-_

Jun:-_ May:-_ Apr:-_ Mar:-_ Feb:-_ Jan:-_

5. Score:- #

"Repeat the memory phrase." - Max error score = 5; Weight = 2

John:-_ Brown:-_ 42:-_ Market Street:-_ Chicago:-_

6. Score:- #

Total Weighted Error Score:-##

APPENDIX O
VISUAL ANALOGUE SCALE

INSTRUCTIONS: Please move the bar on the line to indicate how anxious you feel right now.

Not at all anxious

Very anxious

Use the Left Arrow or Right Arrow key to move the bar on the line.

Please press the Enter key when complete.

APPENDIX P
WORDS USED IN THE VERBAL RECOGNITION MEMORY TASK

Table P1 *Frequency and Length of Target and Distractor Words Used in the Verbal Recognition Memory Task*

Target	Frequency[a]	Distractor	Frequency[a]	Word length
beach	61	bench	35	5
breath	53	wreath	8	6
broom	2	chief	119	5
bug	4	pit	14	3
candle	18	handle	53	6
cap	27	fig	72	3
creek	14	bloom	12	5
dream	64	dread	9	5
dusk	70	dust	70	4
flour	8	floor	158	5
fowl	1	foal	2	4
gate	37	crow	2	4
hall	152	poet	99	4
kite	1	mast	6	4
lamb	7	limb	5	4
lion	17	link	16	4
loot	3	fear	127	4
male	37	mall	3	4
maze	6	haze	7	4
minor	58	manor	5	5
month	130	paper	157	5
mouth	103	south	240	5
pane	3	pans	3	4
peace	198	peach	3	5
pier	3	pies	5	4
pole	18	test	119	4
rein	3	boot	13	4
rite	8	bulb	7	4
roll	35	path	44	4
sail	12	salt	46	4
sheet	45	goals	60	5
ski	5	hog	3	3
snake	44	gloom	14	5
son	166	leg	58	3
soul	47	soil	54	4
stair	2	lunch	33	4
tent	20	rent	21	4
tune	10	tone	78	4
vein	25	moth	1	4
waist	11	punch	5	5

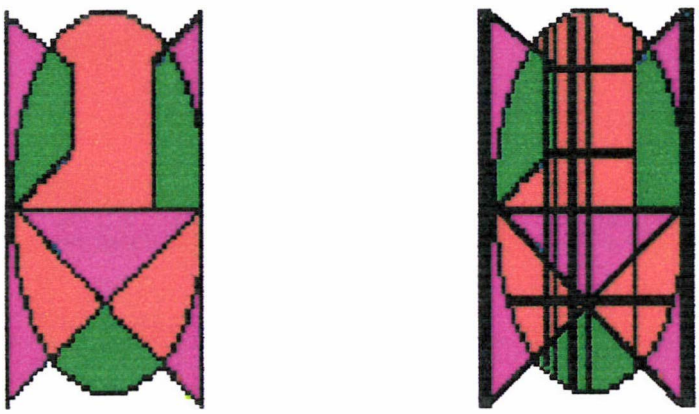
[a]Kucera-Francis frequency is the number of occurrences in a corpus of 1,014,232 words.

APPENDIX Q

ABSTRACT DRAWINGS USED IN THE NONVERBAL RECOGNITION MEMORY TASK

Shown below are two examples of target-distractor pairs used in the NRMT. The first is an example of a visually similar pair and the second is a visually dissimilar pair. During the recognition phase, a target was always paired with a unique corresponding distractor; the side on which the target appeared and the order of similar/dissimilar stimuli was randomised for each participant.

Visually similar target-distractor pair



Visually dissimilar target-distractor pair

APPENDIX R
SCREENING QUESTIONNAIRE

Have you ever been diagnosed with any of the following:

1. A medical condition with known central nervous system complication:- (Y/N)

If yes, please describe:-

---

2. A neurological disorder (e.g., Alzheimer's disease, stroke, brain tumour, head trauma with loss of consciousness greater than 1hr):- (Y/N)

If yes, please describe:-

---

3. A hereditary disease (e.g., Wilson's disease, Huntington's disease):- (Y/N)

If yes, please describe:-

---

4. Alcohol abuse:- (Y/N)

5. Major depression:-(Y/N)

If yes, when were you diagnosed:-

---

6. Have you ever had a neurosurgical operation:- (Y/N)

APPENDIX S
ANTIPARKINSONIAN MEDICATION AVAILABLE IN NEW ZEALAND

Anticholinergic Agents:

Orphenadrine Hydrochloride:	Disipal, 50mg
Benztropine Mesylate:	Cogentin, 2mg

Dopaminergic agents:

Levodopa-Benserazide:	Madopar, CR 100/25mg (controlled release)
	Madopar, 100/25mg
	Madopar, 200/50mg
Levodopa-Carbidopa:	Sinemet, CR 200/50mg (controlled release)
	Sinemet, 100/25mg

Dopamine agonists:

Bromocriptine Mesylate:	Parlodel, 2.5mg
Lisuride Hydrogen Maleate:	Dopergin, .2mg

Monoamine Oxidase-B Inhibitor:

Selegeline Hydrochloride:	Eldepryl/Deprenyl, 5mg

Other Medications:

Amantadine Hydrochloride:	Symmetrel, 100mg

APPENDIX T

CONFIDENCE RATINGS DATA

Table T1 *Recognition Memory Task Confidence Ratings for the Parkinson's Disease (PD) and Healthy Control (HC) Groups as a Function of Gender, Task, and Time*

	PD		HC	
Confidence Rating	Males[a]	Females[b]	Males[c]	Females[d]
		NRMT-Hard - Time 1		
Very Confident	.07 (.15)	.08 (.14)	.02 (.05)	.08 (.17)
Confident	.32 (.28)	.29 (.27)	.20 (.19)	.35 (.27)
Somewhat Confident	.31 (.24)	.26 (.19)	.43 (.26)	.38 (.26)
Not at all Confident	.30 (.27)	.38 (.36)	.32 (.28)	.19 (.21)
		NRMT-Hard - Time 2		
Very Confident	.07 (.17)	.07 (.13)	.03 (.05)	.11 (.23)
Confident	.26 (.21)	.31 (.29)	.27 (.21)	.35 (.31)
Somewhat Confident	.31 (.24)	.27 (.30)	.48 (.20)	.38 (.31)
Not at all Confident	.37 (.38)	.36 (.37)	.22 (.21)	.16 (.21)
		NRMT-Easy - Time 1		
Very Confident	.06 (.15)	.03 (.10)	.04 (.09)	.08 (.21)
Confident	.29 (.26)	.34 (.27)	.18 (.21)	.32 (.29)
Somewhat Confident	.32 (.23)	.27 (.19)	.46 (.27)	.38 (.27)
Not at all Confident	.33 (.29)	.36 (.33)	.32 (.29)	.22 (.23)
		NRMT-Easy - Time 2		
Very Confident	.08 (.18)	.10 (.19)	.03 (.08)	.09 (.22)
Confident	.25 (.22)	.30 (.27)	.26 (.21)	.35 (.29)
Somewhat Confident	.35 (.25)	.27 (.27)	.42 (.20)	.40 (.30)
Not at all Confident	.32 (.34)	.37 (.36)	.30 (.23)	.18 (.23)
		VRMT - Time 1		
Very Confident	.20 (.27)	.14 (.21)	.28 (.28)	.25 (.28)
Confident	.42 (.26)	.37 (.25)	.36 (.26)	.45 (.22)
Somewhat Confident	.22 (.16)	.22 (.15)	.18 (.12)	.19 (.16)
Not at all Confident	.16 (.20)	.27 (.19)	.18 (.23)	.11 (.14)
		VRMT - Time 2		
Very Confident	.26 (.28)	.19 (.22)	.36 (.31)	.27 (.32)
Confident	.36 (.36)	.45 (.28)	.35 (.21)	.47 (.26)
Somewhat Confident	.23 (.20)	.23 (.16)	.19 (.17)	.21 (.20)
Not at all Confident	.16 (.19)	.15 (.12)	.11 (.11)	.08 (.08)

Note. The values represent the mean proportion of trials, with SD in brackets. NRMT = Non-verbal Recognition Memory Task; VRMT = Verbal Recognition Memory Task.
[a] $n = 25$. [b] $n = 16$. [c] $n = 23$. [d] $n = 18$.

APPENDIX U
POWER AND META-ANALYSES
PUBLICATION (IN PRESS)

Recognition Memory Impairment in Parkinson's Disease: Power and Meta-Analyses

Craig J. Whittington, John Podd, and Melanie M. Kan

Contrary findings notwithstanding, the prevailing notion is that recognition memory is little affected by Parkinson's disease (PD). Both a power analysis and a meta-analysis were conducted to help clarify the degree of recognition memory deficit associated with PD. The power analysis confirmed that, in general, memory studies of PD participants have been underpowered. This analysis indicated the need to pool study results in a subsequent meta-analysis, the main finding of which was that recognition memory deficits do occur with PD. The largest deficit occurs in demented PD participants. Nevertheless, deficits also occur in nondemented PD participants on medication, but nondopaminergic CNS abnormalities are more likely to underlie this deficit than PD medication itself. Future development of a theory of cognitive dysfunction in PD should take into account these recognition memory deficits which may increase with disease progression.

Most reviews of the Parkinson's disease (PD) literature suggest that people with this neurodegenerative disorder have impaired recall but relatively intact recognition memory (e.g., Beatty, 1992; Brown & Marsden, 1988; Brown & Marsden, 1990; Cummings & Benson, 1988; Dubois, Boller, Pillon, & Agid, 1991; Growdon & Corkin, 1986; Karayanidis, 1989; Knight, 1992; Knight, Godfrey, & Shelton, 1988; Mahurin, Feher, Nance, Levy, & Pirozzolo, 1993; Sagar & Sullivan, 1988; Saint-Cyr & Taylor, 1993; Taylor, Saint-Cyr, & Lang, 1988). This dissociation between recognition and recall has been used to support the view that the impairment is to do with control and retrieval strategies rather than with the encoding of new material (Brown & Marsden, 1988; Mahurin et al., 1993). Others have suggested that memory deficits are only apparent on the more effortful free recall tasks (Knight, 1992; Knight et al., 1988; Taylor et al., 1988; Weingartner, Burns, Diebel, & LeWitt, 1984), at very short retention intervals (Karayanidis, 1989; Saint-Cyr & Taylor, 1993), or when tasks require PD participants "... to scan mentally, manipulate the material, or organise actively a response" (Dubois et al., 1991, p. 204). Some reviews have noted a possible nonverbal recognition memory deficit, but maintain that verbal recognition is relatively normal (Beatty, 1992; Sagar & Sullivan, 1988).

Craig J. Whittington, School of Psychology, Massey University, New Zealand; John Podd, School of Psychology, Massey University, New Zealand; Melanie M. Kan, School of Psychology, Massey University, New Zealand.

This research was supported in part by grants from the Massey University Research Fund (1-0575-67302A), the Massey University Research Equipment Fund (1-0575-98025), and from the School of Psychology, Massey University. We thank O'Reilly Tucker for his help as an independent judge of coding reliablity.

Correspondence concerning this article should be addressed to John Podd, School of Psychology, Massey University, Private Bag 11222, Palmerston North, New Zealand. Email [J.V.Podd@massey.ac.nz].

Recognition memory may be impaired in PD depending on circumstances, however. Because cognitive deficits may vary as a function of many factors, such as age of onset, symptom duration, medication, type and severity of motor symptoms, depression, and task variables (Beatty, 1992; Levin, Tomer, & Rey, 1992; Starkstein et al., 1989), the issue is not easily resolved. Nonetheless, the widely accepted view is that PD causes problems for recall memory but not for recognition memory. The former is thought to stem from a generalized impairment rather than a specific deficit within the memory system (Brown & Marsden, 1990).

Our own review of the PD memory literature revealed a number of problems which made this recall-recognition distinction worthy of further investigation. For instance, reviews of the relevant research have been narrative-based, and often review only a small sample of studies. Qualitative reviews can sometimes reach the wrong conclusions (e.g., see Schmidt, 1996). There have been no meta-analytic reviews assessing the overall effect sizes (ESs) of the deficits in recall or recognition across PD studies. Even where 'nonsignificant' results are reported, one needs to be cautious. Studies with low statistical power (caused mainly by too few participants, small ESs, or both) will have a high probability of committing a type II statistical error; that is, concluding that there was no experimental effect when in fact there was. Interpretation of primary research is particularly problematic in situations of low statistical power. As part of our research, we carried out a statistical power analysis demonstrating that even for large ESs, on average, the probability of correctly rejecting the null hypothesis is surprisingly low.

The recall-recognition distinction is further complicated by the role of dementia in PD. Some studies have excluded participants who showed signs of dementia as assessed by DSM criteria or brief mental status tests (e.g., MMSE), whereas others have not assessed the status of their participants in this regard. The absence of a universal definition of PD dementia has not helped, producing variations in the diagnostic criteria; consequently, the resulting samples vary in their composition. Memory may be differentially affected in demented versus nondemented participants (Mohr, Mendis, & Grimes, 1995). Furthermore, according to Lieberman (1998) "antiparkinsonian drugs may superimpose a delirium on the dementia" (p. 34). Some researchers have gone so far as to say that demented and nondemented PD represents two distinct disorders each with a different clinical picture (e.g., Lieberman et al., 1979). Others have failed to find a clear distinction between nondemented and demented participants, arguing instead for a "...continuous distribution of cognitive deficit, ranging from intact intellectual function to severe dementia, with the majority exhibiting substantial cognitive impairment (Pirozzolo, Hansch, Mortimer, Webster, & Kuskowski, 1982, p. 79). Additionally, the issue of whether PD dementia should be characterised as 'cortical' or 'subcortical' is still being debated (Mahurin et al., 1993).

A frequently cited article used in support of the view that recognition memory is not impaired in PD is that of Flowers, Pearce, and Pearce (1984). Flowers et al., used an elderly sample of 54 PD participants. Little information was given on the severity of PD in this sample, except that participants ranged from 3 to 24 on the Webster scale of symptom severity and clinical disability.

All participants were stabilised on some form of antiparkinsonian medication. No mention was made about screening for possible dementia. A similar number of healthy participants acted as controls matched approximately by age, occupation, and background.

Flowers et al. (1984) assessed recognition memory with the aid of two verbal recognition memory tasks (pictures of common objects, and words and numbers) and two nonverbal recognition memory tasks (black and white histogram shapes, and coloured abstract pictures). The authors found that recognition memory was poorer for the PD participants on all tasks relative to the controls, but that the differences were neither "large" nor "reliable". Furthermore, they suggest that their results indicate that PD participants have normal registration and ability to retain information. Any memory problems must involve "retrieval or some higher level processing stage" (p. 1180). Since these conclusions appear to be based largely on the outcome of statistical significance tests, we re-analysed some of Flowers et al.'s data in terms of statistical power and actual ESs.

In the Flowers et al. (1984) study, comparisons between PD and control participants were made with two-tailed *t*-tests that lacked adequate statistical power. Given an alpha level of .05 and the number of participants used, we found these tests to have no more than 17% power to detect small effects (as defined by Cohen, 1988). Thus, there was very little chance of rejecting the null hypothesis (type II error probability = 100 - power = 83%) even if it was in fact false. Power to detect medium effects was also rather inadequate, ranging from 53 to 69%. Calculating ESs from the reported data, we found that delayed recognition memory deficits ranged from small to medium (Cohen, 1988). However, there was no evidence of impaired immediate recognition memory except for the words and numbers task. It has been suggested that several of the tasks used by Flowers et al., may have produced performance close to floor and ceiling levels, thus reducing the observed ESs further (Sahakian et al., 1988).

Since the research of Flowers et al. (1984), several studies have reported impaired recognition memory in demented PD participants relative to controls (Appollonio et al., 1994; Helkala, Laulumaa, Soininen, & Riekkinen, 1989; Litvan, Mohr, Williams, Gomez, & Chase, 1991; Tierney et al., 1994). Recognition memory deficits have also been found in nondemented PD participants (Bondi, Kasniak, Bayles, & Vance, 1993; Cooper, Sagar, & Sullivan, 1993; O'Sullivan, 1998; Owen et al., 1993; Sahakian et al., 1988). Studies using PD participants unselected for dementia have also reported recognition deficits (Allain et al., 1995; El-Awar, Becker, Hammond, Nebes, & Boller, 1987; Massman, Delis, Butters, Levin, & Salmon, 1990; Reid et al., 1989; Sagar, Cohen, Sullivan, Corkin, & Growdon, 1988; Sullivan & Sagar, 1989; Tweedy, Langer, & McDowell, 1982). But studies such as Flowers et al., and Taylor, Saint-Cyr, and Lang (1986) appear to have been instrumental in establishing the conclusion that recognition memory is normal in PD.

The present investigation aimed to clarify the degree of impairment in PD recognition memory by using statistical power analysis and meta-analytical techniques to provide a comprehensive view of the literature on the topic. Meta-analysis was designed to be a more precise and objective quantitative method for integrating research findings than the traditional narrative review

(Hunter & Schmidt, 1990). We used meta-analysis to isolate possible sources of variability among the studies involving recognition memory and to provide a quantitative estimate of the magnitude of any memory deficit. The need for this analysis was strongly suggested by preliminary informal calculations of statistical power, suggesting that the power available in many studies was too low for recognition memory deficits to be detected. Under such circumstances, even moderate ESs may not yield significant results. Therefore, we decided to run both a meta-analysis and a full-scale power analysis.

Method

Literature Search

The literature search was restricted to journal articles published between the years 1978 and 1997 inclusive. The main reason for only using studies published in this 20-year period was to avoid those that may have used participants who had disorders other than idiopathic PD (e.g., ShyDrager syndrome, and progressive supranuclear palsy). Such disorders were once classified as PD but are now recognised as separate disease entities (Hietanen & Teravainen, 1986). Unpublished studies were not used in the present power and meta-analyses. Meta-analysts usually include unpublished material to reduce publication bias (i.e., bias due to selectively publishing only significant results). However, Collins and Miller (1994) argue that publication bias is unlikely in some cases; their argument applies here. First, the literature on memory deficits in PD contains many inconsistent findings and several published studies have failed to find a memory impairment. Second, in many studies, the main focus was not recognition memory; so it is less likely that a nonsignificant recognition effect would exert a systematic bias on published research. Therefore, given what appeared to be a small risk of publication bias, it was deemed too costly to search and gather all unpublished work.

Several methods were used to search the published literature. First, computer-based information searches on PsycLit and MedLine were conducted. Key words used in the computer searches included *Parkinson's disease*, paired with the descriptors *cognition*, *cognitive deficit*, *cognitive impairment*, *memory deficit*, and *memory impairment.* Second, a manual literature search was conducted on the journals considered most likely to publish studies on cognitive deficits in PD (see Appendix for these journals). These journals were searched issue by issue for the years 1993-1995. In addition, the references of review articles were searched for further articles. From the large pool of references generated we used only those that met a number of inclusion criteria.

General Inclusion Criteria

All studies selected for inclusion had to be published in a peer-reviewed, English-language journal with a publication date between 1978 and 1997. To be included in the power analysis, a study had to compare the performance of people with PD to healthy control participants on a task explicitly described as

measuring memory. To be included in the meta-analysis, the memory task had to measure recognition memory (see recognition task inclusion criteria below). In addition, a study had to report sufficient information to permit ES estimation as described below.

Power Analysis

Exclusion criteria. Preliminary calculations showed that the average number of statistical tests in each article for which power could be calculated numbered about 10. A power analysis for small, medium, and large ESs would have amounted to some 2600 power calculations. Therefore, we developed a set of criteria to limit the number of calculations.

First, all statistical tests were excluded other than *t* and *F* tests. These statistics were by far the most commonly used tests for evaluating the most important hypotheses. Procedures for calculating power for *t* and *F* tests are well established, unlike for some other statistics (e.g., nonparametric tests). Second, tests that were peripheral to the main hypotheses were excluded (e.g., initial tests to determine participant suitability, or participant demographics). Finally, a few studies that failed to provide information such as sample size and degrees of freedom, necessary to calculate power, were also excluded.

A total of 46 from an initial set of 88 studies remained for the power analysis. Where alpha-adjusted procedures were used to control for multiple significance tests, power was calculated as if no adjustment had been made. In such cases, power calculations will overestimate the true power by a small amount (Sedlmeier & Gigerenzer, 1989).

Determining effect sizes. Statistical power for each test statistic was calculated for small, medium, and large ESs as defined by Cohen (1988). There is evidence that Cohen's estimates for these population ESs are reasonably accurate (Rossi, 1990). Previous power analyses (Cohen, 1962; Rossi, 1990; Sedlmeier & Gigerenzer, 1989; Whittington & Podd, 1996) have tended to use Cohen's population estimates rather than those based on the sample data for each investigation for several reasons. First, ES estimates based on sample data can be perturbed by error inherent in all measures of behaviour. Such error tends to result in underestimates of the true population ES. Second, several of the studies in the present power analysis had small sample sizes. In general, the smaller the sample the less reliable the ES estimate. Third, published research often fails to report enough information to calculate a sample ES. A fourth and final factor that persuaded us to follow the conventional approach to estimating population ESs was that data from the present study lends support to the estimates of population *d* values given by Cohen (1988). The ordered *d* values from our entire data set were divided into three equally sized groups yielding average *d* values of .18 (small), .52 (medium), and 1.09 (large). These values are in good agreement with Cohen's (1988) estimated population *d* values of .20, .50, and .80, respectively.

Power calculations. Statistical power was calculated using the GPOWER computer program developed by Erdfelder, Faul, and Buchner (1996). Power can be calculated for a range of univariate test statistics, including *t* and *F*. The program calculates power using Cohen's (1988) small, medium, and large ESs

as the default. However, any ES value can be entered. For each power calculation, the appropriate values for alpha, ES, and *N* were entered and power was then calculated by the program for 1380 test statistics. Where a study used a particular statistic more than once (almost always the case), the mean power level was calculated separately for small, medium, and large ESs for that study.

Meta-analysis

A total of 39 journal papers were identified; 17 were not used in the meta-analysis because they did not use a nonneurologically impaired control group (Helkala et al., 1989; Pillon, Deweer, Agid, & Dubois, 1993), or did not report sufficient information (Allain et al., 1995; Direnfeld et al., 1984; Jacobs et al., 1995; Reid et al., 1989; Tweedy et al., 1982), or used a possibly biased measure of recognition (Bondi & Kaszniak, 1991; Buytenhuijs et al., 1994; Daum et al., 1995; Hartikainen, Helkala, Soininen, & Riekkinen, 1993; Pillon et al., 1996), or used a recognition test that did not meet the criteria for inclusion in the present analysis (Dubois et al., 1987; Huberman, Moscovitch, & Freedman, 1994; Levin, Llabre, & Weiner, 1989; Sagar et al., 1988; Tsai, Lu, Hua, Lo, & Lo, 1994).

Variables coded from each study. Coding was carried out by the principal author and a research assistant. The first coder evaluated the entire data set, and reliability of coding was assessed on a randomly selected subset of 33% of the sample. Discrepancies in coding were measured with Kappa's coefficient and Pearson's correlation for categorical data and non-categorical data, respectively (Bryman & Cramer, 1994). Overall, the coding was very reliable (Mean r = .89, SD = .19). Where low reliabilities were found, after discussion between coders, the particular variable was re-coded by the principal author using all reports. Studies were not weighted on the basis of their research quality.

General information coded from each study included: year of publication and geographical location of sample (North America, United Kingdom, Europe, Australia, New Zealand, or other).

Participant characteristics for each study's PD and control samples were coded as follows: gender (expressed as the percentage of females); handedness (expressed as the percentage of right-handers); and, average age of participants.

The following characteristics were coded for the PD sample only: Intelligence measure and mean score; premorbid intelligence measure and mean score; depression measure and mean score; affective state of sample (depressed, nondepressed, or unselected); medication status (medicated, withdrawn from medication, never medicated, or unselected); laterality of disease symptoms (unilateral, bilateral, or unselected); age of onset of symptoms (early, less than 45 years old; late, greater than 60 years old; middle, between 45 and 60 years old; or, unselected); physical symptoms measure and mean score; number of participants at each Hoehn and Yahr stage (Hoehn & Yahr, 1967); stage of disease (early, late, or unselected); disease duration (years); dementia measure and mean score; and cognitive status of sample (demented,

nondemented, or unselected). In addition, we noted whether the PD group and control group were matched (yes/no), and if so, the variables used to match.

The following task characteristics were also coded: Name of task; recognition performance measure (e.g., percent correct, *d'*, discriminability); task type (yes-no, *n*-alternative); modality (verbal, nonverbal[1]); delay (delay, no delay); recall (recall before recognition, recognition only); and administration method (computer, manual).

Finally, all statistical data needed to conduct the meta-analysis were recorded (e.g., sample sizes, mean scores and standard deviations, inferential statistics, degrees of freedom, significance levels).

Recognition task inclusion criteria. A wide variety of tasks are used to test recognition memory (see Murdock, 1974). In order to ensure a homogeneous set of tasks, only those testing item information using yes-no or forced-choice tests with accuracy as the dependent variable were included. Thus, tasks using batch testing, or latency and confidence ratings were excluded.

Meta-analytic procedures. The basic approach was modelled on the meta-analytic techniques of Hunter & Schmidt (1990). Most computations were made with the help of a spreadsheet or Schwarzer's (1989) Meta-Analysis Programs (Version 5.3). Schwarzer's software was not used for all computations because modifications made by Hunter and Schmidt (1990) to some of their formulae have not yet been incorporated into the program (R. Schwarzer, personal communication, 1998).

Calculation of effect sizes. The standardised mean difference, *d* (Cohen, 1988) was used as the estimate of ES. The *d* statistic can be defined as the difference between the group means divided by the within-group standard deviation: $d = (Y_e - Y_c) / S_W$, where Y_e is the mean of the experimental group, Y_c is the mean of the control group, and S_W is the within-group standard deviation. The latter is defined as the square root of the within-group variance (V_W), where $V_W = [(n_e - 1)\ V_e + (n_c - 1)\ V_c] / [(n_e - 1) + (n_c - 1)]$. V_e is the variance for the experimental group, and V_c is the variance for the control group, and n_e and n_c are the sample sizes for the experimental group and control group, respectively.

Most ESs were calculated directly from means and standard deviations. Where this was not possible, *d* was transformed from *t*-values or *F*-values, where, $d = t\,[(n_e + n_c) / (n_e n_c)]^{1/2}$, and $t = \sqrt{F}$. For *F* values ≤ 1, a *d* value of zero was assumed. Where only the significance level was reported, this was converted to *r* and then transformed to *d*, where $d = (1/(pq)^{1/2})\ r / (1-r^2)^{1/2}$ and *p* and *q* are the proportions of persons in the two groups. With regard to calculating *r* from significance levels, when significance or nonsignificance was reported without further information, the normal (and conservative) approach was taken. That is, when the results were reported as significant, $p = .05$ was assumed and when nonsignificance was reported, $p = .50$ was assumed.

Multiple effect sizes from single studies. There is no generally agreed method for dealing with multiple ESs from a single study. Including all separate ESs regardless of their interdependence will actually lead to an underestimate

[1] Stimulus material that can not be readily labelled.

of the degree of homogeneity across studies, but will not produce any systematic effect on the mean ES (Hunter & Schmidt, 1990). Where ESs come from different study characteristics that act as real and substantial moderators, these ESs can be entered into the meta-analysis as independent outcome values (Hunter & Schmidt, 1990). However, others argue that ESs should be independent with each study represented only once in an analysis (e.g., Bangert-Drowns, 1986). In the present analysis there was little basis for determining, *a priori*, whether a variable was a real moderator or not. Therefore, initially a weighted average ES was calculated for each study and entered into the overall meta-analysis. (Note that a weighted average will be an underestimate of the value that would have been obtained from an overall composite variable if one could be formed; Rosenthal & Rubin, 1986.) There was one exception to this rule: Where a study reported separate results for two or more independent PD groups (e.g., demented and nondemented participants), a single ES was computed for each group individually. Separate control groups were not required as long as sufficient information was available to calculate ESs for each PD group.

Two papers used some of the same participants to obtain recognition data from different stimulus modalities (Sagar et al., 1988, and Sullivan & Sagar, 1989). For present purposes the two reports were treated as one. Overall, the procedure used in our analysis produced more ES estimates than the number of papers from which the data were taken. In total, the 22 published reports yielded a total of 32 independent PD groups. Therefore, in the overall meta-analysis there were 32 independent ESs.

Summary analysis of effect sizes. For the overall analysis, the weighted mean and variance were computed based on formulas reported by Hunter & Schmidt (1990): mean $d = \sum w_i d_i / \sum w_i = D$ and observed variance of $d = \sum w_i [d_i - D]^2 / \sum w_i$, where D is the weighted average of d, and w_i is the sample size of each group. A weighted average ES was used since large N studies have less sampling error than small N studies and therefore deserve more weight.

The population (residual) variance, s^2_{res}, was then computed by subtracting the sampling error variance, s^2_e, from the observed variance, s^2_r, where the sampling error variance was computed using the formula $s^2_e = [(N - 1) / (N - 3)] [(4 / N) (1 + D^2 / 8)]$, N being the average sample size across all groups.

The estimate of the population variance served as the multiplier in the formula for the 95% confidence interval (CI): $d - 1.96 (s_{res}) < ES < d + 1.96 (s_{res})$. Hedges and Olkin (1985, p. 80) show that d has a small sample bias. By defining a new estimator d^*, they remove this bias: $d^* = d / a$, where $a = 1 + .75 / (N - 3)$. Hedges and Olkin use the symbol "d" for their approximately unbiased estimator. To avoid confusion, we will use d^* to denote this estimator. The correction can be applied either study by study or, as in the present analysis, after the meta-analysis has been done using the average sample size (Hunter & Schmidt, 1990).

To reliably interpret the estimated population ES, the underlying data set should be homogenous. Hunter and Schmidt (1990) suggest that a data set can be considered homogeneous when most of the observed variance is explained by artifacts (e.g., sampling error and error of measurement in the dependent

and independent variables). Artifacts such as measurement error will produce both a systematic reduction in the mean ES and a systematic increase in the variance of ESs (Hunter & Schmidt, 1990). It is possible to correct a meta-analysis for the effect of artifacts, provided sufficient information exists about the artifact. However, substantial evidence suggests that sampling error is responsible for as much as 85% of this artifactual variance (Stoffelmayr, Dillavou, & Hunter, 1983). Therefore, we decided to correct only for sampling error. Furthermore, it would be difficult to correct for other artifacts since little information is available in PD research with which to make the correction.

Once the effect of sampling error is removed from the data, the question remains: To what extent can any residual variance be explained as due to artifacts not corrected for? Hunter and Schmidt (1990) suggest that if 75% of the observed variance is accounted for by artifacts, then the data set can be considered homogeneous. We used Hunter and Schmidt's rule of thumb, but the actual percentage of variance explained by sampling error is also reported to allow readers to judge for themselves.

Some authors (e.g., Rosenthal & Rubin, 1982) advocate the use of a statistical test based on the chi-square distribution to help determine whether there is heterogeneity in a data set. However, this method suffers from all the usual problems associated with statistical significance testing. Most problematic is the power of the test. In data sets containing a large number of studies, the test will have high power and almost any residual variance will produce a significant result. Conversely, in situations of low power, even large amounts of residual variance may not be detected leading to the conclusion that the data set is homogeneous when in fact it is not.

After the effect of artifacts has been removed from a data set, heterogeneity may remain as a result of moderating variables. To examine the influence of possible moderators in the present analysis, the data set was subdivided as a function of the potential moderator and each subset was subjected to further analysis. At this level, a single study may contribute ESs to each of the subsets provided that a separate ES can be calculated as a function of the moderator. Two requirements had to be met in order for a variable to be classified as a moderator. First, the population ES had to vary between subsets. Second, the residual variance in each subset had to be smaller than that seen in the data set as a whole (Hunter & Schmidt, 1990). It is legitimate to search for moderators even after the conclusion has been made that all of the ES variability is due to sampling error (Rosenthal, 1995). We took this approach when we suspected further moderators.

Hunter and Schmidt's (1990) procedure using the least squares method for estimating the mean and variance of a distribution of ESs is based on the assumption that the data do not contain outliers. They suggest that even a single outlier can greatly distort the observed standard deviation, and to a lesser extent the mean. Therefore, outliers were removed from each analysis by excluding values greater than three standard deviations from the mean. Subsequently, stem and leaf graphs were used to detect other possible outliers.

Results

Power analysis

Results of the power analysis. Table 1 summarises the results of the statistical power analysis, based on 48 studies and 1360 power calculations. The median power values are in good agreement with those obtained by other researchers (Cohen, 1962; Sedlmeier & Gigerenzer, 1989) in the area of abnormal psychology. For small ESs, none of the 48 studies we examined had sufficient average power for *t* and *F* tests. Cohen suggests a minimum power value of 80%, but even for a medium ES only 15 of the studies reached this level. Assuming (unrealistically) a large ES for recognition memory deficits in PD, 12 (25%) of the studies had an average power value of less than 80%.

Table 1 *Summary Statistics for the Statistical Power Analysis*

	Effect size[a]		
	Small	Medium	Large
Mean	20 [b]	63	85
Median	16	67	94
Standard deviation	12	24	19
Minimum	6	15	32
Maximum	55	100	100
95% confidence interval	17–23	56–70	80–90

[a]Cohen's (1988) conventions for effect size (e.g., small: d = .20, medium: d = .50, large: d = .80). [b]Power values are expressed as a percentage.

In summary, most studies of memory deficits in PD have adequate power to detect only large ESs; the following results of our meta-analysis strongly suggest that large ESs occur only in certain subsets of participants in this field of research.

Meta-analysis

General characteristics. Data from a total of 32 independent PD groups taken from 22 studies were used in the analysis.[2] At the study level, 77% of the studies were published between 1988 and 1997, with the remainder published between 1978 and 1987. At the sample level, 47% (15) of the PD samples were taken from North America and 44% (14) from the UK. The remaining were sourced from Europe (9%). The vast majority of PD samples (94%) were matched with a healthy control group, most commonly on age and gender. The remaining samples (6%) were compared to an unmatched control group.

PD group characteristics. At the sample level, 84% of the PD groups contained less than 25 participants, while only 3% exceeded 50. The mean sample size was 18.53 (*SD* = 14.83). From the reports that provided sufficient information, most groups (70%) contained 40% or less females. In most groups

[2] A description of all PD groups, their associated control groups, the recognition tasks used, and ES data are available from the authors upon request. Please send a 3.5 inch high density diskette.

(81%), the average age of participants was greater than 60 years. Across all groups, the mean age was 63.41 (*SD* = 4.68). A majority of the PD samples (75%) contained predominantly nondemented participants, while 16% were demented, and 9% were either unselected or insufficient information was given to determine cognitive status.

In all cases where cognitive status was reported, it was confirmed with the use of a quantitative measure (e.g., MMSE). With regard to depression, 28% of the studies contained predominantly nondepressed participants while no study contained predominantly depressed participants. However, 72% were either unselected or this information was not reported. Sixty-nine percent of groups contained participants on antiparkinsonian medication, while 22% contained participants who had never received medication, the remainder being either unselected or this information was not reported. Just under half of groups (48%) had a mean disease duration of between 5-10 years, 24% had less than five years, and 15% had a mean duration over 10 years. However, this information was not reported for 12% of the groups used in the present analysis. Furthermore, for many groups (34%), information about disease stage was not reported. However, in those reports that did provide this information, 33% contained participants in the early stages of Parkinson's, while 24% were made up of participants in the later stages of PD. Nonreporting of participant characteristics, or using unselected groups were factors in many studies. Such characteristics included laterality of symptoms, age of onset of symptoms, IQ, and estimated premorbid IQ.

Control group characteristics. The mean sample size of the 23 independent control groups was 24.61 (*SD* = 13.91). The mean age of controls was 61.91 years (*SD* = 7.37). Most (67%) control groups consisted of 40 to 60% males.

Results of the meta-analysis. Table 2 presents the results of the meta-analysis that addressed the issue of whether recognition memory is impaired in PD. The overall analysis indicated that substantial heterogeneity existed in the data set. Therefore, the data were subdivided according to the cognitive status of the participants; that is, demented, nondemented, and unselected samples.

Table 2 *Cumulated Effect Size (ES) Estimates and Residual Variation as a Function of Cognitive Status After Accounting for Sampling Error*

Moderator	*k*	*N*	*d**	S^2_e	S^2_{res}	95% CI	% Var
Overall	32	1401	0.32	0.10	0.12	-0.35–1.01	44
Demented[a]	4	130	1.30	0.16	0.00	1.30–1.30	100
Nondemented[b]	23	1039	0.16	0.09	0.00	0.16–0.16	100
Unselected	3	157	0.52	0.08	0.05	0.11–0.95	64

Note. *k* = number of ESs; *N* = total sample size; *d** = Hedges and Olkin's (1985) unbiased ES statistic; S^2_e = variance due to sampling error; S^2_{res} = residual variance; 95% CI = 95% confidence interval for *d**; % Var = percentage of variance attributable to sampling error.
[a]One study (Heindel, Salmon, Shults, Walicke, & Butters, 1989) produced an ES that was considered an outlier and so was excluded from the analysis. [b]One study (Sahakian et al., 1988) produced an ES that was considered an outlier and so was excluded from the analysis.

(We expected that cognitive status would be a moderator since memory impairment is a principle component of dementia. Studies that used demented participants were included in the meta-analysis so as to quantify the recognition deficit for all participants.) The analysis was then repeated for each subset.

In the demented subanalysis, the 95% confidence intervals for each of the ESs overlap (see Figure 1). In addition, the ES for every study was greater than .80. ESs of this magnitude can be described as large effects (Cohen, 1988). The meta-analysis indicated homogeneity in this subset since sampling error accounted for 100% of the observed variance (see Table 2). The magnitude (and 95% confidence interval) of the estimated population ES (d^* =

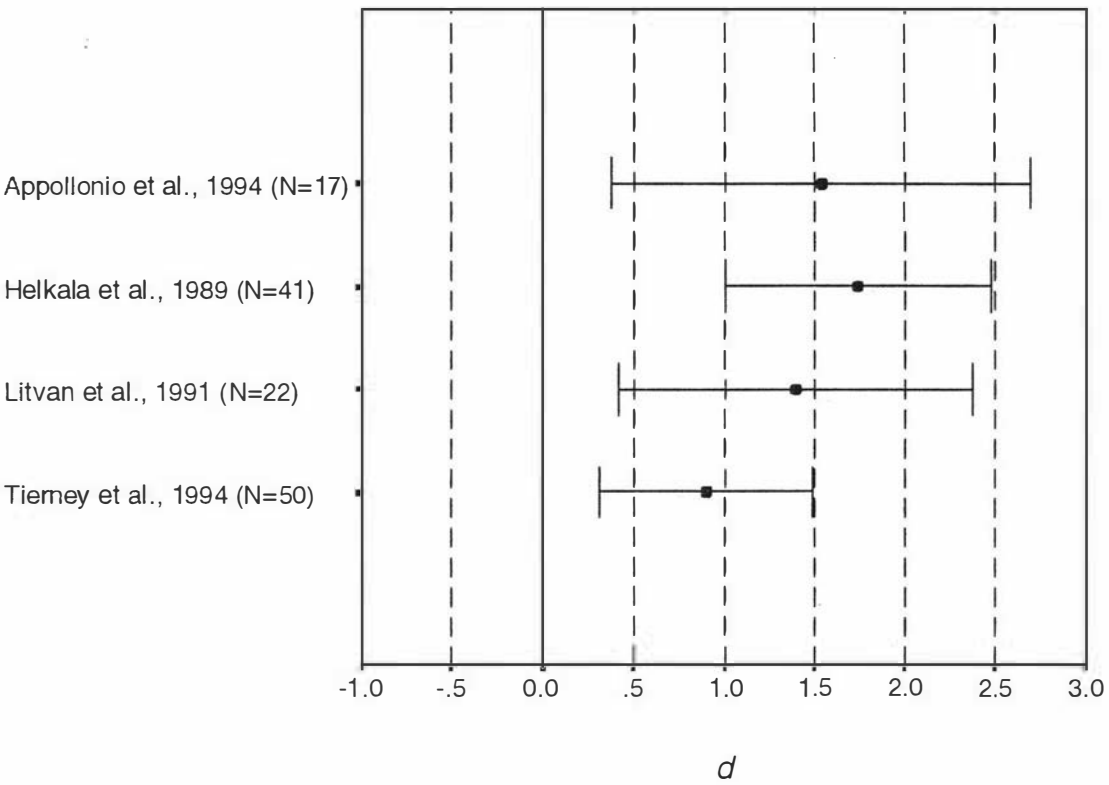


Figure 1. Meta-analysis plot of the effect sizes for each study entered into the demented subgroup analysis. The mean *d* for each study is shown by the points, and the horizontal lines show the 95% confidence interval.

1.30) provides evidence that, at least in this sample, demented PD participants suffered from a large recognition memory deficit.

In the nondemented subanalysis, the confidence intervals also overlap (see Figure 2) and the analysis indicates homogeneity (see Table 2). The estimated population ES was .16. An ES of this magnitude suggests a small recognition deficit in nondemented PD participants.[3] As could be expected from

[3] This is likely to be an underestimate of the population ES as several papers (Breen, 1993; Cooper et al., 1993; Heindel et al., 1989; Taylor, Saint-Cyr, & Lange, 1987) reported only that their results were nonsignificant. Therefore, an ES of zero was entered into the meta-analysis, whereas the actual ES may have been greater than zero.

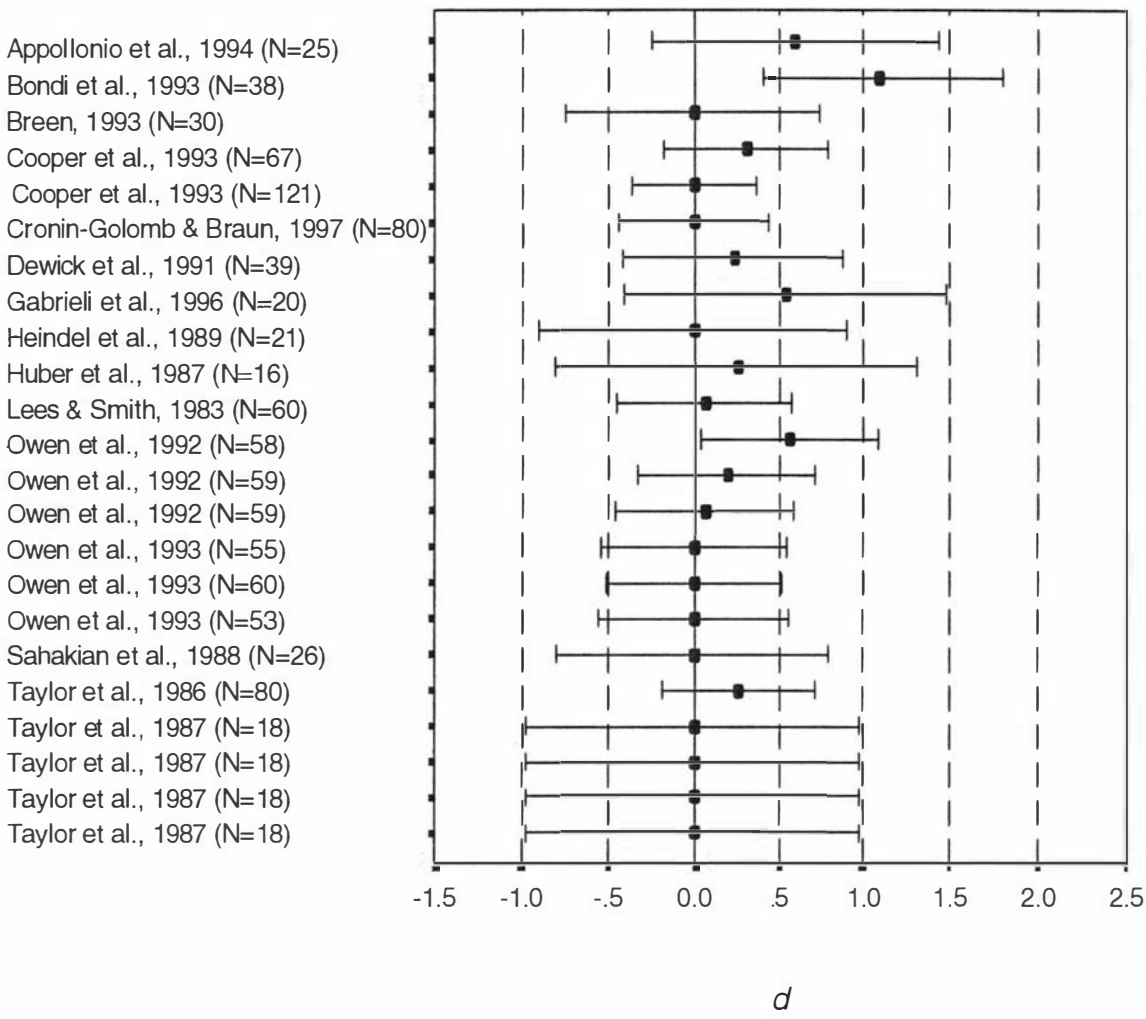


Figure 2. Meta-analysis plot of the effect sizes for each study entered into the nondemented subgroup analysis. The mean *d* for each study is shown by the points, and the horizontal lines show the 95% confidence interval. Four studies have multiple entries because they used independent groups of PD participants.

the separate analyses of demented and nondemented cases, the third subset of unselected participants was not homogenous, although the ES was of moderate magnitude. This analysis showed that the estimated population ES varied between subsets and the residual variance in each subset was smaller than that seen in the overall analysis. Therefore, cognitive status is a probable moderator.

Further analysis of the nondemented subset indicated that just 35% of the ESs were greater than .20, 17% of them were greater than .50, and only one exceeded .80. Moreover, 30% of the PD groups consisted of patients who were newly diagnosed and had not yet received medication, whereas the remainder were receiving antiparkinsonian medication. These *de novo* participants had a much smaller mean disease duration (M = 1.96 years, SD = 0.54) than the medicated participants (M = 8.55 years, SD = 2.15), d = 3.58, 95% CI = 2.20–4.97). In order to assess whether *de novo* and medicated PD participants differed with regard to recognition memory, the meta-analysis was

repeated on these groups separately. As can be seen in Table 3, the ES varied between subsets and the variance in each subset was explained by sampling error suggesting that medication status acts as a moderator of recognition memory deficits in PD. Medicated PD participants had poorer performance relative to controls, whereas *de novo* participants did not. This was indicated by an ES of .23 for the medicated groups, a population value that the associated 95% CI suggests can be viewed with high confidence. There was little evidence of a difference in recognition memory performance between *de novo* PD participants and controls (d^* = .05).

Task characteristics within cognitive status subsets. Within the demented subset, all of the studies used a verbal recognition task. In half of the studies, the task involved recall before the recognition phase. In three out of the four studies, the task had a forced delay between the presentation phase and the recognition phase. Also, in three of the studies, a yes-no task was used.

Within the nondemented subset, the majority (70%) of the studies used a verbal task and a majority (61%) also used tasks involving only recognition. In 39% of the studies the task had a forced delay, while in a further 43% there

Table 3 *Cumulated Effect Size (ES) Estimates and Residual Variation as a Function of Medication Status After Accounting for Sampling Error*

Moderator	*k*	*N*	*d**	S^2_e	S^2_{res}	95% CI	% Var
Nondemented							
Medicated[a]	16	675	0.23	0.10	0.00	0.23–0.23	100
De novo	7	364	0.05	0.08	0.00	0.05–0.05	100

Note. *k* = number of ESs; *N* = total sample size; d^* = Hedges and Olkin's (1985) unbiased ES statistic; S^2_e = variance due to sampling error; S^2_{res} = residual variance; 95% CI = 95% confidence interval for d^*; % Var = percentage of variance attributable to sampling error.
[a]One study (Sahakian et al., 1988) produced an ES that was considered an outlier and so was excluded from the analysis.

was no delay.

The remaining studies used both delay and non-delay tasks, but the results were averaged across this factor for the purpose of the meta-analysis. Finally, in the majority (65%) of the studies a *n*-alternative, forced-choice task was used, while in 31% a yes-no task was used, and in one study the results were averaged across this factor.

Discussion

Power Analysis Results

The statistical power analysisof studies on memory deficits in PD confirmed that for at least small and medium ESs, power was inadequate. Maximum power for a small ES was just 55%, and only 15 of 48 studies had power ≥ 80% for a medium ES. The median power values in the present study (16, 67, and 94%) are consistent with power analyses conducted in the area of abnormal

psychology (Cohen, 1962: 17, 46, and 89%; Sedlmeier & Gigerenzer, 1989: 14, 44, and 90%).

The general picture emerging from a number of power analyses in a wide range of subdisciplines in psychology is that many studies simply do not have the power to detect the effects they may predict, especially when ESs are small. Under such circumstances, it is impossible to decide whether an acceptance of the null hypothesis is due to there truly being an ES of zero, or due to a too high type II error probability (due to low power).

If it is generally true that studies producing small ESs are underpowered, why not remedy the situation by increasing power? The reasons, we suspect, are largely of a practical nature. Power is completely determined by the ES, *N*, and the alpha (*p*) level. Increases in any one or more of these will increase power (Cohen, 1988). Researchers (and journal editors) seem reluctant to adjust the alpha level upwards, despite the fact that most well-known statistical handbooks (e.g., Keppel, 1991) advocate setting the alpha level only after careful consideration of the expected ES and the costs and values associated with drawing the wrong conclusion from a set of results. In regard to the ES, researchers probably try to the best of their ability to eliminate experimental noise that may have the effect of causing the sample ES to underestimate the true population ES. Thus, there is probably little room to manoeuvre here, in terms of increasing power. The remaining variable is *N*, the number of participants. However, for small to moderate ESs, sample sizes for most common statistical tests have to number in the hundreds to produce around 80% power. The cost of such a large N is usually prohibitive.

These difficulties in obtaining sufficient power, especially for small to medium ESs, make it highly likely that studies of recognition deficits in PD will be underpowered. That is, because the probability of a type II statistical error is high, it is unlikely that small memory deficits will produce significant results. As our power analysis shows, power levels have generally been inadequate to detect small to medium ESs. Thus, we have little faith in a conclusion that suggests that recognition memory deficits do not occur in PD.

In summary, the statistical power in many investigations of memory deficits in PD has been too low. There are strong reasons for believing that researchers would find it difficult in practice to increase power to a satisfactory level, even if they wanted to. We share the view of Schmidt (1996) that trying to improve power in individual studies may be too difficult. A better approach may be to increase statistical power by combining the results from several studies in a meta-analysis.

Meta-Analysis Results

The results of the meta-analysis clearly demonstrate that in the sample of studies reviewed, as expected, demented PD participants had impaired recognition memory relative to controls. The initial analysis also revealed evidence of a much smaller recognition deficit in nondemented PD participants. Further analysis of the nondemented cases indicated that medication was a moderating variable; that is, there was a small recognition memory deficit in medicated PD participants, but little evidence of impaired recognition in *de novo* participants.

Previous reviews of memory functioning in PD have generally made little distinction between demented and nondemented participants with regard to recognition memory. The present analysis makes it clear that global cognitive status needs to be taken into consideration, especially at the primary research level. Nevertheless, the concept of PD dementia has been under debate for some time and so will not be covered here. Therefore, the remainder of the present discussion will focus on the analysis involving nondemented participants.

The meta-analysis provides evidence for recognition impairment in only medicated participants. It is therefore possible that antiparkinsonian medication contributes to the observed impairment in these participants. But this seems unlikely, because a number of studies have shown that dopaminergic therapy has no effect on memory performance, or may actually improve it (Cooper et al., 1993; Crowdon, Corkin, & Rosen, 1990; El-Awar et al., 1987; Lange et al., 1992; Mohr et al., 1989). In addition, Huber, Shulman, Paulson, and Shuttleworth (1987) found that delayed recognition memory was not affected by the absolute dopamine level, but was state dependent. In other words, changes in dopamine level between learning and recognition phases produced recognition deficits in PD participants that were not evident when dopamine was held constant. The issue is more complicated for anticholinergic medication since it has been found to have both no affect (Appollonio et al., 1994; Bondi & Kaszniak, 1991; Levin, Llabre, Reisman, Weiner, & Brown, 1991; Sahakian et al., 1988) and a detrimental affect (Dubois et al., 1987; van Spaendonck, Berger, Horstink, Buytenhuijs, & Cools, 1993) on recognition memory. A recently completed study involving 41 nondemented PD participants (Whittington & Podd, 1999[4]) failed to find a relationship between levodopa dose and recognition memory (verbal or nonverbal). This finding was replicated with the same participants after an interval of six months. Furthermore, there was little or no difference in recognition memory between participants receiving anticholinergic medication and those receiving only levodopa preparations. Therefore, there is no compelling evidence that PD medication *per se* brings about recognition memory deficits.

Other differences between medicated and *de novo* patients may provide an explanation for the difference in recognition performance. For instance, the *de novo* patients generally had shorter disease duration and were at an earlier stage than the medicated participants. Although the relationship between motor disability and cognitive decline is not straightforward (Starkstein & Robinson, 1991), there is some evidence that late stage participants show greater recognition memory deficits than those participants in the early stages (Lees & Smith, 1983; Owen et al., 1992; Sahakian et al., 1988; Whittington & Podd, 1999). Owen et al. found only a small pattern recognition memory deficit both in *de novo* participants (d = .06, 95% CI = -.46–.58) and in medicated early PD participants (d = .19, 95% CI = -.33–.71), but a larger impairment in medicated advanced PD participants (d = .56, 95% CI = .02–1.09). Lees and Smith also found only a small recognition memory deficit in *de novo* PD participants with mild symptoms (d = .06, 95% CI = -.46–.57). (It should be noted that the

[4] The raw data are available from the authors upon request. Please send a 3.5 inch high density diskette.

calculations of the *d* and CI values are ours for the above two papers and do not appear in the original publications.) Similarly, Whittington and Podd found that, relative to healthy controls, early PD participants had smaller nonverbal recognition memory deficits (*d* = .47, 95% CI = -.03–.96) than advanced PD participants (*d* = .97, 95% CI = .41–1.53). These results were replicated after an interval of six months and also with a verbal recognition memory task.

Depression is another factor that has been related to the severity of cognitive impairment in PD (Starkstein et al., 1989). It has been suggested that depression and cognitive impairment in PD may interact, accelerating the progression of overall impairment (Starkstein, Bolduc, Mayberg, Presziosi, & Robinson, 1990). Alternatively, Starkstein et al. (1990) suggest that "there may be two forms of PD: one with depression and rapid cognitive decline and one without depression and a gradual cognitive decline" (p. 597). Either of these scenarios is potentially problematic for the present analysis because nearly three-quarters of the studies sampled do not appear to have excluded participants with depression or controlled for differences in level of depression. Complicating the issue further is the evidence that depression may have little impact on recognition memory in PD (Owen et al., 1993) or in healthy elderly people (Boone et al., 1995). Whittington and Podd (1999) examined the contribution of depression to recognition memory impairment by dividing their PD group into nondepressed and depressed subgroups according to a cut-off score of 10 on the Geriatric Depression Scale (Yesavage et al., 1983). The results showed that depressed individuals had a greater nonverbal recognition memory deficit (*d* = .82, 95% CI = .26–1.37) than nondepressed PD participants (*d* = .54, 95% CI = .02–1.06). However, when tested six months later, although overall ESs had decreased, there was little evidence of a distinction between depressed and nondepressed PD participants (*d* = .28, 95% CI = -.24–.79, and *d* = .31, 95% CI = -.23–.85, respectively). A similar pattern of results was found for verbal recognition memory.

Taken together, these data suggest that medication and affective symptomology are unlikely to account for the clear distinction between *de novo* and medicated PD participants with regard to recognition memory deficits. The dissociation in memory dysfunction is more likely related to motor disability; however, it should not be inferred that both memory and motor deficits share a common pathology. While most of the cardinal signs of PD result from dysfunction of dopaminergic mechanisms, there is little correlation between these signs and cognitive performance (Pillon et al., 1989). However, cognitive dysfunction is more likely to be related to motor disability that is unresponsive to levodopa therapy (e.g., gait disorder), in the sense that these two impairments may share a common nondopaminergic neurochemical pathology (Pillon et al., 1989; Sagar, Atchison, Doherty, Ball, & Cooper, 1995). Moreover, impairment in recognition memory "...may be mediated by nondopaminergic mechanisms such as alterations of cholinergic, noradrenergic, or serotoninergic projections to the neocortex and hippocampus" (Lange, Paul, Robbins, & Marsden, 1993, p. 477). Support for this line of reasoning comes from research indicating that there may be two subgroups of PD (Zetusky, Jankovic, & Pirozzolo, 1985). Zetusky et al. suggested that the first presents with predominant tremor, earlier age at onset, and relatively normal functional and cognitive status. The other presents with predominant postural instability and

gait disorder, later age at onset, greater functional and cognitive impairment, dysarthria, dysphagia, and more rapid progression.

In summary, nondemented, medicated PD participants appear to be impaired on tests of recognition memory, but participants early in the course of the disease who have never received medication appear to have normal recognition memory. It is unlikely that medication alone can account for the difference. Rather, it seems that nondopaminergic CNS abnormalities underlie both the recognition memory deficit and motor disability not alleviated by levodopa therapy. Such a view is consistent with the finding that recognition deficits are greater in PD patients with more severe motor disability. Finally, PD is a complex disease and the consensus of opinion suggests that memory disturbances are the result of degeneration in multiple systems (Brown & Marsden, 1990; Hammond-Tooke & Pollock, 1992).

Theoretical Implications

Many commentators have discussed the dissociation between recognition and recall and its theoretical implications for explaining the mechanisms leading to memory deficits in PD. It has been suggested that this dissociation demonstrates that PD does not affect the encoding of new material, but disrupts control and retrieval strategies (Brown & Marsden, 1988; Mahurin et al., 1993). Others have suggested that the difference may simply be an artifact of task difficulty (Appollonio et al., 1994; Bondi & Kaszniak, 1991; Breen, 1993; Tsai et al., 1994). Thus, if the recognition task was made more difficult or 'effort-demanding' PD participants should begin to show a deficit relative to controls (Breen, 1993). To our knowledge, only one study (Whittington & Podd, 1999) has tested this theory directly by varying the difficulty level within a single recognition task. However, Owen et al. (1993) reported the results of a supplementary analysis which lend support to this position. When we re-analysed some of their data in terms of actual ESs, we found that set 2 of their pattern recognition memory task was more difficult than set 1 of the same task. Interestingly, medicated participants with severe PD showed signs of impaired memory relative to controls on set 2 (d = .37, 95% CI = -.17–.91), but not on set 1 ($F < 1$). Owen et al. reported nonsignificant results both for set 1 and set 2, probably because they had insufficient statistical power to detect all but the largest effects (9, 34, and 70% power for small, medium, and large ESs, respectively). Whittington and Podd used a nonverbal recognition memory task specifically to examine the issue of task difficulty. A preliminary analysis of their results indicates that PD participants demonstrated greater impairment at the hard level of difficulty than at the easy level, relative to controls. Therefore, making the recognition task more demanding appears to increase the deficit.

Our present results do not support the view that recognition memory deficits are absent in PD. Thus, the simple theory that Parkinsonians have little or no difficulty encoding new information (intact recognition system) while having problems with information retrieval (faulty recall system; Ruberg & Agid, 1988), needs modification. In fact, a general theory of cognitive dysfunction in PD must take into account not only that a recognition deficit does occur in this

neurodegenerative disorder, but that this deficit may grow with disease severity.

Research Implications

While meta-analysis is not without criticism (see Hunter & Schmidt, 1990), it has many advantages over narrative reviews. However, any meta-analysis is limited by the available database. Unfortunately, many potential moderators of the observed recognition impairment could not be tested in the present analysis. This was due in part to a lack of research findings, but also because the aims of some studies did not necessitate providing information we required for the meta-analysis. Potential moderators include task variables such as modality and delay, and participant variables such as age of onset of symptoms, symptom duration, type of motor symptoms, and depression (e.g., Beatty, 1992; Karayanidis, 1989; Levin et al., 1992; Sagar & Sullivan, 1988; Saint-Cyr & Taylor, 1993; Starkstein et al., 1989). If any sense is to be made of these moderators, investigators need to consider publishing sufficient information to allow their findings to be used in future meta-analyses. We also strongly encourage investigators to consider the implications of low statistical power, demonstrated by our analysis.

Some PD research has indicated that recognition may be impaired at very short retention intervals but not after a delay period (Cooper et al., 1993; Sagar et al., 1988; Sullivan & Sagar, 1989). Support has come from other studies that have found normal delayed recognition in nondemented medicated PD participants (e.g., Appollonio et al., 1994; Dewick, Hanley, Davies, Playfer, & Turnbull, 1991; Gabrieli, Singh, Stebbins, & Goetz, 1996; Huber et al., 1987; Taylor et al., 1986). However, when we calculated ESs for these studies we found small to medium effects indicating possible delayed recognition memory deficits. It is not surprising that these studies reported nonsignificant findings since in no study was there more than 14% power to detect small effects.

Conclusions

The power analysis revealed that, in general, studies of memory deficits in PD (as in many other research areas) have not had sufficient power to detect small to medium ESs. Yet there is no doubt, for both practical and theoretical reasons, that it is important to demonstrate the presence of these smaller ESs. The least controversial way of increasing power is to use larger participant numbers, but for small ESs, very large (and often unattainable) numbers are required. A very acceptable alternative is to pool the results of several small studies in a meta-analysis. When we did this, we found compelling evidence that, as might be expected, demented PD participants suffer from a large recognition memory deficit relative to controls. Nondemented Parkinsonians have a much smaller deficit while *de novos* appear to have normal recognition. Researchers in this area now need to consider the mechanisms responsible for both impaired recall and impaired recognition memory in PD.

Finally, we strongly urge all researchers to ensure that their published data contains sufficient information to allow them to be included in meta-analytic reviews. It seems unlikely (and perhaps even undesirable) that most

individual studies will ever be able to achieve satisfactory statistical power levels for analysing small- to medium-sized effects. Hence, meta-analyses will be required to help researchers reach general conclusions. These meta-analyses can only be as good as the data they are based on, most of which is obtained from individual research papers.

References

References marked with a single asterisk indicate studies included only in the meta-analysis. References marked with a double asterisk indicate studies included only in the power analysis. A triple asterisk indicates studies that were included in both the meta-analysis and the power analysis.

**Agniel, A., Celsis, P., Viallard, G., Montastruc, J. L., Rascol, O., Demonet, J. F., Marc-Vergnes, J. P., & Rascol, A. (1991). Cognition and cerebral blood flow in lateralised Parkinsonism: Lack of functional lateral asymmetries. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 783-786.

Allain, H., Lieury, A., Quemener, V., Thomas, V., Reymann, J., & Gandon, J. (1995). Procedural memory in PD. *Dementia*, *6*, 174-178.

***Appollonio, I., Grafman, J., Clark, K., Nichelli, P., Zeffiro, T., & Hallet, M. (1994). Implicit and explicit memory in patients with Parkinson's disease with and without dementia. *Archives of Neurology*, *51*, 359- 367.

**Auriacombe, S., Grossman, M., Carvell, S., Gollomp, S., Stern, M. B., & Hurtig, H. I. (1993). Verbal fluency deficits in Parkinson's disease. *Neuropsychology*, *7*, 182-192.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin*, *99*, 388-399.

Beatty, W. W. (1992). Memory disturbances in Parkinson's disease. In S. J. Huber & J. L. Cummings (Eds.), *Parkinson's Disease: Neurobehavioral aspects* (pp. 49-58). Oxford: Oxford University Press.

**Blonder, L. X., Gur, R. E., Gur, R. C., Saykin, A. J., & Hurtig, H. I. (1989). Neuropsychological functioning in hemiparkinsonism. *Brain and Cognition*, *9*, 244-257.

**Bondi, M. W., & Kaszniak, A. W. (1991). Implicit and explicit memory in Alzheimer's disease and Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, *13*, 339-358.

***Bondi, M. W., Kasniak, A. W., Bayles, K. A., & Vance, K. T. (1993). Contributions of frontal system dysfunction to memory and perceptual abilities in Parkinson's disease. *Neuropsychology*, *7*, 89- 102.

Boone, K. B., Lesser, I. M., Miller, B. L., Wohl, 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.

**Boyd, J. L., Cruikshank, C. A., Kenn, C. W., Madeley, P., Mindham, R. H. S., Oswald, A. G., Smith, R. J., & Spokes, R. G. S. (1991). Cognitive impairment and dementia in Parkinson's disease: a controlled study. *Psychological Medicine*, *21*, 911-921.

***Breen, E. K. (1993). Recall and recognition memory in Parkinson's disease. *Cortex*, *29*, 91-102.

Brown, R. G., & Marsden, C. D. (1988). Subcortical dementia: The neuropsychological evidence. *Neuroscience*, *25*, 363-387.

Brown, R. G., & Marsden, C. D. (1990).Cognitive function in Parkinson's disease: From description to theory. *Trends in Neuroscience*, *13*, 21-29.

Bryman, A., & Cramer, D. (1994). *Quantitative data analysis for social scientists* (Rev. ed.). London: Routledge.

Buytenhuijs, E. L., Berger, H. J. C., Van Spaendonck, K. P. M., Horstink, M. W. I. M., Borm, G. F., & Cools, A. R. (1994). Memory and learning strategies in patients with Parkinson's disease. *Neuropsychologia*, *32*, 335-342.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: a review. *Journal of Abnormal and Social Psychology*, *65*, 145-153.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, *116*, 457-475.

**Cooper, J. A., & Sagar, H. J. (1993a). Encoding deficits in untreated Parkinson's disease. *Cortex*, *29*, 261-265.

**Cooper, J. A., & Sagar, H. J. (1993b). Incidental and intentional recall in Parkinson's disease: an account based on diminished attentional resources. *Journal of Clinical and Experimental Neuropsychology*, *15*, 713-731.

***Cooper, J. A., Sagar, H. J., & Sullivan, E. V. (1993). Short-term memory and temporal ordering in early Parkinson's disease: effects of disease chronicity and medication. *Neuropsychologia*, *31*, 933-949.

**Cooper, J. A., Sagar, H. J., Doherty, S. M., Jordan, N., Tidswell, P., & Sullivan, E. V. (1992). Different effects of dopaminergic and anticholinergic therapies on cognitive and motor function in Parkinson's disease. *Brain*, *115*, 1701-1725.

**Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S., & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain*,*114*, 2095-2122.

*Cronin-Golomb, A., & Braun, A. E. (1997). Visuospatial dysfunction and problem solving in Parkinson's disease. *Neuropsychology*, *11*, 44-52.

Cummings, J. L., & Benson, D. F. (1988). Psychological dysfunction accompanying subcortical dementias. *Annual Review of Medicine*, *39*, 53-61.

Daum, I., Schugens, M. M., Spieker, S., Poser, U., Schonle, P. W., & Birbaumer, N. (1995). Memory and skill acquistion in Parkinson's disease and frontal lobe dysfunction. *Cortex*, *31*, 413-432.

*Dewick, H. C., Hanley, J. R., Davies, A. D. M., Playfer, J., & Turnbull, C. (1991). Perception and memory for faces in Parkinson's disease. *Neuropsychologia*, *29*, 785-802.

Direnfeld, L. K., Albert, M. L., Volicer, L., Langlais, P. J., Marquis, J., & Kaplan, E. (1984). Parkinson's disease: The possible relationship of laterality to dementia and neurochemical findings. *Archives of Neurology*, *41*, 935-941.

Dubois, B., Danze, F., Pillon, B., Cusimano, G., Lhermitte, F., & Agid, Y. (1987). Cholingergic-dependent cognitive deficits in Parkinson's disease. *Annals of Neurology*, *22*, 26-30.

Dubois, B., Boller, F., Pillon, B., & Agid, Y. (1991). Cognitive deficits in Parkinson's disease. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 5, pp. 195-240). Amsterdam: Elsevier.

**Dubois, B., Pillon, B., Legault, F., Agid, Y., & Lhermitte F. (1988). Slowing of cognitive processing in progressive supranuclear palsy. *Archchives of Neurology*, *45*, 1194-1198.

**El-Awar, M., Becker, J. T., Hammond, K. M., Nebes, R. D., & Boller, F. (1987). Learning deficit in Parkinson's disease: Comparison with Alzheimer's disease and normal aging. *Archives of Neurology*, *44*, 180-184.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: a general power analysis program. *Behaviour Research Methods, Instruments, & Computers*, *28*, 1-11.

*Flowers, K. A., Pearce, I., & Pearce, J. M. S. (1984). Recognition memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 1174-1181.

*Gabrieli, J. D. E., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence for the role of a frontostriatal system in working strategic memory. *Neuropsychology*, *10*, 322-332.

**Grossman, M., Carvell, S., Peltzer, L., Stern, M. B., Gollomp, S., & Hurtig, H. I. (1993). Visual construction impairments in Parkinson's disease. *Neuropsychology*, *7*, 536-547.

**Grossman, M., Carvell, S., Stern, M. B., Gollomp, S.,& Hurtig, H.I. (1992). Sentence comprehension in Parkinson's disease: the role of attention and memory. *Brain and Language*, *42*, 437-384.

Growdon, J. H., & Corkin, S. (1986). Cognitive impairments in Parkinson's disease. In M. D. Yahr & K. J. Bergmann (Eds.), *Advances in neurology* (Vol. 45, pp. 383-392). New York: Raven.

Growdon, J. H., Corkin, S., & Rosen, T. J. (1990). Distinctive aspects of cognitive dysfunction in Parkinson's disease. *Advances in Neurology*, *53*, 365-376.

**Gurd, J. M., & Ward, C. D. (1989). Retrieval from semantic and letter initial categories in patients with Parkinson's disease. *Neuropsychologia*, *27*, 743-746.

Hammond-Tooke, G. D., & Pollock, M. (1992). Depression, dementia, and Parkinson's disease. In A. B. Joseph & R. R. Young. *Movement disorders in neuropsychiatry* (Eds.). Cambridge, MA: Blackwell.

**Harrington, D. L., Haaland, K. Y., Yeo, R. A., & Marder, E. (1990). Procedural memory in Parkinson's disease: impaired motor but not visuoperceptual learning. *Journal of Clinical and Experimental Neuropsychology*, *12*, 323-339.

Hartikainen, P., Helkala, E. L., Soininen, H., & Riekkinen, P. (1993). Cognitive and memory deficits in untreated Parkinson's disease and amyotrophic lateral sclerosis patients: A comparative study. *Journal of Neural Transmission*, *6*, 127-137.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

***Heindel, W. C., Salmon, D. P., Shults, C. W., Walicke, P. A., & Butters, N. (1989). Neuropsychological evidence for multiple implicit memory

systems: a comparison of Alzheimer's, Huntington's, and Parkinson's disease patients. *The Journal of Neuroscience*, *9*, 582- 587.

**Helkala, E. L., Laulumaa, V., Soininen, H., & Riekkinen, P. J. (1989). Different error pattern of episodic and semantic memory in Alzheimer's disease and Parkinson's disease with dementia. *Neuropsychologia*, *27*, 1241- 1248.

**Hietanen, M., & Teravainen, H. (1986). Cognitive performance in early Parkinson's disease. *Acta Neurologica Scandinavica*, *73*, 151-159.

**Hietanen, M., & Teravainen, H. (1988). The effect of age of disease onset on neuropsychological performance in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *51*, 244-249.

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality, *Neurology*, *17*, 427-442.

**Huber, S. J., Freidenberg, D. L., Shuttleworth, E. C., Paulson, G. W., & Clapp, L. E. (1989). Neuropsychological similarities in lateralised Parkinsonism. *Cortex*, *25*, 461-470.

***Huber, S. J., Shulman, H. G., Paulson, G. W., & Shuttleworth, E. C. (1987). Fluctuations in plasma dopamine level impair memory in Parkinson's disease. *Neurology*, *37*, 1371-1375.

**Huber, S. J., Shuttleworth, E. C., & Paulson, G. W. (1986). Dementia in Parkinson's disease. *Archives of Neurology*, *43*, 987-990.

**Huberman, M., Moscovitch, M., & Freedman, M. (1994). Comparison of patients with Alzheimer's and Parkinson's disease on different explicit and implicit tests of memory. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *7*, 185-193.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Jacobs, D. M., Marder, K., Cote, L. J., Sano, M., Stern, Y., & Mayeux, R. (1995). Neuropsychological characteristics of preclinical dementia in Parkinson's disease. *Neurology*, *45*, 1691-1696.

Karayanidis, F. (1989). Parkinson's disease: A conceptualization of neuropsychological deficits within an information-processing framework. *Biological Psychology*, *29*, 149-179.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Knight, R. G. (1992). *The neuropsychology of degenerative brain diseases*. Hillsdale, NJ: Lawrence Erlbaum.

Knight, R. G., Godfrey, H. P. D., & Shelton, E. J. (1988). The psychological deficits associated with Parkinson's disease. *Clinical Psychology Review*, *8*, 391-410.

Lange, K. W., Paul, G. M., Robbins, T. W., & Marsden, C. D. (1993). L-Dopa and frontal cognitive function in Parkinson's disease. In H. Narabayashi, T. Nagatsu, N. Yanagisawa, & Y. Mizuno (Eds.), *Advances in Neurology*, *Vol. 60*. New York: Raven.

**Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A. M., & Paul, G. M. (1992). L-Dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology*, *107*, 394-404.

*Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. *Brain*, *106*, 257-270.

Levin, B. E., Llabre, M. M., Reisman, S., Weiner, W. J., & Brown, M. C. (1991). A retrospective analysis of the effects of anticholinergic medication on memory performance in Parkinson's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *3*, 412-416.

**Levin, B. E., Llabre, M. M., & Weiner, W. J. (1989). Cognitive impairment associated with early Parkinson's disease. *Neurology*, *39*, 557- 561.

Levin, B. E., Tomer, R., & Rey, G. J. (1992). Clinical correlates of cognitive impairment in Parkinson's disease. In S. J. Huber & J. L. Cummings (Eds.), *Parkinson's disease: Neurobehavioral aspects*. Oxford: Oxford University Press.

Lieberman, A. (1998). Managing the neuropsychiatric symptoms of Parkinson's disease. *Neurology*, *50 (Supplement 6)*, 33-38.

Lieberman, A., Dziatolowski, M., Kupersmith, M., Serby, M., Goodgolc, A., Korein, J., & Goldstein, M. (1979). Dementia in Parkinson Disease. *Annals of Neurology*, *6*, 355-359.

*Litvan, I., Mohr, E., Williams, J., Gomez, C., & Chase, T. N. (1991). Differential memory and executive functions in demented patients with Parkinson's and Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 25-29.

Mahurin, R. K., Feher, E. P., Nance, M. L., Levy, J. K., & Pirozzolo, F. J. (1993). Cognition in Parkinson's disease and related disorders. In R. W. Parks, R. F. Zec, & R. S. Wilson (Eds.), *Neuropsychology of Alzheimer's disease and other dementias* (pp. 308-349). Oxford: Oxford University Press.

***Massman, P. J., Delis, D. C., Butters, N., Levin, B. E., & Salmon, D. P. (1990). Are all subcortical dementias alike? Verbal learning and memory in Parkinson's and Huntington's disease patients. *Journal of Clinical and Experimental Neuropsychology*, *12*, 729-724.

Mohr, E., Fabbrini, G., Williams, J., Schlegel, J., Cox, C., Fedio, P., & Chase, T. N. (1989). Dopamine and memory function in Parkinson's disease. *Movement Disorders*, *4*, 113-120.

Mohr, E., Mendis, T., & Grimes, J. D. (1995). Late cognitive changes in Parkinson's disease with an emphasis on dementia. In W. J. Weiner & A. E. Lang (Eds.), *Advances in Neurology* (Vol. 65, pp. 97-113). New York: Raven.

**Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *51*, 757-766.

Murdock, B. B., Jr. (1974). *Human memory: Theory and data.* Potomac, MD: Lawrence Erlbaum.

O'Sullivan, M. (1998). *Memory deficits in Parkinson's disease.* Unpublished Master's thesis, Massey University, Palmerston North, New Zealand.

**Ogden, J. A., Growden, J. H., & Corkin, S. (1990). Deficits on visuospatial tests involving forward planning in high-functioning Parkinsonians. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *3*, 125-139.

*Owen, A. M., Beksinska, M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Sahakian, B. J., & Robbins, T. W. (1993). Visuospatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia*, *31*, 627-644.

***Owen, A. M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Lange, K. W., & Robbins, T. W. (1992). Fronto-striatal cognitive deficits at different stages of Parkinson's disease. *Brain*, *115*, 1727-1751.

**Owen, A. M., Sahakian, B. J., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1995). Dopamine-dependent frontostriatal planning deficits in early Parkinson's disease. *Neuropsychology*, *9*, 126-140.

**Oyebode, J. R., Barker, W. A., Blessed, G., Dick, D. J., & Britton P. G. (1986). Cognitive functioning in Parkinson's disease: In relation to prevalence of dementia and psychiatric diagnosis. *British Journal of Psychiatry*, *149*, 720-725.

Pillon, B., Deweer, B., Agid, Y., & Dubois, B. (1993). Explicit memory in Alzheimer's, Huntington's, and Parkinson's disease. *Archives of Neurology*, *50*, 374-379.

Pillon, B., Dubois, B., Cusimano, G., Bonnet, A. M., Lhermitte, F., & Agid, Y. (1989). Does cognitive impairment in Parkinson's disease result from non-dopaminergic lesions? *Journal of Neurology, Neurosurgery, and Psychiatry*, *52*, 201-206.

**Pillon, B., Dubois, B., Lhermitte, F., & Agid, Y. (1986). Heterogeneity of cognitive impairment in progressive supranuclear palsy, Parkinson's disease, and Alzheimer's disease. *Neurology*, *36*, 1179-1185.

Pillon, B., Ertle, S., Deweer, B., Sarazin, M., Agid, Y., & Dubois, B. (1996). Memory for spatial location is affected in Parkinson's disease. *Neuropsychologia*, *34*, 77-85.

Pirozzolo, F. J., Hansch, E. C., Mortimer, J. A., Webster, D. D., & Kuskowski, M. A. (1982). Dementia in Parkinson disease: A neuropsychological analysis. *Brain and Cognition*, *1*, 71-83.

**Randolph, C., Braun, A. R., Goldberg, T. E., & Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: dissociation of storage and retrieval failures. *Neuropsychology*, *7*, 82-88.

Reid, W. G. J., Broe, G. A., Hely, M. A., Morris, J. G. L., Willimson, P. M., O'Sullivan, D. J., Rail, D., Gence, S., & Moss, N. G. (1989). The neuropsychology of de novo patients with idiopathic Parkinson's disease: The effects of age of onset. *International Journal of Neuroscience*, *48*, 205-217.

**Robbins, T. W., James, M., Owen, A. M., Lange, K. W., Lees, A. J., Leigh, P. N., Marsden, C. D., Quinn, N. P., & Summers, B. A. (1994). Cognitive deficits in progressive supranuclear palsy, Parkinson's disease, and multiple system atrophy in tests sensitive to frontal lobe dysfunction. *Journal of Neurology, Neurosurgery, and Psychiatry*, *57*, 79-88.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*, 183-192.

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin*, *92*, 500-504.

Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin*, *99*, 400-406.

Rossi, J. S. (1990). Statistical power of psychological research: what have we gained in 20 years? *Journal of Consulting and Clinical Psychology*, *58*, 646-656.

Ruberg, M., & Agid, Y. (1988). Dementia in Parkinson's disease. In L. Iversen, S. D. Iversen, & S. H. Snyder (Eds.), *Handbook of psychopharmacology: Vol. 4. Psychopharmacology of the aging nervous system.* New York: Plenum.

Sagar, H. J., Atchison, L., Doherty, S. M., Ball, M. C., & Cooper, J. A. (1995). The contribution of non-dopaminergic pathology to cognitive, motor and affective disability in Parkinson's disease. In I. Hanin et al. (Eds.), *Alzheimer's and Parkinson's diseases* (pp. 69-76). New York: Plenum.

Sagar, H. J., Cohen, N. J., Sullivan, E. V., Corkin, S., & Growdon, J. H. (1988). Remote memory function in Alzheimer's disease and Parkinson's disease. *Brain*, *111*, 185-206.

Sagar, H. J., & Sullivan, E. V. (1988). Patterns of cognitive impairment in dementia. In C. Kennard (Ed.), *Recent advances in clinical neurology: Number 5.* (pp. 47-86). Edinburgh: Churchill Livingstone.

**Sagar, H. J., Sullivan, E. V., Cooper, J. A., & Jordan, N. (1991). Normal release from proactive interference in untreated patients with Parkinson's disease. *Neuropsychologia*, *29*, 1033-1044.

***Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain*, *111*, 695-718.

Saint-Cyr, J. A., & Taylor, A. E. (1993). Cognitive dysfunction in Parkinson's disease. In J. S. Schneider & M. Gupta (Eds.), *Current concepts in Parkinson's disease research.* Kirkland, WA: Hogrefe and Huber.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods*, *1*, 115-129.

Schwarzer, R. (1989). *Meta-analysis programs*. Durham, NC: Duke University.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin*, *105*, 309-316.

Starkstein, S. E., Bolduc, P. L., Mayberg, H. S., Presziosi, T. J., & Robinson, R. G. (1990). Cognitive impairments and depression in Parkinson's disease: A follow up study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *53*, 597-602.

Starkstein, S. E., Preziosi, T. J., Berthier, M. L., Bolduc, P. L., Mayberg, H. S., & Robinson, R. G. (1989). Depression and cognitive impairment in Parkinson's disease. *Brain*, *112*, 1141-1153.

Starkstein, S. E., & Robinson, R. G. (1991). Dementia of depression in Parkinson's disease and stroke. *The Journal of Nervous and Mental Disease*, *179*, 593-601.

Stoffelmayr, B. E., Dillavou, D., & Hunter, J. E. (1983). Premorbid functioning and recidivism in schizophrenia: A cumulative analysis. *Journal of Consulting and Clinical Psychology*, *51*, 338-352.

*Sullivan, E. V., & Sagar, H. J. (1989). Nonverbal recognition and recency discrimination deficits in Parkinson's disease and Alzheimer's disease. *Brain*, *112*, 1503-1517.

**Sullivan, E. V., Sagar, H. J., Cooper, J. A., & Jordan, N. (1993). Verbal and nonverbal short-term memory impairment in untreated Parkinson's disease. *Neuropsychology*, *7*, 396-405.

***Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1986). Frontal lobe dysfunction in Parkinson's disease. *Brain*, *109*, 845- 883.

*Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1987). Parkinson's disease. *Brain*, *110*, 35-51.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1988). Idiopathic Parkinson's disease: Revised concepts of cognitive and effective status. *Canadian Journal of Neurological Science*, *15*, 106-113.

**Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1990). Memory and learning in early Parkinson's disease: evidence for a "frontal lobe syndrome". *Brain and Cognition*, *13*, 211-232.

*Tierney, M. C., Nores, A., Snow, W. G., Fisher, R. H., Zorzitto, M. L., & Reid, D. W. (1994). Use of the Rey auditory verbal learning test in differentiating normal aging from Alzheimer's and Parkinson's dementia. *Psychological Assessment*, *6*, 129-134.

Tsai, C. H., Lu, C. S., Hua, M. S., Lo, W. L., & Lo, S. K. (1994). Cognitive dysfunction in early onset parkinsonism. *Acta Neurologica Scandinavica*, *89*, 9-14.

Tweedy, J. R., Langer, K. G., & McDowell, F. H. (1982). The effect of semantic relations on the memory deficit associated with Parkinson's disease. *Journal of Clinical Neuropsychology*, *4*, 235-247.

van Spaendonck, K. P. M., Berger, H. J. C., Horstink, M. W. I., Buytenhuijs, E. L., & Cools, A. R. (1993). Impaired cognitive shifting in Parkinsonian patients on anticholinergic therapy. *Neuropsychologia*, *31*, 407-411.

Weingartner, H., Burns, S., Diebel, R., & LeWitt, P. A. (1984). Cognitive impairment in Parkinson's disease: Distinguishing between effort-demanding and automatic cognitive processes. *Psychiatry Research*, *11*, 223-235.

Whittington, C. J., & Podd, J. V. (1999). *Recognition memory impairment in Parkinson's disease.* Manuscript in preparation, Massey University, Palmerston North, New Zealand.

Whittington, C. J., & Podd, J. V. (1996). Human performance and physiology: A statistical power analysis of ELF electromagnetic field research. *Bioelectromagnetics*, *17*, 274-278.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37-49.

Zetusky, W. J., Jankovic, J., & Pirozzolo, F. J. (1985). The heterogeneity of Parkinson's disease: Clinical and prognostic implications. *Neurology*, *35*, 522-526.

Appendix

Journals considered most likely to publish studies on cognitive deficits in PD: *Acta Neurologica Scandinavica; Advances in Neurology; Annals of Neurology; Annals of the New York Academy of Sciences; Annual Review of Medicine;*

Archives of Clinical Neuropsychology; Archives of Neurology; Brain and Cognition; Brain and Language; Brain Research; British Journal of Psychiatry; Clinical Neuropsychologist; Clinical Psychology Review; Cognitive Neuropsychology; Cortex; International Journal of Neuroscience; Journal of Clinical and Experimental Neuropsychology; Journal of Nervous and Mental Disease; Journal of Neurology, Neurosurgery, and Psychiatry; Neurology; Neuropsychiatry, Neuropsychology, and Behavioral Neurology; Neuropsychobiology; Neuropsychologia; Neuropsychology; Psychological Assessment; Psychological Medicine; Psychological Reports; and, *Trends in Neurosciences.*