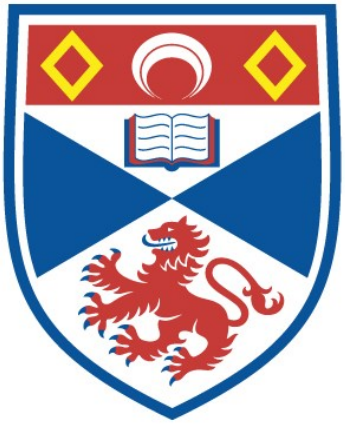
THE COGNITIVE AND MOTOR DEFICITS OF PARKINSON'S DISEASE.

Jeremy Gauntlett-Gilbert

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1999

The cognitive and motor deficits of Parkinson's disease

Jeremy Gauntlett-Gilbert

Submitted to the University of St. Andrews for the degree of Ph.D.

August 1998

Declaration

i) I, Jeremy Gauntlett-Gilbert, hereby certify that this thesis, which is approximately 75,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date: 19/2/99 Signature of Candidate:

ii) I was admitted as a research student in October 1995 and as a candidate for the degree of Ph.D. in October 1995; the higher study for which this is a record was carried out in the University of St. Andrews between 1995 and 1998.

Date: 19/2/99 Signature of Candidate:

iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: Signature of Supervisor:

Copyright

Unrestricted

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to the copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Date: 19/2/99 Signature of Candidate:

Acknowledgments

These acknowledgments were going to be long; now they're short. Thanks to Verity, of course. Verity has supported me and educated me with astonishing good humour. This thesis exists because of her, and she made it a pleasant process. Any mistakes that remain in this thesis are, of course, entirely her fault. Thanks also to Eric for being a second supervisor who was present but not intrusive. Marie and Derek (and some others) offer reassuring evidence that not all academics have a personality disorder. Thanks to all of them.

At the clinical end, thanks to Richard Roberts for referring patients. Also, my gratitude to everybody who took part in the studies in this thesis. A lot of ill people traveled a long way on public transport to take part in my studies - they have my thanks and respect. In particular, JG and GM (Study 6) made an incredible effort at a particularly stressful time.

I'm blessed with a lot of funny, rude, delightful good friends. I can't list them all without this section looking like the inside of an album cover. It may seem rude to credit members of staff but not my friends; however, I rest assured that the people that matter don't give a flying damn about my Ph.D. thesis. Very special people; to my mum, thanks for so many great chats, and so much encouragement. I'm terribly sorry about the phone bill. And to Sarah, I don't know where to start. Thanks for everything. Finally, heartfelt thanks to Duke Ellington.

This thesis is dedicated to my Grandpa, Grandma and my Dad.

Contents

List of Tables and Figures

Papers derived from this thesis

- Gauntlett-Gilbert J., Roberts R.C., Brown V.J. Mechanisms underlying attentional set-shifting in Parkinson's disease. *Neuropsychologia,* in press.
- Gauntlett-Gilbert J., Brown V.J. Reaction time deficits and Parkinson's disease. *Neuroscience and Biobehavioural Reviews,* in press.

Abbreviations used frequently in the thesis

CANTAB	Cambridge Neuropsychological Test Automated Battery
CAPIT	Core Assessment Program for Intracerebral Transplants
DSM III-R	Diagnostic and Statistical Manual of the American Psychiatric Association, Volume 3, Revised
ED	Extradimensional
EMG	Electromyography
GDS	Geriatric Depression Scale
ID	Intradimensional
MCST	Modified Card-Sorting Test
MDS 2.0	Modified Dyskinesia Rating Scale, Version 2.0
MMSE	Mini-Mental State Examination
NART	National Adult Reading Test
PD	Parkinson's disease
RAVLT	Rey Auditory-Verbal Learning Test
RMT	Recognition Memory Test (Warrington)
RT	Reaction Time
WAIS	Wechsler Adult Intelligence Scale
WCST	Wisconsin Card-Sorting Test

Abstract

Deficits of attentional and motor set that are seen in patients with PD were investigated in the research reported in this thesis. A deficit of attentional set is a failure of selective attention to one aspect of a complex stimulus. Deficits of motor set manifest as an inability to form a state of motor 'readiness' that can speed movement initiation. Attentional set was investigated with tasks that require patients to shift attention between perceptual dimensions (extradimensional - ED - shifting tasks), and motor set was studied using reaction time (RT) tasks.

Study 1 rejected the hypothesis that the mechanism of 'learned irrelevance' causes ED shift deficits in patients with PD. Studies 2 and 3 confirmed that learned irrelevance also plays no role in determining the difficulty of ED shifts in healthy subjects. The experimental manipulation used in Study 4 succeeded in creating changes in scores that resembled those seen in patients with PD. Thus, it appears that patients with PD may have a deficit similar to that induced by the experimental manipulation - that is, an inability to attend to all dimensions present when hypothesis testing after an ED shift.

A quantitative analysis of past RT studies of PD showed that the ability to speed movement initiation when given advance information about an upcoming movement - a form of motor set - is intact in patients with PD. In contrast, the motor set that underlies rapid simple RT performance is consistently dysfunctional in patients with PD. Study 5 investigated temporal and spatial motor set in PD, finding that these two mechanisms are functionally separate and that temporal motor set is intact in PD. A final study (Study 6) investigated the cognitive consequences of a novel neurosurgical treatment for PD, finding it to be largely a neuropsychologically 'safe' procedure.

1.1 Parkinson's disease: its nature, prevalence and impact

1.1.1 Nature

Idiopathic Parkinson's disease (PD) is a chronic, neurodegenerative disorder of late middle life. PD is a movement disorder that is associated with dysfunction of the extrapyramidal system and causes a progressive impairment of the ability to make voluntary movements. Patients experience a range of symptoms including tremor, slowness of movement and muscular rigidity (for a full description, see Section 1.2) that can that can leave individuals immobile and unable to care for themselves. The onset of symptoms is spontaneous and the cause of the disease is not known. PD is progressive in character, and whilst the rate of disease progression varies greatly between individuals the majority of patients experience an unremitting increase in the severity of symptoms. Although the most disabling aspect of PD is its impairment of movement, the disease can also cause cognitive deficits and is associated with a high level of affective disturbance.

1.1.2 Prevalence

It is estimated that 120,000 people in the UK and 1,500,000 people in the U.S.A. have PD (Parkinson's Disease Society, 1998; National Parkinson Foundation Inc. 1996). The mean age of onset is around 60 (Quinn, 1995) and prevalence increases rapidly with advancing age (50-59 years - 0.08%; 70-79 years - 0.83%, Mutch et al., 1986). More men are diagnosed with PD than women (a ratio of 1.3 : 1 - Ben Shlomo, 1996) and in western countries more white people have the disorder than black people (Paddinson and Griffith, 1974; Pearce, 1992). It appears that PD is more common in western countries than in China or Africa, but it is unclear whether this impression is genuine or created by differences in diagnostic criteria and the age of sample groups between studies in different countries (Pearce, 1992). Otherwise, there are no known effects of socioeconomic status, level of education or occupation on the prevalence of the disease (Ben-Shlomo, 1996).

1.1.3 Impact

PD represents an increasing challenge to health and social services, as demographic trends and improvements in health care ensure that the number of people with severe PD will grow over the next decade. The growth in the U.K.'s elderly population (an extra 100,000 individuals over the age of 80 in the next 10 years - Department of

Health, 1997) will result in more cases of PD. Also, as health care improves, more patients with PD will survive to experience the more advanced stages of the disease. There is no known cure for PD and whilst symptomatic medical and surgical treatments are beneficial for a limited time they lose effectiveness in the long term and seldom stop PD's progressive course. Therefore, the next decade will see a growth in the number of individuals who have advanced PD and few therapeutic options.

This thesis describes research into specific cognitive and motor symptoms experienced by patients with PD. The past literature relating to the study of these symptoms will be reviewed in detail in Sections 1.5 and 1.6; the following two sections provide brief descriptions of the onset and progression of motor and cognitive symptoms in PD. Whilst PD is a multifactorial disease that often causes simultaneous disturbances of movement, cognition and affect, it is useful to describe the effects of PD on these domains separately.

1.2 PD as a movement disorder

The four 'cardinal' symptoms of PD - tremor, rigidity, bradykinesia and postural problems - are described below, followed by an account of their onset and progression. Not all of these symptoms will be present in any individual patient and some symptoms are more common at certain stages of the disease. For example, disturbances of gait and posture are more common in advanced PD and 25% of patients do not have a resting tremor. Nonetheless, these are the primary symptoms of the disease; in order for a confident diagnosis of PD to be made, it is recommended that at least two of the following four symptoms should be present, at least one of which must be tremor or bradykinesia (Langston et al., 1992).

1.2.1 Cardinal symptoms

Tremor

Patients commonly show a rhythmical resting tremor in their hands and arms. This tremor oscillates at a rate of 4-6 Hz and is caused by alternating activity of flexor/extensor and pronator/supinator muscles. Tremor of the thumb and forefinger is often seen early in the disease and is referred to as a 'pill-rolling' tremor; tremor spreads to the whole hand and arm with disease progression and can ultimately result in a postural tremor. Although the tremor seen in PD is predominantly active when the affected limb is at rest, a less severe tremor is also sometimes seen with movement.

Rigidity

The limbs of a patient with PD are often 'rigid' and resistant to passive movement. Muscular rigidity affects the neck and shoulders early in the disease and is usually more prominent in the arms than the legs. Rigidity is caused by simultaneous contraction of agonist and antagonist muscles. However, flexor muscles are usually more active than extensor muscles and as a consequence patients' limbs are often bent at the elbow and knee. This contributes to the abnormal posture seen in patients with PD (see below). 'Cogwheel' rigidity is often cited as characteristic of PD. This is observed when an examiner manipulates a patient's limb whilst it is at rest; the limb is felt to move freely and then 'stick' repeatedly and in rapid succession. This phenomenon is thought to be caused by an interaction of rigidity and tremor.

Bradykinesia

The most disabling aspect of PD is its impact on patients' ability to make voluntary movements. Patients are impaired in the initiation and execution of voluntary movements (akinesia and bradykinesia, respectively) and make a diminished number of spontaneous movements (hypokinesia). Patients report that the amount of 'effort' required to make a movement is drastically increased by PD. Movements that were previously fluid and automatic come to require deliberate conscious effort. Simultaneous bimanual movements are particularly impaired (Benecke et al., 1986), as well as movements that form part of a sequence (Benecke et al., 1987a). Patients also have problems with fine finger movements, which can lead to difficulties with such tasks as fastening clothing and handwriting. Bradykinesia is independent of rigidity; thalamic lesions can abolish rigidity but leave bradykinesia unchanged (Marsden, 1990).

Posture and gait

PD forces patients to adopt an abnormal posture. As the disease advances, patients' knees, arms and wrists become flexed due to rigidity and the trunk is bent forward. Rigidity also stops patients' arms from swinging freely whilst walking. Patients take short, shuffling steps and in more advanced disease they sometimes 'freeze' and are unable to take further steps. There is also an impairment of balance that arises from the failure of normal 'righting and rescue' reflexes (Martin, 1967) that compensate for sudden shifts in an individual's centre of gravity. As a consequence, patients frequently fall. Loss of balance also contributes to the 'festinating' gait seen in patients with PD

(Parkinson, 1817/1997); patients shuffle forward rapidly, apparently trying to avoid falling forward.

Other physical symptoms

Patients often lose facial expression through muscular rigidity and their voices can lose volume and inflection. Constipation and bladder problems are common and over half of patients have sexual disorders.

1.2.2 Natural history of PD, before and after levodopa

Onset of symptoms and disease progression - before levodopa

The onset and progression of symptoms in untreated PD was documented by Hoehn and Yahr in 1967. Their five-stage classification of the progression of PD is still frequently used to categorise disease severity. Stage I lasts on average three years after the onset of the disease and is characterised by the presence of one or more of tremor, rigidity or bradykinesia on one side of the body. These symptoms spread to the previously unaffected side of the body in Stage II of the disease, which occurs on average between four and six years after onset. Although symptoms remain bilateral after this stage, patients usually experience more severe symptoms on one side of the body. In Stage III, on average seven years after onset, impairments of postural stability become apparent. Stages IV and V represent increasing levels of disability that are caused by worsening severity of existing symptoms. Individuals in Stage IV (nine years after onset) require much assistance in the activities of daily living but can still stand and walk without assistance; in Stage V (14 years after onset), patients are confined to bed or a wheelchair unless aided.

Onset and progression of symptoms - after levodopa

Whilst Hoehn and Yahr's (1967) description of the progression of PD is accurate, contemporary patients with PD do not experience the disease in this way. The underlying pattern of the disease is obscured by the symptomatic relief given by medical therapy, in particular by levodopa. Levodopa therapy provides good relief of mild to moderate symptoms in a majority of patients with PD and it prevents the unremitting decline into disability described by Hoehn and Yahr (1967). However, levodopa retains its clinical efficacy only for a limited time. Within the first five years of levodopa therapy, over half of patients will experience interruptions of the previously smooth clinical response to the drug and possibly also side-effects (Quinn, 1995). The

twin problems of fluctuating response to medication and dyskinesias (abnormal, involuntary movements) characterise the movement status of contemporary patients with advanced PD, rather than the persistent rigid akinesia seen in patients who were treated before the introduction of levodopa.

Fluctuations

Medicated patients with advanced PD often switch abruptly from periods of mobility in which their medication is effective ('on' periods) to periods in which they receive no benefit of medication and are relatively immobile ('off' periods). On taking a new dose of levodopa, a patient can rapidly revert from an 'off' state to an 'on' state. These sudden changes in clinical status represent a major shortcoming of medical therapy for PD, particularly as 'on' periods are often associated with the presence of dyskinesias (see below). Patients experience the level of motor disability associated with their 'underlying' unalleviated PD in 'off' phases, and the transition into an off phase is associated with feelings of dysphoria and panic in two-thirds of patients (Nissenbaum et al., 1987).

Dyskinesias

One of the side-effects of levodopa therapy is to cause patients to make a range of involuntary movements. These movements are usually choreic or dystonic in nature, though in some severe cases (usually younger patients - Quinn, 1995) they can be ballistic. Dyskinesias are associated either with the point at which a dose of levodopa has the greatest clinical effect ('peak-dose dyskinesia') or with the beginning and end of a dose ('diphasic dyskinesia'). In some patients with advanced disease, the thresholds for clinical effectiveness of a dose of levodopa and for the generation of dyskinesias are the same; thus, these patients are either in an 'off' phase and immobile or in an 'on' phase and dyskinetic. Severe dyskinesias not only interfere with voluntary movement but can also be physically exhausting.

1.3 PD as a disorder of cognitive function

Charcot (1875, cited in Lees and Smith 1983) recognised that patients with PD often experience cognitive deficits, contradicting Parkinson's (1817/1997) original assertion that PD left "the senses and intellect uninjured" (p.1). However, it is only through recent clinicopathological studies (such as that of Hughes et al., 1993) that it has been confirmed that PD alone - in the absence of other neuropathology - can cause serious cognitive deficits. Although the cognitive deficits of 10 to 15% of patients with PD can be so severe as to be described as dementia (Brown and Marsden, 1984), in the majority of patients the deficits are less severe. As a consequence, the cognitive changes of PD are nearly always of lesser concern to patients than their motor symptoms. However, the fact that cognitive deficits are masked by even more disabling motor symptoms does not make the cognitive problems of PD any less real; if there are currently 120,000 people with PD in the U.K., at least 12,000 of these will either currently have dementia or will develop it as a consequence of their condition.

It is beyond the scope of this section to provide a comprehensive review of all of the research findings on the cognitive deficits of PD, and the specific deficits under investigation in this thesis will be reviewed in detail in Section 1.6. This section instead aims to provide context for the discussion of specific cognitive deficits by describing the cognitive changes of PD in similar terms to the motor symptoms. It will be shown that cognitive deficits are present in early, unmedicated PD (Section 1.3.1), that they become more severe and encompass more domains of cognition as the disease progresses (Sections 1.3.2 and 1.3.3) and that they can be disabling (Section 1.3.4).

1.3.1 Cognitive deficits are present in unmedicated patients with early PD

Cognitive deficits can be found in patients with PD as early as 16 months after the appearance of motor symptoms (Cooper et al., 1991). The study of cognitive deficits in unmedicated patients with early PD is informative, as it shows that such deficits are an integral part of the disorder, rather than a consequence of widespread neuropathology in advanced disease. Also, some antiparkinsonian medication - particularly anticholinergics - can cause cognitive deficits (Van Spaendonck et al., 1993) and the study of unmedicated patients eliminates this factor. Table 1.1 summarises the findings of 16 studies that have investigated cognitive deficits in unmedicated patients with early PD. It is clear from this table that studies have frequently found areas of spared cognitive function in such patients. This is unsurprising, as most of the patients studied were early in the course of what is primarily a movement disorder. The fact that

cognitive deficits are frequently found in such mildly affected patients is more striking; also, these deficits appear to affect a wide range of cognitive domains.

Selective attention

Inspection of Table 1.1 shows that deficits of selective attention are a consistent finding in unmedicated patients with early PD. A total of eight studies describe deficits in the performance of the Wisconsin Card-Sorting Test (WCST - Milner (1963) or Nelson (1976) version), extradimensional shifts or the Stroop task. The one study that found preserved card sorting (Pillon et al., 1997) used the Nelson (1976) version of the WCST, which is less demanding than Milner's (1963) version and may measure different abilities (DeZubicaray and Ashton, 1996).

Memory and strategy

Recall of verbal material has also been shown to be frequently impaired in early unmedicated PD by the Rey Auditory-Verbal Learning Test (RAVLT) and subtests of the Wechsler Memory Scales (WMS). However, it is unclear whether impaired memory processes cause these test deficits. Buytenhuis et al. (1994) have shown that deficits on a memory test similar to the RAVLT can be partially explained by the failure of patients with PD to use appropriate strategies to organise material. This 'strategy' problem extends to spatial memory deficits. Pillon et al. (1997, 1998) constructed a spatial memory test that appeared to be more sensitive to deficits caused by PD than the spatial working memory test of the CANTAB battery (see Table 1.1). In fact, deficits on Pillon et al.'s test were largely due to a failure of 'strategic' processes (Pillon et al., 1998), indicating that this test was more sensitive to strategy use than the CANTAB test, rather than being more sensitive to spatial memory deficits. These findings are significant, as they demonstrate that high-level cognitive deficits, such as problems with 'strategic' or 'executive' function, can influence performance in a wide range of cognitive domains.

The research findings relating to other cognitive domains listed in Table 1.1 are more equivocal; this is perhaps unsurprising, as subject numbers vary widely between studies and the subject groups used vary in their characteristics. For example, the mean age of the patient groups in these studies varies widely, and neuropsychological performance has been shown to be worse for patients who are older when they experience the onset of symptoms (Reid et al., 1989). As a consequence, studies that examine deficits across a range of cognitive domains in a single group of patients can give a clearer indication of which domains are more vulnerable to PD. Cooper et al.'s (1991) study is an example of this approach, testing a group of 60 patients on 15 neuropsychological tests. Cooper et al. (1991) found deficits in immediate verbal

memory, verbal fluency, the WCST and cognitive sequencing; language function other than fluency was intact, as was memory after a delay, visual perception and forward digit span. Whilst Table 1.1 indicates the range of cognitive function that can be impaired by early PD, Cooper et al.'s results are a better guide to the neuropsychological tests that are most sensitive to early PD.

Domain	Test	Significant impairment	No impairment
Selective attention to aspects	*WCST, MCST (any measure)*	1, 8, 16	13
	ED shifting (CANTAB)	4, 9, 11	
	Stroop	7, 13	
	Odd-man-out		13
	Fimm set-shifting test (with external cue)	5	
Verbal Fluency	*Letter*	8	5, 13, 16
	Category	1*, 5, 8	1*, 13
Miscellaneous 'frontal / executive'	*Trails test*	6, 7	16
	Tower of London (CANTAB)		9
	Cognitive estimates		8
	WAIS picture arrangement	1	
	Proverb interpretation		3
	Backward digit span & digit ordering		1
Memory: Spatial	*Spatial Working Memory - (CANTAB)*		9, 10, 12
	Spatial Span (CANTAB)		9
	Pillon visuospatial learning test	13, 14	
Memory: Verbal	*RAVLT*	15, 16	
	Pillon verbal associates test	14	
	Pillon verbal learning test		13
	WMS logical memory	1, 7	16
	WMS Associates	1, 7	
	Brown-Peterson Test	1, 2	
	Owen verbal memory		12
Memory: Visual	*WMS visual reproduction*	1, 7	16
	Warringon RMT - faces		8
	Benton Visual Recognition	15	
	Recall of Rey Figure	1	
	Pattern recognition (CANTAB)		9, 10
	Gollin incomplete pictures - recognition:	1	
	recall:		1
Visuospatial / visuoconstructive	*WAIS Block design*	7	6
	Benton form discrimination & Benton line orientation		13
	Map test & mental rotation		16
	Matchsticks test		1
	Hartikainen visuoconstructive tests	6	
Language	*WAIS similarities*		3
	Reporters test	1	
	Boston Naming test & token test		1

* - subjects showed a deficit on one category, but no deficit on the other category

Table 1.1 (previous page): Significant deficits and spared function in untreated patients with PD. Unique tests are named after the first author of the article e.g. 'Owen verbal memory test'. Abbreviations: WCST - Wisconsin Card-Sorting Test; MCST - Modified Card-Sorting Test; WAIS - Wechsler Adult Intelligence Scale; CANTAB - Cambridge Neuropsychological Test Automated Battery; RAVLT - Rey Auditory-Verbal Learning Test; WMS - Wechsler Memory Scale; RMT - Recognition Memory Test
Studies included: 1 - Cooper et al. (1991); 2 - Cooper and Sagar (1993); 3 - Cronin-Golomb et al. (1994); 4 - Downes et al. (1989); 5 - Fimm et al. (1994); 6 - Hartikainen et al. (1993); 7 - Hietanen and Teravainen (1986); 8 - Lees and Smith, (1983); 9 - Owen et al. (1992); 10 - Owen et al. (1993a); 11 - Owen et al. (1993b); 12 - Owen et al. (1997); 13 - Pillon et al. (1997); 14 - Pillon et al. (1998); 15 - Reid et al. (1989); 16 - Taylor et al. (1987)

The studies summarised in Table 1.1 all compare mean cognitive performance for a group of patients with that of a group of controls. This gives no information about the distribution of deficits in the PD population; for example, are all patients impaired or do a subgroup have spared cognitive function? The only study that addresses this issue is that of Reid et al. (1989). The authors of this study plot the percentage of patients with a score greater than two standard deviations below the control mean for each of the tests used. This gives a clear picture of the proportion of patients with more serious impairment in each domain of cognitive function.

In summary, unmedicated patients with early PD have been shown to have cognitive deficits in a number of domains. Deficits of selective attention and memory are common; basic language skills and visual perception are generally spared (Cooper et al. 1991). Disturbances of higher-level cognitive function may account for deficits in a number of cognitive domains. For example, deficits of strategy use can cause problems of recall of verbal and spatial material (Buytenhuis et al., 1994; Pillon et al., 1998).

1.3.2 Cognitive impairment increases with disease progression

The existence of gross cognitive deterioration as a consequence of the progression of PD is well-established and a number of longitudinal studies have attempted to find factors that predict the development of parkinsonian dementia (e.g. Jacobs et al., 1995; Palazzini et al., 1995; Mahieux et al, 1998). The existence of more subtle decline in cognitive performance has also been shown by longitudinal studies. For example, Capparos-Lefebvre et al. (1995) have shown that cognitive deficits that are typical of unmedicated patients with early PD - that is, executive function (including the WCST) and memory - were significantly worse three years after a baseline test. This worsening appeared to be specific to the progression of PD, as it did not depend on patients' age at the followup stage. Capparos-Lefebvre et al. also found evidence of a more global

decline in cognitive function; group performance on the modified Mini-Mental State Examination (mMMSE) was worse at the three year followup. Such a decline in mMMSE performance was also noted by Starkstein et al. (1992) after only one year in depressed patients with PD.

A more common method of demonstrating the relationship of cognitive deficits to disease progression is to correlate cognitive performance with measures of disease severity in a single group of patients at a particular time. Both Reid et al. (1989) and Cooper et al. (1992) have found a number of significant correlations between cognitive and motor status using this method. However, the relationship between cognitive status and disease severity is not simple; for example, Cooper et al.'s (1992) study was preceded by a study in 1991 (Cooper et al., 1991) that used the same cohort of patients; in 1991, when these patients were earlier in the course of the disease and unmedicated, the relationship between cognitive and motor symptoms was much less strong. However, this may have been due to the fact that both the cognitive and motor deficits seen in the untreated patients with early PD were mild (Cooper et al., 1992). It also appears that the association between cognitive and motor status may only hold true for certain cognitive functions and particular disease variables. For example, Van Spaendonck et al. (1996) found that verbal fluency was not associated with any measure of disease severity; in contrast, card sorting performance was associated with rigidity, but not with bradykinesia or tremor. Thus, it appears that whilst cognitive impairment does increase with disease progression, individual domains of cognition are likely to degenerate at different rates that are related to the progression of certain motor symptoms.

1.3.3 Cognitive impairment spreads to new domains with disease progression

It is well-established that some of the movement symptoms of PD can be absent when the disease is in its early stages but can appear with disease progression. For example, postural instability and postural tremor are most commonly seen in advanced PD. The cognitive symptoms of PD can also progress from the relatively selective set of deficits seen in early PD (Section 1.3.1) to a global loss of function that can be described as dementia (see Section 1.3.4). Owen and colleagues have provided the most convincing demonstration of this 'spread' of deficits in a series of studies of patients at different stages of PD (Owen et al., 1992, 1993a, 1995, 1996c). Using a cross-sectional design, Owen et al. have assessed the performance of groups of patients at different stages of PD (e.g. unmedicated mild PD, medicated mild PD, medicated severe PD) on a number

of tests of planning and memory. These studies have shown that patients with PD are impaired in the performance of a number of relatively 'simple' memory tasks that those with less severe PD can perform normally. For example, Owen et al. (1992) found that performance on a simple measure of spatial 'span' was normal in patients with mild PD, but impaired in those with severe PD. Similarly, Owen et al. (1993a) found a range of deficits in patients with severe PD that were not present in mild PD, including deficits of spatial recognition memory, visual matching-to-sample and learning of paired associates.

This 'spreading' of cognitive deficits also affects the performance of more complex memory tasks that require use of a strategy to attain optimal performance. Owen et al. (1997) found that memory deficit that selectively impaired recall of one type of material could 'spread' to affect the recall of a range of different materials. Owen et al. found that medicated patients with mild PD showed a deficit on a working memory task that required recall of spatial information, but not on structurally similar tasks that required recall of verbal and pictorial information. However, in patients with severe PD the memory deficit had 'spread' such that memory for all of these items was impaired. Owen et al. (1993, 1997) account for this spread of memory deficits in terms of the pattern of degeneration seen in the caudate nucleus of patients with PD. It is often noted that the cognitive deficits of mild and moderate PD resemble those seen after frontal lobe damage (Taylor et al., 1986). Owen et al. argue that this is due to loss of dopamine in the anterior regions of the head of the caudate nucleus, which is the part of the striatum that receives afferents from parts of the prefrontal cortex. This would account for deficits on such tasks as the CANTAB Spatial Working Memory test, which are seen in both mild PD and after frontal lobe damage (Owen et al., 1992, 1990 respectively). Owen et al. suggest that the nonspatial working memory deficits seen could either arise from dopamine depletion in more posterior areas of the caudate nucleus, or could be non-striatal or non-dopaminergic in origin. A subsequent study in monkeys (Levy et al., 1997a) has shown that the performance of nonspatial working memory tasks is associated with activity in the body of the caudate nucleus, suggesting that the memory deficits seen in Owen et al.'s (1993, 1997) studies are striatal in origin.

Deficits of visual perception have also been shown to emerge with disease progression. Flowers and Robertson (1995) compared the performance of medicated patients with mild, moderate and severe PD on a number of tests of visual perception. In the more severely affected patients, deficits were present on those tests that required the use of 'higher-level', 'interpretive' perceptual abilities. Patients with mild PD were

unimpaired in the performance of both the Embedded Figures test and the 'pattern comprehension' and 'progression' sections of Raven's Progressive Matrices, whereas patients with 'moderate' or 'severe' PD were impaired in the performance of all of these tests. Again, this demonstrates that the progression of PD can lead to the emergence of entirely new cognitive deficits; the progression of PD does not just cause existing deficits to worsen. This is significant as the cognitive deficits seen in patients with early PD are relatively selective, and only a 'spreading' of these deficits can account for how PD can cause the global deterioration of cognitive function known as 'dementia'.

1.3.4 The cognitive deficits of PD can be disabling

As noted above, 'dementia' as defined by DSM III-R criteria can be caused by the neuropathology of PD alone (Hughes et al., 1993). The increased risk of this dementia above that of the general population has been estimated at 10-15% (Brown and Marsden, 1984). However, the nature and characteristics of dementia caused by PD are not well understood, and are certainly less well understood than the more mild cognitive changes caused by PD. It is presumed that dementia results from the increasing severity of the cognitive deficits found in PD combined with the 'spread' of deficits to a wide range of functions described in Section 1.3.3. Two factors have inhibited understanding of the characteristics of the dementia caused by PD. First, in an attempt to exclude patients suffering from neuropathology other than PD, most studies of the cognitive deficits of PD have excluded subjects who were thought to have dementia. This necessarily excludes patients who have dementia caused solely by PD and may lead to studies underestimating the severity of cognitive change in PD. Second, studies that investigate dementia caused by PD seldom carry out post-mortem examination to verify the diagnosis of PD and to rule out the possibility of other neuropathology such as Alzheimer's disease being present. Despite these limitations, attempts have been made to describe the dementia caused by PD and most of these relate to showing that PD causes a 'subcortical' rather than a 'cortical' dementia.

'Subcortical dementia' is a label that has been applied to the disabling cognitive changes caused by diseases such as progressive supranuclear palsy, Huntington's disease and PD. It is thought to differ from 'cortical dementia' - for example, Alzheimer's disease - by the absence of cognitive deficits such as agnosia, apraxia and aphasia that are associated with cortical pathology. In contrast, subcortical dementia is thought to be associated with mood changes, cognitive deficits similar to those seen

after frontal lobe damage, and the clinical impression of 'slowed thought' or bradyphrenia (Darvesh and Freedman, 1996). However, the distinction between cortical and subcortical dementia remains controversial, and 'cortical' deficits may not in fact be absent from patients with dementia caused by PD. Reid et al. (1996) found that dementia in patients with PD resembled 'subcortical dementia' initially but it progressed to include 'cortical' functions and was eventually indistinguishable from dementia caused by Alzheimer's disease. Similarly, Kramer and Duffy (1996) found that the presence of agnosia, aphasia and apraxia were of little use in distinguishing Alzheimer's dementia from dementia caused by PD and normal pressure hydrocephalus. Thus, the classification of dementia caused by PD as a 'subcortical' dementia may be premature. However, neither Reid et al.'s nor Kramer and Duffy's studies carried out post-mortem examination to exclude the possibility that their patients with PD were also suffering from a coexistent, more 'cortical' pathology.

The controversy over the distinction between 'cortical' and 'subcortical' dementia is significant for the consideration of the anatomical substrate of the cognitive deficits of a range of disorders. However, for the purposes of this section it is sufficient to note that PD can cause disabling and relatively global impairments of cognitive function.

1.4 PD and the concept of 'set'

Researchers have often attempted to account for a number of the cognitive and motor deficits of PD in terms of the breakdown of a single mechanism (for example, Brown and Marsden, 1990). The concept of 'set' has frequently been cited and 'inflexibility', 'breakdown' or 'instability' of set has been held responsible for deficits in both cognitive and motor domains (Flowers and Robertson, 1985; Robertson and Flowers, 1990). This section will define the concepts of 'motor set' and 'attentional set' and will review the literature that relates the breakdown of these mechanisms to the deficits seen in PD

1.4.1 Definition of 'set'

The term 'set' is defined as follows in one of the most widely-used modern dictionaries of psychology: "Any condition, disposition or tendency on the part of an organism to respond in a particular manner" (Reber, 1985, p. 689). Thus, 'set' has been used to describe almost any state that is 'internal' to an organism that influences its behaviour in a fashion independent of external cues or stimuli. This usage is clearly too broad to be useful, but it reflects the long history of the term; 'set', with various prefixes such as 'perceptual' or 'motor', has been used to describe a range of psychological phenomena that have little in common. The unregulated use of the term 'set' is not a new phenomenon; Gibson reviewed the use of the term in 1941, finding that "the underlying meaning is indefinite, the terminology chaotic, and the usage by psychologists highly individualistic" (p. 781). Gibson found that 'set' had been used to describe phenomena as diverse as visual hallucinations and Pavlovian conditioning. He concluded that no common meaning could be discerned for the term, but he defined a number of areas of research in which the term 'set' was used to describe a relatively consistent set of psychological phenomena. Two of these areas are the topics of this thesis.

Gibson defined one use of the term 'set' as follows: "an intention to react by making a specific movement, or not so to react (reaction time and conditioning experiments)" (p. 811). This thesis focuses on the 'intention to react' that is seen in reaction time (RT) studies and this 'intention' will be termed 'motor set'. Gibson noted that reaction times (RTs) are faster when subjects are ready to react and anticipating the stimulus. This is demonstrated by the manipulation of foreperiod, that is, the interval between a warning stimulus and an imperative stimulus in an RT task. Both very short and very long foreperiods lead to relatively slow RTs due to the failure of a subject to

be optimally 'ready' for the imperative stimulus. Similarly, RTs tend towards zero when the foreperiod is constant and the imperative stimulus can be anticipated; in all of the cases, the warning signal imperative signal and response requirements are the same, but the internal 'set' of the subject alters RT. Thus, the term 'motor set' is used in this thesis to refer to the speeding of RT that occurs as a consequence of a subject's anticipation of the imperative stimulus and readiness to respond.

The term 'attentional set' will be used to describe another of Gibson's categories. He defines this type of process thus: "a prearoused expectation of stimulus objects, qualities or relations (perception experiments)" (p. 811). Gibson describes work by Külpe (1904) to illustrate this type of set. Külpe presented sets of coloured letters tachistoscopically, having instructed subjects to report a particular aspect of these stimuli. The subjects' perceptions of the stimuli varied greatly according to the instructions that they had been given. For example, if subjects had been instructed to count the number of letters that had been displayed, they often could not report the colour of these letters. Subjects were attending selectively to one aspect of a stimulus, to the detriment of the other attributes of that stimulus. This ability is explicitly tested by a number of widely-used clinical neuropsychological tests such as the WCST and the Stroop test, and is often referred to as 'attentional set'. Studies of the ability of patients with PD to attend selectively to one of a number of possible stimulus attributes will be reviewed in Section 1.6.

1.5 PD and motor set

Patient with PD have frequently been shown to have an RT deficit (see Section 4.0). The cause of this deficit is unknown, but it is widely accepted that peripheral factors such as muscular activation cannot account for such slowing of RTs. It is known that sometimes patients with PD fail to initiate a movement because muscular activity (measured by electromyography - EMG), whilst present, is of insufficient magnitude to move a limb (Hallett, 1990). However the main problem in RT tasks appears to be that the onset of patients EMG activity is later than that of controls, rather than insufficient (Yokochi et al., 1985). This is not caused by slowed information transmission from motor cortex; studies using stimulation of motor cortex have shown that corticospinal conduction times are normal in patients with PD (Thompson et al., 1986). Thus, the RT deficit seen in patients with PD appears to arise from the slowing of some 'central' process that must occur before the execution of the motor program.

A number of lines of evidence implicate a failure of 'motor set' in the RT deficit of patients with PD. Studies have suggested that patients with PD are unable to use 'advance information' about a movement to speed their RTs; that is, when patients are informed of the nature of the response required before it is prompted by the imperative stimulus they are less able than controls to use this information to speed their movements (Bloxham et al., 1984; Sheridan et al., 1987). This can be described as a failure to attain the appropriate 'set' to maximally speed RTs. However, there is contradictory evidence showing that patients can use advance information - for a review, see Section 4.2. It has also been suggested that a failure of a kind of 'set' causes the simple RT deficit seen in patients with PD. Goodrich et al. (1989) suggested that healthy subjects are able to use a form of 'set' that is not accessible to patients with PD to speed their RTs in simple RT tasks. Because response selection is not required in a simple RT task, Goodrich et al. suggested that healthy subjects are able to 'attentionally focus' on the appearance of the imperative stimulus. Goodrich et al. demonstrated the role of this 'focus' or 'set' by requiring patients and controls to perform a secondary task whilst engaged in a simple RT task. The secondary task impaired controls' RTs such that they performed at the same level as patients; in contrast, the secondary task did not impair patients' performance. Goodrich et al. argued that the secondary task disrupted the controls' focusing, or 'set', but that the patients' performance was unchanged because they had never attained a 'set'. Both the studies relating to the use of 'advance information' and 'attentional focusing' suggest that patients with PD are in

some sense unprepared for the imperative signal in an RT task; they have not attained the appropriate 'set'.

The slowing of RT seen amongst patients with PD is often seen as a facet of their 'akinesia', the general problem of movement initiation experienced by patients with PD. A further reason to implicate a failure of motor 'set' in the PD RT deficit comes from clinical observations of akinesia. Set is defined as an internal predisposition or state that is not dependent on external cues, and it has often been noted that patients' akinesia is worst when they are required to move without external cues or prompting. Numerous clinical reports indicate that akinesia can be overcome, at least temporarily, if patients are cued strongly enough (Schwab, 1972). For example, if lines are painted on the floor for patients to step over, patients' gait improves markedly (Martin, 1967). Thus, akinesia is worst when patients have to generate movements 'internally'; a failure of motor 'set' could contribute to akinesia and by extension to the PD RT deficit.

Section 4 of this thesis includes a quantitative review of past studies of RTs in patients with PD. This review aims to establish the conditions under which patients with PD show an RT deficit and also to resolve controversies surrounding the effects of medication on RTs, the consistency of the simple RT deficit and patients' ability to use advance information to speed their movements. Section 5 describes a study of a form of 'motor set' that is seen in RT tasks that use a variable foreperiod.

1.6 PD and attentional set

The hypothesis that patients with PD have impaired attentional set has arisen from the finding that they are impaired in the performance of a number of tasks that require selective attention to one aspect of a complex stimulus. Patients' impairment on the WCST has already been noted (see Section 1.3.1) and patients also show consistent impairments on such tasks as the Stroop test, the Odd-Man-Out test (OMO, Flowers and Robertson, 1985), extradimensional (ED) shifting (Downes et al., 1989) and the Embedded Figures test. Some of these tests incorporate 'shifts' of attentional set and some have a problem-solving element (for example, the WCST). Other tasks, such as the Stroop test, only test the ability to form and maintain an attentional set. However, all require subjects to attend to one aspect or attribute of a complex stimulus on the basis of an internal 'disposition' rather than by using information in the stimulus. This section will review the literature describing the performance of patients with PD on these tests. Tests that index the basic ability to form and maintain a set will be reviewed first, followed by an account of the 'classic' attentional set-shifting deficit of patients with PD, the WCST deficit. Finally, studies that attempt to clarify the nature of the WCST deficit will be described.

1.6.1 Formation and maintenance of set

Figure-ground perception

There is evidence that some patients with PD have problems in the basic ability to form and use an attentional set. A number of studies show that patients with PD do not perform optimally when required simply to attend to one aspect or attribute of a stimulus whilst ignoring other competing aspects. Two studies show that some patients have problems distinguishing 'figure' from 'ground' in perceptual displays. Talland (1962) showed that patients were less able than controls to control their perception of the Necker cube. When patients attempted to sustain one view of the Necker cube their perceptions spontaneously reverted to the other 'view' more often than controls. Also, patients were less able than controls to voluntarily switch between the two views of the cube. A deficit has also been reported on the Embedded Figures test (Flowers and Robertson, 1995). This test requires subjects to 'pick out' the outline of a simple shape that is embedded in a more complex figure. As such, the test requires selective attention and the ability to ignore irrelevant aspects of a stimulus. Flowers and Robertson (1995) found a deficit on this test in moderately and severely impaired patients with PD,

though not in mildly impaired patients. This test was also included in Taylor et al.'s (1986) study, which also reported no deficit in patients with early PD. Thus, it appears that the ability consistently to distinguish figure from ground is impaired only in patients with more advanced PD.

Stroop effect

Two paradigms have been used to test patients' ability to attend selectively to one of two perceptual 'dimensions'. The first is the widely-used Stroop test, the second a sophisticated visual attention paradigm devised by Maddox et al. (1996). Both of these tests tax subjects' ability to attend to one dimension in the presence of another competing dimension; neither (in their original forms) test subjects' ability to 'shift' or 'switch' attentional set. The Stroop task tests subjects' ability to attend to a dimension in the face of strong interference. Subjects are required to report the colour of the ink in which a number of words are printed. The 'Stroop effect' is seen when two conditions are compared (though the exact procedure often varies - Lezak, 1995); in one condition, the colour word and the ink colour are congruous (e.g. the word 'RED' printed in red ink), in the other condition they are incongruous (e.g. 'BLUE' printed in yellow ink). The Stroop effect is the increase in either latency or errors seen in the 'incongruous' condition compared to the 'congruous' condition. The increase in difficulty is caused by the 'automatic' reading of the colour word and the interference of this with naming the ink colour. Thus, the Stroop test can be seen as a strong test of selective attention to the dimension 'ink colour'.

Patients with PD have frequently been shown to have a deficit on the Stroop task; that is, they show an increased 'Stroop effect', measured either by increased latency or errors (Brown and Marsden, 1988; Henik et al., 1993; Hietanen and Teräväinen, 1986; Meco et al., 1996; Pillon et al., 1996; Stam et al., 1993). Only two studies have shown intact Stroop performance in patients with PD (Cools et al., 1984; Van Spaendonck et al., 1995). Van Spaendonck et al.'s failure to find an effect may be explained by subject characteristics; whilst the mean age for the patient groups in which deficits were found was 64.1 years, the mean age of Van Spaendonck et al.'s patient group was 53.9 years implying that a deficit was not seen due to the patient group's early age of onset (Reid et al., 1989). There is no obvious reason for the failure of Cools et al.'s (1984) study to find a deficit on the Stroop test but this single study does not detract from the overall finding of a deficit on the Stroop test in patients with PD.

Some studies have attempted to investigate the mechanisms underlying the parkinsonian Stroop deficit. Brown and Marsden (1988, 1991) have suggested that

patients with PD have limited attentional 'resources' and that a Stroop deficit is present when patients' 'resources' are exceeded. They have supported this argument by showing that the presence of cues ameliorates the Stroop deficit seen in patients with PD and claiming that this is due to the cues reducing the 'resource requirements' of the task (Brown and Marsden, 1988). They have also showed that patients' (but not controls') performance of the Stroop task is impaired by the presence of some secondary tasks and have argued that this is because the secondary task increases 'resource demands' beyond the patients' limits. This account of the Stroop deficit seen in patients with PD is attractive, but cannot account for all of the data. A study by Henik et al. (1993) shows that an alteration of attentional function is more likely to cause the Stroop deficit than a nonspecific 'reduction of resources'. Henik et al. compared the speed of ink colour naming when the word was congruent with the ink colour (e.g. 'RED' in red ink) or irrelevant to the ink colour (e.g. 'TIGER' or 'XXXX' in red ink). The colours of the congruent words were named more quickly than those of the irrelevant words for all subjects, but this benefit of congruence was greater for patients with PD than for controls. This facilitatory effect cannot be accounted for in terms of 'resource depletion' and appears to arise from a 'beneficial' inability to inhibit word reading in patients with PD.

The Maddox paradigm

A recent study by Maddox et al. (1996) provides further evidence that some patients with PD have deficits of selective attention. Maddox et al. displayed either horizontal or vertical lines, or both, and required subjects to categorise these lines by their length (e.g. 'long' or 'short') according to a simple criterion that was clearly displayed throughout the study. Patients with PD could successfully categorise vertical lines that were displayed on their own and also performed normally when required to categorise according to the relative lengths of the horizontal and vertical lines. However, when patients were required to categorise according to the length of the vertical line and to ignore the horizontal line, a proportion of patients with PD did not perform optimally. This is a clear demonstration of an inability to attend selectively in the presence of competing stimuli which is particularly convincing due to the simple nature of the task.

To summarise, Maddox et al. (1996) found a deficit of selective attention in patients with PD that had also been indicated by studies using the Stroop test. However, both of these paradigms demonstrate deficits under extreme experimental conditions. The Stroop test is a curiosity because few perceptual dimensions interfere as competitively as written words, and Maddox et al.'s test is so simple to perform that

specialised mathematical modeling must be used for each subject's data to detect a deficit. In contrast, the 'paradigmatic' attentional set-shifting deficit in PD is usually considered to be poor performance on a test that requires repeated shifting between simple dimensions - the WCST

1.6.2 Shifting set - studies of the WCST

The WCST is a widely-used test in clinical neuropsychology, favoured for its sensitivity to frontal lobe damage (Milner, 1963). To carry out this test, subjects are given a pack of 'response' cards and must place them next to one of four 'stimulus' cards (see Figure 1.1) according to a rule. Subjects select the appropriate 'stimulus' card for their individual 'response' card on the basis of one of three rules; subjects must either match colour, shapes or number of shapes. For example, when sorting to the rule 'colour', a response card showing a yellow shape must be placed next to stimulus card with a yellow shape. If the rule is 'number', a response card with two shapes on it must be placed next to the stimulus card that has two shapes. Further examples of this sorting process can be seen in Figure 1.1.

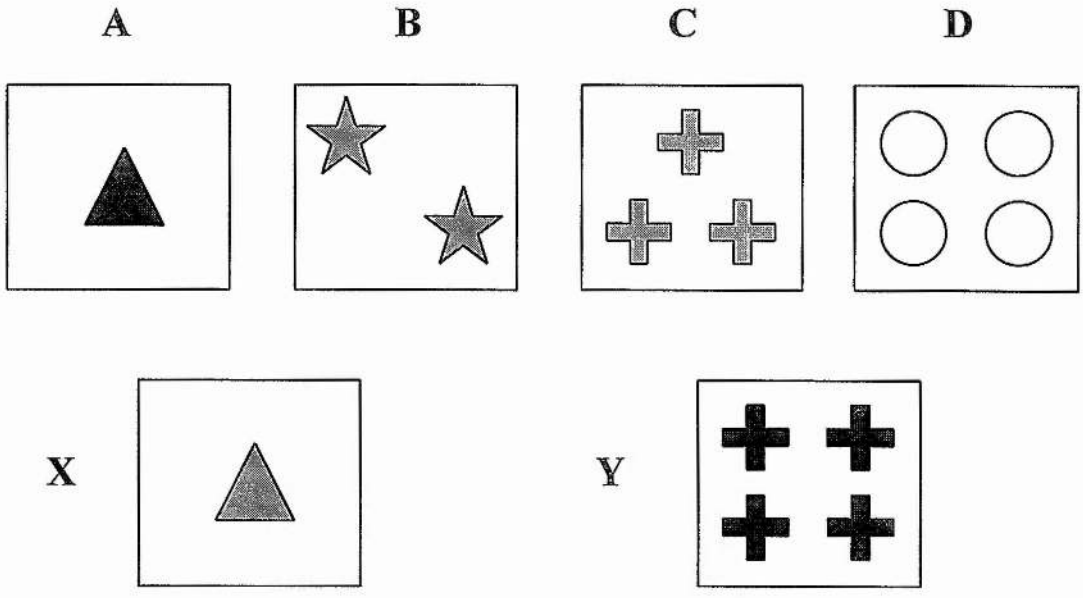


Figure 1.1: This figure depicts the four 'stimulus cards' used in the WCST (A to D) and two sample response cards (X and Y). If the rule is 'shape', then card X must be placed next to A, and card Y next to C. Alternatively, if the rule is 'colour' (here shown in gray tones) X must be placed next to B and Y next to A.

However, the exact procedure by which the WCST is administered varies greatly between the two versions of the test. Milner (1963) developed the original version of the WCST, but a simpler version was created by Nelson (1976) who noted that Milner's procedure was found to be too difficult and frequently aversive by many patients. The principal differences between the two tests are as follows:

(a) Selection of rules: In Milner's version, the 'rule' that subjects must use is decided by the tester and subjects must work it out on the basis of feedback ('correct' / 'wrong') given by the tester. In the Nelson version of the WCST subjects are invited to select any rule and then to follow it.

(b) Rule shifts: In Milner's version, the rule is changed without informing the subject after ten consecutive correct responses to the previous rule. In contrast, subjects are invited to 'find another rule' after six consecutive correct responses in Nelson's version.

(c) Feedback: Ambiguous feedback is possible in Milner's version of the WCST. For example, if a subject makes a correct response by placing a response card depicting two red triangles next to a stimulus card showing one red triangle, then the subjects could conclude that either of the rules 'colour' or 'shape' are correct. In Nelson's version, response cards that can yield ambiguous feedback are not used.

Both versions of the WCST yield a number of performance measures; the number of rules successfully used (or 'categories achieved') is recorded, and the number of different types of error are also recorded.

A review of the literature reveals 17 studies that compared the performance of patients with PD and controls on the WCST and reported either comparison statistics or raw scores. Different studies used different versions of the WCST. Milner's and Nelson's versions of the WCST are quite different and there is some consensus that they should not be regarded as testing the same abilities (DeZubicaray and Ashton, 1996). As a consequence, studies of PD that used the Milner version are reviewed separately from those that used the Nelson version.

Milner version - review

Nine studies were found that used the Milner version of the WCST. The only outcome measure reported by all of these studies was the 'number of categories achieved', so studies are reviewed in terms of this measure. Patients achieved fewer categories than controls in six out of nine studies (Bowen et al., 1975; Lichter et al., 1988; Meco et al., 1996; Taylor et al., 1986, 1987; Tröster et al., 1995). Three studies found that patients performed normally (Cooper et al., 1991, 1992; Dalrymple-Alford et al., 1994) but in all

of these studies the patients tested had mild PD. Cooper et al. (1991) studies newly-diagnosed, unmedicated patients with PD and their 1992 study tested the same cohort of patients four month later whilst medicated. Similarly, Dalrymple-Alford et al. (1994) found no WCST deficit in a group of patients described as having "mild PD of relatively recent onset" (p. 361). It would be desirable to review all nine studies in terms of a more sensitive measure than 'categories', such as 'number of perseverative errors', but only five out of nine studies report this measure and even fewer report 'non-perseverative' errors. However, the 'categories' measure still shows a PD deficit in the majority of studies and in all studies of more advanced PD.

Nelson version - review

Of the eight studies found that used the Nelson version of the WCST, five reported reduced category achievement in patients with PD (Brown and Marsden, 1988; Gotham et al., 1988; Lees and Smith, 1983; Partiot et al., 1996 (my calculations); Pillon et al., 1986). Three studies report no deficits in patients with PD (Canavan et al., 1989; Pillon et al., 1996, 1997) but two of these studies used patients with mild PD, a group that has already been shown to sometimes perform normally on the Milner version of the WCST (see above). Pillon et al. (1997) used 'de novo' unmedicated patients and Canavan et al. (1989) described their patients as being in the 'early stages' of PD. However, Pillon et al.'s (1996) patients were medicated and had moderate PD, so disease severity cannot account for a failure to find a WCST deficit in category achievement in this group. Although these patients achieved the same number of categories as controls, their performance was not normal as they made more perseverative errors than controls. Thus it appears that moderately or severely impaired patients have a consistent deficit in some aspect of performance of the Nelson version of the WCST. It is interesting to note that although the tasks differ substantially, the Nelson and Milner versions of the WCST seem equally sensitive to PD and equally insensitive to mild PD.

This review has confirmed the widely-held opinion that a deficit in WCST performance is characteristic of patients with PD, although it appears that some patients with early PD may be unimpaired. The WCST deficit is one of the most consistent cognitive deficits that has been found in PD. As noted in Section 1.3.2, it appears that this deficit increases in severity with disease progression (Caparros-Lefebvre et al., 1995). It is also of interest that the WCST deficit seen in patients with PD is not ameliorated by levodopa treatment. Bowen et al. (1975) found that a group of levodopa-treated patients achieved no more categories on the WCST than untreated patients despite having better disease status due to medication. Gotham et al. (1988), Kulisevsky

et al. (1996) and Starkstein et al. (1989) have all assessed WCST performance in patients undergoing withdrawal of levodopa medication. Neither Gotham et al. or Starkstein et al. found any differences in WCST performance 'on' or 'off' medication. Kulisevsky et al's findings were similar, but they also noted a transient *decrease* in performance one hour after the administration of levodopa in a group of patients who showed a 'wearing-off' response to levodopa. Thus, it appears that patients' WCST deficit is not eliminated by medication and may even be worsened by medication in patients who have lost their smooth response to levodopa.

Although the WCST deficit of patients with PD is well-established, the impaired cognitive mechanism that causes this deficit has not been characterised. Successful performance of the WCST requires a range of cognitive abilities; for example, Downes et al. (1989) have argued that subjects must be able to use a matching-to-sample principle, learn conditional visuospatial contingencies and shift attentional set. Analysis of WCST error scores has indicated little about the cognitive abilities involved in this task other than the fact that more than one ability is required. Factor analysis of WCST performance has revealed two factors (Greve et al. 1997; Paolo et al., 1995). The first factor has been termed "conceptualisation", "problem-solving" (Paolo et al.) and "undifferentiated executive function" (Greve et al.)' the second had been termed "failure to maintain set" (Paolo et al.) and "attentional function" (Greve et al.). These broad and ill-defined definitions are clearly unhelpful for explaining the deficit seen in patients with PD and it seems unlikely that further analysis of the WCST in its classic form will clarify the cognitive abilities that underly successful performance. Instead, a number of researchers have elected to design new tasks that are aimed at pinpointing the aspect of the WCST that challenges patients with PD. These tasks are reviewed in the following sections.

1.6.3 The Odd-Man-Out (OMO) test

Many studies have attempted to replicate and extend the PD WCST deficit using different paradigms, or have devised tests with the intention of isolating the particular cognitive demand that causes the WCST deficit in patients with PD. However, these studies have often taken the central ability underlying successful WCST performance to be 'set-shifting', which is often loosely defined as the ability to switch from almost any behaviour to another behaviour. As such, these studies have often used paradigms that do not test attentional set. An example of this is a series of studies by Van Spaendonck et al. (1993, 1995, 1996). This group has attempted to extend the WCST 'set-shifting'

deficit by using structurally similar card-sorting tasks that require subjects to categorise different types of stimulus material. For example, Van Spaendonck et al. have created the 'Animals Sorting Test'. In this test, subjects are given a pack of cards with animal names on them, such as 'giraffe', and are required to sort these onto two piles on the basis of a rule which is subsequently changed. An example of such a rule might be 'native vs. exotic' or alternatively 'herbivore vs. carnivore'. This may be an interesting test, but is does not test visual selective attention to one aspect of a complex stimulus. There are many other examples of studies that claim to test 'set-shifting' but test abilities unrelated to attentional set (e.g. verbal fluency - Downes et al., 1993; priming - McDonald et al. 1996; problem-solving - Cronin-Golomb et al., 1994). However, two paradigms have been developed that clearly test visual selective attention and each had been used in a number of studies of PD. These tests are the Odd-Man-Out (OMO) test (Flowers and Robertson, 1985) and the attentional shifting test from the CANTAB battery (Roberts et al., 1988).

Flowers and Robertson (1985) state that they developed the OMO test due to dissatisfaction with the Nelson version of the WCST. However, the demands of the OMO test are quite similar to those of Nelson's WCST. Subjects are shown a card that depicts three shapes; one shape differs from the other two in terms of one dimension (such as 'size') and another shape differs from the others in a different dimension (such as 'shape'). Subjects are invited to 'pick the odd one out' and are allowed to select a rule relating to either perceptual dimension. Subjects must then select the odd one out according to the same rule for new cards until that pack is finished, when a new pack is started and subjects are required to choose a different rule. The main difference between the OMO test and Nelson's WCST (apart from the exact nature of the sorting task) is that the requirement to keep sorting to the end of a pack of cards allows assessment of maintenance of set. Indeed, Flowers and Robertson (1985) found that patients with PD were impaired in this 'set-maintenance' rather than 'set-shifting' between rules. However, this finding has not been replicated. Both Richards et al. (1993) and Pillon et al. (1996) found shifting deficits on the OMO test in patients with PD. Pillon et al.'s (1997) study even contradicts these two studies, showing normal performance in patients with PD. However, this study tested 'de novo' unmedicated patients who were also shown to be unimpaired in the performance of the WCST.

The principal problem with the OMO test is not the inconsistency of the results that have been found with it, but that the mechanisms underlying the performance of this task are no clearer than those required by the WCST. Another paradigm is

noteworthy for allowing the assessment of visual selective attention alone, uncontaminated by other cognitive demands. This is the attentional set-shifting component of the CANTAB battery, described below.

1.6.4 Extradimensional (ED) shifting

The completion of a rule shift on the WCST requires subjects to shift their attention from one perceptual dimension (for example, colours) to another (for example, shape). An ED shift is a relatively pure measure of this ability to shift attentional set; it is also well researched, and comprehensible in terms of learning theory. Roberts et al. (1988) have devised a test that includes a test of ED shifting as well as other abilities that are required for the performance of the WCST. This test can be completed, with some modification, by both humans and marmosets and the methodology used has benefited from much previous work (see Slamecka, 1968).

Roberts et al.'s paradigm involves a number of different tests of discrimination learning that increase in difficulty as the test progresses. Subjects move from one stage of the test to the next when they reach a criterion of six consecutive correct responses; if they do not achieve this criterion after 50 trials the test is terminated. The initial stage of the test is a simple two-choice visual discrimination task where subjects are required to work out which of two black shapes is 'correct' on the basis of feedback. A reversal stage follows this. After the reversal stage, a new 'perceptual dimension' is introduced; white lines are superimposed on the black shapes, creating a compound stimulus. However, the black shapes continue to be relevant for the purposes of discrimination learning. The two most significant stages of this test take place whilst both 'dimensions' are present, and are termed the intradimensional (ID) and extradimensional (ED) shifts. In both of these stages, black shapes and white lines are present before and after the shift, but the individual shapes and lines change in form (a 'total change' design - Slamecka, 1968). At the ID shift, the stimulus containing one of the new black shapes will be correct after the shift. In contrast, after an ED shift subjects must cease responding to the black shapes that were relevant before the shift and instead attend to the white lines. This shift of attention from one dimension to the other is thought to parallel the rule shift of the WCST. The relevant comparison that isolates shifts of attentional set in the Roberts et al. paradigm is the comparison between ID and ED shifting performance. After both ID and ED shifts, subjects are confronted with novel stimuli and must work out the new correct response on the basis of feedback. The only difference between the two shifts is the influence of subjects' attentional set; other

factors such as the requirement to carry out hypothesis testing are equivalent in the two shifts. If subjects were selectively attending to black shapes before the shifts, this aids ID shifting as black shapes remain relevant, but retards ED shifting as black shapes are now irrelevant and must be ignored. As a consequence, both healthy humans and marmosets acquire ID shifts more rapidly than ED shifts (Roberts et al. 1988).

Patients with PD show deficits of ED shifting with spared or only mildly impaired ID shifting (Downes et al., 1989; Lange et al., 1992; Owen et al., 1992; Robbins et al., 1994). That is, patients are able to maintain attention to a dimension, but are impaired in switching their attention between dimensions. ED shift deficits have been found in patients at all stages of PD and in both medicated and unmedicated patients. No study has yet found intact performance of the standard ED shift in patients with PD; in contrast to the WCST, ED-shift deficits are found in unmedicated patients with early PD (Downes et al., 1989; Owen et al., 1992). Although only four studies have found ED shift deficits in patients with PD the inclusion of more than one group of patients in some studies means that deficits have been found in a total of seven separate groups of patients.

Like the WCST deficit, the ED shift deficit appears to increase in severity with disease progression; Owen et al.'s (1992) 'medicated - severe' group performed more badly at the ED shift than the 'medicated - mild' and 'unmedicated' groups. It also appears that levodopa medication does not reduce the ED shift deficit. Downes et al. (1989) and Owen et al. (1992) compared unmedicated patients with medicated patient with mild PD, but found no difference in ED shift performance between these groups. However, in both of these studies the 'medicated - mild' groups had more advanced PD than the unmedicated groups, so it is possible that any cognitive improvement caused by levodopa might have been 'canceled out' by disease progression. The best evidence relating to the effects of medication comes from a study by Lange et al. (1992) in which patients performed ED shifts whilst medicated and after 13 hours of medication withdrawal. The same number of patients failed to acquire the ED shift when on or off medication.

This brief review has shown that an ED shift deficit is present even at the earliest stages of PD when motor symptoms are minimal. The deficit increases in severity with disease progression and is not ameliorated by levodopa. The presence of intact ID shift shows that the cognitive problem in patients is one of visual selective attention only. Studies 1 to 4 of this thesis (Sections 2.0 and 3.0) are aimed at clarifying

what causes the ED shift deficit, which is one of the most consistent cognitive deficits seen in PD.

1.7 Anatomical substrates of the cognitive and motor deficits of PD

The cognitive and motor deficits described above are likely to have different neural substrates, as deficits in cognitive and motor domains are of a different magnitude and are sometimes uncorrelated. This section aims to provide a brief review of current theories relating to the anatomy underlying the cognitive and motor deficits of PD. In order to do this, it is necessary to describe both the neuropathology of PD and the anatomy and connectivity of the basal ganglia. The principal neuropathological lesion seen in PD is a loss of cells in the substantia nigra pars compacta (SNc) that leads to a massive loss of dopamine in the striatum. However, this cannot be seen as a discrete 'lesion' and it would be incorrect to attribute the deficits seen in PD solely to striatal or nigrostriatal pathology. The striatum is part of a number of larger neural circuits, or 'loops', that involve other areas of the basal ganglia as well as thalamus and cortex; disruption of striatal function alters the function and 'balance' of these loops. Thus, the ultimate functional consequences of striatal dopamine loss can result from disruption of one of the structures in a 'loop' that is physically distant from the damaged striatum. The first section of this review (1.7.2) will describe the neuropathology of PD, emphasising the points at which this pathology affects the basal ganglia 'loops'. This will be followed by a description of the anatomical connections within these loops (Section 1.7.3). Contemporary theories of basal ganglia function cite an imbalance in these intrinsic circuits as causing the symptoms of PD. These theories will be summarised in Section 1.7.3, as they provide the rationale for the surgical treatment of PD that is described in Section 6.0.

1.7.1 A note on the cognitive deficits of PD

The following sections assume that loss of dopamine in the putamen is the principal cause of the movement deficits of PD, particularly bradykinesia. This is uncontroversial. Levodopa is known to have its beneficial effects on movement symptoms by acting in the putamen (Pearce, 1992) and a number of imaging studies have found a correlation between putaminal dopamine metabolism and motor status in patients with PD (e.g. Holthoff-Detto et al., 1997; Morrish et al., 1995). Some movement symptoms of PD may have extrastriatal origins - for example, there may be a cerebellar component to tremor (Caparros-Lefebvre et al., 1994; Deiber et al., 1993) - but the role of the putamen in the generation of bradykinesia is not disputed. However, the following sections will also attribute the attentional set-shifting deficits of PD (and

many other cognitive deficits) to loss of dopamine in the head of the caudate nucleus. In contrast to the consensus surrounding motor symptoms, the neurochemical and neuroanatomical basis of the cognitive deficits of PD is a topic of much debate (see Dubois and Pillon, 1995). PD causes disruption of a range of neurotransmitter systems (see Section 1.7.2) and investigators have often asserted that the cognitive deficits of PD are extrastriatal or non-dopaminergic in origin (e.g. Cooper et al., 1991). Thus, the association of cognitive deficits with caudate dysfunction must be justified.

Some of the cognitive deficits of PD are almost certainly independent of caudate dopamine loss. For example, Stern et al. (1984) and Mayeux et al. (1987) have found that deficits of vigilance in PD (as measured by the Continuous Performance Test) are associated with CSF levels of noradrenaline metabolites but not dopamine metabolites. However, there is also substantial evidence that many of the cognitive deficits of PD are a consequence of caudate dysfunction or of disruption of the cortical projection sites of the caudate nucleus. The caudate was originally implicated in the cognitive deficits of PD because of the resemblance of these deficits to those seen after frontal lobe damage (Taylor et al., 1986). The caudate is part of a loop that receives projections from and projects to the dorsolateral prefrontal cortex (DLPFC), and animal studies have shown that lesions of the head of the caudate and its projection sites in frontal cortex can have similar behavioural consequences (Battig et al., 1962). The caudate is also involved in the performance of cognitive tasks that have been shown to be sensitive to the effects of PD. Performance of the Tower of London task in humans and a spatial working memory task in monkeys has been shown to be associated with caudate metabolism (Owen et al., 1996a; Levy et al., 1997a, respectively). Most directly, a recent imaging study has shown that performance of a memory task is correlated with caudate dopamine metabolism in patients with PD (Holthoff-Detto et al., 1997).

There is no direct evidence that links deficits of attentional set-shifting to caudate dysfunction. However, indirect evidence is plentiful. ED-shifting deficits are present in patients with very early PD, when neuropathology is principally confined to dopamine loss in the striatum (see Section 1.3.1). Damage to the frontal lobes, particularly to the lateral frontal areas that are reciprocally connected to the caudate, results in ED-shift deficits (Dias et al., 1996; Owen et al., 1990). However, it is possible that the 'frontal-like' attentional set-shifting deficits seen in patients with PD are due to the loss of intrinsic dopamine in the frontal cortex (see Section 1.7.2) rather than to dopamine loss in the caudate. This explanation was ruled out by a study by Roberts et

al. (1994) which showed that dopamine depletion of the frontal cortex in monkeys *enhanced* ED shift performance. Thus, it seems most likely that the ED shift deficits seen in patients with PD are due to caudate dysfunction. Although the ED shift deficits seen in patients with PD and those with frontal damage are not identical (see Owen et al., 1993b and Section 2.4) there is more evidence to connect the ED shift deficits seen in PD to dopamine depletion of the caudate nucleus than to depletion of any other area or neurotransmitter system. The following sections proceed from this premise.

1.7.2 Neuropathology of PD

PD causes both structural changes and the loss of a range of neurotransmitters; Agid et al. (1990) provide a detailed review of these changes. This section will briefly summarise the nature and location of the principal neurochemical changes seen in PD. Loss of striatal dopamine has been cited above as the most important pathological change of PD, and this will be described first. The functional significance of changes to other neurotransmitter systems is less clear, and these will be described subsequently.

Dopamine systems - Nigrostriatal

Pigmented cells are lost in the SNc of patients with PD, which leads to degeneration of the nigrostriatal bundle. Dopaminergic innervation of the striatum is lost, and striatal dopamine levels are drastically reduced. The degeneration of the SNc and striatum varies widely between individuals (Kish et al., 1988), but often follows a similar general pattern; the ventrolateral tier of SNc cells is affected first, leading to depletion of dopamine in the posterior putamen that subsequently spreads rostrally to into the anterior putamen and the caudate nucleus (Sawle, 1995). Dopamine depletion has been found to be around 85-90% in the putamen and 75% in the caudate (Marsden, 1992). It is widely accepted that motor symptoms do not appear until putaminal dopamine depletion reaches around 80% and there is thought to be a 'preclinical' period in which dopamine levels are falling but no motor symptoms are apparent (Morrish et al., 1995).

Mesocorticolimbic

The mesocorticolimbic dopamine system arises from the ventral tegmental area (VTA) and projects to a range of subcortical structures, including the nucleus accumbens, hypothalamus, amygdala and hippocampus. The VTA also innervates frontal, cingulate and entorhinal cortices, and the parolfactory gyrus. Tyrosine hydroxylase (a marker of dopamine synthesis) is reduced by 50-70% in the VTA (Agid et al., 1990) and dopamine depletion in the nucleus accumbens can reach 75% (Marsden, 1992).

Other dopaminergic systems

Loss of dopamine is seen in the globus pallidus and there is also dopamine depletion in the hypothalamus that cannot be accounted for by dysfunction of the mesocorticolimbic pathway. It is thought that hypothalamic dopamine depletion is due either to loss of intrinsic dopamine degeneration of nigro-hypothalamic fibres (Agid et al., 1990). Dopamine levels are also halved in two brainstem structures - the locus coeruleus and the area postrema.

Noradrenaline

Pigmented cells are lost in the locus coeruleus, leading to a degeneration of dorsal ascending noradrenergic pathways. Cortical and subcortical noradrenaline and its metabolites are depleted.

Acetylcholine

There is widespread loss of cholinergic neurons in the Nucleus Basalis of Meynert in PD, which leads to a loss of acetylcholine in frontal, temporal and entorhinal cortices as well as the hippocampus. Cholinergic cell loss in the NBM has been shown to be as severe as that seen in Alzheimer's disease (Candy et al., 1983). 40% of cholinergic neurons are also lost in the pedunculopotine tegmental nucleus (PPTG). In contrast, levels of choline acetyltransferase (CAT, a marker of cholinergic neurons) are normal in the caudate nucleus and SNc.

Serotonin

Serotonin and its metabolite are reduced in the basal ganglia, hypothalamus, hippocampus and frontal cortex of patients with PD. This results from loss of cells in the raphe nuclei and ascending serotonergic pathways.

GABA

The impact of PD on GABAergic pathways is unclear, as the marker for GABA (glutamic acid decarboxylase, GAD) is affected by a number of conditions that are associated with prolonged terminal illness, such as anoxia. GAD levels are reduced in many areas of the brains of patients with PD at post-mortem, but in contrast GABA receptors are present in normal numbers in all areas but the SN (Agid et al., 1990).

Neuropeptides

Five neuropeptides have been consistently found to be depleted in patients with PD - cholecystokinin, leu- and met-enkephalin, somatostatin and substance P.

1.7.3 Functional neuroanatomy of PD

This section describes the impact of striatal pathology described above on the function of the basal ganglia in PD. Much contemporary research on the functional neuroanatomy of the basal ganglia has been guided by models proposed by Alexander et al. (1986) and Alexander and Crutcher (1990). These models describe the anatomy of the basal ganglia at two levels of organisation. The first model (Alexander et al., 1986) describes the extrinsic connections of the basal ganglia and introduces the concept of cortico-basal ganglia-thalamocortical 'loops'. Alexander et al. (1986) proposed that there are a number of 'loops' passing from and to cortical areas through the basal ganglia; two of these 'loops' are thought to subserve motor and cognitive ('executive') processes respectively and are described in detail in Section 1.7.3.1. The second model (Alexander and Crutcher, 1990) describes the connectivity of structures within the basal ganglia. Alexander and Crutcher (1990) proposed that there is a common neuroanatomical 'layout' within the basal ganglia that is shared by all of the 'loops' described above. Using data from neurophysiological studies of parkinsonian monkeys, Alexander and Crutcher (1990) have created a model of how the weights of these intrinsic basal ganglia connections are changed by PD and how these changes generate the symptoms of PD. This model will be described in Section 1.7.3.2.

Alexander and Crutcher's (1990) model has been extremely influential, but it has been updated by subsequent research. The principal changes to the model are described in Section 1.7.3.3. However, the majority of the recent research on basal ganglia neuroanatomy has been aimed at refining understanding of the 'motor' loop that passes through the putamen. The 'cognitive' loop is much less well understood, and research clarifying its function will be summarised in Section 1.7.3.4.

1.7.3.1 Extrinsic connections of the basal ganglia - the 'loops'

It is well-established that the basal ganglia receive projections from a wide range of cortical areas and that the basal ganglia also project to cortex via the thalamus (Parent, 1990). Alexander et al.'s (1986) widely-cited model emphasises that these cortical connections are arranged as a 'loop'. The basal ganglia are thought to receive connections from a range of cortical areas and to project to a more restricted subset of the same cortical areas. Within the basal ganglia, inputs from cortex are received by the striatum, and efferents leave the basal ganglia via the internal section of the globus pallidus (GPi) and the substantia nigra pars reticulata (SNr) which are regarded as

output nuclei. The processing that occurs between the striatum and the output nuclei is the topic of the following section (1.7.3.2). Overall, the loops are thought to proceed thus: cortex - striatum - GPi/SNr - thalamus - cortex.

The 'motor' loop is the most clearly understood of the basal ganglia loops and it has been the focus of much research aimed at clarifying the causes of the motor symptoms of PD. The motor loop 'begins' with a range of cortical areas, including the supplementary motor area (SMA), motor cortex (MC), somatosensory cortex (SC) and the arcuate premotor area (APA, see Figure 1.2). These all project to the putamen which, as noted above, is the principal site of dopamine depletion in PD. The motor loop passes from the putamen to the 'output' nuclei of the basal ganglia, passing though the ventrolateral GPi and caudolateral SNr. Projections are then sent back to the SMA via the ventrolateral nucleus of the thalamus.

The basal ganglia loop that involves the DLPFC and passes through the caudate nucleus is known as the 'dorsolateral prefrontal' loop (Alexander et al., 1986) and has been most closely associated with cognitive processes. This loop will be referred to as the 'cognitive' loop for brevity; this is not intended to imply that other basal ganglia loops are devoid of cognitive function. For example, the anterior cingulate and lateral orbitofrontal loops may well subserve some cognitive processes, just as the oculomotor loop serves 'motor' processes relating to eye movements. The 'cognitive' loop arises from three cortical areas - the dorsolateral prefrontal cortex (DLPFC), posterior parietal cortex (PPC) and APA (see Figure 1.2). These areas project to the dorsolateral sector of the head of the caudate nucleus. As for the motor loop, the cognitive loop leaves the basal ganglia through the GPi/SNr and projects back to the DLPFC via the thalamus. However, it is a feature of Alexander et al.'s (1986) model that the cognitive and motor loops are thought not to overlap in these areas (see Figure 1.2). Cognitive and motor loops also pass through different areas of the striatum, but both the caudate and the putamen are dopamine-depleted in PD so both loops may be affected despite their anatomical segregation. The majority of recent research into PD focuses on the consequences of striatal dopamine depletion for information processing within the basal ganglia. The following section describes Alexander and Crutcher's (1990) and DeLong's (1990) theories relating to how striatal pathology generates the motor symptoms of PD.

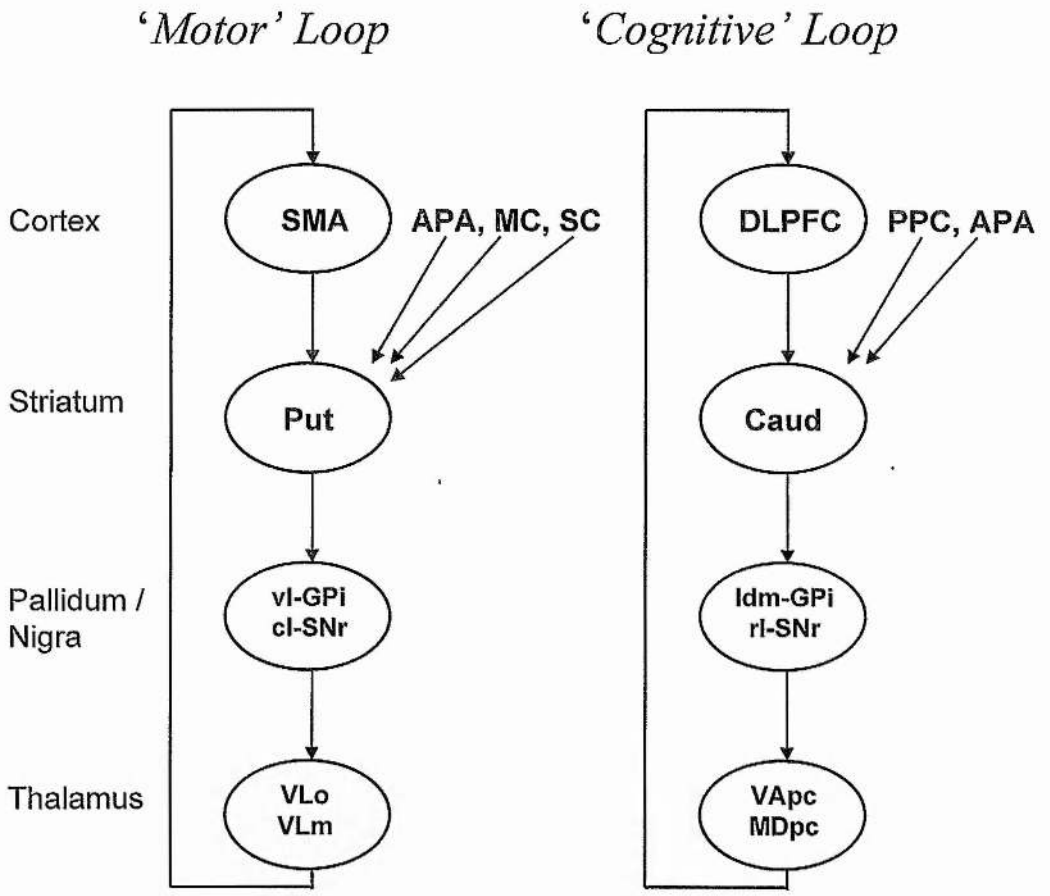


Figure 1.2: The organisation of the 'motor' and 'cognitive' loops. Abbreviations are as follows: SMA - supplementary motor area; APA - arcuate premotor area; MC - motor cortex; SC - somatosensory cortex; Put - putamen; vl-GPi - ventrolateral GPi; cl-SNr - caudolateral SNr; VLo - ventrolateral nucleus of the thalamus pars oralis; VLm - ventrolateral nucleus of the thalamus pars medialis; DLPFC - dorsolateral prefrontal cortex; PPC - posterior parietal cortex; Caud - caudate nucleus; ldm-GPi - lateral dorsomedial GPi; rl-SNr - rostrolateral SNr; VApc - ventral anterior nucleus of the thalamus, pars parvocellularis; MDpc - mediodorsal nucleus of the thalamus, pars parvocellularis. Figure adapted from Alexander et al. (1986).

1.7.3.2 Intrinsic basal ganglia circuits

Figure 1.3 is an adaptation of Alexander and Crutcher's (1990) diagram of the internal connections of the basal ganglia. According to Alexander and Crutcher, this organisation is common to all of the basal ganglia 'loops'. Parts of this diagram have already been described above; for example, the connection of the cortex to the striatum

and the efferent projections from GPi/SNr to cortex via the thalamus. The diagram also depicts the dopaminergic innervation of the striatum by SNc.

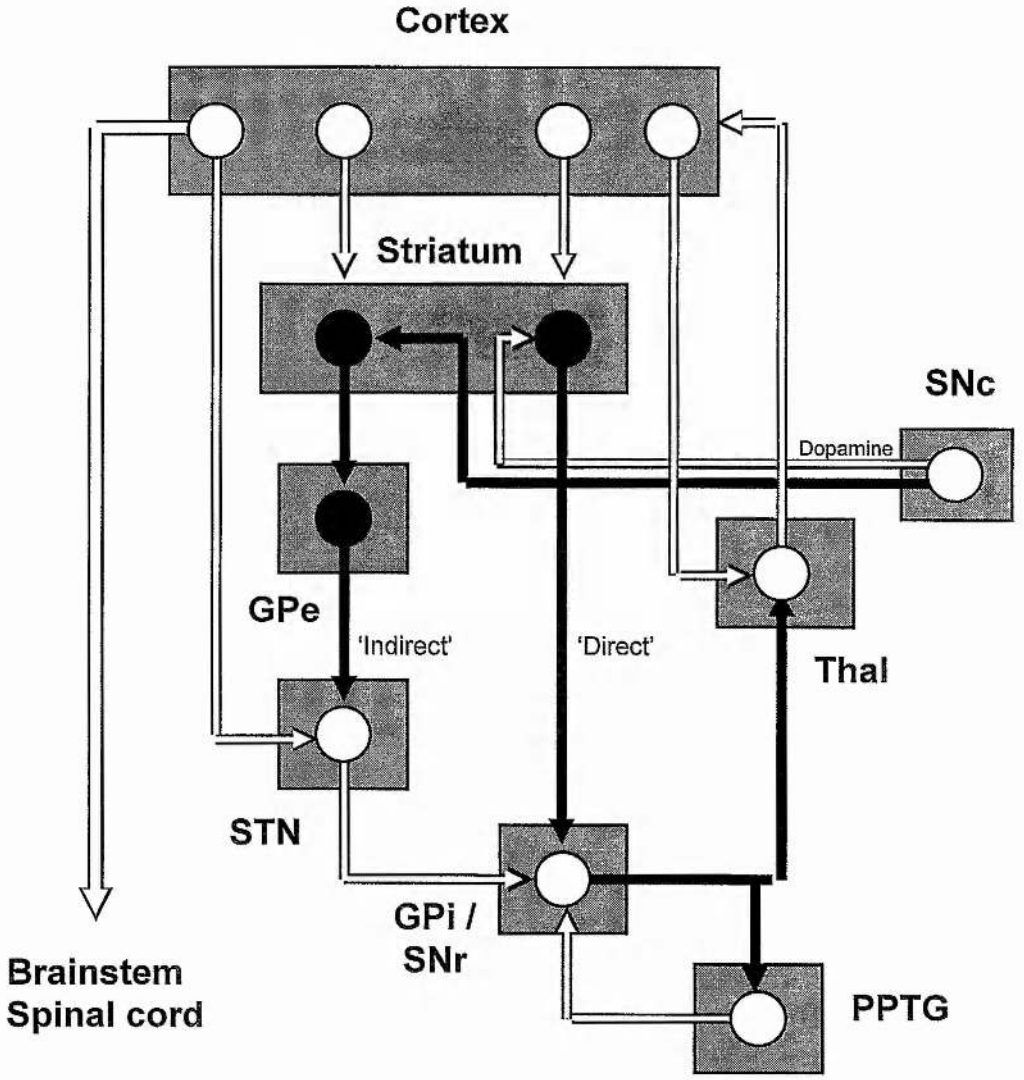


Figure 1.3: The intrinsic connections of the basal ganglia. This diagram depicts the two routes by which the striatum is thought to be able to influence the GPi/SNr - the 'direct' pathway (striatum - GPi/SNr) and the 'indirect' pathway (striatum - GPe - STN - GPi/SNr). Filled arrows represent inhibitory connections and unfilled arrows represent excitatory connections. These are primarily GABAergic and glutamatergic respectively, with the exception of the dopaminergic SNc-striatum connection (labeled 'dopamine'). New abbreviations: GPe - external segment of the globus pallidus; STN - subthalamic nucleus; PPTG - pedunculopontine tegmental nucleus. Figure adapted from Alexander and Crutcher (1990).

The new aspects of Figure 1.3 are the depiction of the connections between the striatum

and GPi/SNr , which include the external segment of the globus pallidus (GPe) and the subthalamic nucleus (STN). Also, the connections in this diagram are labeled as excitatory (unfilled arrows) or inhibitory (filled arrows). Excitatory connections are glutamatergic and inhibitory connections are primarily GABAergic, with the exception of the SNc-striatum connections which are dopaminergic. Figure 1.3 indicates that there are two routes by which the 'input' regions of the basal ganglia (the striatum) can influence the 'output' regions (the GPi/SNr) - the 'direct' and 'indirect' pathways.

Activity of the direct and indirect pathways has opposing effects on the activity if GPi/SNr neurons, and the two pathways are differentially affected by dopamine in the striatum. Dopamine has an excitatory effect on the direct pathway and an inhibitory effect on the indirect pathway. If the direct pathway is active, its inhibitory nature suppresses activity of GPi/SNr neurons. GPi/SNr inhibition of the thalamus is therefore decreased, disinhibiting the thalamus and allowing it to send a strong excitatory signal to cortex. Activity of the indirect pathway has the opposite effect on GPi/SNr and thalamocortical neurons. GPe neurons are inhibited by the striatum, leading to a reduction of the GPe's inhibitory projection to the STN. The STN is disinhibited, and sends a strong excitatory connection to the GPi/SNr, increasing the activity of neurons in this structure. The GPi/SNr inhibits the thalamus, leading to underactivity of thelamocortical neurons and a weak excitatory connection to cortex. In the motor loop, it is thought that that activity of the indirect pathway inhibits movement by suppressing the activity of thalamocortical neurons, whereas activity of the direct pathway facilitates movement by having the opposite effect.

The effects of PD

The effects of PD on the function of the 'direct' and 'indirect' pathways has been studied in monkeys that have been rendered parkinsonian by the protoxin MPTP. These animals show motor symptoms and neuropathology that closely resembles those of PD in humans, and electrophysiological studies of these monkeys is the basis of the following model (DeLong, 1990). This model sets out to describe the implications of PD for the motor loop; the effects of PD on the cognitive loop were not considered by Alexander and Crutcher (1990) or DeLong (1990), and this will be discussed further in Section 1.7.3.4.

DeLong (1990) argued that dopamine depletion of the striatum causes the direct pathway to become underactive and the indirect pathway to become overactive, in accordance with the proposed differential effects of nigrostriatal dopamine on these two pathways. Underactivity of the direct pathway leads to a loss of striatal inhibition of

GPi/SNr neurons, and overactivity of the GPi/SNr results. The GPi/SNr therefore strongly inhibits the thalamus, suppressing thalamocortical activity. Overactivity of the indirect pathway also leads to hyperactivity of GPi/SNr neurons. The striatum inhibits GPe, leading to a failure of GPe to inhibit STN. The disinhibited STN sends a strong excitatory connection to GPi/SNr. Thus, the overall effect of the changes in both pathways is to overexcite GPi/SNr neurons; the GPi/SNr therefore inhibits the thalamus and suppresses normal thalamocortical activity. This tonic inhibition of thalamocortical neurons is thought to be the primary cause of the movement deficits of PD. The precise mechanism by which this change generates movement symptoms in unknown. DeLong (1990) speculated that the increased tonic activity of the GPi/SNr could 'wash out' or diminish the phasic changes of activity in this structure that are associated with movement. Alternatively, the suppression of thalamocortical neurons could simply result in the underactivation of the cortical projection areas of the motor loop. There is some evidence for this; the SMA has been found to be underactive in patients with PD performing self-generated movements (e.g. Jahanshahi et al., 1995; Playford et al., 1992).

1.7.3.3 Modifications of Alexander and Crutcher's model

Subsequent studies of the motor loop in primates and rodents have indicated that some modifications must be made to Alexander and Crutcher's (1990) original model. Reviews of this research by Chesselet and Delfs (1996) and Levy et al. (1997b) have concluded that whilst Alexander and Crutcher's model was principally correct, the nature of the 'indirect' pathway and the functions of the GPe and STN must be reinterpreted. For example, one piece of evidence suggests that the 'indirect' pathway may not in fact constitute a link between the striatum and GPi/SNr. Parent and Hazrati (cited in Graybiel, 1995) have shown that neurons from the GPe do not synapse onto neurons that project from STN to GPi/SNr. Instead, these GPe neurons connect with neurons that project from STN back to GPe. However, there is neurophysiological evidence that is even more problematic for the traditional picture of the indirect pathway. Recent research has confirmed that the STN and GPi/SNr are overactive in animals with experimentally induced parkinsonism. However, it appears that GPe is not underactive in either patients with PD or in animal models of PD (Levy et al., 1997b). It was a feature of Alexander and Crutcher's (1990) model that overactivity of the STN was caused by a loss of inhibition from the underactive GPe. If the GPe is not underactive, it is clear that STN hyperactivity must derive from a source that was not

included in the original indirect pathway. There are two possible sources of STN overactivity; STN is heavily innervated by cortex and also receives dopaminergic input from the SNc. No study has yet verified which of these two connections is important.

The research reviewed by Chesselet and Delfs (1996) and Levy et al. (1997b) has also led to a major re-evaluation of the functional role of the STN. In Alexander and Crutcher's model, the STN was only a relay station in the indirect pathway. However, the STN is now regarded as a major input structure of the basal ganglia. Cortical input to the STN was noted in Alexander and Crutcher's model (see Figure 1.3), but it was accorded no functional significance. This view has changed, partly because cortical innervation may be a source of STN hyperactivity and partly because this connection allows cortex to influence the GPi/SNr as rapidly as the traditional direct pathway - that is, with only two synapses. The direct and rapid nature of this connection is thought to imply functional significance (Mink and Thach, 1993). Despite confusion about its primary sources of input, the role of the STN in driving the hyperactivity of GPi/SNr is still thought to be important. Contemporary approaches to the surgical treatment of PD have involved lesioning either the GPi or the STN in an attempt to reduce inhibition of thalamocortical neurons (Goetz, 1993; Limousin, 1995). Both approaches are beneficial and it would appear that the STN lesion could only have a positive effect by its influence on GPi/SNr activity.

In summary, it is still accepted that overactivity of the STN and the GPi/SNr are pathological hallmarks of PD. However, it appears that the overactivity of the STN was not caused by disinhibition by GPe via the indirect pathway. Rather, STN hyperactivity may be caused either by cortex or by SNc. The STN is now regarded as highly functionally significant and a major input area of the basal ganglia.

1.7.3.4 The specific neural substrates of 'set' and the status of the cognitive loop

The sections above have given an extensive description of the functional neuroanatomy of the 'motor' loop. The role of this loop in the generation of many of the motor symptoms of PD is undisputed, but two questions are left unanswered by the descriptions of neuroanatomy above. First, what is the neural substrate of motor 'set' (as defined in Section 1.4.1) as opposed to the other motor symptoms of PD? The second question concerns the cognitive loop. The functional anatomy of this loop is much less well defined than the motor loop and it has not been mentioned at all in the section immediately above. What evidence is there that the cognitive deficits of PD are

dependent on dysfunction of the cognitive loop?

There is substantial evidence that the motor loop described above subserves deficits of motor 'set' as well as the other symptoms of PD. The beneficial motor effects of GPi lesions include alleviation of akinesia (Iacono et al., 1995) which is a problem of movement initiation and consequently of motor set. It is clear that pallidotomy has its effects on the motor loop, as although it is a subcortical intervention it results in restoration of bloodflow to the SMA (Davis et al., 1997; Grafton et al., 1995; Samuel et al., 1997) which is the cortical projection site of the motor loop. Further evidence for the involvement of the motor loop in motor set was described in Alexander and Crutcher's (1990) original paper. Alexander and Crutcher explicitly suggested that the motor loop subserved preparatory motor 'set' due to the finding of 'set'-related neurons in the putamen and SMA. These neurons become active with the onset of a signal that cues the direction of an upcoming movement. This activity persists until the movement is prompted by the imperative signal, and thus these neurons appear to be holding the future direction of the movement 'online' until it is initiated. Thus, there is evidence from PET studies of patients with PD to single-neuron studies in monkeys that the motor loop subserves the deficits of motor 'set' seen in PD.

Evidence for the involvement of the cognitive loop in the cognitive deficits of PD is less direct. It is clear that the cognitive loop can be dysfunctional in PD. Bloodflow in the DLPFC (the cortical target of the cognitive loop) has been shown to be reduced in PD (e.g. Jahanshahi et al., 1995). Also, pallidotomy has been shown to restore bloodflow to the DLPFC (Ceballos-Baumann et al., 1994; Eidelberg et al., 1996; Samuel et al., 1997); in order for there to be restoration of bloodflow to the DLPFC there must presumably have been preoperative hypometabolism. However, the linkage of this cortical hypometabolism to the cognitive deficits of PD is tenuous. Both Jahanshahi et al.'s (1995) study and all of the studies of pallidotomy cited above have showed a reduction of DLPFC bloodflow whilst patients have been performing motor tasks. The only study that related the function of the cognitive loop to the cognitive deficits of PD is that of Owen et al. (1998).

Owen et al. (1998) found consistent differences in bloodflow in the right GPi between patients and controls when they were performing spatial working memory and planning tasks. This result might appear uninformative, as both cognitive and motor loops pass through the GPi. However, the right GPi was ipsilateral to the side of the movement in Owen et al.'s (1998) tasks, ruling out the involvement of the motor loop. Owen et al. also found differences in caudate bloodflow between patients and controls

when they were performing a spatial memory task, which also strongly implicates the cognitive loop. Although Owen et al.'s study is only preliminary and did not find differences in DLPFC bloodflow between patients and controls, the results of this study provide direct evidence of cognitive loop involvement in the cognitive deficits of PD that confirms the suggestive evidence cited in Section 1.7.1.

1.7.3.5 Surgical treatment for PD

In previous sections, allusions have been made to surgical treatment of PD by the creation of lesions in GPi ('pallidotomy'). This is not currently a common treatment for PD. Pallidotomy was widely used before the introduction of levodopa, with surgeons placing lesions in the anterodorsal and ventrolateral pallidum (Iacono et al., 1995). However, the results of these procedures were variable, and when levodopa became available in the 1960s, pallidotomy was largely discontinued. Pallidotomy was reintroduced by a group of Swedish neurosurgeons in the late 1980s (Laitinen et al., 1992). At the time, GPi hyperactivity had just been reported in parkinsonian monkeys, and Laitinen et al. (1992) noted that abolition of this hyperactivity could have been the source of the clinical efficacy of pallidotomy. Thus, Laitinen et al. carried out pallidotomies that were aimed at the ventrolateral GPi on the basis of evidence from functional neuroanatomy in monkeys, using stereotaxic procedures guided by CT imaging. Initial reports of the effectiveness of this procedure were highly optimistic - Laitinen et al. (1992) claimed that pallidotomy could 'abolish all parkinsonian symptoms' (p. 14). The current consensus on the effectiveness of pallidotomy is more modest. Most investigators accept that pallidotomy is highly effective at controlling levodopa-induced dyskinesias, but that its effect on the negative symptoms of PD (e.g. bradykinesia) are less strong and are inconsistent (Favre et al., 1996; Samuel et al., 1998). Pallidotomy is currently still regarded as an experimental treatment. A number of studies have been published that assess the consequences of pallidotomy for motor function (e.g. Baron et al., 1996; Dogali et al., 1995), but the cognitive consequences of this ablative technique are less well understood. Section 6.0 reviews the literature to date that describes the cognitive consequences of pallidotomy, and also describes a followup study of two patients undergoing this technique.

2.0 Study 1: Mechanisms underlying set-shifting in Parkinson's disease

2.1 Introduction

Patients suffering from Parkinson's disease show a deficit in the performance of the Wisconsin Card-Sorting Test (WCST, Grant and Berg 1948 - see Section 1.6.2). A deficit on the WCST is usually regarded as evidence of frontal lobe dysfunction (Milner, 1963), but the poor performance of patients with Parkinson's disease does not always resemble that of patients with frontal lobe lesions. Frontal lobe damage leads to gross deficits in the performance of the test, characterised particularly by high numbers of perseverative errors and achievement of fewer categories than controls (Milner, 1963). The deficits of patients with PD are consistent but usually less severe (Brown and Marsden, 1990), and some studies have even shown normal levels of category achievement (see Section 1.6.2) and no increase in perseverative errors (Bowen et al., 1975; Cooper et al., 1991; Lees and Smith, 1983).

The processes that underlie successful performance of the WCST are poorly understood. It is thought to test "attentional set-shifting", the ability to stop attending and responding to one perceptual dimension and acquire a response to a new one. As described in Section 1.6.4, Roberts *et al.* (1988) have developed an attentional set-shifting paradigm that is designed specifically to dissociate the different cognitive requirements of the WCST, in a sophisticated, multi-stage test. Basic discrimination and reversal learning ability are tested initially, followed by the introduction of compound stimuli created from more than one dimension (as are WCST cards) and intra- and extradimensional shifts. Intradimensional shifts (ID shifts) require subjects to acquire a response to a novel exemplar of a previously correct dimension. Extradimensional shifts (ED shifts) resemble the rule shift of the WCST; subjects are required to cease attending to the previously correct dimension and to respond to a dimension that was previously irrelevant. Patients with PD show impaired ED shift performance on Roberts *et al.*'s test, in the presence of spared ID shift performance (Downes et al., 1989; Lange et al., 1992; Owen et al., 1992). This pattern of results is also seen in patients with frontal lobe lesions (Owen et al., 1991).

The parkinsonian ED shift impairment does not appear to be due to perseverating in a response to the last correct dimension. Downes et al. (1989) found that measures of lose-stay behaviour (i.e. perseverating on the previously relevant dimension) after an ED shift did not differ between patients and controls. An alternative explanation is that patients were failing due to a tendency to ignore a previously irrelevant dimension, which, after the shift, became relevant. Subjects that failed the ED

shift reported generation of a number of elaborate hypotheses about the correct response after the shift, but failed to consider the previously irrelevant dimension as an option.

The difficulty in making the shift might be due to 'learned irrelevance', which is defined in the animal learning literature as retardation of conditioning to a stimulus that has previously been experienced as uncorrelated with reward (Mackintosh, 1983). It is intuitively plausible that patients with Parkinson's disease suffer from enhanced learned irrelevance, as it is closely related to latent inhibition (the slowing of conditioning to a preexposed stimulus) which is known to be increased by dopamine antagonists such as haloperidol (Feldon and Weiner, 1991; Williams et al., 1994).

Owen et al. (1993b) formally tested the role of perseveration and learned irrelevance in the ED shift deficit using two modified versions of Roberts et al.'s (1988) ED shift. The 'perseveration' shift required subjects to switch their response to a novel dimension whilst the previously relevant dimension became irrelevant. Learned irrelevance cannot affect the learning of this shift (the old irrelevant dimension is replaced by the novel dimension) and therefore poor performance would indicate a perseverative deficit. The converse of this is a shift ('learned irrelevance') in which the previously relevant dimension is replaced by a novel one (so preventing perseveration), and subjects must acquire a response to the previously irrelevant dimension. Enhancement of learned irrelevance would retard the acquisition of this shift. Downes et al.'s (1989) conjectures were confirmed by Owen et al. (1993b). Medicated patients with Parkinson's disease showed a deficit on the 'learned irrelevance' shift, but not on the 'perseveration' shift. Medication appeared to have the effect of reducing perseveration, as unmedicated patients showed an additional perseverative deficit. By contrast perseveration was pronounced amongst a group of patients with frontal lobe lesions, who showed no learned irrelevance deficit. This pattern of results may explain why WCST deficits are a consequence of both Parkinson's disease and frontal lobe damage, but increases in perseverative errors are not always seen in patients with Parkinson's disease.

Van Spaendonck et al. (1995) followed up Owen et al.'s (1993b) results, reasoning that increased learned irrelevance should cause difficulties in the rule-shifting phase of a card-sorting test, but not in the initial rule-acquisition phase; in the acquisition phase, the relevant dimension has not been previously exposed as irrelevant. However, patients with Parkinson's disease showed a deficit in the first rule-shift, but not the second. The salient dimension after the second shift had been preexposed as irrelevant for longer than its counterpart in the first shift, so exaggerated learned

irrelevance should have caused a greater deficit in the second shift. Van Spaendonck et al. concluded that increased learned irrelevance is not the source of the parkinsonian card-sorting deficit, and suggested that a "reduced self-generation of problem-solving strategies" caused the deficit found in their study. This form of deficit has frequently been postulated in Parkinson's disease (Taylor and Saint-Cyr, 1995), but does not explain why there should be a learned irrelevance deficit but no perseverative deficit in Owen et al.'s (1993b) study; their perseveration shift does not appear to require less self-generation of strategies than the learned irrelevance shift.

Owen et al.'s (1993b) hypothesis of a parkinsonian enhancement of learned irrelevance remains attractive due to its specificity and the ease with which it can be tested. Previous explanations of the parkinsonian deficit in attentional set-shifting have been highly general, and have attempted to account for performance on a wide range of cognitive tasks. Cools et al. (1984) proposed that patients have a general deficit in "shifting aptitude" that affects both cognitive and motor domains. Flowers and Robertson (1985) suggested that patients have a problem with maintenance of set, resulting in errors across all phases of a test rather than exclusively the shifting phases. Two theories suggest that the source of the attentional set-shifting deficit lies specifically in dysfunctional attentional mechanisms. Brown and Marsden (1990) have formulated one of the most ambitious and general theories to date, suggesting that patients have depleted attentional resources, and as a result are abnormally dependent on external cues to aid their performance. ED shifts are harder than ID shifts for healthy controls; the parkinsonian ED shifting deficit could be due to the extra attentional demands of ED shifting exceeding the limited processing resources available to patients. Maddox et al. (1996) have found evidence for a more specific parkinsonian deficit in visual selective attention. A proportion of the patients in their study could not optimally attend to one perceptual dimension in the presence of a similar, irrelevant dimension. Such an attentional deficit would certainly cause problems in the performance of an attentional set-shifting task.

The hypothesis of an enhancement of learned irrelevance contrasts with the theories above. It deals with a much more specific cognitive process which has been well-defined in experimental psychology research. Learned irrelevance is a widely-studied mechanism in the literature on animal learning (Mackintosh, 1983), where it applies to a discrete stimulus, and learned irrelevance to a perceptual dimension has been verified as playing a role in ED shifting in humans (Whitney and White, 1993). The purpose of this study is to provide a strong test for the hypothesis that there is a

parkinsonian enhancement of learned irrelevance but not perseveration. To do this, a replication of Owen et al.'s (1993) study was undertaken, with the addition of two conditions in which an enhancement of learned irrelevance or perseveration should *improve* performance. Additionally, preexposure to the relevant and irrelevant dimensions is equalised for patient and control groups. Animal studies have shown that the strength of learned irrelevance is dependent upon the amount of experience of the irrelevant stimulus (Bennett et al., 1995a). A patient deficit in the early stages of Owen et al.'s (1993b) study (such as on simple discrimination or reversals) could lead to increased preexposure to either or both dimensions, enhancing learned irrelevance or perseveration in the absence of abnormal attentional mechanisms. A decision time measure is also included to provide a measure that is possibly more sensitive than errors to criterion, and less prone to ceiling effects; in Owen et al.'s 'learned irrelevance' condition, controls were performing almost optimally in terms of errors to criterion (the shift could be solved using perfect hypothesis testing in a maximum of two errors, and controls recorded an average of 2.8).

2.2 Method

2.2.1 Subjects

2.2.1.1 Patients with Parkinson's disease

Ten patients took part in the study. All patients were outpatients at Dundee Royal Infirmary, and were diagnosed as having idiopathic PD in the absence of dementia by a consultant neurologist (Dr. Richard C. Roberts, University of Dundee). All patients showed a strong clinical response to levodopa, and none had a history of any neurological disorder other than PD or any neurosurgical intervention. Disease severity varied between Hoehn-Yahr stages I and III, with an average of 2.25 (Hoehn and Yahr, 1967). All patients were taking levodopa, and additionally seven were taking benzhexol, four were taking dopamine agonists (bromocriptine/ropinirole) and two selegiline. Neuropsychological data for the patients are contained in Table 2.1 (National Adult Reading Test (NART, Nelson 1982), Mini-Mental State Examination (MMSE, Folstein et al., 1975), and Geriatric Depression Scale (GDS, Yesavage et al., 1982). MMSE and GDS scores were not available for two patients.

2.2.1.2 Control subjects

Twelve controls were tested, having been recruited from the St. Andrews branch of the Royal British Legion and by advertisement. None had a history of neurological disorder, and none were taking drugs known to affect CNS function. Controls did not differ significantly from patients in age or NART-estimated premorbid verbal IQ. Patients reported significantly more affective disturbance than controls on the GDS, and scored significantly lower on the MMSE than controls (see Table 2.1 for means). All subjects gave informed consent before taking part in the study.

	Patients	Controls
Age	65.8 (2.5)	68.3 (2.3)
MMSE	27.6 (0.8)	29.8 (0.1)
GDS	8.9 (2.0)	3.6 (1.1)
NART	109.1 (3.3)	115.7 (2.5)
Hoehn-Yahr stage	2.25 (0.2)	-
Male : Female	7 : 3	10 : 2
Handed - R : L	9 : 1	10 : 2

Table 2.1: Subject characteristics for patients and controls. Figures in parentheses are standard errors.

2.2.2 Materials and procedure

Testing took place in a quiet office, either in Dundee Royal Infirmary or in the School of Psychology at St. Andrews. The test program was run on a portable computer (Datalux databrick), using a touchscreen (MicroTouch TruePoint) to record responses. Subjects were introduced to the touch screen and the basic format of the test procedure using a practice program. This was a simple two-choice discrimination, in which subjects were required to deduce which shape (either a circle or a cross) was 'correct', using the feedback that they were given after each response.

A trial was initiated by the subject holding down the space bar, and a response was registered when the subject touched one of the two boxes containing stimuli on the screen. Subjects used their preferred hand. When the space bar was pressed, a variable delay was initiated (300, 500, 700, or 900 ms), and subsequently one stimulus appeared on the left and one on the right of the screen. Subjects were required to decide which stimulus was correct, and to release the space bar and reach out and touch the correct stimulus. Response time was defined as the time between appearance of the stimulus and release of the space bar. As soon as the space bar was released, the stimuli disappeared; this ensured that subjects did not make their decision after releasing the space bar (a problem in previous studies - Zimmermann et al., 1992). After the subject touched a box on the screen, both boxes immediately disappeared and a message appeared in the centre of the screen saying "Correct", followed by a tick, or "Wrong", followed by a cross. Movement time was defined as the latency between releasing the key and touching the screen. The practice program terminated after thirty trials.

The format described above for the practice program was also used for the test program. The principal differences were that complex, compound stimuli were used and subjects were required to deduce the rule that defined which stimulus was correct. Also, the stimuli changed periodically (a 'shift' - see below), and this was usually accompanied by a change in rule. The test program was initiated with the following instructions:

"Hold down the key and you will see two shapes [experimenter demonstrates]. There is a rule that tells you which one is correct. The rule will be something like "always the black one is correct" or "always the triangle is correct". There will be a lot of different shapes throughout the test - some of them will have spots or stripes, or there may be more than one shape, perhaps of different colours - but there is always one simple rule that tells you which one is correct, and you have to work this out. As before in the practice program, you will be guessing for the first couple of tries. Every so often, the shapes will change, and this may be followed by a shift in the rule. You will then have to work out the new rule. Also, the computer will sometimes tell you how you are doing - for example, it will say "well done, you've got the rule", or "bad luck, the rule is in fact...". This doesn't mean that there's been a shift in the rule, it's just to help you along. You can start now. Go at your own pace, and if you feel tired, take a rest."

After the subject had performed 30 trials, their performance was analysed online by the computer. If they had achieved a criterion of eight or more out of the last ten trials correct, a message appeared on the screen congratulating them and confirming that they had guessed the correct rule. If they had scored less than eight out of the previous ten correct, a message appeared that explicitly informed them of the correct rule. This criterion was chosen to reflect the fact that subjects might have learned the rule, and yet still make isolated mistakes. Between eight and twelve trials later, a similar criterion was applied, and if the subjects were above criterion, a shift occurred. Using this procedure, we hoped to equalise exposure to the dimensions, as most subjects should shift after 30 + 8-12 trials. Explicitly informing subjects of the correct rule also ensured that they were responding stably to the correct rule before the subsequent shift. If subjects scored below eight out of ten when the criterion was checked a second time, another message appeared and informed them of the correct exemplar, and another eight to twelve trials elapsed. If they had reached criterion at this point, a shift occurred, otherwise that block of the test terminated.

2.2.3 Stimuli and types of shift

The presence of a single feature in one of the two stimuli presented defined it as correct, and this was the 'rule' that the subjects were required to learn. The stimuli were made from two 'dimensions', and a pair of stimuli always differed from each other on both dimensions; the dimensions used were 'shape', 'colour', 'pattern', and 'number'. For example, if stimuli were created from the dimensions shape and colour, two possible stimulus pairs would be red square/green circle, or green square/red circle. The rule in this case could be "always touch the red [or green] stimulus, irrespective of its shape", or "always touch the circle [or square], irrespective of its colour". The particular examples of a dimension that were used to make up a stimulus were referred to as ' exemplars'. For example, red and green are exemplars of the dimension 'colour'. The dimension from which the correct feature was taken was referred to as the 'relevant' dimension, and the dimension that was uncorrelated with reward was called the 'irrelevant' dimension.

Stimuli changed when a shift occurred. Six different types of shift were used (see below and Table 2.2) and a 'total change' design (Slamecka, 1968) was used in all but one case ('cued nonshift', see below). In a total change design, if a dimension is present both before and after a shift, the individual examples of that dimension change. For example, if the dimension 'colour' was represented by red and green before a shift, these examples would be replaced by two other colours, such as blue and yellow, after the shift.

All shifts are described below, and represented graphically in Table 2.2. There were two pairs of 'deficit' and 'improvement' shifts. If patients had shown enhanced learned irrelevance, they should have generated more errors than controls on the learned irrelevance shift (deficit), and fewer errors on the learned irrelevance PLUS shift (improvement). Similarly, increased perseveration by patients should have resulted in a deficit on the perseveration shift, and an improvement on the perseveration PLUS shift. The intradimensional, learned irrelevance and perseveration shifts are all structurally identical to those used by Owen et al. (1993b).

Intradimensional shift (ID):

In this shift, the same dimensions were relevant and irrelevant before and after the shift. As a consequence, subjects did not have to switch their attention to a different dimension after the shift. The only requirement was to acquire a response to a new exemplar of the relevant dimension. Although patients at various stages of PD have shown variable ID shift performance in past studies, the patient groups that most closely

resemble the patients in this study (medicated patients with mild PD) have been found to be unimpaired in ID shift performance (Downes et al., 1989; Owen et al., 1992, 1993b).

Learned irrelevance (LI)

After this shift, the previously irrelevant dimension became relevant, and as a consequence subjects were required to acquire a response to a dimension that they had previously been ignoring. If subjects showed an enhanced tendency to ignore the previously irrelevant dimension (that is, enhanced learned irrelevance), this would have slowed the acquisition of the new rule, and caused errors. The dimension that had been relevant before the shift disappeared, and a novel irrelevant dimension was introduced. Any enhanced tendency to attend to the previously relevant dimension (that is, perseveration) could not affect the performance of this shift, as the previously relevant dimension was not present after the shift.

Learned irrelevance PLUS (LI+):

After this shift, the dimension that had been irrelevant remained irrelevant. The previously relevant dimension was replaced with a novel relevant dimension. If enhanced learned irrelevance was present in patients, they would have persisted in ignoring the dimension that was irrelevant after the shift. As a consequence, they would have been more likely to sample the novel relevant dimension as they searched for the rule. This would have reduced the numbers of errors committed by patients.

Perseveration (P):

The previously irrelevant dimension disappeared after this shift, so learned irrelevance could not have affected the performance of this shift. The previously relevant dimension became irrelevant after the shift. If patients had shown an enhanced attentional bias to the previously relevant dimension, they would have persisted in searching this dimension for the rule after the shift. This would have resulted in the generation of errors. The new rule was taken from a novel relevant dimension introduced after the shift.

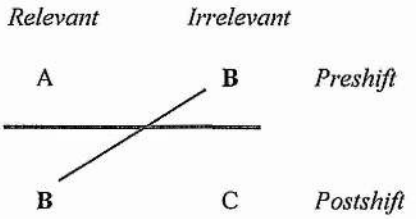
Perseveration PLUS (P+):

This shift was similar to the ID shift, as the dimension that was relevant before the shift remained relevant after the shift. There was a change in the irrelevant dimension at the shift, however; the previously irrelevant dimension was replaced by a novel irrelevant dimension. Perseveration would facilitate the performance of this shift, as subjects would preferentially attend to the relevant dimension after the shift and would not be perturbed by the change of the irrelevant dimension.

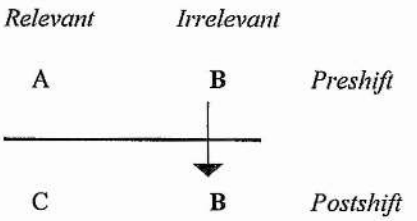
Cued Nonshift (Cnon):

This was a change in stimuli that appeared to be a shift but in fact involved no change in rule. The structure of this change was similar to that of the P+ shift, described above. The previously irrelevant dimension was replaced by a novel, irrelevant dimension. The previously relevant dimension remained relevant, and there was *no change* in the exemplars of this dimension. The same exemplar defined the correct rule after the change. Thus, the appearance of the stimuli changed due to the introduction of a new irrelevant dimension, but the rule remained the same. Previous studies have shown patients with PD to be more distractible than controls (Sharpe, 1990), and also more likely to make errors when a simple stimulus becomes compound even though the rule does not change (C_D stage, Owen et al., 1992). The Cnon shift should reveal any tendency to shift when the appearance of the stimuli changes.

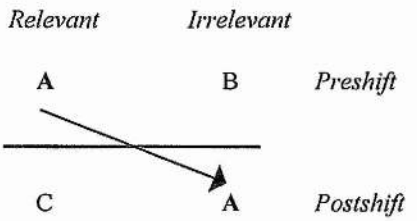
Learned Irrelevance (**LI**):



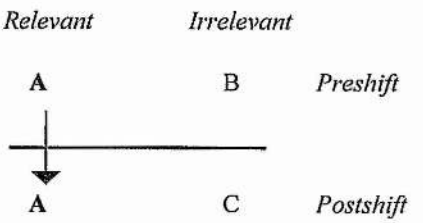
Learned Irrelevance PLUS (**LI+**):



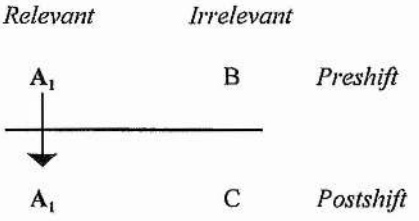
Perseveration (**P**):



Perseveration PLUS (**P+**):



Cued Nonshift (**Cnon**):



Intradimensional shift (**ID**):

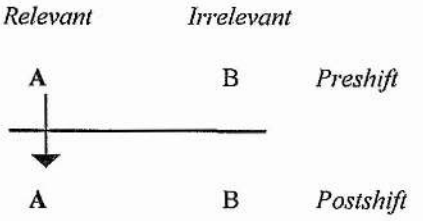


Table 2.2: Each letter represents a dimension present before (top row) and after (bottom row) a shift. The letters in the left column of each shift represent the relevant dimension from which the rule was taken. In all cases shown , the relevant dimension before the shift is represented by 'A', and the irrelevant dimension before the shift is represented by 'B'. As an example, in the learned irrelevance shift (LI), the previously irrelevant dimension (B, top line) becomes relevant after the shift (B, bottom line). The previously relevant dimension (A) disappears, and is replaced by a novel dimension (C). In the Cnon shift, 'A1' represents a specific exemplar of the dimension 'A' which is correct both before and after the shift.

A typical testing session consisted of two separate blocks of trials. Each block included an 'acquisition' phase, where subjects learned an initial rule, and then between

two to four shifts. There was a break between the two blocks in which other neuropsychological tests were completed in order to minimize the effects of fatigue. Subjects were tested in two separate sessions on different days.

The ID shift was always performed first, and only occurred once. All other shifts were performed twice over the course of the study. An example of the sequence of shifts experienced in a typical testing session can be seen in Table 2.3. The order of presentation of shifts was arranged such that when a shift was performed for a second time, it was preceded by a different shift to that which preceded it the first time. It was important to ensure that no shift was associated with a particular set of dimensions or rules. If some dimensions are harder to 'shift to' than others, it could confound results if these dimensions were consistently associated with particular shifts. To avoid this, there were two versions of the test; half of the subjects completed the 'normal' version, and half the 'counterbalanced' version in which dimensions and rules were introduced in a different order.

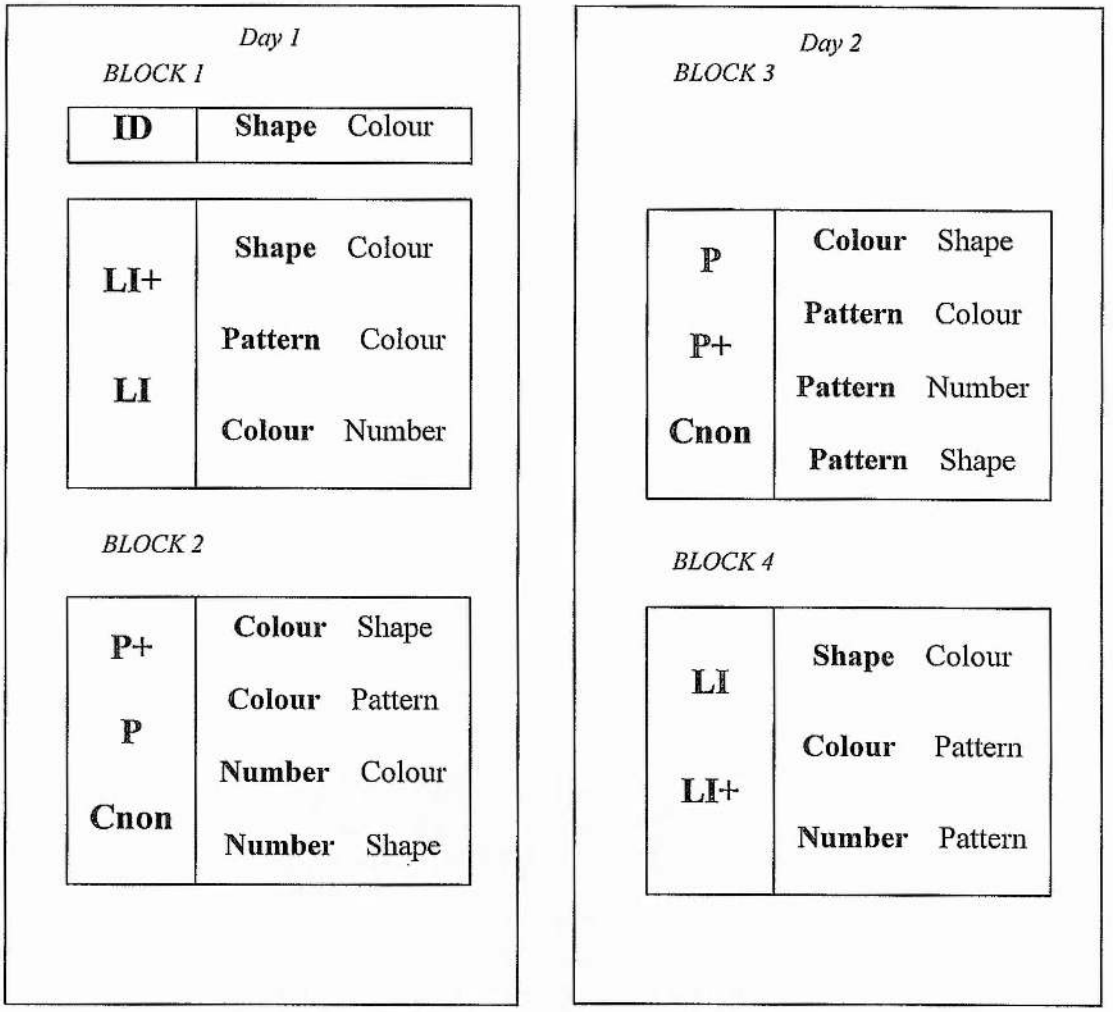


Table 2.3: This table represents the sequence in which one subject experienced the shifts. Subjects were tested over two days, performing two blocks each day. The left column for each block represents the shifts experienced in that block. The right column represents an example of the dimensions that might be associated with each shift, with the relevant dimension in bold type. The experiment always started with an ID shift, but thereafter the order of presentation of the blocks was counterbalanced. The order of shifts within the blocks remained constant. Half of the subjects were in a 'counterbalanced' group, in which the dimensions were introduced to each block in roughly the opposite order.

2.2.4 Data analysis

2.2.4.1 Errors to criterion

The principal dependent measure was errors to criterion. Subjects were judged to have reached criterion when they had performed six correct responses in a row after a shift. As each subject performed each shift twice, their average score was used in calculations. The criterion differed from that used to determine progression to a shift (eight out of ten; see above), and was chosen as it was the criterion used by Owen et al. (1993b), allowing these results to be directly comparable to theirs. If subjects did not reach this criterion by the thirtieth trial after the shift, the total number of errors committed up to that point was used.

2.2.4.2 Response times

Two measures of response time were calculated. 'Decision time' was a measure intended to reflect the average 'thinking time' for trials before subjects learned the rule. This was the mean response time from the second trial after a shift (the first was always a guess) to the trial where subjects reached criterion. After the rule was learned, the task was essentially a two-choice reaction time. 'Reaction time' was the average response time for the all of the trials after criterion was reached. Post-hoc cutoffs were applied to the response time data; trials were only included if response time was greater than 150 ms, less than 15000 ms, and movement time was less than 3000 ms. A movement time of over 3000 ms was thought to reflect a trial where the subject had failed to make his or her decision whilst holding down the space bar.

2.3 Results

Predictions

If patients with PD have increased learned irrelevance or perseveration, they should show an increase in errors on the 'deficit' shift (LI or P, respectively), and fewer errors on the 'improvement' shift (LI+, P+).

2.3.1 Deficit and improvement shifts

Visual inspection of errors to criterion for the LI and LI+ shifts revealed that patients did not show the predicted pattern of deficit on the LI shift and superiority on the LI+ shift (see Figure 2.1, Panel A). Rather, the patients show elevated errors on both shifts, ruling out the possibility of a parkinsonian deficit in learned irrelevance. Inspection of Panel B of Figure 2.1 suggests that a perseverative deficit might be present amongst patients, however; patients make more errors than controls on the P shift, and marginally fewer on the P+ shift. It is, however, unlikely that this crossover provides good evidence for the existence of a perseverative deficit. The P+ shift is functionally the converse of the P shift, but differs from it in another important fashion; the P+ shift is a type of ID shift, whereas the P shift is an ED shift. The P+ shift does not require subjects to acquire a response to a different dimension after a shift (see Table 2.2), whereas this is a requirement of the P shift. Given that patients with PD have previously been shown to have a selective ED shifting deficit with preserved ID shifting (Downes et al., 1989), it is parsimonious to assume that the pattern of results reflects this, rather than being illustrative of a perseverative deficit.

2.3.2 Intradimensional and extradimensional shifts

The pattern of results in this study is remarkably consistent when the shifts are categorised as either ED or ID shifts. The LI, LI+ and P shifts all require subjects to acquire a response to a different dimension (see Table 2.2) and therefore are examples of ED shifts. The ID and P+ shifts are both ID shifts (Table 2.2). (Cnon is not a shift at all - the rule does not change - and must be considered separately.) When the three ED shifts are analysed, there is a clear overall parkinsonian deficit [main effect of group $F(1,19) = 11.38, p < 0.005$]. The shifts differ in difficulty, indicating that although all are ED shifts they are not homogeneous [main effect of shift $F(2,38) = 3.85, p < 0.05$]. The parkinsonian deficit is consistent across all shifts, however, as can be seen from Figure 2.1 and the absence of a group by shift interaction [$F(2,38) = 0.04, p > 0.1$]. When a similar analysis is performed for the two ID shifts (P+, ID) no parkinsonian deficit is found [main effect of group $F(1,20) = 1.07, p > 0.1$]. Although the P+ and ID shifts do not differ overall in difficulty [main effect of shift $F(1,20) = 0.77, p > 0.1$]

there is a group by shift interaction [$F(1,20) = 4.37, p < 0.05$], caused by the decrement in control performance on the ID shift. This was unexpected, but can be explained by the outlying poor performances of two control subjects (making 10 and 15 errors respectively; no other subject made more than 4 errors at the ID shift). The ID shift was only performed once, and is thus more vulnerable to isolated incidences of bad performance; all other shifts were performed twice, and the two error scores were averaged to give each subjects' performance.

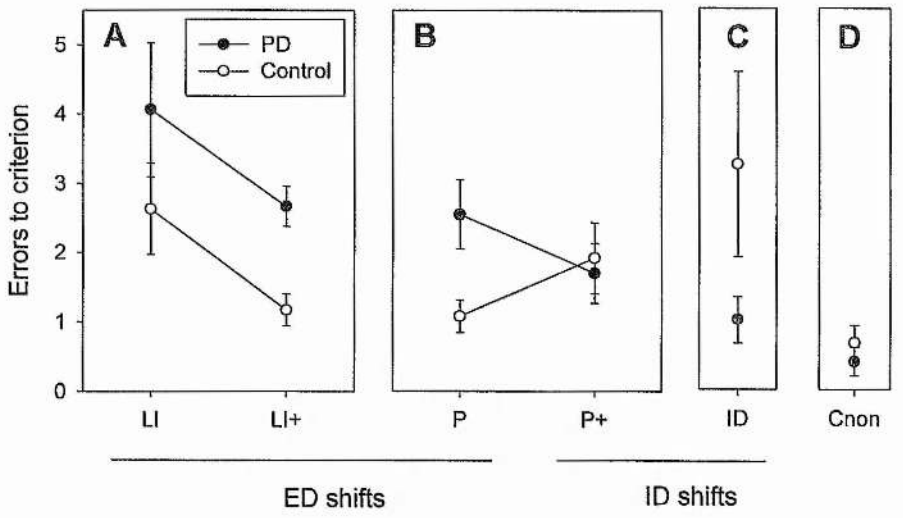


Figure 2.1: Mean errors to criterion for patients and controls across the six shifts. The shifts can be seen as pairs of 'deficit' and 'improvement' shifts (LI/LI+; P/P+), or as examples of ED shifts (LI, LI+, P) and ID shifts (P+, ID). Error bars represent the standard error of the mean.

Performance can also be analysed in terms of 'optimal' hypothesis testing. All shifts in this experiment could be successfully solved whilst making only 0, 1 or 2 errors, if a systematic deductive approach was taken (0, of course, representing a 'lucky guess'). Table 2.4 shows the percentage of ED and ID shifts solved optimally for patients and controls. There is parity between groups in the performance of ID shifts, but a clear parkinsonian decrement in ED shift performance. Furthermore, there is a consistent parkinsonian deficit on all ED shifts, irrespective of which mechanism (learned irrelevance or perseveration) the shift had been designed to test.

	PD	Control
ED shifts	45.3	77.1
ID shifts	79.3	83.3

Table 2.4: Numbers represent the percentage of ID and ED shifts solved in 0, 1 or 2 errors by patients and controls. Solving a shift in under 3 errors was thought to represent optimal hypothesis testing.

2.3.3 Spontaneous shifting and loss of set

Subjects from both groups occasionally made errors after reaching criterion, indicating that they had 'lost set'. Mean errors after criterion per block of testing was calculated for each subject; patients with PD made more errors after criterion than controls [$t(20) = 3.37, p < 0.01$]. Patients with PD did not make any more errors on the Cnon shift than controls [$t(20) = 0.80, p > 0.5$]. Thus, a change in the appearance of the stimuli did not perturb the patients' performance any more than that of controls.

2.3.4 Response times

Performance did not differ significantly between the six shifts for either decision or reaction time, and there were no main effects of group or interactions for either measure [all $p > 0.1$]. The absence of a parkinsonian deficit on any of the response time measures was surprising. In fact, the PD group was on average 130ms slower than controls at the "reaction time" stage of the task but this difference did not reach significance.

2.3.5 Movement Times

Although no parkinsonian deficit was found on any response time measure, patients showed significantly slower movement times than controls [$t(20) = 2.91, p < 0.01$].

2.3.6 Correlational analysis

An association has previously been found between GDS score and the difference between ID and ED errors on a similar task (Downes et al., 1989). In order to test for such an association in these data, the procedure used in Downes et al.'s (1989) study was replicated. Mean ID and ED shift error scores were calculated for each subject, and the difference between the two was entered into a Spearman's correlational analysis with the variables age, NART, GDS and MMSE. The ID-ED difference did not

correlate significantly with any of the demographic or neuropsychological variables (all $p > 0.05$).

2.3 Discussion

This study set out to provide a strong test for the hypothesis that patients with PD might have exaggerated learned irrelevance or perseveration. An enhancement of either learned irrelevance or perseveration in patients should have produced a distinctive pattern of results; namely, deficits on the LI and P shifts, and improvements on the LI+ and P+ shifts, respectively. The results did not fit either predicted pattern. Instead, patients with PD showed a clear deficit on those shifts that demanded acquisition of a response to a different perceptual dimension (ED shifts). Patients' performance was spared on shifts that did not require a shift in response to a new dimension (ID shifts). There appears to be a global ED shift deficit amongst patients with PD that is unrelated to the mechanism (learned irrelevance or perseveration) that each shift was designed to test. Patients also made more errors after criterion, indicating that they 'lost set' more frequently than controls. Distractibility or a tendency to spontaneously shift do not appear to account for this result, as patients showed no tendency to switch rule when the appearance of the stimuli changed in the manner of a shift in rule.

These results have clear implications for Owen et al.'s (1993b) claims of enhanced learned irrelevance in medicated patients with PD. If increased learned irrelevance causes a deficit on the LI shift, it should cause an improvement on the LI+ shift; in fact, patients showed greater errors on both shifts. It can be concluded that enhanced learned irrelevance, as indexed by the LI shift, is not present in medicated patients with PD. A similarly strong conclusion cannot be drawn about a perseverative deficit. This is because the P and P+ shifts are not absolutely comparable, as the P shift is an ED shift, whereas the P+ shift is an ID shift. The increase in errors on the P, but not P+, shift could be taken as evidence of a perseverative deficit, but is better accounted for in terms of the well-established selective parkinsonian ED-shifting deficit.

The conclusion that must be drawn is that there is a global ED-shifting deficit in patients with PD. This is a straightforward extension of the past literature. Furthermore, the present results indicate that an ED-shifting impairment can arise from a mechanism that is not learned irrelevance or perseveration. There are some discrepancies between the results of this study and that of Owen et al. (1993b). Although the patients in this study were all medicated, nevertheless their performance resembled that of Owen et al.'s non-medicated PD group in showing a deficit at both LI and P shifts. This study failed to replicate Owen et al.'s finding of normal P shift performance in medicated

patients with PD. This discrepancy could be explained if the patients in this study had been experiencing fluctuations in their response to medication; if they were experiencing an 'off' period at the time of testing, it might explain why their results resemble those of non-medicated patients. However, the patients in this study were all well-stabilised on their medication, and they were all tested whilst 'on'. Fluctuations in response to medication become more common with increasing severity of PD. As the patients in this study and Owen et al.'s medicated group did not differ substantially in disease severity (mean Hoehn-Yahr stage of 2.25 vs. 2.65 respectively), it is unlikely that they differ in their response to medication. A difference in experimental design between the two studies is most likely to account for the difference in results.

There were also differences in overall numbers of errors between this study and that of Owen et al. (1993b). This is undoubtedly due differences in the intrinsic 'difficulty' of the stimuli used in the two studies. However, in addition, there were differences in relative difficulty for the P and LI shifts in our study and that of Owen et al.. The LI and P shifts are structurally identical across the two studies, yet errors to criterion show that controls found the P shift harder than LI in Owen et al.'s study, making nearly three times as many errors on the P shift than on the LI shift. In the present study, we find the opposite - controls make twice as many errors on the LI shift than on the P shift. The control groups from the two studies are comparable in terms of age and NART-estimated verbal IQ. The relative difficulty of the P and LI shifts also differs for patients. In Owen et al.'s study, the medicated PD group made a similar number of errors at the LI and P shifts, whereas in this study the pattern of patients' performance resembles controls - more errors on the LI shift than on the P shift. It is possible that the methodological differences that account for these differences in relative difficulty may also explain the failure to replicate normal performance on the P shift for patients with PD in this study.

This study differs methodologically from that of Owen et al. (1993b) in three significant ways. First, the studies differed in the way in which individual shifts were associated with particular dimensions. It is clearly important to avoid associating any given shift with the same dimensions in every subject, as it is easier to acquire a response to some dimensions than to others. Owen et al. avoided this problem by using unique dimensions for the LI and P shifts, and counterbalancing the assignment of dimensions to shifts. Subjects encountered entirely different dimensions at the P and LI shift, thus subjects never experienced a dimension twice. Counterbalancing was also used in this study, although it consisted of using the same dimensions in a different

order (thus associated with different shifts) rather than using completely different dimensions. In the present study, there were also the additional shifts (P+, LI+, Cnon). Thus, to enable the repeated measures design, dimensions had to be used more than once. Therefore, dimensions were used repeatedly, but subjects were each tested on each shift twice, with the pre/post-shift dimensions never identical. This further diminished the association between a specific shift and particular dimensions.

Secondly, preexposure to the various dimensions before a shift was not controlled for in Owen et al.'s (1993b) study. All subjects received more exposure to the dimension that was relevant before the ED shift than its irrelevant counterpart. The extra preexposure to the relevant dimension might affect the LI and P shifts differentially, and may account for the elevated errors in the P shift in Owen et al.'s controls. Preexposure to stimulus dimensions was equalised across groups and shifts in this study.

Finally, Owen et al.'s (1993b) dependent measure (ED_{errors}) differs from that used in this study. They count errors to a criterion of six consecutive trials correct, as in this study, but also add the errors committed at the subsequent reversal stage (EDR) to create their 'ED_{errors}' measure. Thus, it is impossible to derive the number of errors committed at the ED shift from the scores given. With hindsight, this combination of scores is inadvisable, as the performance of ED shifts and reversals is likely to have different neural substrates (Dias et al., 1996). It is unfortunately impossible to point to any single one of the above methodological differences as a clear explanation for this study's failure to replicate the normal performance found on the P shift in Owen et al.'s (1993b) medicated PD group. The performance of the patients in this study is more similar to that of Owen et al.'s non-medicated PD group, who show a clear deficit at both the P and LI shifts.

This study set out to test one hypothesis about the source of the ED-shifting deficit in patients with PD. The hypothesis that learned irrelevance is enhanced in medicated patients with PD has not been supported. Instead, the presence of a global ED-shifting deficit in patients with PD has been replicated and is shown here to extend to a number of new variants of the classical ED shift. The cause of the ED-shifting deficit remains unclear, however, and hypotheses other than that of learned irrelevance must be considered as potential explanations.

There is no global deficit of 'shifting aptitude', as proposed by Cools et al. (1984), in the patients studied in this experiment. Both ID and ED shifts required subjects to shift response to a new rule, but no ID shifting deficit is seen in patients in

this study. The patients in this study also display no general problems of increased distractibility, as they show no tendency to shift spontaneously when the stimuli changed in the manner of a shift (Cnon shift). Flowers and Robertson (1985) suggested that patients with PD suffer from a problem in the maintenance of set that is not caused solely by distractibility or perseveration. This general description of the parkinsonian deficit fits the results of this study well. Errors after criterion are elevated, indicating a loss of set that would be ameliorated by simple attentional perseveration. Flowers and Robertsons' proposals effectively describe some of the results of this study, but they give no insight into the processes underlying either loss of set or the ED-shifting deficit in patients with PD. In contrast, Brown and Marsden (1990) refer to specific psychological models in their theory of attentional resource depletion (central executive - Baddeley, 1986; supervisory attentional system - Norman and Shallice, 1980). Their theory can predict a selective cognitive deficit that affects ED shifting but spares ID shifting by suggesting that ED shifts 'overload' the limited attentional resources available to patients in a way that ID shifts do not. Unfortunately, as Brown and Marsden (1990) have noted, there is no way of predicting in advance what will overload resources. Thus, the suggestion that ED shifts exceed available resources is *post hoc*, which weakens the explanatory power of Brown and Marsden's theory. Also, Brown and Marsden would not predict increased errors after criterion, in patients, as this 'reaction time' phase of the test is presumably not very attentionally demanding.

Most recently, Maddox et al. (1996) have identified a parkinsonian deficit in visual selective attention that would certainly alter shifting tasks that make use of two perceptual dimensions. Maddox et al. demonstrated that their patients could attend to a single dimension and combine information from two dimensions as effectively as controls. Patients with PD no longer performed optimally when required to attend selectively to one dimension whilst ignoring an irrelevant dimension. Whilst this deficit is a plausible cause of a number of problems in an attentional set-shifting task, it cannot account for a selective ED shifting deficit in the presence of spared ID shifting. The analysis of the nature of ED and ID shifts below will explain why this is the case, as well as ruling out some potential explanations of the parkinsonian ED-shifting deficit as impossible.

ED and ID shifts make identical demands on subjects when considered in terms of pure hypothesis testing. After both types of shift, two new exemplars of both dimensions are present, and one of these must define the new rule. The subject must deduce which one of the four features defines the stimulus as correct. However, if

subjects are preferentially attending to one dimension before and immediately after a shift, this will have opposing effects on the ED and ID shifts. Subjects will first consider the previously relevant dimension as a possible source of the new rule. This will speed ID shifting, as the new rule will indeed be chosen from the previously relevant dimension. If subjects are performing an ED shift, however, their attentional bias towards the previously relevant dimension will work against them, as the dimension they consider first has become irrelevant after the shift. Thus, the finding that humans or animals find ED shifts harder than ID shifts has always been used as evidence of the presence of selective attention (Slamecka, 1968). In patients with PD, the ID-ED difference is exaggerated. This suggests that the problem experienced by patients with PD is one of exaggerated selective attention, not a global failure of selective attention. The deficit found by Maddox et al. (1996) would elevate errors at both ED and ID shifts, and after criterion. Similarly, a basic perceptual problem or 'reduced self-generation of problem-solving strategies' (Van Spaendonck et al., 1995) would exaggerate errors at both ED and ID shifts. An explanation for the selective ED shifting deficit must postulate an inflexible exaggeration of selective attention rather than a global breakdown of selective attention. One of the merits of Owen et al.'s (1993b) proposed mechanisms of learned irrelevance and perseveration is that they are exaggerations of attention and can therefore account for a selective ED-shifting deficit.

The existence of elevated errors after criterion appears inconsistent with the presence of exaggerated selective attention; an increased attentional bias towards the relevant dimension should reduce errors after criterion rather than increase them. Flowers and Robertson (1985) offer a possible explanation for this contradiction, suggesting that old response sets 'intrude' over new ones in patients with PD. It is possible that errors after criterion represent the intrusion of an excessively strong attentional bias that was present before the shift. This proposal is easy to test; there should be more errors to and after criterion at an ED shift than at an ID shift amongst patients with PD. Intrusions of the set present before the shift would cause errors after criterion at an ED shift due to an attentional switch to the irrelevant (previously relevant) dimension. Such a switch would not generate errors after an ID shift, as the same dimension is relevant before and after the shift. Also, errors after criterion should not be elevated after an LI shift. After such a shift, the previously relevant dimension is no longer present, so the set relating to attention to that dimension cannot interfere with discrimination learning. Unfortunately, there is insufficient data in this study to test these hypotheses.

This study allows some conclusions to be drawn about the nature of the parkinsonian ED-shifting deficit, but the exact cause of this problem remains unclear. ED-shifting deficits have been found in patients with schizophrenia (Elliott et al., 1995), Huntington's disease (Lange et al., 1995) and after frontal lobe damage in humans and monkeys (Dias et al., 1996; Owen et al, 1991, 1993b). This study partially replicates Owen et al.'s (1993b) finding of a dissociation between the performance of patients with Parkinson's disease and those with frontal lobe damage. We found increased errors on the LI shift in patients with Parkinson's disease, which is in marked contrast to the normal LI shift performance seen in patients with frontal lobe damage (Owen et al., 1993). This finding provides further evidence for the existence of dissociable mechanisms underlying ED-shift performance, and it is clinically important to understand these mechanisms. Learned irrelevance could cause an ED-shifting deficit (Whitney and White, 1993), but a deficit on the LI shift is not necessarily evidence for enhanced learned irrelevance. The absence of an improvement on the LI+ shift in this study indicates that the LI shift deficit in patients with PD cannot be due to enhanced learned irrelevance.

3.0 Investigations of attentional set-shifting in healthy subjects

3.1.1 Introduction

Humans and some other animals are able to attend selectively to specific aspects of compound visual stimuli. This ability is demonstrated when a subject is required to respond to one element of a complex stimulus and ignore irrelevant aspects of that stimulus. Studies have repeatedly shown that humans can both attend to one aspect, or 'dimension', of a stimulus and also shift their attention between dimensions; that is, they can stop responding to one dimension and acquire a response to a dimension that had previously been irrelevant and unattended. Such a shift is known as an extradimensional (ED) shift. Although performance on ED shifts has been extensively studied (Wolff, 1967), the mechanisms that govern selective attention to one stimulus dimension amongst many and the subsequent shift to another dimension are not well understood. Ignorance of the mechanisms underlying ED shift performance has clear clinical consequences, as a number of patient groups have been found to show ED shifting deficits. Patients with Parkinson's disease (Downes et al., 1989), frontal lobe damage (Owen et al., 1991), schizophrenia (Elliot et al., 1995), Huntington's disease (Lange et al., 1995) and multiple systems atrophy (Robbins et al., 1992) all fail to make ED shifts more often or make more errors on ED shifts than healthy controls. Breakdown of a single cognitive mechanism could be responsible for the deficits seen in all these patient groups, or different groups may fail the shift for different reasons.

Owen et al. (1993b) proposed that there are two mechanisms which are involved in the ability to attend selectively to one of two dimensions in an ED shifting task that involves two dimensions. An exaggeration of either of these mechanisms could cause ED-shifting deficits. The first mechanism, known as 'perseveration', governs the active maintenance of attention to the dimension that is relevant. Whilst perseveration is essential for the performance of a discrimination task that requires selective attention, it slows the successful completion of an ED shift. If strong perseveration is present at the time of the ED shift, attention will be maintained to the dimension that was relevant before the ED shift but becomes irrelevant after the ED shift. A primary requirement of successful ED shifting is that subjects must stop attending to the previously relevant dimension; perseveration slows this. The second mechanism proposed by Owen et al. (1993b), known as 'learned irrelevance', is the process whereby subjects actively inhibit attention to the irrelevant dimension. As with perseveration, this mechanism is probably an essential component of the simple selective attention required in a discrimination task, as it prevents information from the irrelevant dimension from interfering with

responses to the relevant dimension. Thus, enhanced learned irrelevance would also retard ED shifting. Successful completion of an ED shift requires that subjects acquire a response to a dimension that was previously irrelevant; the presence of learned irrelevance ensures that this dimension will be unattended immediately after the ED shift, thus making a shift of attention to this dimension more difficult.

The mechanisms of learned irrelevance and perseveration can explain the maintenance of attention to a single dimension by excitatory and inhibitory attentional mechanisms that are similar to those found in work on spatial attention (Posner and Petersen, 1990). Owen et al. (1993b) provided evidence for the existence of learned irrelevance and perseveration by means of a double dissociation. Two novel versions of the ED shift were created, that were thought to be differentially susceptible to disruption by increased learned irrelevance and perseveration. Patients with frontal lobe damage were found to have exaggerated errors only on the 'perseveration shift', and medicated patients with Parkinson's disease showed elevated errors only on the 'learned irrelevance shift'.

Subsequent research has placed Owen et al.'s (1993b) hypothesis in doubt. An experiment with patients with Parkinson's disease (Study 1) showed that Owen et al. (1993b)'s 'learned irrelevance shift' may not adequately measure learned irrelevance, which casts doubt on the empirical evidence for the dissociation between learned irrelevance and perseveration. However, intuitively, it seems to be a reasonable hypothesis and there is some empirical support. Whitney and White (1993) used generalisation gradients to analyse the transfer of attention between dimensions at the ED shift. They concluded that learned irrelvance was the principal mechanism that slowed ED shifting in young adults. Also, ED-shifting tasks closely resemble some paradigms used to induce a phenomenon termed latent inhibition in adult humans. Latent inhibition paradigms consist of two phases, a 'preexposure' phase and a 'test' phase that are equivalent to the period before and after an ED shift in a dimensional shifting task. At the preexposure phase, subjects' attention is engaged by some form of cognitive task, known as the 'masking' task, that is designed to divert the subject's attention away from an irrelevant stimulus that is repeatedly presented without consequence. This phase can be seen as equivalent to the phase before an ED shift, where subjects respond to a relevant dimensions and ignore an irrelevant dimension. The 'test' phase of the latent inhibition paradigm is similar to an ED shift, as subjects are required to cease attending to the stimuli involved in the masking task and to acquire a response to the stimulus that had previously been irrelevant. This task is then

compared with a condition in which the stimulus that becomes relevant at 'test' was not experienced or preexposed as irrelevant. Preexposure of a stimulus as irrelevant has been shown to retard acquisition of a response to this stimulus in a number of studies (Lubow and Gewirtz, 1995). This effect is known as latent inhibition and closely resembles Owen et al.'s (1993b) proposed mechanism of learned irrelevance, which consists of slowing of acquisition of a response to a perceptual dimension that has been preexposed as irrelevant. Indeed, the separation of learned irrelevance and latent inhibition into two separate mechanisms may be false; Bonardi and Hall (1996) have shown that effects attributed to learned irrelevance in rats can be accounted for in terms of the mechanism of learned irrelevance. The similarity of latent inhibition to learned irrelevance and of adult human latent inhibition paradigms to ED shifting adds face validity to the possibility that learned irrelevance can affect ED shifting.

Nearly all of the experimentation on ED shifting in clinical populations has used 'standard' ED shifts in which the same dimensions are present before and after the shift. However, Owen et al. (1993b) used novel versions of the ED shift in an attempt to index learned irrelevance and perseveration. These shifts involved introduction of a novel dimension after the ED shift, and Study 1 has cast doubts on the ability of these shifts to accurately index learned irrelevance and perseveration. Ideally, it would be possible to study the effects of learned irrelevance and perseveration on the 'standard' ED shift, as this would give results that would be most relevant to work with clinical populations. An experimental design has been developed that permits such a study, using a method that resembles a latent inhibition paradigm. This method permits the selective preexposure of either the relevant or irrelevant dimension, and allows the effects of this preexposure on ED-shifting performance to be studied in healthy subjects. As learned irrelevance and latent inhibition closely resemble each other and may not be different mechanisms (Bonardi and Hall 1996), preexposure of the irrelevant dimension should retard ED-shifting.

This preliminary study will only examine learned irrelevance, as this mechanism is implicated in ED-shifting by Whitney and White's (1993) study and it closely resembles the well-established phenomenon of latent inhibition. Both young and elderly healthy subjects were tested, as unpublished data (Salzman 1993, cited in Lubow and Gewirtz, 1995) suggests that elderly subjects may show a greater effect of latent inhibition than young subjects. This study compared 'preexposed' (PE) with 'non-preexposed' (NPE) groups in both young and elderly subjects. Care was taken to select levels of preexposure that would be appropriate to demonstrate an effect of

learned irrelevance in the PE group. As no previous experiments have attempted to induce learned irrelevance in an ED shifting task through preexposure, there are no published precedents that indicate which levels of preexposure are appropriate. Therefore, the levels of preexposure used in this study were selected on the basis of a study of latent inhibition (De la Casa et al., 1993) and an older literature relating to 'overlearning' in ED shifting tasks.

De la Casa et al. (1993) carried out the only study to date that has related levels of preexposure to latent inhibition effects in humans. They found no latent inhibition effect after 3 minutes of preexposure, but found a clear effect of latent inhibition after 6 and 15 minutes of preexposure. Using De la Casa et al.'s (1993) study as a guideline, the PE group in this study received 200 trials (around 10 minutes) of preexposure. The selection of a level of preexposure group was more complex, as the level chosen had to account for a well-established effect of prior learning on ED shifting known as 'overlearning' (Wolff, 1967). 'Overlearning' refers to the *facilitatory* effect of low levels of preexposure of both dimensions on ED shifting. Studies from the overlearning literature usually compare the ED shift performance of two groups; the first group encounters an ED shift immediately after learning the rule (reaching 'criterion') at the preceding discrimination phase, whereas the second group is allowed 10 or 20 'extra' discrimination trials after reaching criterion before the ED shift occurs. The group that did not experience the extra 'overlearning' trials performs less well at the ED shift. Thus, there appears to be a performance deficit associated with experiencing very low levels of preexposure; that is, performance is poorer when no overlearning is experienced. It is clearly important to avoid any deficits in performance in the NPE group in this study, as they could cancel out deficits due to learned irrelevance in the PE group. Therefore, it was necessary to give subjects in the NPE group sufficient overlearning to ensure that they had no performance 'deficits' caused by excessively brief preexposure. Subjects in the NPE group received 40 trials (around 2 minutes) of preexposure. This level was designed to allow at least 20 trials of overlearning in the NPE group, whilst corresponding to a level of preexposure that had previously been shown not to induce latent inhibition (De la Casa et al.'s '3 minutes' group). Any remaining performance differences seen between the PE and NPE groups in this study should be due to learned irrelevance alone and should not be minimised by excessively brief preexposure in the NPE group.

Previous experimental work has indicated that young healthy subjects perform close to 'ceiling' in ED shifting tasks when simple, easily discriminable stimuli are used

- for example, squares and circles and colours such as red and green (Study 1, unreported data). They tend to make only one or two errors, and their response times are rapid. In this study, it was felt to be important to avoid 'ceiling' and 'floor' effects in both young and elderly subjects, as maximum sensitivity to any effect of preexposure was needed. It was accepted that any effect of preexposure might be small, particularly as there was no way of determining the amount of preexposure that would be optimal. Pilot testing was carried out to ensure that stimuli presented an appropriate level of difficulty, and a measure of response time was taken to provide a more sensitive measure than errors to criterion.

3.1.2 Method

3.1.2.1 Subjects

Young subjects

Eighteen young subjects were tested, drawn from the undergraduate and postgraduate community at the University of St. Andrews. This group consisted of ten males and eight females, with a mean age of 21.2 years (range 18-31).

Elderly subjects

Eighteen elderly subjects were recruited by advertisement. Six males and twelve females were tested, with a mean age of 63.4 years (range 54-73). All denied a history of major psychiatric or neurological problems, and scored above 24 on the Mini-Mental State Examination (Folstein et al., 1975).

All young and elderly subjects claimed to have intact colour vision. Informed consent was obtained from all subjects prior to the experiment, and subjects were paid at a rate of £3.50 per hour. Experimental design and procedure were almost identical for young and elderly subjects throughout, with minor variations in stimuli and display time described below.

3.1.2.2 Materials and procedure

The experimental test was run on a portable computer (Datalux Databrick), and responses were registered by a touch screen (Microtouch Truepoint). Before starting the main test, subjects completed a practice program that illustrated the format of the experimental task and introduced the response requirements in a graded fashion. The final stage of the practice program was a two-choice visual discrimination task. Subjects were required to work out which of two shapes (a circle or a cross) was 'correct' on the basis of feedback ('correct' / 'wrong') given after each response.

The experimental program was a two-choice visual discrimination in which subjects were required to work out a 'rule' that defined which of two complex stimuli was correct. A stimulus was correct when a certain visual feature was present in it. This feature could be an aspect of the shape of the stimulus, its colour, a striped pattern on the stimulus or the background colour behind the stimulus. As in the practice program, subjects were required to use feedback to work out the rule. During the course of the test, the rule that defined the correct shape changed periodically. A trial was initiated when the subject pressed and held down the space bar of a keyboard with his or her preferred hand. After a delay of variable length (300, 500, 700 or 900 ms) elapsed, one 'box' appeared on the left, and one on the right of the screen, each containing one

stimulus. The subject was required to release the space bar and to reach out and press one shape with his or her preferred hand. If one second elapsed before a response was made (two seconds for elderly subjects), the stimuli disappeared and subjects were required to press one of the empty boxes. As soon as a response was made, feedback was delivered in the form of a graphical 'tick' or 'cross' followed by the words 'correct!' or 'wrong...' respectively. 'Response time' (RT) was defined as the time elapsed between the appearance of the stimuli on the screen and the release of the space bar. The time between the release of the space bar and a response on the screen was termed 'movement time' (MT).

Before starting the main test, subjects were given comprehensive instructions which are reproduced here in full: "This is the main program. The format is basically the same as the practice program that you've just completed. You press down the space bar and two shapes appear - one is right, one is wrong, and your job is to press the one that is correct. There is a rule telling you which shape is correct. This rule will be something like 'always press the red shape', or 'always press the triangle'. A number of different rules are possible, as the shapes can differ in four ways. They can be different colours or different shapes, they can have different striped patterns on them, or there can be a different colour behind the shape. The shapes can differ from each other in more than one way at once, so they may look complicated. However, there is always one simple rule that tells you which shape is correct. Your job is to work out this rule and press the correct shapes.

Although I have said that the rule is simple, this does not mean that the task is easy. The differences between the shapes may be very subtle, and quite hard to see. For example, there may be a small difference in the outline of the shape, or a slight difference in the shading of the colour. As a consequence, I suggest that at first you hold down the space bar for long enough to give yourself time to look at the shapes. After a few trials, you will be confident that you know what the rule is, and then you can speed up and treat this as a reaction time task.

There are two additional factors that make this task difficult. First, the rule will change as you go through the task. For example, the shapes will change and you will realise that the old rule can no longer apply. If this happens, I suggest that you slow down and give yourself time to look at the shapes, and when you are confident that you know the rule you can speed up again and treat it as a reaction time task. The second difficulty is that you only ever get about one second ["two seconds" for elderly subjects] to look at the shapes. That is, after the shapes have been on the screen for one

second ["two seconds"] they will disappear. If you have not decided which shape to press by this time, you will have to press one of the boxes anyway. One final point - there are no hidden rules. The rule always has something to do with the shapes, colour, stripes or background colour, although it might be hard to see. You can start the test now."

3.1.2.3 Stimulus composition and types of shift

A pair of complex stimuli appeared on each trial, one of which was 'correct'. These stimuli could differ from each other in colour, shape, superimposed pattern or background colour. These general visual features were known as 'dimensions', and stimuli always differed from one another on two dimensions. For example, a pair of stimuli might differ from one another in colour (blue or yellow) and shape (triangle or cross), creating two possible stimulus pairs - blue triangle / yellow cross, or yellow triangle / blue cross. (The stimuli used in this study were not as simple as this - see Section 3.1.2.5 and Figure 3.1.) The presence of a single visual feature from one of the two dimensions defined the stimulus as correct, and this was the 'rule' that subjects were required to work out. An example of such a rule might be 'always press the blue shape', or 'always press the triangle'. The two instances of each dimension were known as 'exemplars'; for example, 'blue' and 'yellow' were both exemplars of the dimension 'colour'.

The rule changed, or 'shifted', either four or six times in the course of the experiment. Two types of shift were important for this study: intradimensional (ID) shifts and extradimensional (ED) shifts. In both ID and ED shifts, the same dimensions were present before and after a shift. However, when the shift occurred, the particular exemplars associated with these dimensions were changed. For example, the colours blue and yellow might be replaced with red and green, and the shapes triangle and cross might change to square and circle. In an intradimensional shift, the new correct exemplar was chosen from the dimension that was relevant before the shift. If the colour 'blue' had been correct before the shift, the correct exemplar would be either the colour 'red' or the colour 'green' after the shift. In contrast, after an ED shift the new correct exemplar was taken from the previously irrelevant dimension. As a consequence, subjects were required to ignore the dimension that they had previously attended to (colour) and acquire a response to a new, previously irrelevant dimension (shape - either 'square' or 'circle').

3.1.2.4 Design

The principal aim of this experiment was to examine the influence of preexposing the previously irrelevant dimension on the performance of the ED shift. Subjects were therefore randomised into preexposed (PE) and non-preexposed (NPE) groups in a between-subjects design. As ED shift performance was being compared between groups, the last two shifts completed by both groups were an ED shift and an ID shift. In order to allow comparison between groups, the stimuli and rules associated with these two shifts were identical. However, the groups differed in their experience of dimensions before the ED shift. It was anticipated that subjects in the PE group would make more errors at the ED shift or have longer decision times than the NPE group.

Both groups completed the same number of trials before the ED shift. The NPE group received 160 trials of preexposure to two dimensions that were not involved in the ED shift, as represented in Table 3.1. This group then received 40 trials of preexposure to the relevant and irrelevant dimensions to allow acquisition of this discrimination before the ED shift ('acquisition' (AQ) shift, Table 3.1). In contrast, the PE group experienced a single dimension ('pattern' - Table 3.1) as irrelevant from the beginning of the test up until the ED shift, entailing 200 trials of preexposure. Eighty trials elapsed after the ED shift for both groups in order to allow acquisition of this comparatively difficult shift.

The effect of prior ID shifting on ED shift performance was also investigated as a supplementary variable. Young subjects in both PE and NPE groups were randomised into two groups; one group received three ID shifts whilst responding to the two dimensions present in the 'preexposure' phase ('3ID' group), and the other group received only one ID shift in this phase ('1ID' group). These two groups only differed in the number of ID shifts that they experienced; both groups performed the same number of trials before the ED shift.

	NON-PREEXPOSED		PREEXPOSED	
	Relevant	Irrelevant	**Relevant**	Irrelevant
	Colour	Background	**Colour**	Pattern
1 or 3 ID shifts				
	Colour	Background	**Colour**	Pattern
AQ shift				
	Shape	Pattern	**Shape**	Pattern
ED shift				
	Pattern	Shape	**Pattern**	Shape
ID shift				
	Pattern	Shape	**Pattern**	Shape

Table 3.1: The left column of this table represents the shifts that were experienced by both groups in the order in which they were presented. The rest of the table depicts the dimensions that were experienced as relevant and irrelevant before and after these shifts for the NPE and PE groups.

3.1.2.5 Stimulus generation and pilot testing

In past experiments, it has been shown that young subjects perform close to 'ceiling' when required to make ID and ED-style shifts with simple stimuli and unlimited display times (Study 1, unreported data). These subjects made an average of around one error to criterion on all shifts, indicating that they are equally able to use optimal hypothesis testing to solve all types of shift. In this study, the aim was to create ID and ED shifts which were solvable, but on which young subjects generated errors. Display time and stimulus discriminability were manipulated to achieve this. The features that differentiated the stimuli were subtle and hard to detect when stimuli were only displayed briefly.

Stimuli for use with young subjects were piloted as pairs of exemplars from a single dimension - for example, two shapes. These pairs of exemplars were presented to young subjects in a two-choice visual discrimination task in which display time was restricted to one second. As pilot testing proceeded, exemplar pairs that were 'too easy' or 'too hard' were noted and adjusted accordingly. Piloting stopped when it was judged that most subjects could solve each discrimination within fifteen trials, yet seldom

solved them with only one error. The finalised exemplars were combined to form the compound stimuli used in this study. A pair of stimuli created from the dimensions 'shape' and 'pattern' are shown in Figure 3.1, Panel A.

Stimuli for the elderly subjects were developed from the stimuli described above. The exemplars developed for young subjects were altered by slightly exaggerating the salient difference between each pair of exemplars. Pilot testing was then carried out with elderly subjects in an identical fashion to that described above, but with a stimulus display time of two seconds. An example of compound stimuli made from the dimensions 'shape' and 'pattern' can be seen in Figure 3.1, Panel B.

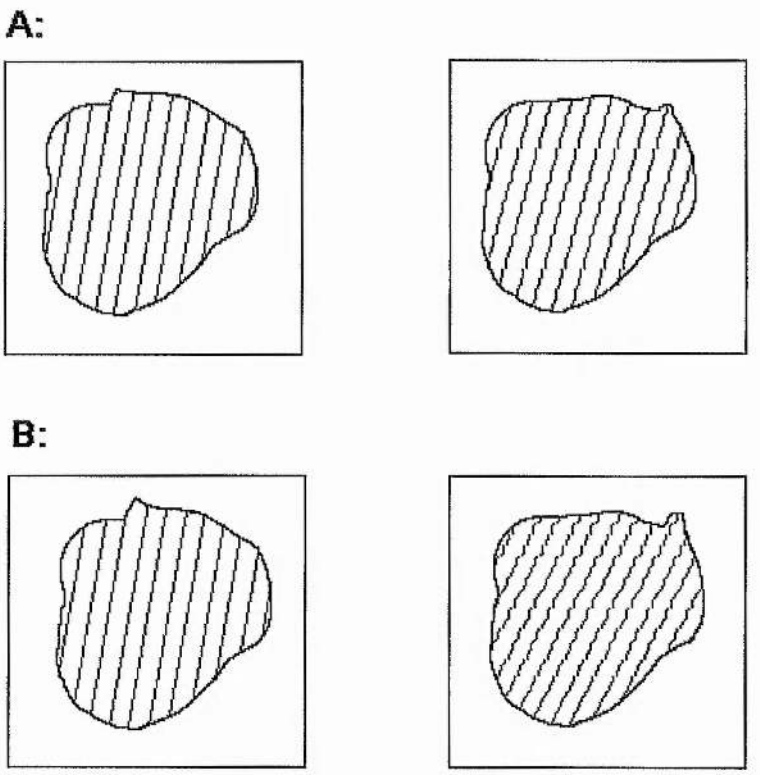


Figure 3.1: Representative stimulus pairs for young and elderly subjects. Each pair differs in both shape and pattern. The salient differences are the postition of the protruberance on the top edge of the shape (to the left or right) and the angle/smoothness of the pattern (either towards vertical and smooth or away from vertical and crinkled).

3.1.2.6 Data analysis

Subjects were excluded from the analysis if they made fewer than eight correct responses in the ten trials before the ED shift. These individuals were excluded because it is meaningless to measure 'attentional set-shifting' unless it can be demonstrated that subjects have not successfully formed an attentional set before the shift.

Subjects were considered to have learned a rule, or 'reached criterion', when they had made ten correct responses in a row. In this study it was not uncommon for

subjects to make six correct responses and subsequently 'lose' the rule again, therefore a criterion of six consecutive correct responses (as used in Study 1) was not used. Both 'errors to criterion' and 'trials to criterion' were calculated for the ED shift. Errors to criterion was considered to be the principal dependent variable, and a measure of trials to criterion was required for Life Tables analysis (see below). Both of these measures were bimodally distributed, as subjects tended either to learn the rule early after a shift or to fail to learn it at all. Thornton et al. (1996) noted that conventional parametric or nonparametric analyses are inappropriate for these data, due to the fact that neither 'errors' nor 'trials' to criterion take any account of whether or not a subject ultimately succeeds in learning a rule. For example, if the rule is not learned after an ED shift in the 80 trials available, the 'trials to criterion' score will be 80, despite the fact that criterion was never reached. To overcome this problem, Thornton et al. (1996) used a Cox's Proportional Hazards Regression procedure that allowed entry of both trials to criterion data and a variable indicating whether or not the subject ultimately reached criterion. In this study, a more simple approach was used, involving the creation of Life Tables (SPSS for Windows v. 6.0, Norusis 1994). This allows entry of trials to criterion data (values between 10 and 85), and also a variable indicating whether criterion was eventually achieved (0 or 1), and yields a Wilcoxon (Gehan) statistic. This procedure was used in all of the between-group analyses in this study. Subjects who did not reach criterion at the ED shift were allocated trials to criterion scores of 85 to differentiate them from subjects who achieved criterion on the last trial of that shift (who would score 80).

It is possible that two subjects could require the same number of trials to solve an ED shift, yet one subject could have required much more 'thinking' time on each individual trial to achieve the solution. In order to account for this, 'decision time' was also used as a dependent variable. This was the mean response time for all of the trials completed before the subject was judged to have learned the rule. Mean response time was calculated from the second trial after the ED shift (the first was always a guess) until the trial where subjects achieved criterion. Trials were excluded in which RT was under 150 ms or over 5000 ms. If an MT of over 3000 ms was recorded, this was thought to reflect a subject's failure to make a decision before releasing the space bar; such trials were also excluded.

3.1.3 Results

3.1.3.1 Absolute levels of errors

Figure 3.2 depicts mean errors to criterion for the ED shift for young and elderly subjects. Optimum performance of the ED shift would have resulted in group mean error scores of 1. Alternatively, if all subjects in a group failed to acquire a shift the group score would have been around 40. Mean error scores for the young and elderly PE and NPE groups all fell between 12 and 20 errors to criterion, indicating that subjects' performance should not have been affected by 'floor' or 'ceiling' effects.

3.1.3.2 Effect of preexposure

Learning to criterion

Visual inspection of Figure 3.2 indicates that there is no difference in errors to criterion between the PE and NPE groups in either young or elderly subjects. This was verified using a Life Tables analysis which revealed no effect of preexposure for either young (Wilcoxon (Gehan) $p = 0.96$) or elderly subjects (Wilcoxon (Gehan) $p = 0.47$). Preexposure did not exaggerate errors for the PE groups; indeed, the absolute error scores for the PE groups were lower than for the NPE groups. Survival functions were also plotted for these data, which represent the proportion of subjects in each group that had not yet learned the rule at each of the 80 trials after the ED shift (Figure 3.3). This allows inspection of the distribution of the trials to criterion data, and confirms that there was no overall difference in rate at which subjects in the PE and NPE groups solved the ED shift.

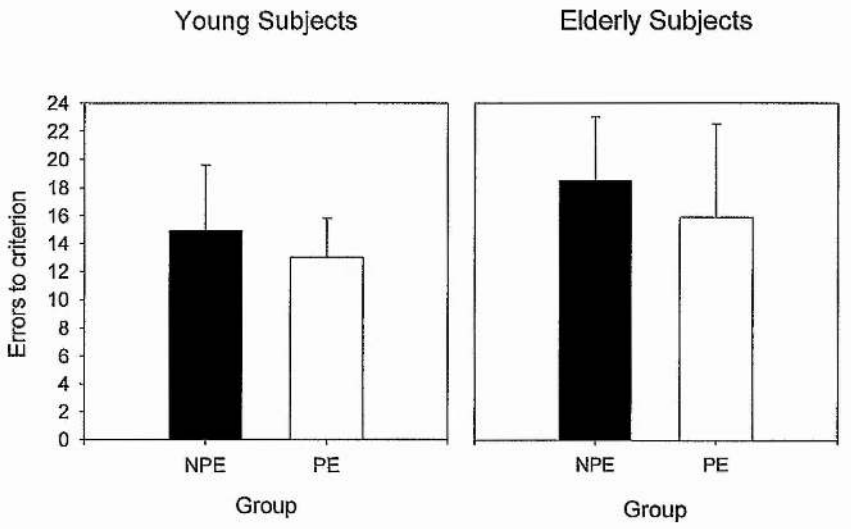


Figure 3.2: Mean errors to criterion for the young and elderly subjects at the ID and ED shifts. Error bars represent the standard error of the mean.

Decision times

Decision times were calculated for each subject. There was no effect of preexposure on decision times for either young ($t(16) = 0.22, p = 0.83$) or elderly ($t(16) = 1.09, p = 0.29$) subjects. However, the use of short display times in this study may have restricted the variability of the RT data. If subjects had not made a decision after one second (two seconds for elderly subjects), the stimuli disappeared and subjects could gain no further information by delaying their response. In practice, when subjects had not yet learned the rule, they tended to respond immediately after the disappearance of the stimuli. This would act to reduce any differences between decision time scores for subjects in PE and NPE groups, and indicates that errors and trials to criterion remain the most important dependent variables.

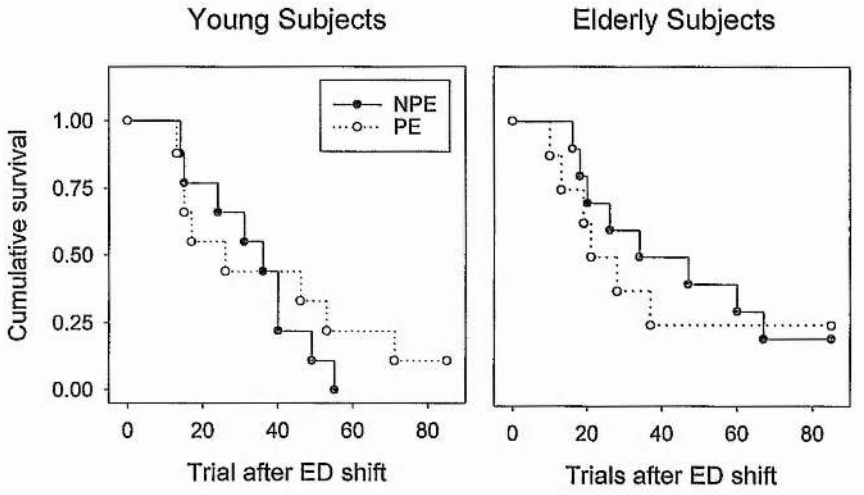


Figure 3.3: Survival functions for the PE and NPE groups at the ED shift. Some survival functions do not reach zero at 80 trials; this reflects that fact that some subjects did not reach criterion at this shift.

3.1.3.3 Effects of experiencing prior intradimensional (ID) shifts

ED shift performance did not change as a consequence of the number of prior ID shifts. The 3ID group acquired the ED shift at the same rate as the 1ID group (Wilcoxon (Gehan) $p > 0.5$).

In summary, preexposure of the irrelevant dimension did not retard acquisition of the ED shift in either young or elderly subjects when learning was measured in terms of errors and trials to criterion. Decision times were also the same between PE and NPE groups, although this may have been a consequence of using restricted stimulus display times.

3.1.4 Discussion

The effect of preexposure of the irrelevant dimension on ED shifting was studied by comparing the ED shift performance of two groups that had experienced different levels of preexposure. Greater preexposure of the irrelevant dimension did not cause any change in ED shift performance in either young or elderly subjects. It had been anticipated that greater preexposure would retard ED shifting, either by increasing the number of errors committed before the solution of the ED shift, or by increasing the 'thinking time' that subjects required on each trial. However, there was no trend towards increased errors or trials to criterion in the data; mean errors for the two PE groups were actually lower than for respective NPE groups. Decision times were also found to be equal across PE and NPE groups.

The failure to find any effect of preexposure on ED shifting was not anticipated, and must be accounted for. In particular, it is important to determine whether the manipulation of preexposure was adequate.

3.1.4.1 Can experimental factors account for the failure to find an effect of preexposure?

A number of experimental factors could have limited the ability of this study to detect a preexposure effect, even though the study was designed to minimise these. Randomisation into groups ensured that the PE and NPE groups were demographically homogeneous within young and elderly subjects, indicating that subject factors could not account for the failure to find a preexposure effect. A substantial manipulation of preexposure was used (the preexposed group received five times as much preexposure as the non-preexposed group), and the size of this manipulation was based on that used in a past study of latent inhibition (De la Casa et al., 1993). Any effect of preexposure could have caused an increase in errors without this measure encountering a 'floor' effect, and the decision times measure should have been sensitive to any minor change in ED shift difficulty. As there were no obvious experimental factors preventing the detection of an effect of preexposure, other interpretations must be considered.

3.1.4.2 Were the correct levels of preexposure used?

Prior studies have shown that some levels of preexposure are ineffective at producing an effect of latent inhibition. Only one study has explicitly manipulated preexposure to study its effects on latent inhibition in humans (De la Casa et al., 1993). De la Casa et al. (1993) found no latent inhibition after 3 minutes of preexposure, but an effect of latent inhibition was present after 6 minutes of preexposure and was maximal after 15 minutes of preexposure. The current study used a similar ratio of preexposure between

non-preexposed and preexposed groups (1 : 5), although the absolute times used were around 2 and 10 minutes, respectively. If the size of a learned irrelevance effect increases continuously with mounting preexposure (as for latent inhibition in De la Casa et al.'s (1993) study) the levels of prexposure used should have differed sufficiently to detect a learned irrelevance effect. As no other studies exist that relate level of preexposure to the size of a learned irrelevance or latent inhibition effect in humans, there was no *a priori* reason to assume that the levels used were not optimal.

It is possible that an effect of learned irrelevance may be all-or-nothing in a task, and its magnitude may not vary continuously with preexposure. Thus, one might assume that learned irrelevance could have occurred after as few as 40 trials (as in the NPE group) or perhaps it had not yet occurred even after 200 trials (as in the PE group). As a consequence, learned irrelevance might be present or absent in both the PE and NPE groups, in which case a difference in performance between these two groups would not be anticipated. Learned irrelevance must, by definition, depend on prior experience of a dimension as irrelevant, but this does not imply that the size of the effect varies as a function of the amount of preexposure received. However, both learned irrelevance and latent inhibition are usually described as the loss of associability that a stimulus incurs when it is presented either in the absence of reward or uncorrelated with reward (Mackintosh, 1983). This description implies that associability is lost incrementally and that as a consequence the disruption of learning (in this case, ED shifting) should increase continuously with increasing amounts of preexposure. Prior experiments on latent inhibition support this prediction. The effects of latent inhibition found by De la Casa et al. (1993) in humans and by DeVietti et al. (1987) in rats both increased continuously (though not necessarily in a linear fashion) with preexposure. Latent inhibition was found after 6 minutes of preexposure in De la Casa et al.'s (1993) study, indicating that it is unlikely that there was too little preexposure in the PE group in this study (which experienced 10 minutes of preexposure). The 2 minutes of preexposure experienced by the NPE group in this study was well below the 6 minutes required to induce latent inhibition in De la Casa et al.'s (1993) study, so it is also unlikely that latent inhibition / learned irrelevance was present in the NPE group.

3.1.4.3 Is perseveration the more important mechanism in slowing ED shifting?

It was anticipated that learned irrelevance would retard ED shifting in this study, but it is possible that perseveration is a more important determinant of ED difficulty than learned irrelevance. In this case, only exaggerated preexposure of the relevant dimension would retard ED shifting. The current study was therefore followed up with

an experiment in which both the irrelevant and relevant dimensions were preexposed (Study 2). More subjects were tested in this followup study in order to increase statistical power, and therefore to maximise the likelihood of finding an effect of preexposure.

3.2.1 Introduction

Study 1 failed to find any effect of preexposure of the irrelevant dimension on ED shifting. However, it remains possible that preexposure of the relevant dimension could retard ED shifting by means of a mechanism such as perseveration. In order to test this hypothesis, a study that was similar to Study 1 was carried out, but in which a non-preexposed group was compared to a group that had experienced preexposure of both the relevant and irrelevant dimensions. Comparison of the results of this study with those of Study 1 allows examination of the role of preexposure of the relevant dimension in ED shifting. Preexposure of both dimensions also maximised the possibility of finding any effect of preexposure at all, by exaggerating the effects of both of the mechanisms that are postulated - perseveration and learned irrelevance. Thus, this study can answer the more general question of whether preexposure of dimensions can affect ED shifting when the amount of preexposure is similar to that used in latent inhibition experiments. This study also extends Study 1 by including manipulation checks that verify that subjects are successfully using selective attention in the experimental task. The ED shift and the subsequent ID shift were compared, as slower acquisition of an ED shift compared to a similar ID shift is considered to be due to the differential effect of selective attention on the two shifts, with selective attention benefitting an ID shift and retarding an ED shift.

3.2.2 Method

3.2.2.1 Subjects

34 subjects were tested, drawn from the undergraduate and postgraduate community at the University of St. Andrews. This group consisted of 21 females and 13 males with a mean age of 21 years (range 18-28). All subjects claimed to have intact colour vision. Informed consent was obtained from all subjects prior to the experiment, and subjects were paid at a rate of £3.50 per hour. Some of the data used in this study was originally collected for use in Study 2, as the NPE group in this study completed identical shifts to the NPE group in Study 2. Data from nine subjects in the NPE group in this experiment were previously used in Study 2.

3.2.2.2 Materials and procedure

The materials and procedure used in this study were identical to those used in Study 2.

3.2.2.3 Stimulus composition and types of shift

Stimuli and shifts were structurally identical to those used in Study 2.

3.2.2.4 Design

This study used an identical between-subjects design to that used in Study 2; subjects were randomised into 'preexposed' (PE) and 'non-preexposed' (NPE) groups, which were then compared to assess the effects of preexposure on ED shifting. The sequence of shifts and the allocation of dimensions to shifts used in this study was almost identical to that used in Study 2 (see Table 3.2). As in Study 2, the last two shifts that the subjects in both groups experienced were an ED shift and an ID shift that were identical across groups. Thus, the PE and NPE groups only differed in their experience of dimensions before the ED shift.

The NPE group experienced exactly the same dimensions, shifts and stimuli in the same order as the NPE group in Study 2. The first 160 trials experienced by the NPE group involved two dimensions that were not involved in the ED shift (see Table 3.2). This was followed by 40 trials of exposure to the dimensions involved in the ED shift to allow acquisition of the discrimination before the ED shift ('acquisition' shift, see Table 3.2). In contrast, the PE group experienced the dimensions involved in the ED shift from the beginning of the study, entailing 200 trials of preexposure. Thus, the PE groups experienced five times more preexposure of both the relevant and irrelevant dimensions than the NPE group before completing an identical ED shift.

As in Study 2, the number of ID shifts completed before the ED shift was also manipulated. Subjects were assigned a '1ID' or a '3ID' group that received either one or three ID shifts before the ED shift respectively. The total number of trials completed before the ED shift did not differ between these groups.

	NON-PREEXPOSED		PREEXPOSED	
	Relevant	Irrelevant	**Relevant**	Irrelevant
1 or 3 ID shifts	**Colour**	Background	**Shape**	Pattern
AQ shift	**Colour**	Background	**Shape**	Pattern
ED shift	**Shape**	Pattern	**Shape**	Pattern
ID shift	**Pattern**	Shape	**Pattern**	Shape
	Pattern	Shape	**Pattern**	Shape

Table 3.2: The left column of this table represents the shifts that were experienced by both groups in the order in which they were presented. The rest of the table depicts the dimensions that were experienced as relevant and irrelevant before and after these shifts for the NPE and PE groups.

3.2.2.5 Data analysis

Data from this study were analysed in a similar fashion to that used in Study 2. All of the dependent measures used in this study were the same as in Study 2; errors and trials to a criterion of ten consecutive correct responses were used, and decision times were calculated using the same *post hoc* RT and MT cutoffs. Life Tables analysis was performed on trials to criterion data for all between-groups comparisons; within-group analyses were carried out on errors to criterion data using the Wilcoxon Signed-Ranks test. This study differed from Study 2 due to the fact that shifts other than the ED shift were examined, and different exclusion criteria were used for subjects and individual data points.

Subjects were excluded from data analysis if they did not reach criterion (ten consecutive correct responses) before the ED shift. In previous studies (Studies 1 and

2), subjects were included if they scored eight or more correct out of the ten trials before the ED shift. However, when data from this study were inspected, it became apparent that some subjects had achieved eight correct responses out of the ten trials before the ED shift whilst having never attained a criterion of ten correct responses in a row. These subjects were excluded, as ten consecutive correct responses was thought to be good evidence of learning, and ED shifts were considered to be invalid of subjects had not acquired the rule before the shift.

The final three shifts completed by subjects were analysed in this study: the 'acquisition' (AQ) shift, the ED shift and the subsequent ID shift. As for the ED shift, the AQ and ID shifts were also only thought to be valid if subjects had acquired the rule preceding the shift. Data from these shifts were excluded if subjects had not reached criterion in the phase before the shift. This led to different numbers of data points being analysed at each shift; the number of points included in each analysis is recorded in the text.

3.2.3 Results

Figure 3.4 represents mean errors to criterion for the last three shifts that the PE and NPE groups completed. As in Study 2, the absolute level of ED shift errors ensures that this measure should not be affected by floor or ceiling effects. Visual inspection of the results for the ED shift indicates two main findings; more errors are committed at the ED shift that at the AQ and ID shifts, and there is no effect of preexposure on errors committed at the ED shift.

3.2.3.1 Relative difficulty of ED and ID shifting

If an ED shift is found to be more difficult than a comparable ID shift, this is usually taken as evidence that subjects are selectively attending to the relevant dimension before both shifts; such attention facilitates the performance of ID shifts but retards the performance of ED shifts (Slamecka, 1968). In this study, subjects made more errors at the ED shift than at the ID shift (n for ED = 34, for ID = 31; $z = 2.68$, $p < 0.01$).

3.2.3.2 Effect of preexposure on the ED shift

However, despite this evidence for the presence of selective attention, the difficulty of the ED shift was not altered by preexposure. The PE and NPE groups acquired the ED shift at the same rate (Wilcoxon (Gehan) $p > 0.5$), as indicated by the survival function (Figure 3.5). No effect of preexposure was seen in the RT data; decision times did not differ for PE and NPE groups ($t(32) = 0.25$, $p > 0.5$). The greater difficulty of the ED shift compared to the ID shift is thought to represent the difficulty of shifting attention between dimensions, as in all other respects ED and ID shift make the same cognitive demands; these data indicate that ease of shifting attention between dimensions is not affected by preexposure of those dimensions.

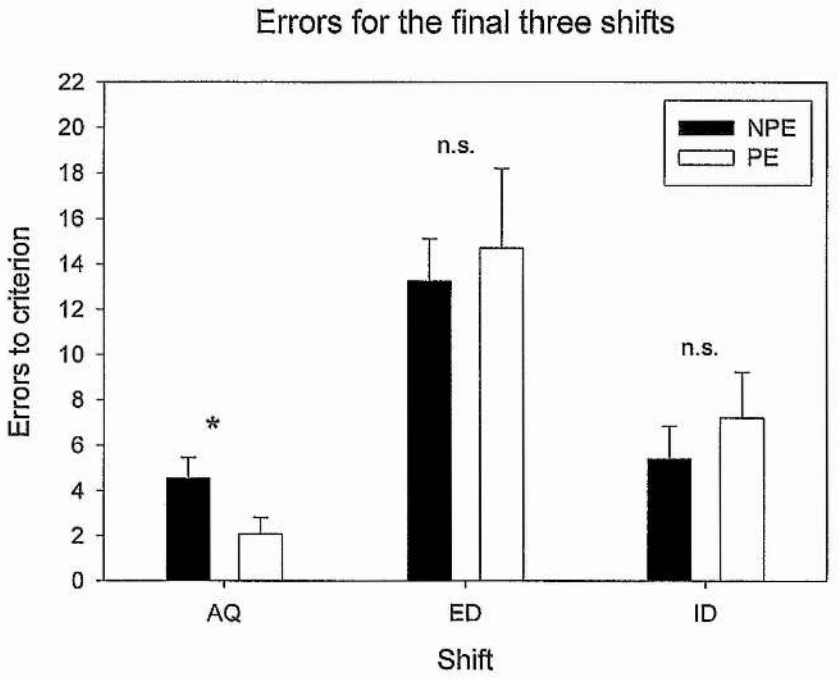


Figure 3.4: Mean errors to criterion for the last three shifts completed for the PE and NPE groups. '' signifies a significant difference between PE and NPE groups at $p > 0.05$. Error bars represent the standard error of the mean.*

3.2.3.3 Effect of preexposure on the AQ shift

Further evidence for the presence of selective attention comes from examination of performance of the AQ shift. Subjects in the PE group acquired this shift more rapidly than those in the NPE group ($n = 24$; Wilcoxon (Gehan) $p < 0.05$, Figure 3.5 Panel B). The PE and NPE groups were both required to shift to the same exemplars and stimuli at this shift, but differed in the dimensions that they were responding to before the shift (see Table 3.2). The PE group were responding to the same dimensions before and after the shift, making the AQ shift an intradimensional shift. The same dimensions were relevant and irrelevant before and after the shift, and subjects were only required to acquire a response to a new exemplar of a dimension that was already relevant. In contrast, the dimensions that the NPE group encountered after the AQ shift were completely different to the dimensions that they had been responding to before the shift. Subjects were required to learn the new rule with no indication as to which of the two new dimensions might be relevant. The fact that the PE group was already responding

to the dimension that was relevant after the AQ shift explains this group's superior performance at that shift.

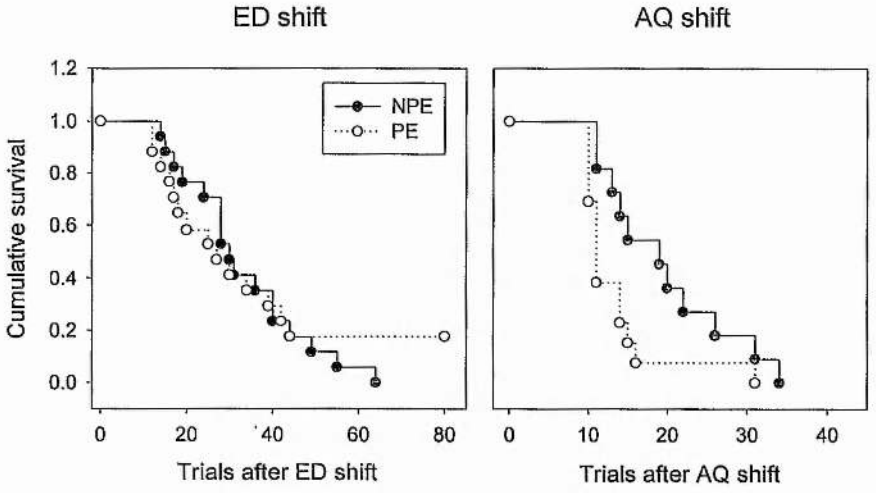


Figure 3.5: Survival functions for the PE and NPE groups at the ED and AQ shifts. Where the curve does not reach zero at 80 trials, this signifies that some subjects failed to reach criterion.

3.2.3.4 Effect of prior ID shifting on ED shift performance

In contrast to Study 1, ED shift performance was affected by the number of ID shifts completed before the ED shift. A greater number of prior ID shifts facilitated performance; the 3ID group solved the ED shift more rapidly than the 1ID group (Wilcoxon (Gehan) $p < 0.01$).

3.2.3.5 Summary

Evidence from both the AQ shift and the comparison of ID and ED shifts shows that subjects could selectively attend to one of two dimensions in this study. However, comparison of the PE and NPE groups indicates that preexposure did not affect the strength of this selective attention.

3.2.4 Discussion

This study recorded the consequences of increased preexposure of the relevant and the irrelevant dimension for the performance of an ED shift. Increased preexposure did not slow learning or increase decision times at the ED shift. It was hypothesised that an increase in preexposure would cause an exaggeration of the mechanisms of learned irrelevance and perseveration, which are thought to govern selective attention to a perceptual dimension. These mechanisms are thought to retard acquisition of the new rule after an ED shift by maintaining attention to the previously relevant dimension and preventing a shift of attention to the newly relevant dimension. However, ED shift performance was not disrupted by increased preexposure, despite evidence from the examination of other shifts that a strong effect of selective attention was present. It was not the case that ED shift performance was generally insensitive to the effects of prior learning, as greater ID shift experience prior to the ED shift facilitated ED shift performance; the manipulation of preexposure alone failed to change ED shift performance. Thus, it appears either that learned irrelevance and perseveration were not affected by the manipulation of preexposure that was used, or that these mechanisms do not play a significant role in determining the difficulty of an ED shift.

ED shifts were originally used to demonstrate that organisms could pay attention to perceptual dimensions as well as acquiring responses to discrete stimuli (Slamecka, 1968). More recently, ED shifts have been used to investigate impairments of attentional shifting in clinical populations (e.g. Owen et al., 1991, 1992; Elliott et al., 1995; Lange et al., 1995). It is clear that humans can shift their attention between perceptual dimensions, but despite the large amount of research on ED shifting, the cognitive mechanisms underlying such shifts have not been characterised. A plausible account was given by Owen et al. (1993b), and this study aimed to extend Owen et al.'s (1993b) hypothesis by inducing learned irrelevance and perseveration in healthy young adults.

3.2.4.1 Choice of method

The method that was used to enhance learned irrelevance and perseveration in this study was based on that used to induce latent inhibition. As learned irrelevance is a better-understood mechanism than perseveration, and closely resembles latent inhibition, it is the failure of preexposure to enhanve learned irrelevance in this study that must principally be accounted for. The decision to use a preexposure paradigm similar to that

used to induce latent inhibition was based on three factors. First, dimensional shifting paradigms involving ED shifts closely resemble the paradigms used to induce latent inhibition in adult humans. Second, subjects' accounts of ED shifting strongly imply that some mechanism of 'learned inattention' (Lubow 1997) - either learned irrelevance or latent inhibition - is at work. Subjects reported that a failure to consider the newly-relevant dimension was a major factor in slowing their acquisition of the ED shift; typical statements included "I just wasn't looking at the stripes [the newly-relevant dimension] after the shift", and "it's quite easy when you know what you're looking for - as soon as you see the difference in the stripes, you've got it". The final factor in selecting a latent inhibition-like paradigm was that recent animal work suggests that learned irrelevance effects in animals may in fact be misattributed latent inhibition effects - there may be no separate mechanism of 'learned irrelevance' (Bonardi and Hall, 1996). If this were the case, any effect of 'learned irrelevance' in this study should act in an identical fashion to latent inhibition. Similarly, in the literature on latent inhibition on humans, Lubow (1997) now refers to latent inhibition as an instance of a more general concept of 'learned inattention', which would subsume both learned irrelevance and latent inhibition.

In this study, subjects were able to selectively attend to a dimension, as shown by the greater difficulty of ED shifts over ID shifts, and the superiority of the PE group in performing the AQ shift. As described above, subjects reported the effects of 'learned inattention' on their ED shift performance, yet increased preexposure did not change ED shift performance and therefore did not enhance learned irrelevance. As in Study 1, either an account must be given of why the levels of preexpoosure chosen were not appropriate to detect learned irrelevance, or it must be concluded that learned irrelevance was not a significant mechanism in determining ED shift performance in this study.

3.2.4.2 Was the correct amount of preexposure used?

De la Casa et al. (1993) found a maximal effect of latent inhibition after 15 minutes of preexposure, and Graham and McLaren (1998) induced a latent inhibition effect in students with 20 minutes of preexposure. From these instances, it could be argued that the level of preexposure used in the PE group in this study was too low, being only around 10 minutes. However, a brief survey of the literature show that latent inhibition to visual stimuli can also be induced by relatively brief preexposure - for example, 15 trials (Salzman et al. 1993, cited in Lubow 1997), 30 trials (Dressler et al., cited in

Lubow 1997) and 80 trials (Lubow et al. 1992, Expt 2, cited in Lubow and Gewirtz 1995), compared to the 200 trials used in this study. These studies indicate that it is unlikely that the PE group in this study was 'under-preexposed', but they raise another possibility; perhaps a latent inhibition effect was present in both the NPE (40 trials of preexposure) and the PE (200 trials of preexposure) groups. 40 trials of preexposure may have delivered sufficient preexposure for the creation of a latent inhibtion effect in the NPE group; if a latent inhibition effect were present in both the NPE and PE groups, a difference in ED shift performance between these groups would not be expected. If this were the case, a comparison of the NPE group with a group that had received even less preexposure might reveal poorer ED shift performance in the NPE group due to latent inhibition in the NPE group. However, a consideration of the literature on the effects of overlearning on ED shifts indicates that it is unlikely that there was an effect of latent inhibition in the NPE group. Subjects in the NPE group experienced on average 24 trials between reaching criterion at the AQ shift and experiencing the ED shift. Studies of overlearning indicate that these 'extra' 24 trials are more likely to have facilitated ED shifting due to an overlearning effect than to have retarded ED shifting through latent inhibition. Wolff's (1967) excellent review of the overlearning literature cites studies that show that between 10 and 48 trials of overlearning facilitate ED shifting in college students. One study in particular shows that the level of preexposure experienced by the NPE group in this study should aid ED shifting. Guy et al. (1966, cited in Wolff 1967) showed that 20 trials of overlearning (compared to 24 in the NPE group) after a criterion of 10 consecutive correct respsonses (as in the NPE group) facilitated ED shifting in college students. This represents a beneficial effect of preexposure in an experimental condition that is highly similar to that of the NPE group and it therefore indicates that latent inhibition is unlikely to have been present in the NPE group. Thus, there was sufficient preexposure in the PE group to generate latent inhibition, but the NPE group should have been free from latent inhibition. Therefore, the choice of inappropriate levels of preexposure does not appear to account for the absence of a learned irrelevance / latent inhibition effect in this study.

3.2.4.3 Implications of the overlearning literature for learned irrelevance and perseveration

It is informative to consider the implications of the literature reviewed by Wolff (1967) for the proposed mechanisms of learned irrelevance and perseveration. Overlearning has consistently been shown to facilitate ED and reversal performance, and in one study

it improved performance on ED and ID shifts that used a 'total change' of exemplars (Eimas, 1966). It is impossible to account for a facilitation of ED *and* ID shifting in terms of changes in the strength of learned irrelevance and/or perseveration. Increases or decreases in the strength of these mechanisms can only have opposing effects on the difficulty of ID and ED shifts. Indeed, the attractiveness of the learned irrelevance / perseveration hypothesis was that it could account for deficits in ED shift performance in the presence of relatively intact ID shift performance in patients with frontal lobe damage and Parkinson's disease (Owen et al. 1991, 1992). Consideration of the overlearning literature complicates any account of mechanisms underlying ED shifting, as it appears that ED and ID shifts can be influenced in the same direction by a simple experimental manipulation (overlearning) or separately by brain damage (selective ED shifting deficits).

3.2.4.4 Implications of this study combined with the overlearning literature

When the overlearning literature is considered alongside the results of this study, it becomes clear that neither learned irrelevance nor latent inhibition can affect ED shifting. Latent inhibition is usually described as the loss of associability that a stimulus incurs each time it is presented without consequence. Learned irrelevance, if it is in fact a separate effect, is a similar reduction in associability that occurs whenever a stimulus is presented uncorrelated with reward. Both effects are incremental. Neither of these effects can account for an improvement of ED shifting performance with small amounts of preexposure (as shown by the overlearning literature) that is followed by stable ED shift performance despite further preexposure (40 vs. 200 trials of preexposure in this study). This study combined with the overlearning literature shows that at no point does increased 'preexposure as irrelevant' retard ED shifting; indeed, at brief preexposure levels ED shifting is facilitated. Learned irrelevance can therefore have no effect on two-dimension ED shifting tasks.

3.3.1 Introduction

The results of Study 3 cast doubt on the hypothesis that the mechanisms of learned irrelevance and perseveration affect ED shifting performance. Whilst Study 3 cannot 'prove' that learned irrelevance does not affect ED shifting (indeed, neither can any study), there is now a strong case for considering other accounts of the parkinsonian ED shifting deficit. Any new hypothesis must be able to account for the presence of an ED shifting deficit in the context of unimpaired ID shifting, as this is the pattern of results most commonly seen amongst patients with PD (Downes et al. 1989, Lange et al. 1992, Owen et al. 1993b, Study 1). One study has found impaired ID shifting amongst patients with PD (Owen et al., 1992), but even in this study the magnitude of the deficit seen at the ID shift was smaller than that present at the ED shift in two out of three groups of patients with PD. In general, a selective ED shift deficit is seen in patients with PD rather than a general problem of 'shifting aptitude' (Cools et al. 1984); the cause of this deficit is currently unknown.

Channon et al. (1993) carried out a study of hypothesis-testing and discrimination learning that revealed a deficit amongst patients with PD that could account for a selective ED shifting deficit. The experimental task used in Channon et al.'s (1993) study involved working out a 'rule' in a similar fashion to the requirements of an ED shift, but the overall methodology was sufficiently different to that used in Studies 1 to 3 to require detailed description. Channon et al. (1993) modified a task developed by Levine (1970) for use in their study. Subjects were required to work out a rule that defined one of two stimuli as correct, as in a classical dimensional shifting task. The rule was always to choose one exemplar taken from one of four dimensions; the dimensions were size (large / small), letter (A / B), colour (black / white) or position (right / left). Channon et al. (1993) were able to observe which rule subjects were responding to by using a 'blank trials' method developed by Levine (1970). This technique involves requiring subjects to respond for a number of trials without feedback. During this 'no-feedback' period, the location of stimuli are arranged such that response to a particular rule will be associated with a unique and identifiable pattern of responses to the left or right stimulus. For example, if subjects were responding to the rule 'black', their responses might be left - right - left - right; if they were responding to the rule 'large' their pattern of responses would be quite different but also unique and identifiable. Observation of these response patterns allows investigators to analyse how subjects respond to the information received on 'feedback'

trials, particularly on those trials at which feedback is negative. Channon et al. (1993) required patients with PD and healthy elderly controls to complete eight discrimination problems whilst analysing their responses to feedback using the 'blank trials' method.

If healthy subjects receive negative feedback whilst responding to the rule 'small', they will reject this rule and test another hypothesis (Levine, 1970). However, as well as rejecting the rule 'small', they will reject rules relating to other aspects of the stimulus that incurred negative feedback. For example, if a 'small' 'black' 'A' on the 'right' was shown to be an incorrect stimulus, a subject who was hypothesis-testing perfectly would reject all of these rules. In practice, only a minority of subjects show this 'perfect' hypothesis-testing, but most subjects reject more than one rule after a negative feedback trial (Levine, 1970). Channon et al. (1993) showed that the response of patients with PD to negative feedback was subtly different to that of controls. After a negative feedback trial, patients correctly rejected the rule that they had been responding to as often as controls. However, patients were more likely to go on to test one of the other rule that had been ruled out by negative feedback. For example, if patients were responding to the rule 'large', and received negative feedback after pressing a 'large' 'black' 'A' on the 'right', they would correctly reject the rule 'large' but would be more likely than controls to go on to test either the rule 'black', 'A' or 'right' which had also been shown to be incorrect by the feedback. This result implies that patients with PD are less able to gather or use information about 'untested' dimensions than controls.

If the deficit seen in Channon et al.'s (1993) study were present in a two-dimension ED shifting task, it would impair performance at an ED shift but not at an ID shift. Immediately after both and ED and an ID shift, subjects report testing rules that relate to the dimension that had been relevant before the shift. In the case of an ID shift, the dimension that subjects are testing will be the dimension from which the new correct rule is taken. In this case, the ability to gather information about the other, 'untested' dimension does not affect performance as this dimension is and remains irrelevant. In contrast, this ability can aid ED shifting. Immediately after an ED shift, subjects will test a rule from the dimension that was relevant before the shift. However, the new correct rule will not be taken from this dimension; rather, it will be taken from the other, 'untested' dimension. Therefore, the ability to gather information about this 'untested' dimension will be an advantage, as it allows subjects to acquire information about the newly-relevant dimension. If patients with PD were less efficient at gathering

information from the 'untested' dimension than controls, this would lead to impaired performance at the ED shift, but unchanged performance at the ID shift.

In this study, an attempt was made to induce a 'deficit' in normal subjects similar to that found by Channon et al.'s (1993) in patients with PD. An experimental manipulation was created that aimed to reduce subjects' ability to acquire information from the 'untested' dimension. If the manipulation used leads to a selective ED shifting deficit in young subjects, this implies that the deficit found be Channon et al. (1993) is a good candidate for the cause of the parkinsonian ED shifting deficit. If, however, the manipulation impairs both ID and ED shifting it is unlikely to be the cause of the problems displayed by patients with PD. This study compared one group that completed ID and ED shifts in normal conditions ('Standard' group) with another in which the ability to acquire information about the 'untested' dimension was restricted by the experimental manipulation ('Mask' group).

The ability to acquire information about the 'untested' dimension was restricted in this study by reducing stimulus display time and including a masking stimulus. This manipulation was developed on the basis of subjects' descriptions of the process of hypothesis-testing after a shift. Subjects report that they explicitly test only one hypothesis at a time after a shift, and there is evidence for this from Levine's (1970) studies. However, it is clear that subjects can reject more than one rule when they encounter negative feedback, as described by Levine (1970). This apparent contradiction is resolved by considering subjects' descriptions of the process of hypothesis-testing. After an ED shift, a subject might test the rule 'green' by pressing a green square. If a subject receives negative feedback as a consequence of this response, s/he will reject the rule 'green' and will also recall the stimulus (a green square) and from this reconstruction of the stimulus will reject both the rules 'green' and 'square'. The manipulation used in this study acts at two stages of this hypothesis-testing process. Shortening display time forces subjects to direct their attention towards the dimension that they believe to be relevant. The stimuli used in this study were not easy to discriminate, and if attention was not directed towards the salient features of these stimuli then discrimination within the brief display time allowed would have been impossible. Thus, shortened display time minimised the time available to inspect the 'untested' dimension and derive information from it by forcing subjects to attend to the dimension that they assumed was relevant. The introduction of a masking stimulus was intended to disrupt the reconstruction of a stimulus after feedback that allowed retrospective conclusions to be drawn about the untested dimension. The masking

stimulus resembled the salient stimuli, and was intended to decrease the clarity of any mental reconstruction of the stimuli.

To summarise, young subjects were tested in this study to test the hypothesis that masking and restricted display time would lead to a selective ED shifting deficit, rather than an global slowing of ED and ID shifting due to increased task difficulty. If such a selective deficit were found to be present, it would add plausibility to the hypothesis that the deficit found by Channon et al. (1993) is also the deficit that causes the selective impairment of ED shift performance seen in patients with PD.

3.3.2 Methods

3.3.2.1 Subjects

40 subjects were tested (29 females and 11 males), all of whom were undergraduates or postgraduates at the University of St. Andrews. The mean age of these subjects was 21.9 years, and all claimed to have intact colour vision. Subjects were paid at a rate of £3.50 per hour, and gave informed consent before taking part in the study. Some data used in this study was originally collected for Study 3, as the PE (3ID) group from Study 3 and the 'Standard' group in this study (see 'Design' section, below) completed exactly the same shifts using the same stimuli. Data from 9 subjects in the 'Standard' group were previously used in the PE group of Study 3.

3.3.2.2 Materials and procedure

Subjects were split into two groups in this study, as described in the 'Design' section below. Subjects in the 'Standard' group experienced the same procedures that were used in Study 3. However, the experimental procedure differed for the 'Mask' group. Stimuli were only displayed for 750 ms to the Mask group, in contrast to a stimulus display time of 1000 ms for the Standard group. For the Standard group, stimuli disappeared after the stimulus display time had elapsed, leaving empty 'stimulus boxes'. In the Mask group, each stimulus was instead replaced with a masking stimulus (described in the 'Stimulus composition' section below) at the end of the stimulus display time. The masking stimulus persisted until subjects made a response by touching the screen; at that point, the screen was cleared and the feedback stimuli were displayed as in Study 3. The instructions given to the subjects in the mask group were altered to reflect the changed experimental procedure; instructions for the Standard group were as for Study 3. The instructions for the Mask group are reproduced here in full:

"This is the main program. The format is basically the same as the practice program that you've just completed. You press down the space bar and two shapes appear - one is right, one is wrong, and you job is to press the one that is correct. There is a rule telling you which shape is correct. This rule will be something like 'always press the red shape', or 'always press the triangle'. A number of different rules are possible, as the shapes can differ in four ways. They can be different colours or different shapes, they can have different striped patterns on them, or there can be a different colour behind the shape. The shapes can differ from each other in more than one way at once, so they may look complicated. However, there is always one simple rule that tells you which shape is correct. Your job is to work out this rule and press the correct shapes.

Although I have said that the rule is simple, this does not mean that the task is easy. The differences between the shapes may be very subtle, and quite hard to see. For example, there may be a small difference in the outline of the shape, or a slight difference in the shading of the colour. As a consequence, I suggest that at first you hold down the space bar for long enough to give yourself time to look at the shapes. After a few trials, you will be confident that you know what the rule is, and then you can speed up and treat this as a reaction time task.

There are two additional factors that make this task difficult. First, the rule will change as you go through the task. For example, the shapes will change and you will realise that the old rule can no longer apply. If this happens, I suggest that you slow down and give yourself time to look at the shapes, and when you are confident that you know the rule you can speed up again and treat it as a reaction time task. The second difficulty is that you only ever get less than a second to look at the shapes. That is, after the shapes have been on the screen for just under a second they will disappear, and they will be replaced by another shape that is irrelevant and has nothing to do with the task. This shape has a red background to help you to distinguish it from the shapes that are important for the task. If you have not decided which shape to press by this time, you will have to press one of the boxes anyway. One final point - there are no hidden rules. The rule always has something to do with the shapes, colour, stripes or background colour, although it might be hard to see. You can start the test now."

3.3.2.3 Stimulus composition and types of shift

The stimuli used in this study were taken from the set used in Studies 2 and 3, with one exception. One additional stimulus was used in this study; a 'masking' stimulus, experienced only by the Mask group. This was a complex stimulus created by the superposition of a number of exemplars from the dimensions 'shape' and 'pattern' that were used in this study (the masking stimulus is depicted in Figure 3.6 alongside a typical relevant stimulus). The mask was designed to be visually similar to the relevant stimuli, as the replacement of a stimulus with a mask after a brief display period was intended to retard or confuse the memory of that stimulus.

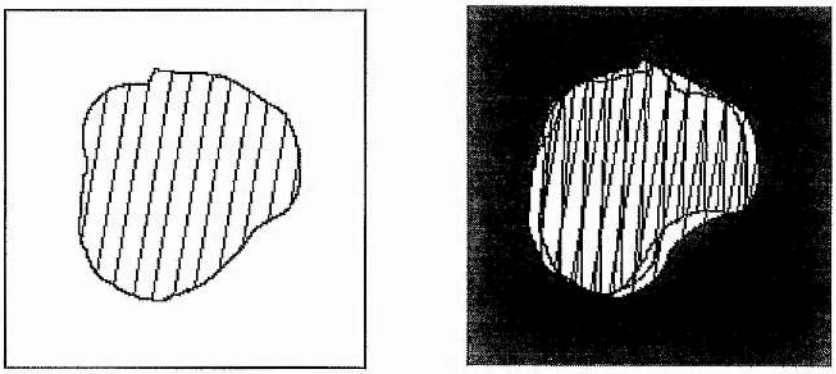


Figure 3.6: The shape on the left is a sample of a typical stimulus from this study. The shape on the right is the masking stimulus - the background colour for this stimulus (here in grey) was red.

3.3.2.4 Design

This study set out to examine the consequences of decreasing display time and introducing a masking stimulus for the performance of both ED and ID shifts. Subjects were therefore randomised into two groups, a 'Mask' group and a 'Standard' group. These two groups experienced exactly the same stimuli and shifts in the same order as each other, the only difference between the groups being shortened stimulus display time and the presence of a masking stimulus in the Mask group.

Subjects experienced the same shifts and stimuli as those in the PE group of Study 3 who experienced three ID shifts before the AQ shift ('3ID' group). Table 3.3 depicts the order of the shifts. Two dimensions, 'shape' and 'pattern' were present throughout the study, 'shape' being relevant at the beginning. After the initial discrimination, subjects experienced four ID shifts in which 'shape' remained the relevant dimension. 40 trials elapsed after each ID shift. These shifts were followed by an ED shift in which 'pattern' became the relevant dimension; 80 trials occurred after the ED shift, which were followed by a final ID shift.

3.3.2.5 Data Analysis

Selection of an appropriate ID shift

As the principal goal of this study was to compare the effects of masking upon ED and ID shifts, it was thought to be important to select an ID shift that involved shifting to stimuli of comparable difficulty to those present after the ED shift. Most ED shifts are, of course, harder than ID shifts due to the influence of subjects' experience before the

shift. However, the stimuli present after both ED and ID shifts can be considered without reference to experience before the shift, and it is clear that some stimuli will be easier to discriminate than others. For example, if the stimuli present after an ID shift have highly salient differences they will be relatively easy to discriminate and as a consequence hypothesis testing will be relatively simple. However, if such stimuli are quite similar, with only subtle differences between exemplars, hypothesis testing will be relatively difficult; exemplars may be confused with each other and there will be greater need to selectively attend to salient features. As a shift's absolute difficulty can vary as a consequence of stimulus difficulty as well as its nature (ID or ED), it was felt to be important that this study did not compare an ID shift involving 'easy' stimuli with an ED shift involving 'difficult' stimuli. If masking caused a selective ED shift deficit in such a case, it could be argued that this result could be explained by reference to task difficulty; the increase in task difficulty caused by masking might have further complicated the 'difficult' ED shift by left the 'easy' ID shift undisturbed. It was therefore decided to select an ID shift that was appropriate for comparison with the ED shift before the experiment began. Subjects completed five ID shifts in this study. It would have been easy, but misleading, to 'trawl' through the data for these ID shifts until an ID shift was found that most closely confirmed the experimental hypothesis. By choosing an ID shift before the study began on principled grounds it was hoped to avoid this inappropriate selection of data.

Shift	**Relevant**	Irrelevant
	Shape	Pattern
ID		
	Shape	Pattern
ID		
	Shape	Pattern
'Comparable' ID		
	Shape	Pattern
ID		
	Shape	Pattern
ED		
	Pattern	Shape
ID		
	Pattern	Shape

Table 3.3: Order of shifts and dimensions for both groups

An appropriate ID shift was selected by examining data from the PE group from Study 3. All of the subjects from this group had completed four shifts that were associated with the same order of presentation and the same exemplars as shifts in this study, and all of these shifts were analysed; two ID shifts before the ED shift, the ED shift itself, and one ID shift after the ED shift. There are no direct measures of the 'difficulty' of the stimuli present after a shift,so in order to select an ID shift that was comparable to the ED shift, a response time measure was calculated that was thought to index the absolute discriminability of the stimuli present after these shifts. After learning the rule after a shift, a subject should respond to easily discriminable stimuli with a more rapid response time than to 'difficult' stimuli that are hard to discriminate. To reflect this, the response time measure used was the mean response time for the eight trials after a subject reached criterion. This measure was calculated for the three ID shifts and the ED shift. For the ED shift, the group mean response time was 691 ms, compared to ID shift response times of 695, 579 and 803 ms. The ID shift with a response time of 695 ms was therefore chosen as the most appropriate comparison for the ED shift. This ID shift was the second shift before the ED shift, and is the shift used in all comparisons with the ED shift.

Dependent measures

Subjects were included in the analysis only if they reached criterion at the shifts that preceded both the ID and ED shifts. The dependent measures that were used in Studies 2 and 3 were also used in this study; errors and trials to criterion were calculated for all shifts and decision times were calculated for the ID and ED shifts. Two additional dependent measures were taken that were not used in Study 3. As masking was expected to have a differential effect on ID and ED shifts, a 'difference' score was calculated for each group by subtracting ID errors from ED errors. Also, 'discrimination times' were calculated for the ID and ED shifts. This measure was the same as that used above (see 'Selection of an appropriate ID shift' above) to index the absolute discriminability of stimuli - the mean response time was calculated for the eight trials after subjects reached criterion. Data were analysed in a similar fashion to Study 3. Wilcoxon (Gehan) statistics were calculated from trials to criterion data for between-group comparisons of performance at the ID and ED shifts. Difference scores were compared using Mann-Whitney tests, and within-group comparisons of errors to criterion scores were made using the relevant nonparametric tests (Wilcoxon or Friedman).

3.3.3 Results

3.3.3.1 Manipulation checks

Errors to criterion were calculated for all of the ID shifts in this study to test the hypothesis that some ID shifts might be more difficult than others due to the presence of less discriminable stimuli after the shift. The four ID shifts that were carried out before the ED shift were analysed; despite the fact that the dimension 'shape' was relevant before and after all of these shifts, performance differed between them (χ^2 (3) = 62.63, p < 0.001). The errors to criterion data for these shifts are plotted in Figure 3.7 alongside the 'discrimination times' for the same shifts. The more difficult shifts appear to be associated with longer discrimintion times, as predicted. However, it appears that the ID shift that was chosen for comparison with the ED shift involved stimuli of comparable difficulty to the ED shift. Discrimination times were calculated for the ID and ED shifts to confirm that the absolute difficulty of the stimuli used in the two shifts did not differ; discrimination times did not differ for the ED and the comparable ID shift ($t(26) = 1.72$, $p > 0.05$). The ED shift was, as anticipated, more difficult than the ID shift; subjects generated significantly more errors at the ED shift ($Z = 4.78$, $p < 0.001$) indicating a strong effect of selective attention.

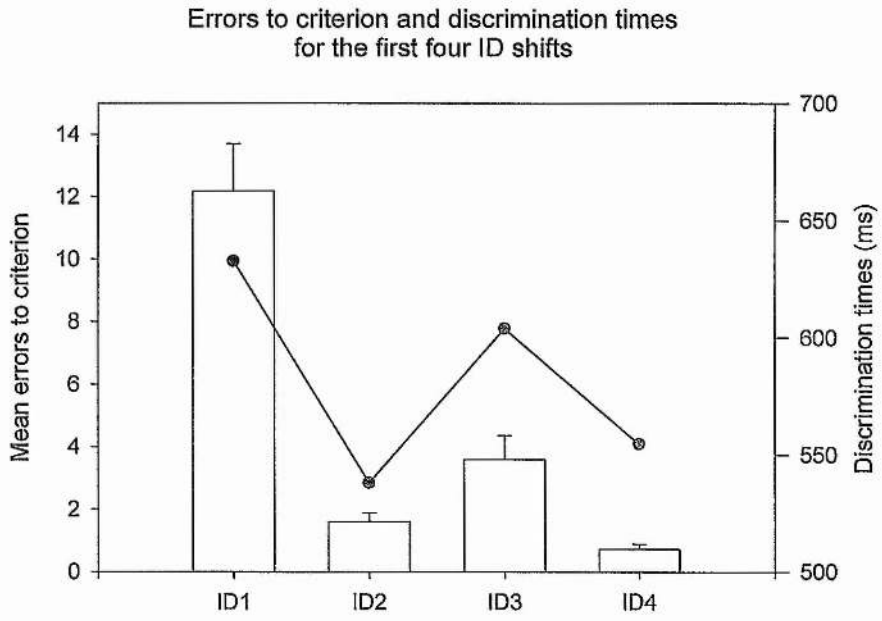


Figure 3.7: Errors to criterion (left axis, bar chart) and discrimination times (right axis, line and scatter plot) for the first four ID shifts (labelled ID1 to ID 4 in order of presentation).

3.3.3.2 Effects of Masking

Learning to criterion

Figure 3.8 shows mean errors to criterion for the Mask and Standard groups at the ED shift and at the 'comparable' ID shift (see Section 3.3.2.5). Inspection of this graph suggests that masking did have an effect on shifting performance, though it appears to have acted by slightly reducing the number of errors generated at the ID shift and by elevating errors at the ED shift. This is confirmed by the ID-ED difference scores, which are significantly larger for the Mask group than for the Standard group ($U = 136.5, p < 0.05$, one-tailed). Although masking had a differential effect on the ED and comparable ID shifts, subjects in the Standard and Mask groups did not differ in the rate at which they solved either the ED or the ID shift (Wilcoxon (Gehan) $p > 0.1$ for both shifts). Survival functions for the ED and ID shifts are plotted in Figure 3.9.

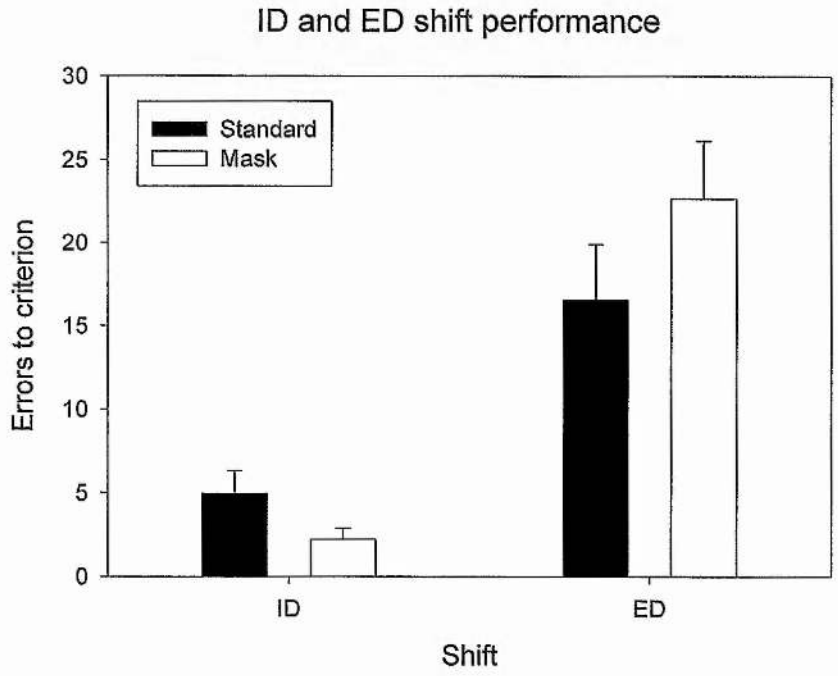


Figure 3..8: Mean errors to criterion for the ED and comparable ID shifts for the Standard and Mask groups. Error bars represent the standard error of the mean.

Decision times

Decision times were also calculated for all subjects, and can be used to assess the differences in 'thinking time' between Standard and Mask groups. However, there is an *a priori* reason to doubt the validity of this measure. Decision times for the ED shift for

both groups were over 800 ms, but stimulus display time for the Mask group was 750 ms; restricted display time in the Mask group might have spuriously reduced decision times for this group. Despite these reservations, decision times did not differ between the Standard and Mask groups for either the ID or ED shift. Presumably, this indicates that subjects in the Mask group made their decisions within the 750 ms display period at the ED shift, the extra time before response initiation being accounted for by movement initiation processes.

In summary, the effect of masking in this study was to increase the difference in errors between the ED and ID shifts by slightly increasing errors at the ED shift and decreasing errors at the ID shift.

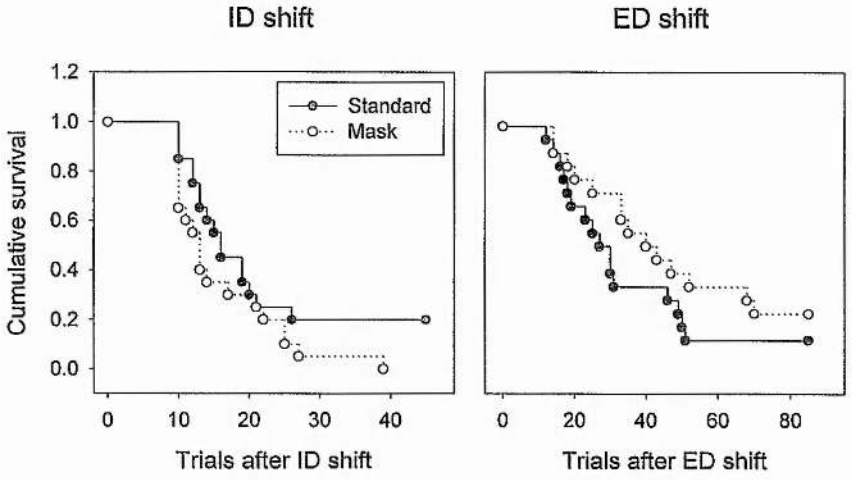


Figure 3.9: Survival functions for the Standard and Mask groups for the ED and comparable ID shifts. Curves which do not reach zero represent groups in which not all subjects reached criterion.

3.3.4 Discussion

This study compared the effect of masking and restricted display time on the performance of ED and ID shifts. Masking increased the difficulty of the ED shift relative to the ID shift. However, this effect was achieved both by a small reduction of errors at the ID shift and by increased errors at the ED shift. Masking influenced performance at the ID and ED shift in opposite directions, despite the fact that the ID shift was chosen such that the stimuli present after this shift were of a similar difficulty to those present after the ED shift. The effect of masking found in this study was small, but it influenced ID and ED shifts in a way that was predicted, rather than causing a global increase in the difficulty of both the ID and ED shifts. The change in performance caused by masking resembled that which would be caused by an increase in selective attention; a speeding of ID shifting and slowing of ED shifting.

3.3.4.1 Influence of stimulus 'difficulty' on shift performance

Comparisons of ID and ED shift difficulty have been important in the literature on dimensional shifting and selective attention. The greater difficulty of ED shifting over ID shifting was originally used to demonstrate that subjects could attend to perceptual dimensions as well as discrete stimuli (Slamecka, 1968). More recently, the presence of selective ED shifting deficits amongst patients with PD has been used to argue that these patients do not have general problems with shifting or hypothesis testing, as ID shifting is preserved, but rather that their deficit is confined to shifts of selective attention. However, this study has shown that the absolute difficulty of different ID shifts can vary substantially, even when these shifts involve the same dimensions (see Figure Q.1). The difference in the difficulty of these shifts appears to be due to the differing 'difficulty' of the stimuli that subjects are required to discriminate after the shift. Stimuli that are more similar, and therefore less discriminable and more easy to confuse, are more difficult and subjects will make more errors at shifts that involve such stimuli. It follows that ED shifts should be compared with ID shifts that involve stimuli of similar difficulty, to avoid confounding an ID-ED difference that is due to selective attention to one that is caused by stimulus difficulty.

3.3.4.2 Effect of masking

In this study ED shift performance was compared to performance on an ID shift that had been selected as an appropriate comparison shift before the study began. Masking

had opposite effects on these comparable shifts, as opposed to increasing the difficulty of both shifts.

A significant deficit in ED shift performance was not found in the Mask group, although the results are supportive of the experimental hypothesis in that masking had differential effects on ID and ED shifts. It is possible that an experimental manipulation that caused a greater increase in the ID-ED difference could magnify this difference to the extent that a statistically significant ED shift deficit would be seen in the Mask group. There is reason to believe that the manipulation used in this study was not strong. For example, subjects' decision times indicate that selection of a response after the ED shift occurred well within the display time of 750 ms. Also, the presence of a masking stimulus did not ensure that subjects could not engage in mental reconstruction of the stimulus after feedback. The masking stimulus should have disrupted any visual afterimage of the stimuli, but stimulus reconstruction could only be reliably prevented if subjects were required to engage in an attentionally demanding task after a response was made. A future experiment that compared display times that differed by more than 250 ms - for example, 700 ms vs. 1400 ms - and required subjects to complete a distracting task after making a response might create a larger effect and achieve a significant selective ED shift deficit in the masking group.

3.3.4.3 Comparison with Study 1

Further supportive evidence for the experimental hypothesis can be obtained by comparing the results of this study with those of the healthy young subjects in Study 1. The results for this group were not reported but can be briefly summarised; 11 students were tested whose mean errors to criterion at both ID and ED shifts were all very close to 1. This can be compared with ID and ED shifts errors scores of 4 and 20 respectively in this study. The principal methodological differences between Study 1 and this study were that Study 1 used simple, easily discriminable stimuli and imposed no restriction on stimulus display time; the subjects for both studies were drawn from the same population (undergraduate and postgraduate students). Thus, restricting display time and using 'difficult' stimuli in this study elevated ID shift errors from 1 to 4 and ED shift errors from 1 to 20. It can be argued that restricting stimulus discriminability and display time in this study acted in the same fashion as masking was thought to act, and affected ED shifting much more than ID shifting. The first stage in this argument is to explain why there was no ID-ED difference in Study 1.

Young subjects in Study 1 were performing at a 'ceiling' for both ID and ED shifts. It appears that they could carry out perfect hypothesis-testing at both types of shift. For example, a subjects could test the rule 'green' by pressing a green square that s/he had studied for an unlimited time; after receiving negative feedback s/he could easily reject the rules 'green' and 'square', recalling the stimulus if necessary. Subjects could acquire information from both dimensions and if they were using this 'perfect' method they should have been able to solve any shift in 0, 1 or 2 errors. 96.4% of shifts were completed in 2 errors or less by young subjects in Study 1, providing empirical confirmation that perfect hypothesis-testing was present for both ID and ED shifts. In comparison to Study 1, this study shows a small increase in ID shift errors (from 1 to 4) and a large increase in ED shift errors (from 1 to 20). The increase in ID shift errors in this study can be attributed to subjects occasionally confusing the 'difficult' exemplars and to the fact that lapses of attention cause errors when stimulus display time is limited. However, both of these factors also apply to the ED shift and therefore cannot account for the mean error score of 20 at this shift. The increased errors at the ED shift are likely to have arisen from the increased difficulty of carrying out perfect hypothesis-testing at the ED shift in this study, due to restricted display time and stimulus difficulty acting similarly to the 'masking' manipulation in this study. These conditions restricted attention to the dimension that was being explicitly tested and made mental reconstruction of stimuli difficult. In summary, subjects in this study showed a selective ED shift deficit compared to those in Study 1.

3.3.4.4 Accounting for the absolute level of ED shift errors

The introduction to this study (Section 3.3.1) suggested that an inability to gather information from the 'untested' dimension could account for a selective ED shift deficit An analysis of the process of hypothesis-testing reveals that whilst such a deficit would slow ED shifting, it cannot account for the performance of a subject who makes 20 errors before reaching criterion. As discussed above, if subjects are using perfect hypothesis-testing they can solve a shift in 0, 1 or 2 errors depending upon which stimulus pairs appear and what responses they make. This 'perfect' performance depends upon subjects being able to gather information from both dimensions that are present, that is, to be able to draw conclusions about the rules 'green' and 'square' after responding to a green square. If subjects could only draw conclusions about one dimension (for example, 'green' from a green square) this would indeed be a disadvantage for ED shifting. However, even in this case all ED shifts could be solved

within 3 errors. After an ED shift, there are 4 exemplars that could be present - for example, green, red, square and circle - and even if subjects tested the correct exemplar last they would only commit 3 errors. Therefore, an inability to learn about the untested dimension cannot account for an ED shift error score of 20.

Analysis of the data from this study cannot reveal the source of the 'extra' errors that elevate group mean error scores to 20. However, subjects' accounts of hypothesis-testing after the ED shift were consistent and point to a probable explanation. It appears that restricted display time and the use of difficult stimuli caused some subjects not to notice that there was variation in the dimension that they were not testing after the ED shift. For example, subjects might respond to the dimension 'shape' (the previously relevant dimension) after an ED shift but not notice that there was any difference between the two 'patterns' being displayed. This is particularly surprising as subjects were informed of all the possible dimensions before testing started. Pilot testing ensured that the two patterns on display were consistently discriminable, but subjects had to 'notice' that they differed before these exemplars could be included in hypothesis-testing. The 'extra' ED shift errors found in this study can be accounted for by the time taken for subjects to notice variation in the untested dimension. This account of the 'extra' errors is speculative and based on the subjective impressions of the participants in this study, but the subjects' descriptions of ED shifting were very consistent. Subjects did not report 'noticing' the difference in the patterns and disregarding it; they reported that they only started to solve the ED shift after noticing variation on the untested dimension.

3.3.4.5 Conclusion

This study found that masking and restriction of stimulus display time had a differential effect on the performance of ID and ED shifts. Masking did not simply cause an increase in task difficulty, as this would have elevated errors at both the ID and ED shifts. A comparison of the results of this study with those of the healthy young group in Study 1 reveals that different experimental conditions can permit qualitatively different styles of hypothesis-testing. In attempting to create ID and ED shifts that were free from 'ceiling' effects in Studies 2 to 4, a task was created that differed substantially to a superficially similar task in Study 1. The comparison of this study with Study 1 did support arguments about the importance of gathering information from the 'untested' dimension, but this cannot explain the level of errors committed at the ED shift in this

study. An account of the generation of these errors has been suggested on the basis of subjects' accounts of their performance of the ED shift.

4.0 A quantitative review of reaction time studies of Parkinson's disease

4.1 Introduction

As noted in Section 1.2, one of the most disabling features of PD is an impairment of voluntary movements. Clinically, this is apparent as "akinesia" (inability to initiate movement) and "bradykinesia" (slowness of movement). In the laboratory, an initiation impairment has been quantified by measuring reaction time (RT) to unpredictable stimuli. Patients with PD have frequently (but not always) been shown to have longer RTs than controls. It has recently become apparent that this slowing of RT is dependent on the particular sensory and motor parameters of the testing. What is not clear is the behavioral aspects of the tasks which make them vulnerable to the effects of the disease. Identifying the necessary and sufficient conditions for parkinsonian impairment (that is to say, the critical parameters of impaired tasks) is an important goal of research into PD in particular and the functions of the basal ganglia in general. In order to do this, a variety of behavioural tasks are required in which some features of testing are held constant while other parameters are varied. These tasks may then be compared to dissociate processes that are spared and impaired in PD. RT tasks take two basic forms. In a choice RT task, a response is selected from two or more alternatives, conditional on the nature of the imperative stimulus. In contrast, the same response is made across all trials in a simple RT task. One of the first studies to use a variety of RT tasks to study patients with PD was that of Evarts et al. (1981). They identified a group of patients with RT deficits but normal movement times and showed that parkinsonian RT impairments were specific to initiation and not merely due to bradykinesia. Despite the promise of using RT tasks, there is little consensus about which tasks are sensitive to deficits in PD. Most investigators would agree that patients with PD show a deficit in simple RT, but deficits in choice RT are unpredictable and the effect of dopaminergic medication on RT measures is also unclear.

In this review, three issues will be addressed that relate to RT deficits in PD. The first is whether there is a deficit in motor preprogramming, the second concerns the effects of dopaminergic medication on RTs and the third is the existence of deficits in simple RT tasks. Studies will be reviewed under each of these headings and a quantitative analysis of the past literature will be carried out.

In the first section, the literature concerned with a deficit in motor pre-programming will be reviewed. It has been suggested that movement initiation difficulties might arise from a failure of "intentional action" (Frith and Done, 1986), which might be manifested as an inability to use advance information about the nature of a movement to

speed its initiation. The idea that patients have a deficit in motor preprogramming - specifically, that they are unable to derive a RT benefit from the provision of information indicating which one of a number of possible responses will be required - has been popular, partly due to its elegant simplicity and the ease with which it might be tested. For example, cues which precede the imperative can convey all, little or no information. In a fully-cued task, information about the nature of the required response (e.g. a left/right button press) is made available before the imperative signal. Control subjects respond more quickly when cues are available and this performance improvement is taken to reflect the benefit of being able to prepare the movement in advance. This movement preparation is, of course, a type of motor set. On both cued and uncued trials, the imperative signal is the same; however, an internal 'predisposition' speeds response initiation on cued trials. It is important to note that, although many investigators have assumed them to be identical, fully-cued RT tasks are equivalent to simple RT tasks on neither theoretical nor empirical grounds (Jahanshahi et al., 1992a). For cued and simple RT tasks to be equivalent, the subject must have had sufficient time to use the cue to prepare the required response and the response must be prepared as completely as if it were the same on each trial. At short cue-imperative delays (for example 250 ms - Bloxham et al., 1984), this may not be the case. Empirically, it has been demonstrated that movement times are faster in a simple RT condition compared to a fully-cued choice condition, implying different processes may be at work (Jahanshahi et al., 1992a). Therefore, in considering the ability of patients with PD to use advance information to prepare a response (the "motor preprogramming" section of this review), only those studies making a comparison of cued and uncued choice RT will be included, rather than a comparison of choice RT with a true "simple" RT. The review will be followed by a quantitative analysis of past results.

In the second section, the effects of medication withdrawal on RT performance is reviewed, with a view to determining which, if any, of the RT deficits are sensitive to levels of dopamine. Studies of medication manipulations can be informative in two ways. First, they may allow identification of dopamine-dependent cognitive and motor processes. Second, such studies are an essential adjunct to conventional experiments on patients with PD: many patients participating in RT studies have been on medication for a number of years and withdrawal studies can reveal how much of their performance is influenced by the drugs that they are taking. In a study of the effects of medication withdrawal, Harrison et al. (1995) conclude that much of the previous literature is methodologically flawed. For example, studies have frequently used low numbers, have failed to control for order-of-testing effects or have not included control groups. Furthermore, medication manipulations

have differed radically, from overnight withdrawal to the study of spontaneous "on-off" motor fluctuations in advanced patients. Harrison et al. (1995) provide a useful summary of mean RTs taken from past studies of medication effects, as well as a valuable methodological critique; this will be expanded into a full literature review and quantitative analysis.

Impairment of simple RT has been cited as a consistent consequence of PD (Jahanshahi et al., 1992a). This deficit has been associated with a failure of a type of motor set in patients with PD ('attentional focussing' - see Section 1.5). Section 4.4 will establish whether simple RT, and by extension attentional focussing, is consistently impaired in PD.

4.2 Motor Preprogramming

4.2.1 Review

In 1981, Evarts et al. compared cued and uncued two choice RT tasks, using visual and kinaesthetic imperative signals. Although they did not invoke a deficit in motor preprogramming to account for their results, this study prompted many of the other studies described here. Evarts et al. found a significant parkinsonian impairment in the cued RT condition (which they refer to as simple RT), but, to their surprise, the patients' RTs in the uncued condition were not selectively longer than in the cued condition. Control subjects were not tested in the uncued condition, but it was noted that the cued-uncued difference for their patients was much less than had been recorded for elderly controls in prior studies (e.g. Talland, 1963), thus, the patients appeared to be impaired in the cued but not the uncued condition. This result was unanticipated, particularly as it contradicted the subtractive logic of Donders (1969), which proposes that a cued condition should be equivalent to an uncued condition, with the stage of "response selection" removed. This anomaly was noted but not investigated further by Evarts et al. (1981).

The finding of Evarts et al. (1981) does appear to rule out the possibility of Donderian "stage processing" - how is it possible that uncued choice RT is not impaired whereas cued choice RT, with all its "stages" in common with the uncued condition, but with fewer stages overall, is impaired? It is logical to assume that cued choice RT is not equivalent to uncued with a stage deleted, but rather is a qualitatively different task. Bloxham et al. (1984) replicated Evarts et al.'s study (testing a control group for all conditions) and proposed a plausible interpretation of their own and the Evarts et al. data. They examined cued and uncued choice RT. Either a cue or a warning signal was given 250 or 2000 ms in advance of the imperative stimulus and patients were found to be impaired relative to controls in the cued, but not the uncued condition. Bloxham et al. interpreted this result as the patients' failure to make use of the cue to prepare (or "preprogram") a response. In a separate tracking experiment, patients were shown to be capable of using advance information to control the direction of an ongoing movement; Bloxham et al. concluded that the parkinsonian deficit was an impaired ability to make use of advance information to initiate a movement, whereas guidance was intact.

Three studies (Bloxham et al., 1987; Heilman at al., 1976; Jahanshahi et al., 1992a) have manipulated the presence of a warning stimulus in their designs, which cued subjects as to the imminent arrival of the imperative stimulus, but gave them no information about the nature of the response required. Patients and controls benefited equally from the

presence of a warning stimulus in all of these studies, implying that Bloxham et al.'s (1984) results were specific to motor preprogramming and not merely an effect of arousal or stimulus anticipation. One inconsistent report has attempted to replicate and extend the effect in a study of the "on" and "off" phases of late PD (Girotti et al., 1986). Patients were impaired on a cued, three-choice RT task, but unlike in Bloxham et al.'s (1984) study they also suffered from a significant deficit in an uncued three-choice task, irrespective of the "on" and "off" phase (Girotti et al., 1986). Instead of showing a motor preprogramming deficit, these results imply a simple, consistent deficit across both tasks; the parkinsonian ("on" phase) deficit is 109 ms in the cued condition and 106 ms in the uncued condition (figures from table 2 of Girotti et al., 1986).

Using Rosenbaum's (1980) partial cueing methodology, Stelmach et al. (1986) assessed the performance of patients when given partial information of the parameters (arm, direction, extent) of a required movement. These parameters were either unspecified (uncued condition), fully specified (cued condition) or one or two of the three parameters were given. The results of this study also did not support the hypothesis of a motor preprogramming deficit. Patients were significantly impaired across all conditions, but benefited from the various levels of precueing to the same extent as controls. This study clearly allowed a more comprehensive analysis of motor preprogramming abilities than the previous work, including that of Bloxham et al. (1984) and the authors concluded that Bloxham et al.'s claims of a motor preprogramming deficit were "somewhat overstated". Again, the possibility of a general RT deficit in PD that does not vary across experimental conditions is raised.

Sheridan et al. (1987) were able to replicate Bloxham et al.'s (1984) results, in a study to investigate the significance of increasing the movement difficulty for patients with PD. An 'index of difficulty' (ID) is defined by Fitts' ratio rule as a function of movement amplitude and target size (Fitts, 1954). The ID was varied across trials and, for some blocks of trials, the required movement was cued prior to target onset. Patients' RTs were significantly longer than controls', but only when the movement was cued in advance. This pattern of impairment is suggestive of a motor preprogramming deficit. However, Robbins and Brown (1990) have criticised this conclusion, noting that the effect of cueing did not vary with the difficulty of the movement whereas one might expect that cueing should provide a greater RT benefit for more difficult movements. They have also suggested that the movement time data in Sheridan et al.'s (1987) study indicate that responses may have been coming under the control of visual feedback after movement initiation, which would suggest that initiation times might not be an accurate measure of motor preprogramming.

Some of the anomalous data on motor preprogramming abilities might be accounted for by medication effects. Pullman et al. (1990) partially replicated Bloxham et al.'s results; RTs (as in Girotti et al., 1986; Stelmach et al., 1986) were lengthened in patients relative to controls in both cued and uncued conditions. Patients show a statistically nonsignificant 49 ms advantage of the cueing, whereas controls show a statistically significant 80 ms benefit of cueing (figures estimated from Figure 1 of Pullman et al., 1990) at clinically optimal ("high") levodopa infusion levels). This result raises the possibility of two dissociable patterns of RT deficit in PD - a deficit in motor preprogramming, which might sometimes be masked by an additional generalised slowing of RT across all conditions. However, this view still does not account for all the data. Lichter et al. (1988) incorporated cued and uncued two-choice RT tasks into their investigation into the relationship between motor, cognitive and computed tomographic variables in PD. No significant parkinsonian deficit was found in the cued tasks, but a deficit was found in the uncued condition. This is clearly contrary to previous results and resistant to any interpretation involving deficits of motor preprogramming. Interestingly, both Alzheimer's disease and closed head injury may result in slowed choice RT, and simple RT has been found to be spared in these conditions (Gordon and Carson, 1990; Miller, 1970). Lichter et al.'s patient group were particularly severely impaired (one third of the patient group were in Hoehn-Yahr stages four and five) and therefore might be more representative of generalised neurological trauma.

Two recent studies have used designs to test explicitly the hypothesis of a deficit in motor preprogramming. Jahanshahi et al. (1992a) controlled for delay between preparatory and imperative signals and also the type of advance information provided. The study included a simple RT task and four choice tasks with uncued, partially cued (2 types) and fully cued conditions. There were also conditions involving a noninformative warning stimulus and all conditions made use of unpredictable foreperiods from 0 to 3200 ms. The results do not support the presence of a deficit in motor preprogramming in PD. The clearest example of this comes from comparing the Parkinsonian deficit in the fully cued, partially cued and uncued four-choice RT conditions, where the group by condition interaction did not reach significance. Jahanshahi et al. (in agreement with Stelmach et al., 1986) conclude that patients can make full use of advance information to preprogram their responses. They suggest that the parkinsonian impairment may be found at a different stage in processing, such as response initiation. Using similar methodology, Brown et al. (1993a) drew the same conclusion in the course of their study of stimulus-response compatibility in PD. Willingham et al. (1995) tested patients with Huntington's disease and those with PD,

to allow comparison between two degenerative disorders of the basal ganglia; both patient groups showed normal use of advance information to speed their reactions, but with a significant overall deficit.

The initial finding of Bloxham et al. (1984), although it has been replicated a number of times, is unlikely to be due to a deficit in motor preprogramming *per se*, as numerous studies designed specifically to reveal such a deficit have failed to do so. Below, alternative possibilities that might account for these data are considered.

4.2.2 Comment

The replications of Bloxham et al.'s results do appear to rule out the possibility of Donderian "stage processing". It is logical to assume that the cued RT is not equivalent to deleting stages from an uncued choice RT but rather is qualitatively different. Once a psychological dissociation is accepted, it is not unreasonable to suggest that there may also be a neural dissociation. Frith and Done (1986) provide one possibility, that there are multiple "routes to action" and that these routes are differentially dependent upon internal control. The 'fast route' governs responses which are completely prespecified (as in a simple RT task): in this case, the required movement is 'held in mind' until the imperative signal. When it is not possible to select the response until after the imperative (a choice RT task) either the 'slow' or the 'direct' route will govern the RT. If the associations between the stimuli and the responses are compatible, such that it is not necessary to consult an internal mapping to select the response, the 'direct route' governs the RT. On the other hand, the slow route will be navigated whenever the stimulus-response associations must be 'held in mind' to be consulted. Thus, in the operation of both the fast and the slow routes, a response or a stimulus-response mapping is 'held in mind', and these routes to action are therefore particularly dependent upon internal control and, consequently, particularly vulnerable to striatal dopamine depletion. In the direct route, the stimulus implies the response and no such internal control is required and, consequently, the operation of this route is not impaired by striatal dopamine depletion. This account is intuitively plausible and fits well with observations both from the clinic (e.g., patients showing less impairments when given visual cues to guide their movements - Martin, 1967; Schwab, 1972) and from neuropsychological experimentation that patients are more impaired when their actions are under "internal control" (Stam et al., 1993). However, one of the problems with this formulation is that it is vulnerable to *post hoc* interpretation of the data: it is not always obvious whether an action is "internal" or "external" and it is tempting to use the presence or absence of a deficit in the patient group as the defining feature. For example, in the

discussion of their Experiment III, Frith and Done (1986) state that "when there was no precue subjects would use the 'slow' route" but "when a precue was given... in advance subjects would use the 'fast' route" (p.173). In either case, both the slow and the fast routes are described as "strategic" and involving "will" or "internal intention". They go on to say (p. 175) that "problems in the control of action observed in patients with PD are consistent with an impairment in the route by which internal acts of will control responses", which would appear to suggest that patients should be impaired in both the slow and the fast routes. However, Bloxham et al. (1984) use what appear to be the same RT tasks, involving the fast (precued condition) and slow (uncued condition) routes. However, they found impairments only in the precued and not in the uncued condition. This implies that the task has been incorrectly assigned a route and the uncued condition is not a slow route task but is a "direct route" task, as Bloxham et al. (1984) claim. This ambiguity indicates that the "routes" approach might not be predictive of a deficit in PD and thus diminishes its explanatory power.

An alternative to the "internal-external" dichotomy has been raised in two papers (Hallett, 1990; Robbins and Brown, 1990). They suggested that the assumption of psychological and neurological dissociation between actions controlled internally and those guided by external stimuli might be premature. Hallett (1990) proposed that one might assume that all RT tasks involve an initiation stage, which has a fixed value. This value would result in a "floor" on the most rapid response (for example, 200 ms for the normal system). This movement initiation stage is the main determinant of minimum RT in simple RT tasks. If the initiation stage were lengthened in PD, this would result in an impairment on all simple RT tasks. In choice RT tasks, Hallett proposed that the initiation stage runs in parallel with processes of response selection. Lengthened movement initiation would therefore result in a choice RT deficit in patients only in those tasks where the intiation stage takes longer than response selection. One would expect to see a choice RT deficit in patients in tasks where response selection can be achieved rapidly, as the movement initiation stage is then the rate-limiting step for patients and controls. A RT deficit in patients with PD would therefore be seen in tasks where RTs are rapid, and approaching the minimum latency imposed by the movement initiation stage. No RT deficit should be seen amongst patients in tasks which result in slower RTs, as here the principal determinant of RT is the process of response selection which is presumed to be unimpaired in patients. Thus, this hypothesis allows for differential deficits in different tasks, but the probability of an impairment in a particular task will depend upon the speed with which the task is normally performed, with faster tasks being more vulnerable to the effects of the disease

than tasks with slower RTs. This would predict the differential slowing of cued choice RTs seen in patients with PD, as cued tasks are usually performed more quickly than uncued tasks. Robbins and Brown (1990) suggested that dopamine is a "limited resource" needed for response initiation, with more dopamine needed for fast RTs. When this resource is depleted, as in the case of PD, fast RTs (requiring more dopamine) will be impaired first with slower RTs impaired as the disease progresses. Conversely, dopamine replacement therapy will restore dopamine levesl sufficient for normal performance of slower RTs and gradually restore the deficits in ever faster RT tasks as dose is increased. This hypothesis would also predict a differential parkinsonian deficit in cued RT tasks, as well as making clear predictions about the effects of medication.

4.2.3 Quantitative analysis

4.2.3.1 General method.

A database of studies which reported the RT performance of patients with PD and control subjects was constructed. Studies were found using online literature databases (*BIDS* 1982-1996, *Medline* 1985-1996), and also by searching leading journals that publish behavioural studies of PD (such as *Brain; Journal of Neurology, Neurosurgery and Psychiatry; Neuropsychologia; Neurology; Movement Disorders*). The following types of study were excluded:

1. Those in which initiation times were obtained from EMG recordings or eye movements, as these studies resulted in RTs for patients and controls which were particularly fast and outliers with respect to the bulk of the data set.

2. Those in which the RTs corresponded to "thinking" times for "cognitive" tasks such as memory scanning, planning etc. These studies also resulted in outliers, being particularly slow, and potentially would have conflated measures of movement initiation with the cognitive deficits of PD.

3. Those which involved serial RT tasks (e.g. Ferraro et al., 1993) or sequence learning or repetition tasks (e.g. Rafal et al., 1987). Although the dependent measures in these studies are speeded responses to an external stimulus, they are dissimilar to the simple, rapid responses to imperative stimuli required in conventional RT studies.

If the studies included data which were collected under conditions of varying medication, only the conditions in which patients were on their normal medication regime were included, except when looking specifically at the effects of medication. 44 studies were included, from which mean or median RT values were taken (measured from

magnified graphs or, where available, from data presented in tables) from as many different conditions as were presented in the individual studies.

The use of raw RT values for our quantitative analysis requires some justification, as meta-analytical techniques usually call for effect sizes to be used when summarising the results of past studies. The use of effect sizes is inappropriate here, due to the well-established fact that variance increases proportionally to mean RT (see Luce, 1986). This effect would lead to spuriously low effect sizes in studies with long RTs, due to increased variance in patient and control groups. This relationship is particularly significant, as past research has led us to suspect a relationship between the size of the parkinsonian RT deficit and control group RT (Brown et al., 1993b; Robbins and Brown, 1990).

4.2.3.2 Are patients impaired in the use of advance information?

Studies which tested the hypothesis that patients are differentially impaired in the use of advance information to select and prepare a response (Appendix 2) were selected. In these studies, a comparison was made between a two (or more) choice RT and a task in which there was a precue to indicate the correct choice prior to the imperative signal. If patients with PD are impaired in using advance information to program their responses, it would be expected that there would be a greater difference between control and patients' RT in the cued compared to the uncued condition.

The studies were carefully screened to ensure that a comparison was made between cued and uncued choice RT and not between simple and choice RTs. Only completely uncued and fully cued conditions were included; partial cueing conditions, or those in which cues were sometimes invalid (as in the "covert orienting" paradigm - Posner, 1978), were not used. Table 4.1 shows the mean group RTs for two choice RT tasks. In support of this hypothesis, the patient groups were, on average, more impaired in cued compared to uncued choice RT.

		Control	PD	PD deficit
CUED				
	Mean	358.00	463.89	105.89
	S.E.M.	(26.63)	(24.31)	
UNCUED				
	Mean	439.33	508.00	68.67
	S.E.M.	(36.67)	(25.82)	
Effect of cue:		81.33	44.61	

Table 4.1: Mean RTs for patient groups and control groups for the cued and uncued conditions of the 2 choice RT studies listed in Appendix 2. The right-hand column show the mean RT deficit for patients with PD in cued and uncued conditions . Values in parentheses represent the standard error of the mean.

However, in some studies, the number of choices used in a choice RT task was greater than two. In these cases, it might be expected that an inability to use advance information would result in patients showing an *even* greater impairment relative to controls in the cued condition as the number of choices increased; presumably the benefits of precueing should be greater as the number of potential responses increases. However, this was not found to be the case. Figure 4.1 shows the effects of advance information on RT of patients and controls as a function of number of choices in the task. It is clear that for the multiple (>2) choice RT tasks the average patient impairment in the uncued condition is at least as great, if not greater, than in the cued condition. Like control subjects, patients show a greater benefit of cueing in tasks that include more than two possible responses. Furthermore there is absolutely no evidence of a greater parkinsonian deficit in these tasks. This was confirmed by an ANOVA on the means of conditions (Cueing x Number of choices x Group, $F(1, 24) = 4.83, p < .05$). This is hard to explain in terms of a deficit in motor preprogramming; patients appear to have an overall deficit in the use of cues, but do not have a greater deficit where cues are more useful (>2 choice tasks).

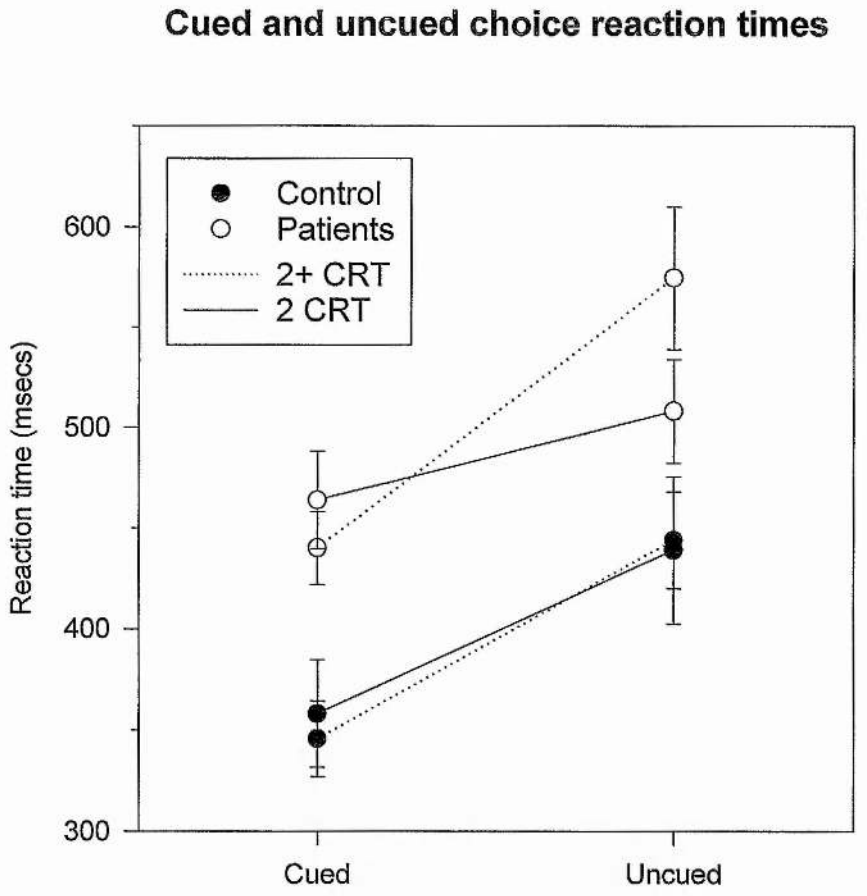


Figure 4.1: Means of patient ("PD") and control group RTs taken from studies in Appendix 2. Open circles are patients and filled circles are controls performing fully cued and uncued RT tasks. The filled lines connect points representing two-choice RT tasks (2 CRT), and points on the dotted lines represent conditions in which there were more than two response choices (2+ CRT).

The data were then examined in a different way: control group RT was plotted against patient RT. A general RT impairment in the patients would result in a linear relationship (with a slope of 1) between control and patient RT, but with the patients having a generally longer RT (i.e., a greater intercept). If, however, there were a differential deficit in the use of advance information, there would be a discontinuity in the data, with cued RTs being more impaired than uncued RT (i.e., parallel lines with slopes of 1 but with a greater intercept for the cued tasks). Neither were found to be the case; when regressions were

fitted to the relationship for both precued and uncued choice RT, the slopes were remarkably similar and both were less than 1. The best predictor of the degree of deficit in the patients was the RT of the control group such that faster control group RTs are associated with greater patient deficits, as shown in the inset graph (regression coefficient: slope = -0.17 ms deficit/control ms, intercept = 163.2 ms, $r^2 = 0.15$, $r(26) = 0.381$, $p < .05$).

Thus, it appears that there is a relationship between RTs of control subjects and patients with PD, such that in tasks where control subjects react most quickly, patients are most likely to be impaired. In tasks with longer control group RTs, patients are less likely to be impaired or the degree of impairment is less severe.

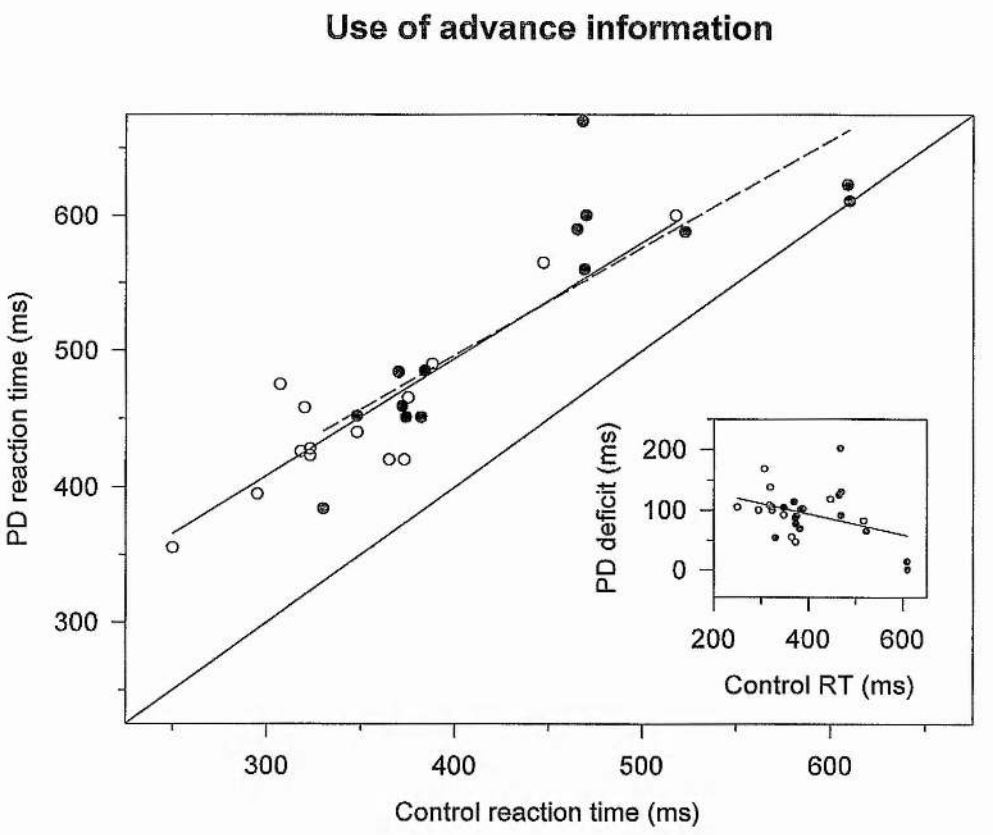


Figure 4.2. Patient ("PD") versus control group RTs, performing cued and uncued choice RT tasks (Appendix 2). Filled circles and the unbroken regression line represent group performance on uncued conditions, open circles and the broken line represent fully cued conditions. INSET: Parkinsonian deficit (patient RT minus control group RT; "PD deficit") plotted against control group RT. The regression line shown is for all data on the plot. Inclusion and symbols are the same as for the main plot.

This significant relationship is all the more remarkable given the fact that the patient groups between studies were heterogeneous: clinical ratings (most often, Hoehn and Yahr) are given in most of the reports and the degree of impairment within and between studies reflects a wide range on this rating scale (patients ranged from Hoehn-Yahr stages one to five). In all of the studies included in Figures 4.1 and 4.2, the patients were optimally medicated, with medication including levodopa alone or in combination with anticholinergics, dopamine agonists and/or MAO-B inhibitors. The effect of levodopa

medication on RT of patients with PD is another controversial area in the literature and therefore this issue is examined in the following section.

4.3 Effects of medication

4.3.1 Review

Patients who participate in studies of RT are usually on long-term dopaminergic medication (in particular, levodopa). It is unclear how this medication affects their performance and this can be studied by manipulating their medication status. Some studies have directly examined the effects of dopaminergic medication on RTs and some have included a medication withdrawal condition in experiments designed to test a different hypothesis. Withdrawal of medication may reveal the mechanisms behind the deficits of PD; levodopa medication does not alleviate all symptoms equally (Velasco and Velasco, 1973) and associations and dissociations of symptoms allow inferences about both underlying processes and the role of dopamine in the symptoms of the disease. These data could also be used to support theories of normal function.

Notwithstanding significant alterations in clinical status after medication administration/withdrawal, a number of studies have found no statistically significant change in RTs. As a result, some authors have argued that RT is a poor index of disease status (Harrison et al., 1995). This raises the question of what is being measured in studies that *do* show a significant RT deficit when medication is withdrawn and also why most studies reveal some measurable, if statistically non-significant, slowing of RT when patients are off medication (see Table 1 of Harrison et al. 1995). The failure to find a significant RT deficit may be due to limitations of these studies. Harrison et al. have made a valuable methodological critique of most of the studies published to date, citing low numbers, failure to counterbalance order of testing and lack of control groups as possible causes of low statistical power and nonsignificant effects.

A further source of error may lie in the medication manipulations chosen. Four methods have been used (and mixed): (a) spontaneous "on-off" fluctuations in motor status, (b) patients going onto medication for the first time, (c) withdrawal of medication and (d) direct intravenous infusion of levodopa. These methods are clearly not equivalent and consequently may not generate comparable results. In this section, the approach will be to review the studies in detail, citing their designs and medication manipulations and then to look for general trends with a quantitative analysis of past results.

4.3.1.1 Studies that have failed to find a significant effect of medication

Rapid fluctuations ("on-off") in motor state are common in advanced PD. As patients fall into the debilitating "off" state, it appears as if the beneficial effects of their medication

have been abruptly removed. Two studies have tested patients in on and off states. Rafal et al. (1984) studied six patients whilst experiencing both on and off states and four before and after taking levodopa for the first time. Use of a covert orienting paradigm (Posner, 1978) allowed separation of cognitive and motor processes and patients showed no disruption of selective attentional mechanisms when off medication. Despite the dramatic alterations of clinical status between on and off states, however, the slowing of RT was nonsignificant. Although Rafal et al. (1984) did counterbalance the order of testing of the on-off patients, their patient group was heterogeneous. The patients with motor fluctuations were presumably in the later stages of PD (although described as cognitively intact) and those never previously medicated were presumably only mildly affected (although two were described as having "early dementia"). Lack of a significant effect on RT may have been due to highly variable subject characteristics, a limited number of subjects (on-off, n = 6; first medication, n = 4) and the mixture of two different medication manipulations.

Girotti et al. (1986) studied 21 patients in on and off states and an elderly control group. Being in an off state produced a nonsignificant slowing of similar size in precued and uncued three-choice RT tasks. Although Girotti et al. counterbalanced the order of testing (on-off, off-on), high variance may have resulted in the failure to find a significant RT deficit: despite the relatively large subject sample, patients' RTs in the off state had a standard deviation of 189 and 156 ms compared to 67 and 82 ms in controls (from Tables 2 and 3 of Girotti et al., 1986). Although on-off fluctuations cause large and frequently distressing changes in motor symptoms, it appears from both sets of results reviewed so far that they do not cause consistent changes in RT (Girotti et al., 1986; Rafal et al., 1987). It is important to note that the majority of patients in both of these studies were still medicated (although Rafal et al. (1987) sometimes delayed a dose in order to keep individuals in an off state) and inferences about dopaminergic function are not based on direct manipulation of medication status.

Two studies have recorded RTs from patients before and after starting levodopa therapy. Velasco and Velasco (1973) recorded simple RTs before and after putting patients on huge doses of levodopa (averaging 5.5g per day). Monthly follow-up assessments were made seven times whilst patients were on levodopa, but despite selecting "the study [follow-up] that showed maximal improvement for each patient" (p. 93), no significant improvement in simple RT was found (a speeding of 66 ms was seen). It is surprising that no significant increment in RT could be found despite large medication doses, selective use of data and no control for practice effects.

Where order of testing cannot be counterbalanced, control groups are clearly needed. Jordan et al. (1992) assessed age-matched controls twice for comparison with off-on medication testing in patients. Patients were randomised to either bromocriptine, benzhexol or levodopa monotherapy. Jordan et al. report significant clinical improvements for all three groups, but only a nonsignificant improvement in simple and go/no-go choice RT. It is impossible to assess the role of levodopa specifically here; no significant differences were found for RTs for patients on the three different therapies, but with only seven patients in each group, this may be due to lack of statistical power. The authors also do not report the interval between testing off and on medication, making comparisons with other studies problematic.

Overnight withdrawal of medication allows substantial time without drugs to elapse with least distress to the patient. Patients are usually asked to delay their morning dose of levodopa until after testing has been completed. A related method is asking patients to skip a dose, as used in one study of simple RT (Bloxham et al., 1987). Patients were tested after a medication fast of, on average, nine hours and, subsequently, one hour after taking medication. Administration of medication after the fast produced only a nonsignificant speeding of RT, which did not interact with the presence of either a temporal warning signal or a secondary task. Controls were tested twice to account for their fixed (off-on) order of testing, which makes the design of this study robust (Bloxham et al., 1987). The authors attribute their failure to find a significant RT effect to their medication manipulation, particularly retesting only one hour after administration of levodopa, an interval which allowed only a "variable" improvement in clinical condition. One other study (Montgomery and Nuessen, 1990) has used a similar manipulation and found a significant RT effect, but it suffers from a number of methodological problems that will be discussed in the following section.

Starkstein et al. (1989) used a longer withdrawal period and more severely disabled patients. All seven patients experienced severe on-off fluctuations and stopped taking medication at least 24 hours before testing. Simple RTs to auditory stimuli were tested after the medication was stopped and again after medication had restored patients to their "best" motor state (established before the study). Changes in clinical status between off and on states were dramatic, but no significant improvement in RT performance was seen.

Two further studies have failed to find any effect of medication on RTs using overnight withdrawal. Jahanshahi et al. (1992b), using the same design as their 1992a study described above, used a relatively long withdrawal period (average 14.4 hours) and compared data from this session with another when patients were fully medicated.

Withdrawal produced no deficits in either simple or choice RT, nor did it affect patients' ability to use advance information to speed movements. This provides a replication of Girotti et al. (1986)'s study, which indicated that patients showed equal benefits of precueing a response in both on and off states. Neither Girotti et al. (1986) nor Jahanshahi et al. (1992a) found a deficit in using advance information, but it is now clear that their prior results were not simply a consequence of testing long-term medicated patients. Jahanshahi et al. used a fixed on-off order of testing (1992b). They did not include a control group and cannot adequately rule out practice effects masking a deficit in the off stage. Labutta et al. (1994) did test control subjects, but only once, which fails to provide any information about practice effects in their on-off design. Patients could successfully store and use advance information about a movement for up to eight seconds, both when medicated and after an overnight fast. Labutta et al. (1994) withdrew medication from only five of their ten subjects and gave no information about the disease severity of these five - the original group of ten varied from Hoehn-Yahr stages 1.5 to 4. Once again, low subject numbers, high subject variability and practice effects might have made even a strong medication effect hard to detect, thus, this negative result cannot be used to draw any strong conclusions.

The work of Pullman et al. (1988, 1990) stands apart for its highly sophisticated manipulations of medication levels. Levodopa was infused intravenously at levels calibrated to provide each patient with maximal (high, on), minimal (low, off) and a medium (mid) clinical response. With increasing levels of levodopa, simple RTs showed a clear, although statistically nonsignificant, speeding and directional 2-choice RTs were significantly faster (1988). Two years later, Pullman et al. (1990) found no significant effects of levodopa on cued and uncued 2-choice RTs using movement amplitude as the response. This led them to propose a distinction between amplitude and directional choice RT tasks, but this hypothesis has since been tested and not supported (Jones et al., 1993). Pullman et al. (1988, 1990) tested only five patients and five controls. Furthermore, the controls were tested only once and there is no information provided about the order in which the patients were tested at the three levodopa levels. These aspects of the design limit the effectiveness of their superior pharmacological manipulation. A greater number of subjects may have resulted in statistically significant effects and avoided the direction/amplitude dissociation.

The studies described above are consistent in that they report that obvious clinical changes caused by medication manipulation are not necessarily accompanied by a statistically significant alteration in RT. Another consistent aspect of these studies is

inadequate design: seven out of ten have ten or fewer patients and six out of ten fail to counterbalance testing order or include two tests of a control group.

4.3.1.2 Studies that have found a statistically significant effect of medication

Pullman et al. reported significant differences in two-choice RT in patients at different levels of plasma levodopa, using the methods described above (1988). There was a large RT deficit for both simple and choice tasks when patients on the low (off) dose were compared with controls; optimal levels of levodopa almost completely "medicated away" the choice RT deficit, but had less of an effect on simple RTs. Two other studies compare the effect of medication on simple and two or more choice RTs. Jahanshahi et al. (1992b) did not show a similar effect; uncued four choice RTs were found to be fractionally faster when patients were off medication. As practice effects were not controlled, this result is ambiguous. Harrison et al. (1995) found no statistically significant differences between the medication effects recorded for simple or two-choice RTs, but although the difference was not statistically significant the withdrawal deficit was almost twice as great in the choice compared to the simple condition (32 ms and 17 ms, respectively).

The remaining five studies all report a RT deficit after medication withdrawal. Zappia et al. (1994) used the most complete withdrawal method yet reported. Medicated patients had their dose of levodopa gradually reduced and ultimately stopped. Patients (including a never-medicated group) were then tested after reaching "baseline" motor conditions, defined as when clinical and RT measurements were similar on two consecutive days. Six-choice RTs were recorded for each hand at baseline, then at 1, 1.5, 2, 4 and 24 hours after administration of a single 200 mg dose of levodopa. RTs were significantly faster than baseline at 1, 1.5, 2 and 4 hours after medication. However, maximal clinical benefit and RT speeding were attained two hours after administration of levodopa. This finding may account Bloxham et al.'s (1987) failure to detect a significant RT benefit of medication, as subjects were tested just one hour after administration in this study. A control group was tested once in Zappia et al. (1994)'s study, but practice effects were eliminated simply by overtraining; the RT task was performed daily during the gradual withdrawal until performance had stabilised at baseline.

Montgomery and Nuessen (1990) studied tradeoffs between speed and accuracy in the movements of PD patients, reporting a significant effect of medication withdrawal on RT. Medication was withdrawn overnight and simple RTs were tested before and one hour after administration of the patients' normal dose. An analysis of movement times led the authors to conclude that the "speed-accuracy operator" was abnormal in patients off

medication. RTs were also longer when patients were off. However, the RTs reported in the paper are unusually fast for whole arm movements; one patient off medication records a mean simple RT of 192.5 ms. Typical simple RTs, involving a button press, of healthy 24 year-olds have been reported as 222 ms (Goodrich et al., 1990). This raises questions about the methods used for recording RT in this experiment.

The last three studies that report RT changes after medication withdrawal do not find gross alterations in mean RT. Significant slowing whilst off medication are only seen in certain conditions. One study investigated two-choice, left/right RTs where the imperative signal had varying degrees of compatibility with the response (Brown et al., 1993b). Medication withdrawal for at least 24 hours produced no significant changes in mean RT, but caused a significant shift in the RT distribution for the tasks with greater stimulus-response compatibility.

Malapani et al. (1994) set out to discover whether the parkinsonian deficit in performing simultaneous motor acts (Benecke et al., 1986) also applies in the cognitive domain. Patients were tested after at least 18 hours of withdrawal and again 90 minutes after administration of levodopa. No withdrawal deficits were seen for basic go/no-go RTs to visual or auditory stimuli. A deficit was seen in the "concurrent" condition; this was a three-choice task required no-go, unimanual or bimanual responses dependent on the appearance of visual and/or auditory stimuli. Both unimanual and bimanual responses to auditory stimuli were significantly slowed by withdrawal and the authors concluded that concurrent processing of information (here visual and auditory) requires adequate dopaminergic transmission. This is premature, however. The slowing of unimanual responses is a sound result, but bimanual responses were not required in any other condition of the experiment and simultaneous motor acts are known to be sensitive to different medication levels (Benecke et al., 1987b). Motor demands have therefore not been equalised across tasks and inferences about cognitive function ("concurrent processing") must be provisional. A medication-dependent, auditory choice RT deficit was certainly present.

Harrison et al. tested in a fixed on-off order after overnight withdrawal; they also test a control group of patients with PD at the same times without withdrawal, these being matched for age and disease duration and severity (1995). This is preferable to testing an elderly control group, as it accounts for the fact that practice and fatigue effects may differ between patients and controls. This study found a simple RT deficit when intervals between the previous response and the stimulus onset were very short (50, 100, 200 ms). A slowing

is usually seen in normal subjects at short response-stimulus intervals ("refractoriness") and this is exaggerated after medication withdrawal in patients.

4.3.2 Comment

The studies that have found significant speeding of RTs with medication administration are characterised by strong medication manipulations and good designs. The withdrawal in study (Zappia et al., 1994) is complete and two other studies (Brown et al. 1993b; Malapani et al. 1994) withdraw for at least 24 and 18 hours respectively. Two studies (Brown et al. 1993b; Malapani et al. 1994) counterbalanced their order of testing and two others (Harrison et al., 1995; Zappia et al., 1994) control for practice effects using a control group and overtraining. The ideal time to test after administration of medication appears to be two hours. The significant effects appear to be task-specific and even modality specific (Malapani et al., 1994). Jahanshahi et al. (1992b) have suggested that clinical changes and reaction and movement time deficits will appear at different levels of dopamine depletion. Results from a number of studies suggest that as dopamine levels fall, clinical deficits will appear first, followed by movement time and then RT slowing (Jahanshahi et al., 1992b). This review is also tentatively suggesting that there may be different thresholds between different RT tasks. There is a trend for choice RT to be affected by medication withdrawal more than simple RTs, and those studies reporting significant medication-dependent RT deficits tend to be using choice RT tasks.

Although correlated with disease status (Zappia et al., 1994), RT is not a measure of the severity of PD. The best index of disease severity is attained by use of a clinical rating scale, such as the Webster (1968) and Universal Parkinson's Disease Rating Scale (UPDRS, Langston et al., 1992). RTs incorporate such nonmotor variables as stimulus registration and identification and response selection and initiation. There is a large literature covering the stages of processing involved in RTs in normal subjects (Luce, 1986) and bringing these concepts to bear on PD is the most productive use of RT studies. If such studies can demonstrate a dissociation between medication effects on simple and choice RTs, a whole new area of research into dopamine's role in decision processes might be initiated.

4.3.3 Quantitative Analysis

4.3.3.1 Effects of medication

Harrison et al. (1995), considering the literature on the effect of levodopa medication on RT performance, concluded that the data were inconclusive, with only 6 of the 13 studies they

reviewed reporting a statistically significant RT benefit of levodopa medication and the remaining studies showed no significant effect of levodopa on RT. However, from the table of means they present (see Table 1 of Harrison et al. 1995), it is apparent that, notwithstanding the failure of some studies to find statistically significant effects, in only one of the 13 studies is there a condition which does not show some improvement in mean RT with medication. If there were no effect of medication on RT performance, one would expect that on average studies would show no benefit and that an equal number of studies would report a decrement in RT as report an improvement.

The literature on the effects of medication on RT was reexamined, including only those studies in which RTs were reported for patients with and without levodopa (medication withdrawn, controlled infusions of levodopa or patients showing "on-off" fluctuations - see Appendix 3) and for controls. Two possible candidates for this analysis were specifically excluded; (a) one combined data for patients going on to levodopa, benzhexol and bromocriptine for the first time, allowing no analysis of the effect of levodopa alone (Jordan et al., 1992), and (b) another in which the graphs are unclear and in which the RTs are anomalous, with some patients having an unmedicated RT (which is not EMG) of under 200 ms (Montgomery and Nuessen, 1990). Studies included in the previous "advance information" analysis were also excluded, (Girotti et al., 1986; Pullman et al., 1990) in order to examine an independent set of data. RT values for the patients were plotted against control group RT for both medicated and unmedicated conditions in Figure 4.3. The regression line for the medicated patients showed a similar slope (slope = 0.77 ms/control ms) as seen previously (Figure 4.2; slope of the overall regression = 0.83 ms/control ms) even though all these studies were unique to this analysis and no study was included in both analyses. However, for the non-medicated conditions, the slope of the regression line approached 1 (0.92), with an intercept of 115.4 ms: that is to say, the deficit in the non-medicated patients was, on average, equivalent to 115 ms, regardless of the control group RT.

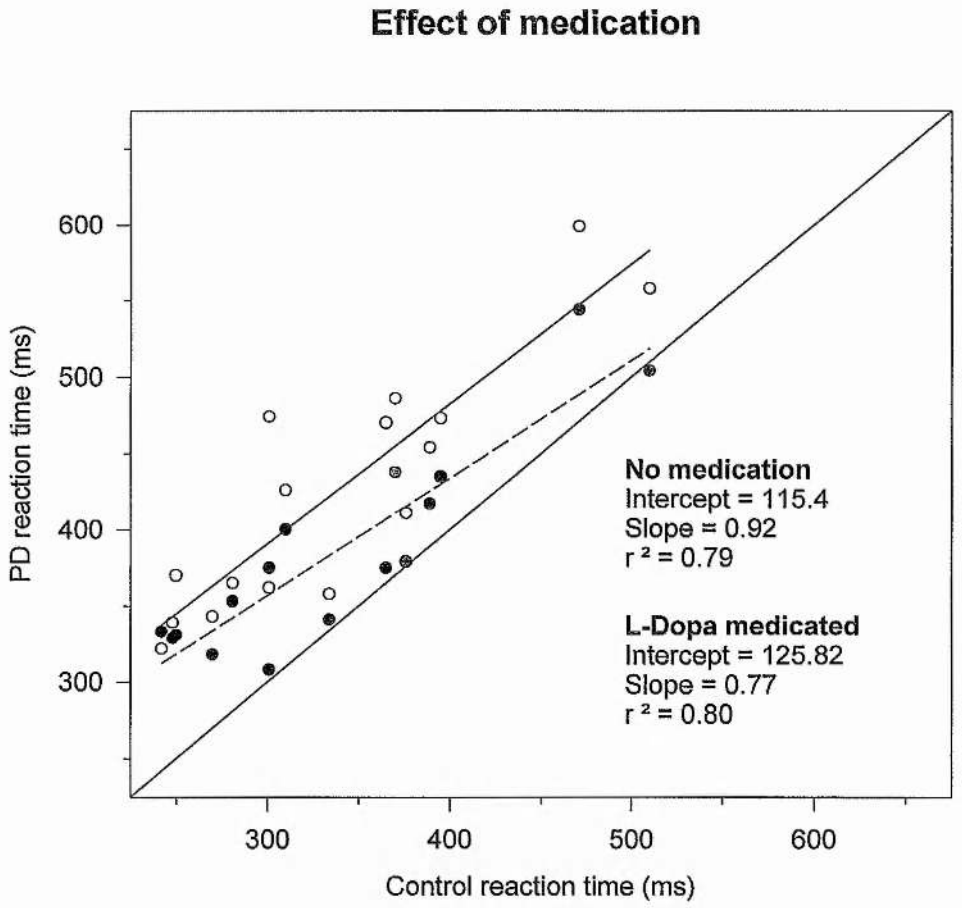


Figure 4.3. Patient ("PD") and control group RTs in studies manipulating medication state (Appendix 3). Filled circles and the broken regression line represent levodopa medicated patients, open circles and the unbroken line show values for the same patients off medication.

There are two particularly important features of Figure 4.3: (a) the relationship demonstrated in Figure 4.2, between control group RT and degree of patient deficit, is replicated in an independent data set for medicated patients; (b) medication does have an effect on RT of patients with PD, but the effect of the medication is dependent upon control group RT, with there being little effect of medication on tasks in which control performance is fast and a greater effect when control performance is slower. Put another way, medication can mask a deficit in RT in those conditions when control group RT is long. Robbins and Brown (1990) have predicted that medication will ameliorate deficits found on "slow" RT tasks first, which is consistent with the data in the literature.

4.3.3.2 Two-choice RTs

Out of twenty studies in which 2-choice RTs are reported, seven studies were specifically examining the use of advance information or the effects of medication and are reviewed above. Thus, there remain a further 13 studies in which 2-choice RT is measured (see Appendix 4). These 13 studies were examined to replicate the relationship between control group RT and parkinsonian deficit on another independent data set. There is a relationship between control and patient RT that is similar to the overall regression of the data in Figure 4.2 (slope = 0.84, compared to 0.83 in Figure 4.2). The relationship is less strong, accounting for only 41% of the variance ($r^2 = 0.41$, compared to $r^2 = 0.80$ in Figure 4.2), but this is not surprising given that these 13 studies are very heterogeneous with regard to the experimental purpose and design and the patients' characteristics. There is a great deal of variation in the types of responses (ranging from key presses and releases to wrist flexion) and in the types of stimuli (visual, tactile etc.) and in their compatibility with responses. Nevertheless, these studies provide an additional demonstration of a relationship between control group RT and patient deficit on a third set of data.

4.4 Simple RTs

4.4.1 Review

Simple RT tasks (a single, predefined response to an imperative signal) provide important information about the parkinsonian deficit. Because they are straightforward, intervening cognitive processes are few and deficits can be examined without the need for complex psychological theories. Jahanshahi et al. (1992a) cite five studies before 1992 that compare the simple RTs of controls and patients with PD; a statistically significant parkinsonian deficit was found in all five. In fact, eleven studies before 1992 and seventeen after, have included such comparisons for conventional, non-EMG tasks (see Appendix 5). Of these, twenty studies have reported a statistically significant simple RT deficit for patients. As is evident from previous sections of this review, this level of consistency is unusual. The remaining eight studies do not report straightforward nonreplications and so these will be described more fully below.

The failure of two studies (Fimm et al., 1994; Reid et al., 1989) to find statistically significant slowing of simple RT amongst patients with PD can perhaps be explained by patient characteristics. The "untreated" group of Fimm et al. (1994) were 6 ms faster than controls on the simple RT task; they were also 5.5 years younger than controls. The "treated" group (age-matched with controls) showed a statistically significant deficit. Reid et al. (1989) studied patients with early- and late-onset PD. The late-onset group showed a clear (50 ms), but nonsignificant, simple RT slowing. The early onset group were only 10 ms slower than controls, but were again on average more than five years younger than controls.

Two studies also report significant simple RT slowing in one, but not all of their patient groups (Mayeux et al., 1987; Talland, 1963). Mayeux et al. (1987) divided their patients (*post hoc*) into subgroups with and without "bradyphrenia" on the basis of scores on the continuous performance test. The "with bradyphrenia" group showed a dramatic simple RT deficit, but the "without" group showed no statistically significant slowing. Without this *post hoc* division, there is a 240 ms deficit for the patient group as a whole. Talland (1963) showed a simple RT speeding for his mildly and moderately affected patient groups (71 and 39 ms, respectively), but a clear slowing for his "severe" group. These speedings are anomalous and hard to explain.

Nakashima et al. (1993) report no statistically significant slowing of premotor (EMG) simple RT. If their measure of "movement time" is added to this figure, however, a standard button press RT latency can be produced and patients show a 114 ms deficit.

Montgomery and Nuessen (1990) find no significant simple RT deficit, but the problems with this study have already been described above in the "medication" section. Only two studies (Harrison et al., 1995; Zimmermann et al., 1992) stand apart by showing no quantitative slowing of any sort, with no obvious methodological reasons for this.

If all of the nonsignificant deficits reported above are averaged, (omitting Montgomery and Nuessen (1990) due to difficulty in obtaining figures from the graphs presented in the paper), a mean slowing of 17 ms is still present. If the anomalous "speeding" conditions of Talland (1963) are dropped, this figure rises to 31 ms. Clearly, these studies do not significantly detract from the overall picture of simple RT slowing in PD. A t-test, comparing means of control and patient simple RT across all studies, confirms a significant patient deficit ($t(116) = 6.29, p < 0.001$); mean patient deficit is 71ms.

An explanation for this consistent deficit has been put forward by Goodrich et al. (1989). They suggest that healthy subjects can recruit some form of attention-demanding process to prepare and speed their simple RT performance. A simple RT task can be completely prepared as all aspects of the response are prespecified, whereas in a choice RT task, similar preparation is not possible as the response remains to be specified by the imperative signal. Goodrich et al. (1989) suggest that this attention-demanding process is not available to patients. They tested this hypothesis using a dual task method; patients with PD and controls performed a simple and a two-choice RT task, in the presence and absence of a secondary task (reading aloud). In the absence of the secondary task, the patients with PD were significantly impaired in the simple RT task, but their choice RT was not significantly longer. The secondary task slowed both simple and choice RTs for both groups, but the controls' simple RTs were slowed much more than those of the patients. In the presence of the secondary task, a simple RT deficit was no longer present in the patients - reading aloud had impaired the controls' performance such that they performed at the same level as patients. Goodrich et al. (1989) argued that the secondary task had removed the controls' ability to recruit extra attentional resources to speed their simple RTs. This study was a replication of a similar experiment by Bloxham et al. (1987) where the secondary task was motor - the authors note that the patients perform "as if they were constantly performing a motor task even though apparently at rest" (p1182, Bloxham et al., 1987).

Presumably, such attentional preparation takes time. Jordan et al. (1992) varied cue-stimulus intervals in their simple and go-nogo choice RTs, hypothesising that a deficit in attentional focusing should lead to a greater slowing of RTs for patients at very short (100 ms) cue-stimulus intervals. At these intervals there is presumably less time to prepare,

or "focus". No such deficit was found and Jordan et al. concluded that the patient SRT deficit must be due to poor attentional focusing within the first 100 m after the cue (100 ms was their shortest delay). Jordan et al. may have misinterpreted the hypothesis, however; Goodrich et al. would predict that short delays (if they interfere with attentional focusing) would impair controls' performance, bringing it down to the level of patients. It is controls who can recruit the extra attentional resources that are inaccessible to patients to speed their simple RTs, so it is their performance that should be disrupted by short delays.

4.4.2 Comment

Jordan et al.'s (1992) data do not falsify the "attentional focusing" hypothesis. It is a potential candidate for explaining the nonspecific simple RT deficit in patients with PD. Goodrich et al. (1990) have found that a number of different secondary tasks selectively disrupt simple RT performance in healthy subjects and have also extended this finding to different stimulus modalities and responses. There is clearly some form of modality-independent supplementary process speeding simple RTs in healthy subjects. The identification of such processes in healthy subjects adds weight to Goodrich et al.'s (1989) hypothesis, as does the fact that a number of other "attentional" processes are disrupted in PD (vigilance, Mayeux et al., 1987; attentional set-shifting, Downes et al., 1989).

4.4.3 Quantitative analysis

Both of the above analyses (Figures 2 and 3), as well as the above quantitative analysis of two-choice RTs, show that the parkinsonian deficit is dependent on control group RT. However, a simple RT deficit is found in nearly all of the studies reviewed, which suggests that simple RT may not be subject to the same relationship with control group RT that is seen for choice RT.

There are 26 studies in which simple RTs are reported for patients and control subjects (Appendix 5). When simple RT for patients is plotted against control group RT (Figure 4.4), there is no relationship between patient deficit and control group RT (See Figure 4.4, inset). In fact, the slope of the relationship between control and patient RT is 0.98 ms/control ms, with an overall RT deficit in the patients equivalent to 76 ms. There is very large range of mean RT values for the control subjects, from 215 to 455 ms. Procedural differences are unlikely to account for this, as simple RT tasks are so basic as to allow for little variation. This range is most likely the result of well-known differences in RTs to different kinds of stimuli or involving different kinds of responses (see Luce (1986), p. 91), for example, the value of 215 ms was a vibrotactile study (a finger-lift response

made to a tactile stimulus) (1989) whereas the value of 455 ms was obtained in a study in which a key release response was made to a visual imperative signal (Jahanshahi et al., 1992a).

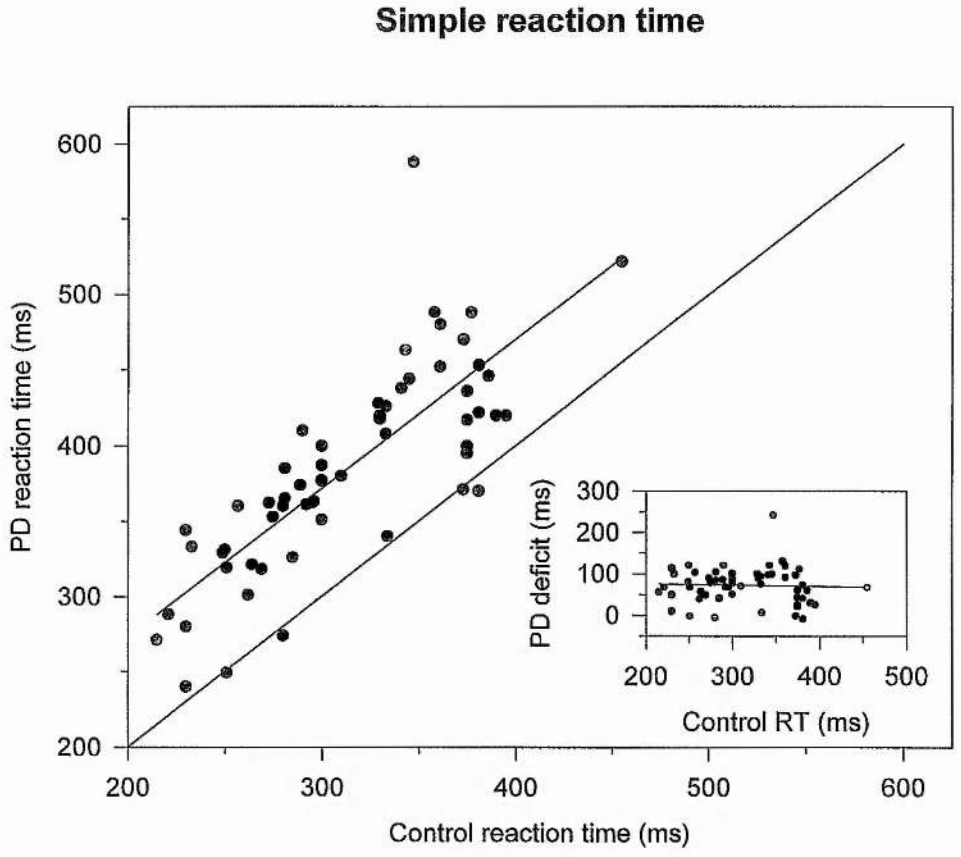


Figure 4.4. Simple RTs for patients ("PD") and controls (Appendix 5). INSET: Parkinsonian deficit (patient RT minus control group RT; "PD deficit") versus control group RT, inclusions are the same as for the main graph.

In the case of choice RT, RT differences between tasks are, at least in part, assumed to be due to central processing differences (for example, compatibility of the stimulus with the response). However, for simple RT (defined as those tasks in which there is a single response and, typically, only one imperative signal) the range of control values is unlikely to represent central processing time, as the cognitive load of all simple RT tasks can be assumed to be equivalent. Therefore, it can be concluded that the relationship between control group RT and patient deficit seen for two-choice RT above is unlikely to be due to

peripheral factors such as signal detection or muscular recruitment and is more likely due to central processing factors.

4.5 General discussion

4.5.1 Conclusions of the literature review

Many investigators have used RT tasks to quantify performance impairments caused by PD. Although RTs are a poor index of disease status, they have nonetheless allowed the identification and investigation of parkinsonian movement initiation deficits that are independent of bradykinesia. The results of the many studies of RT in PD have, however, been equivocal and often contradictory.

This review may have settled the controversy over the existence of a motor preprogramming deficit in PD, as it shows that the best-designed studies have failed to find such a deficit. Methodology and design have also been shown to be critical factors in studies of the effects of medication withdrawal on RT in PD. Specifically, the studies using more complete withdrawal procedures have been the most likely to find significant RT decrements after withdrawal, and it also appears that choice RT performance is more vulnerable than simple RT performance to the removal of medication.

A deficit in simple RT performance is the most consistent finding in the literature. A minority of studies have failed to find a statistically significant simple RT deficit, but these often report a nonsignificant slowing or use less than optimal control groups. A quantitative analysis of past studies has shown that the presence of a simple RT deficit in PD is unrelated to the speed at which controls can perform that task; a deficit is equally likely to be found on tasks that controls perform with fast or slow RTs. This is in marked contrast to the pattern of results found for choice RT studies, implying that Goodrich et al.'s (1989) contention that 'attentional focussing' is exclusive to simple RTs may be correct. In three independent sets of data, a quantitative analysis has shown that patients with PD are more likely to show RT slowing on tasks that controls can perform with fast RTs. This may account for some of the ambiguities and contradictions in the literature (such as the presence of a cued, but not an uncued RT deficit in some studies). The mechanism underlying this effect may be a genuine psychological factor, or it may prove to be a statistical or methodological artefact. Other interpretations of this effect that do not involve central psychological factors will now be considered.

4.5.2 Alternative explanations of the results of the quantitative analysis

4.5.2.1 Scaling factors

As RT increases, so does variance (see Luce, 1986). Hence, as RT increases, the difference in RT required to reach statistical significance also increases. It is therefore possible that

given equivalent patient deficits in two conditions, which result in similar increases in RT, a deficit in the slower condition might not reach statistical significance to due to greater variance. If Figures 4.2, 4.3 and 4.4 plotted effect size against control group RT expect effect size would be expected to decrease with increased control group RT, but in fact they plot group mean RTs, irrespective of variance. Thus, the relationship between control group RT and patient deficit seen cannot be attributed to a gradual increase in variance with RT.

4.5.2.2 Truncation of data

Most investigators using RT measures (and all of the studies reviewed here) report mean values and pay little, if any, attention to the underlying RT distribution. Furthermore, in many studies there is a "window" of RT, with responses outwith this window being considered either "anticipatory" or "late" errors and these RTs are discounted. If such a window were set too low, a proportion of the genuine RTs would be considered late and excluded, so reducing the mean value spuriously. This would be an issue particularly in more complex tasks, in which control performance is slower; a greater proportion of the patients', longer, RTs would be excluded, so reducing the difference between the patient and control groups. The use of RT windows in these studies was examined. Only 12/44 studies explicitly stated that RT windows were used and in these the maxima ranged from 850 to 15000 ms. It is possible that other studies used, but did not report, RT windows, but in the absence of this information, no conclusions can be drawn about the effects of data truncation.

Nevertheless, it seems unlikely that the relationship is spurious and due to truncation of data; the effect of using a RT window is to place a ceiling on RT, but from the data plotted in Figure 4.3, it is clear that the RTs of the medicated patients were not at the level of an artificial ceiling, because non-medicated RTs were yet slower.

4.5.2.3 Spurious relationship

Any correlation of two factors raises the issue, to what extent is the variable a causal factor as opposed to being itself correlated with a third, genuine, causal factor? In the case of RTs, there are many aspects of a task which might determine the speed of RT in that task. The most obvious of these is task complexity - the more complex a discrimination, the longer the RT is likely to be. Thus, if patients are differentially impaired as a function of task complexity per se, it might appear that speed of control performance was important. However, whereas it is not unreasonable to assume that patients would be more impaired in more difficult tasks, justification for the opposite hypothesis (that they would be more

impaired in simpler tasks) is harder to find. When Evarts et al. (1981) first reported the difference between simple and choice RT in patients with PD, their hypothesis had been that the patients were more likely to be impaired in the more complex task and it was with considerable surprise that they reported the opposite finding and abandoned this line of investigation. Therefore, it is unlikely that complexity of the task is the critical variable which underlies the relationship shown here between control group RT and patient deficit.

4.5.3 Conclusions from the quantitative analysis

4.5.3.1 Dopamine as a limited resource

It has been previously suggested (Robbins and Brown 1990) that fast RT performance might be differentially dependent upon dopamine, such that with progressive dopamine depletion there is progressive deficit in the maximum speed of response. Initially, fast RTs will be impaired. As depletion increases, this impairment will spread to slower RTs. Dopamine replacement therapy works in the complementary direction, restoring first the performance in the conditions which need less dopamine (slow RTs). This suggestion is consistent with the results of the quantitative analysis, but it requires refinement as some apparently falsifying instances exist.

4.5.3.2 Qualifications of the hypothesis

A number of individual studies explicitly contradict this "resource limitation" hypothesis; in these studies, control group RT does differ across experimental conditions, but the parkinsonian deficit is not necessarily greatest in the "fastest" condition (e.g. Cooper et al., 1994). However, there are specific circumstances which might account for the apparent lack of consistency in the literature.

If the demands of the task used interact with any of the other cognitive deficits of PD, the relationship seen between control group RT and parkinsonian deficit may no longer hold true. For example, Cooper et al. (1994) show a clear increase in parkinsonian deficit as control group RT increases, but there are multiple cognitive demands in their task that push control group RT to over 600 ms in some conditions. This is better described as measuring "thinking time" and their choice RT condition correlates highly (r=0.44) with digit ordering, a test of executive function. Cooper et al. used a go-nogo RT task in their study, and failed to control for reponse probability across conditions; as cognitive demand increased, response probability decreased. Study 5 in the next section shows that the RTs of patients with PD increase with decreasing response probability (see Section 5.4), and this factor might also account for the increase in patient deficit with control group RT.

The pattern of results described above may also not apply in patients with more severe PD, involving more widespread neuropathology. The study that has included the most severely affected patients (Lichter et al., 1988) has shown a greater impairment for the slower uncued choice RTs than for fast cued RTs (see "motor preprogramming" section above), which resembles the greater slowing of choice compared to simple RT seen in other neurological conditions. Too few studies have been performed on severely affected patients to confirm this and quantitative analysis is precluded due to difficulties in comparing the various clinical rating scales that have been used to quantify disease severity. A crude analysis was performed, dividing our database by median split into more and less affected PD groups in a semi-arbitrary fashion, using rating scale data where available. When patient RT is plotted against control group RT, the regression line for the less affected group shows the characteristic slope of less than 1 (Figures 4.2 and 4.3), whereas the line for the more impaired group has a slope of more than 1 (a greater deficit at longer control group RTs). Given the variety of studies plotted and arbitrary classification, this can only be seen as a tentative result. Also, the measures of disease severity used were group means and each study included patients with a range of different severities.

4.5.3.3 Final conclusions

The quantitative aspect of this review has shown that a significant deficit in the RT of patients with PD is more likely to be found in 'fast' choice RT tasks - those in which controls respond with a fast RT. A deficit is also likely to be found on all simple RT tasks, as these are all instances of 'fast' tasks. This is a description of a phenomenon. Although it has been suggested that this finding could represent a "limitation of resources", in which dopamine required for fast RTs is absent or depleted, the cause of the effect has nevertheless not been investigated. Recognition of the effect is more important than immediate efforts to establish a cause. Future investigators must take care not to postulate complex accounts of the source of parkinsonian RT deficits without first accounting for the relationship of those deficits to control group RT.

5.0 Study 5: The response of patients with Parkinson's disease to temporal and spatial probability information

5.1 Introduction

Section 4.0 has demonstrated that patients with Parkinson's disease (PD) have slowed movement initiation times, which can be shown using a variety of reaction time (RT) tasks. An RT deficit is also seen in rats that have been rendered 'parkinsonian' by dopamine depletion of the dorsal striatum (Carli et al., 1985). The RT deficit seen in patients with PD is specific to movement initiation, and has been shown to be independent of bradykinesia (Evarts et al., 1981). Although the cause of the RT deficit is ultimately dopamine depletion, the functional processes that are impaired as a consequence of dopamine depletion are unknown.

Most RT tasks make use of a variable foreperiod (also known as 'delay' or 'stimulus onset asynchrony') between trial initiation and the arrival of the imperative stimulus. This ensures that subjects cannot speed their responses by anticipating the time of arrival of the imperative stimulus. RTs have consistently been shown to decrease as a function of increasing delay in humans and other animals (Luce, 1986). This speeding of RT with foreperiod ('delay-dependent speeding') is thought to reflect a readiness to initiate movement, or motor set, that is more advanced at longer foreperiods (Näätänen and Merisalo, 1977). Unilateral striatal dopamine depletion impairs this motor readiness for contralateral responses in rats (Brown and Robbins, 1991), greatly reducing the effect of foreperiod on speeding of RT. If patients with PD showed a similar reduction of delay-dependent speeding, this might account for their RT deficits on a range of tasks.

However, two studies which have specifically looked for an impairment in this regard have reported that patients with PD show normal delay-dependent speeding (Bloxham et al., 1987; Jahanshahi et al., 1992a). Patients' response to a warning signal before an imperative stimulus also appears normal (Heilman et al., 1976; Jahanshahi et al., 1992a). In contrast, patients with PD do differ from controls in some aspects of their response to foreperiod. Bloxham et al. (1987) found that the increase in RT caused by a secondary task was related to foreperiod in controls but not in patients. Patients with PD also have more general problems in estimating and reproducing time delays (Pastor et al., 1992), so an altered response to foreperiod in an RT task would not be surprising.

The apparent contradiction of altered parkinsonian response to foreperiod (Bloxham et al., 1987) in the context of spared overall delay-dependent speeding may be resolved by an analysis of a typical RT task. Subjects may be responding to two factors when speeding their responses with delay. As delay increases, there is an increase in the probability of the occurrence of the imperative stimulus; if the

imperative stimulus does not appear at shorter delays, it must appear at the longest possible delay. Thus, RTs would be faster as a function of increasing temporal probability of stimulus occurrence. Alternatively, motor set may be a more simple process - subjects may just become more ready to move as time passes, leading to quick RTs at long delays. These two factors, 'temporal probability' and 'elapsed time' can be separated in an RT task (Frith and Done, 1986). This approach can be used to explain the nature of any parkinsonian deficit. This study aims to show whether RTs vary as a function of elapsed time or probability of stimulus occurrence in healthy human subjects and in patients with PD.

There is evidence that RT varies as a function of the probability of stimulus occurrence at a particular location (Carpenter and Williams, 1995). Carpenter and Williams (1995) varied the likelihood of a left or right target to which an eye movement response was required. They found that median RT to a given target was predicted by the log prior probability of a target appearing at that location. This effect can also be described as the consequence of a type of motor 'set', as subjects' 'internal' knowledge of probability allows them to speed their movements accordingly. As RT varies with the probability of stimulus occurrence at a location ('spatial probability'), it is predicted that RT will change as a function of the temporal probability of stimulus occurrence, rather than with increasing elapsed time. If RT does vary with temporal probability, it is possible that a single process governs the relationship between both spatial and temporal probability and RT. A manipulation of spatial probability has been included in this study to assess whether any deficit in response to temporal probability in patients with PD is also seen in their response to spatial probability.

5.2 Method

5.2.1 Subjects

5.2.1.1 Patients with PD

Eight patients with PD took part in this study. All were outpatients at Dundee Royal Infirmary, under the treatment of a consultant neurologist (Dr. Richard C. Roberts, University of Dundee). Inclusion criteria for the study were as follows: (a) diagnosis of mild or moderate idiopathic PD, (b) absence of dementia, (c) absence of any other neurological problems, (d) no history of neurosurgical intervention. Disease severity, as rated by a consultant neurologist (Dr. Roberts), varied between Hoehn-Yahr stages 1 and 2.5, with an average of 2.2 (Langston et al., 1992). All patients were taking levodopa, and in addition four were taking benzhexol, two bromocriptine and one selegiline. All patients and elderly controls completed the following standard neuropsychological tests: the Mini-Mental State Examination (MMSE, Folstein et al., 1975), the Geriatric Depression Scale (GDS, Yesavage et al., 1982), and the National Adult Reading Test, (NART, Nelson 1982). Subject details are in Table 5.1.

5.2.1.2 Elderly controls

Ten elderly control subjects were recruited from the St. Andrews branch of the Royal British Legion and by advertisement. Subjects were excluded if they reported a history of neurological problems, or if they were taking drugs known to affect CNS function. Elderly controls did not differ significantly from patients in GDS or NART scores, or age. Patients scored significantly lower on the MMSE than elderly controls (U=17.5, p=0.032).

5.2.1.3 Young subjects

Fifteen young subjects were tested, drawn from the undergraduate and postgraduate community at the University of St. Andrews. All subjects gave informed consent before participating in the study.

	Patients	Elderly controls	Young controls
Age	66.3 (2.4)	68.3 (2.2)	24.1 (0.7)
MMSE	27.6 (0.8)	29.6 (0.2)	-
GDS	8.9 (2.0)	7.5 (2.5)	-
NART	110 (4.0)	119 (2.5)	-
Male : Female	5 : 3	8 : 2	5 : 10

Table 5.1: Neuropsychological and demographic data for all subjects. S.E.M. in parentheses.

5.2.2 Materials and procedure

Testing took place in offices either in Dundee Royal Infirmary, or in the School of Psychology at the University of St. Andrews. The experimental tasks were administered using a Macintosh Powerbook computer running a Psyscope (Carnegie Mellon University, 1994) script, using an external button box to record responses (Carnegie Mellon University Mark 5 button box, 1994). After obtaining consent, subjects were interviewed to establish their medical history and demographic variables. They were then introduced to the use of the button box. On top of this box were three coloured buttons 5 cm apart, with an L.E.D. positioned above each button. The central button was the 'home' button, and one 'response' button was situated to the left, and one to the right of the home button.

The experimental task was a two-choice reaction time task. Subjects were required to press and hold down the home button with the index finger of their dominant hand. After a variable foreperiod of either 250, 500, 750 or 1000 ms, an L.E.D. was illuminated next to one of the response buttons. This was the imperative stimulus, and subjects were required to release the home button and press the button next to the illuminated L.E.D. Reaction time (RT) was defined as the time between the appeareance of the imperative stimulus and the release of the home button. Movement time (MT) was the time between the release of the home button and the response to a response button. On some trials ('catch' trials in the Falling block - see below), an imperative stimulus did not appear and subjects were required to maintain pressure on the home button until a tone sounded. The exact instructions to subjects were as follows:

"In this experiment you will use the three keys in front of you. To start each trial, hold down the yellow [central] button. A light will appear which signifies the start of a waiting period. At the end of this time, a light will turn on over the left or the right button. You must release the yellow button and press the button that corresponds to the left or the right light. A light will appear over either the left or right button on every trial, but after varying delays. If you press the correct key the computer will 'beep' but if you get it wrong it will 'boop'. Please respond as quickly and as accurately as possible using the index finger of your writing hand. If at any time you need a break, just don't press the yellow button until you are ready. Please press the centre button to begin testing." In the Falling block, subjects were given the extra instructions: "On some trials a light will not appear over the left or right button. If this happens, keep your finger on the yellow key until the computer beeps, then let go of the key and press it again for the next trial."

The experiment was completed in two blocks of RT trials, separated by an interval in which neuropsychological tests were completed. At the end of the experiment, all subjects were asked whether they had noticed if responses were required more frequently to one button than to the other.

5.2.3 Design

5.2.3.1 Spatial probability

The experimental task differed from a classical two-choice RT task, as the probability of a response being required to the left or right response button was not equal. Twice as many responses were required to one response button than to the other. The location of the 'more probable' button (left or right) was balanced across subjects.

5.2.3.2 Temporal probability

The experiment was completed in two separate blocks of RT trials. The relationship between foreperiod and the probability of imperative stimulus occurrence differed between the two blocks. One block was named the 'Rising' block, and the other the 'Falling' block; the order of presentation of these two blocks was counterbalanced.

(1) Rising block: This block was a typical two-choice reaction time task. A response was required on each trial, and an equal number of responses were required at each foreperiod (250, 500, 750 and 1000 ms). As a consequence, the probability of the imperative stimulus appearing increased as foreperiod lengthened; if the imperative stimulus had not appeared at any of the first three foreperiods, it had to appear at the fourth. If both the rising temporal probability and the bias in spatial probability are

taken into account, the likelihood of the imperative stimulus appearing at each side for each delay can be calculated. These data are in Table 5.2.

(2) Falling block: The 'falling' block used the same foreperiods as the rising block. The number of responses required at each foreperiod was unequally distributed, however, such that the imperative stimulus was most likely to appear at the shortest delay, and the probability of its appearance decreased as foreperiod lengthened. Catch trials were introduced, in which the imperative stimulus did not appear and no response was required. Thus, the majority of responses were required at the shortest foreperiod, and as time increased it became less likely that the stimulus would appear, and more likely that the trial would in fact prove to be a catch trial. The probability and number of responses for each side at each delay is shown in Table 5.2.

Foreperiod (ms):	250	500	750	1000
Rising Block:				
more probable	0.167	0.333	0.500	0.667
less probable	0.083	0.167	0.250	0.333
Falling Block:				
more probable	0.383	0.150	0.033	0.017
less probable	0.192	0.075	0.017	0.008

Table 5.2: Probability of stimulus occurrence at the more or less frequent target location at the four foreperiods (250-1000ms) for the Rising and Falling blocks.

5.2.4 Data analysis

Post-hoc cutoffs were applied to both the RT and MT data. RTs of less than 150ms or more than 1000ms were excluded from the analysis. MTs of above 5000ms were also excluded. An error was recorded when a subject pressed the incorrect button.

5.2.5 Predictions

Following the findings of Frith and Done (1986), it was predicted that RT would vary with temporal probability of stimulus occurrence rather than elapsed time in healthy subjects. It was not possible to make strong predictions about the response of patients with PD to probability of stimulus occurrence due to the inconsistency of the past literature. However, it was strongly predicted that patients with PD would show an overall RT deficit. Figures 4.2 and 4.3 clearly show that patients with PD always have RTs that are similar to or slower than those of controls.

5.3 Results

5.3.1 Young Subjects

RTs for the young control subjects are depicted in Figure 5.1. Visual inspection of this graph indicates that RTs fall as foreperiod lengthens in the Rising block and increase with lengthening foreperiod in the Falling block (Block x Foreperiod interaction ($F(3, 42) = 24.42, p < 0.001$). This indicates that RTs cannot be changing solely with elapsed time when they change with foreperiod, as the temporal distribution of foreperiods was identical across blocks. The temporal probability of stimulus occurrence at each foreperiod is noted in parentheses in Figure 5.1; it is clear that RTs vary with the temporal probability of stimulus occurrence rather than elapsed time. The effect of spatial probability on RT also varied across blocks. RTs were faster to the more probable target in both blocks (main effect of Target $F(1, 14) = 11.2, p < 0.01$), but this effect was more pronounced in the Falling block. (Block x Target interaction $F(1, 14) = 5.36, p < 0.05$). The Falling block also differed from the Rising block in absolute levels of RT; RTs were generally slower in the Falling block than in the Rising block. This was further confirmation of the effect of temporal probability on RT. The probabilities associated with stimulus occurrence at each foreperiod were generally lower in the Falling block than in the Rising block (Table 5.2, and Figure 5.1 values in parentheses). Thus, if RT is speeded by greater probability of stimulus occurrence, slower RTs would be predicted in the Falling block. The RTs of the young control group clearly show that RT varies with both temporal and spatial probability in healthy young humans.

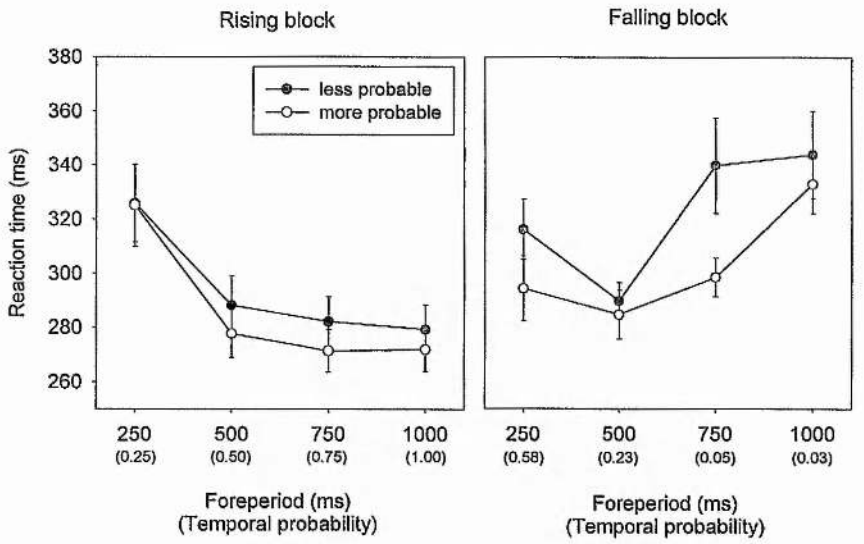


Figure 5.1: Mean RT for young subjects' responses to the more and less probable buttons at each foreperiod. The temporal probability of a stimulus occurring at a given forepeiod is in parentheses below that foreperiod.

5.3.2 Patients with PD and elderly controls

5.3.2.1 Reaction times

Figure 5.2 shows that the RTs of the elderly controls and the patients with PD followed a similar pattern to those of young subjects. As in young subjects, RTs fell with rising temporal probability and lengthened with falling temporal probability, indicating that subjects responded to temporal probability rather than elapsed time (Block x Foreperiod interaction ($F(3, 39) = 33.48, p < 0.001$). Patients' and controls' RTs showed similar changes as a result of changing temporal probability, indicating that the response of patients with PD to temporal probability was normal (Group x Block x Foreperiod interaction $F(3, 39) = 0.62$). However, neither patients with PD nor elderly controls showed faster RTs to the more probable response location in any part of the experiment (main effect of Target and all interactions with Target n.s.). This occured in spite of the fact that all patients and controls correctly reported the location of the more probable target. Although patients with PD did not differ from elderly controls in their response to temporal or spatial probability (all interactions with Group n.s.), their RTs were generally slower than those of the elderly controls. The patients initiated their

movements on average 49 ms more slowly than elderly controls (main effect of Group, $F(1, 13) = 3.4$, $p < 0.05$, one-tailed).

In summary, patients with PD responded normally to temporal probability. Neither patients nor elderly controls showed an effect of spatial probability, in marked contrast to the results found for young subjects. In the absence of an effect of spatial probability, it is impossible to assess the response of patients with PD to this variable. Patients initiated their movements more slowly than elderly controls, showing a global RT deficit.

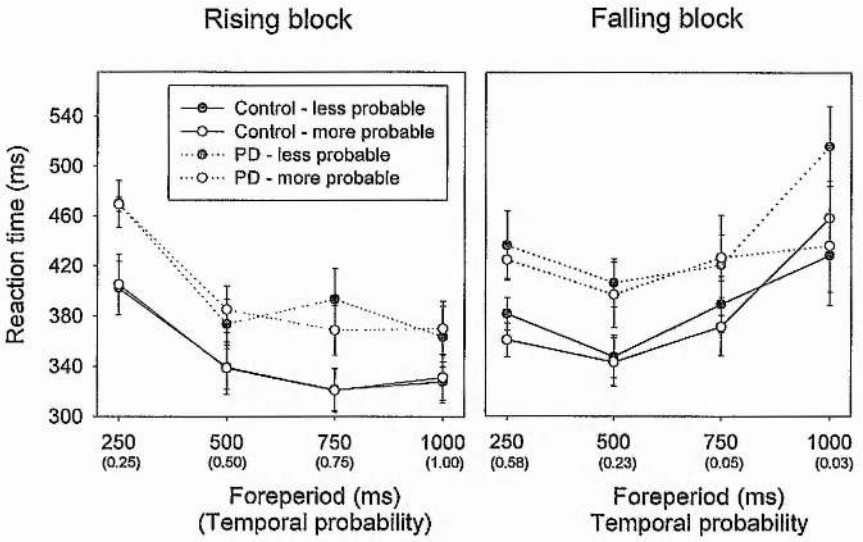


Figure 5.2: Mean RT for patients (broken line) and elderly controls (unbroken line) to the more and less probable response buttons at each foreperiod. Values in parentheses represent the temporal probability of stimulus occurrence at each foreperiod.

5.3.2.2 Movement times and errors

Mean MTs were calculated for each subject. Patients with PD showed significantly lower MTs than elderly controls ($t(8.84) = 3.54$, $p < 0.01$, d.f. corrected for unequal variances). It was not possible to carry out a statistical analysis of the difference in error rates between the two groups, as subjects made very few errors. Controls made a total of 3 errors (out of around 4800 trials), and patients with PD made 16 errors.

5.4 Discussion

This study aimed to establish whether the delay-dependent speeding of subjects' RT is due to their response to the temporal probability of stimulus occurrence, or the amount of time elapsed after trial initiation. Patients with PD were tested to assess their response to these two factors, and to establish whether an altered response to foreperiod could account for the parkinsonian RT deficit. Healthy young subjects' RTs varied as a function of the temporal probability of stimulus occurrence rather than elapsed time, higher probabilities leading to faster RTs. Young subjects also responded to a manipulation of spatial probability, showing faster RTs to a more probable target. Patients with PD and healthy controls responded to temporal probability rather than elapsed time, but showed no response to spatial probability, in contrast to young subjects. Although patients with PD did not differ from elderly controls in their response to temporal probability, they did initiate their movements more slowly than controls, showing an overall RT deficit.

The performance of young subjects in this study replicates Frith and Done's (1986) finding that delay-dependent speeding is due to a response to the temporal probability of the occurrence of the imperative stimulus. The increase in RTs seen with lengthening foreperiod in the Falling block is clear confirmation that subjects' response to temporal probability was far stronger than any response to elapsed time. Variations in spatial probability also altered young subjects' RTs. Carpenter and Williams (1995) observed a similar response to spatial probability to that seen in this study, finding that eye movement RTs were faster to a more probable target. They went on to speculate that increasing temporal probability might also speed RTs, and that this speeding might be governed by the same process that mediates response to spatial probability. The role of temporal probability in determining RT has been confirmed in this study, and the fact that young subjects responded to both spatial and temporal probability raises the possibility that a single process governs RT variation with spatiotemporal probability. However, data from healthy elderly subjects and patients with PD suggest that RT response to spatial and temporal probability are dissociable.

Healthy elderly subjects and patients with PD showed clear changes in RT with changing temporal probability, but showed no response to a manipulation of spatial probability. As young subjects responded to both spatial and temporal probability, the data from elderly subjects suggest that response to spatial probability is differentially vulnerable to aging, and that responses to spatial and temporal probability might be

dissociable. However, this finding is only a single dissociation. It is possible that a single process governs response to spatial and temporal probability, but that this process is diminished or weakened in elderly subjects compared to young subjects. If the manipulation of spatial probability used in this experiment was weak compared to the manipulation of temporal probability, any reduction in a subject's response to probability would abolish the effect of spatial probability first. However, a consideration of the elderly subjects' response to temporal probability suggests that this interpretation of the results is not appropriate. If a single process governing response to both spatial and temporal probability were being diminished by aging, some reduction in response to temporal probability would be caused by aging. However, elderly subjects' response to temporal probability was at least as great as that of young subjects. For example, in the Rising block increasing temporal probability caused healthy elderly subjects' RTs to drop by as much as 84 ms; the same changes in temporal probability caused the RTs of young subjects to drop by a maximum of 54 ms. Similar results are seen in the Falling block (115 ms vs. 54 ms). This comparison of young and healthy elderly subjects indicates that aging causes a loss of response spatial probability in the presence of preserved or accentuated response to temporal probability. This dissociation provides evidence that the processes mediating RT change to alterations in spatial and temporal probability are functionally distinct. Thus, it appears that two types of motor set can be active in an RT task. Both types of set involve the internalisation of information about the probability (spatial or temporal) of a response and the use of this information to speed movement. These two types of motor set will be referred to as 'temporal set' and 'spatial set'.

The response of patients with PD to spatial probability cannot be assessed as neither the patients nor the healthy elderly control group showed any response to a manipulation of spatial probability. However, patients' response to temporal probability - that is, temporal set - was normal. Their response to rising and falling temporal probability mirrored that of elderly control subjects, although the RTs of patients were generally longer than those of controls. The existence of a normal response to temporal probability in the presence of the well-replicated parkinsonian RT deficit indicates that altered response to temporal probability cannot be the source of the parkinsonian RT deficit. Specifically, patients with PD have no problems with acquiring information about the relationship between foreperiod and probability of stimulus occurrence, and this information plays a part in speeding their responses.

In order to maximally speed their RT, subjects must acquire information about the relationship between probability of stimulus occurrence and foreperiod. For example, information about the duration of the longest delay is important; if subjects acted as if the longest delay on this task was 750 ms, speeding at the 1000 ms delay would be attenuated. This learning is presumably implicit, as subjects do not usually report explicit strategies of time and probability estimation. Despite reports of impaired implicit ('procedural') learning amongst patients with PD in serial RT tasks (Ferraro et al., 1993), patients learn the relationship between probability and foreperiod as successfully as controls. Implicit learning of response sequences may be impaired in PD, but learning of temporal probability information is clearly intact. A comparison of the results of this study and of those of serial RT tasks also completes a double dissociation between implicit and explicit knowledge in RT tasks. In this study, elderly subjects and patients with PD could correctly report the bias in spatial probability towards one target, but this did not lead to a RT speeding towards this button. In contrast, individuals performing serial RT tasks show an RT benefit when responding to a repeating sequence of targets despite being unaware of the existence of this sequence (Nissen and Bullermer, 1987).

Once information about time and probability of stimulus occurrence has been acquired, it must be used to speed responses. Some researchers have proposed that patients with PD have problems using advance information to speed movements (Bloxham et al., 1984), or more generally in the use of 'internal' information that is unsupported by external cues (Brown and Marsden, 1990). This study shows that patients with PD have no problems using the kind of 'internal' 'advance' information required for delay-dependent speeding of RT. Intact delay-dependent speeding also implies that patients are successful in estimating time delays, in contrast to previous reports (Pastor et al., 1992). Previous studies have identified deficits of absolute time estimation, however, whereas delay-dependent speeding only requires assessment of relative length of delay.

As the acquisition and use of temporal probability information is intact in patients with PD, the anatomical substrate of these processes is presumably extrastriatal. Putaminal dopamine metabolism correlates with locomotor disability in patients with PD (Holthoff-Detto et al., 1997), and dopamine depletion of the dorsal striatum in the rat causes an RT deficit (Carli et al., 1985). An RT deficit is present in patients in this study, and is presumably striatal in origin - the intact temporal set seen in patients must be anatomically as well as functionally distinct.

In conclusion, this study has shown that the change of RT with foreperiod is due to subjects' response to the temporal probability of the arrival of the imperative stimulus. Although both spatial and temporal probability of stimulus occurrence can alter young subjects' RTs, different processes appear to mediate the reponse to these variables. Patients with PD respond normally to foreperiod, and therefore have normal temporal set. This replicates previous findings of normal response to foreperiod or a warning signal in patients with PD (Bloxham et al., 1987; Heilman et al., 1976; Jahanshahi et al., 1992a), and indicates that an abnormal temporal set cannot be the cause of the parkinsonian RT deficit. The processes that mediate the perception of temporal probability information and its use to speed RT must be intact in PD, and are presumably dependent on areas outwith the dorsal striatum.

6.0 Study 6: The neuropsychological consequences of pallidotomy for Parkinson's disease

6.1 A review of studies up to the end of 1997

The neuropsychological consequences of pallidotomy for PD are not well understood. A rich literature has emerged on the clinical and motor improvements seen after pallidotomy, but pre- and postoperative neuropsychological testing has been neglected in comparison. The current surge of interest in pallidotomy was started by Laitinen et al.'s paper in 1992, yet Baron et al. (1996) were the first to mention neuropsychological testing in a full-length paper, and papers concentrating on neuropsychological followup only started to appear in 1997 (e.g. Soukup et al. 1997), five years and over one thousand operations later. This neglect of pre- and postoperative neuropsychological assessment is shocking. Transient declines in cognitive function were frequently reported after pallidal surgeries for PD performed in the 1950s, which strongly indicates a need for neuropsychological assessment of the pallidotomies performed in the 1990s (McFie 1960; Riklan et al. 1960). Even if this literature did not exist, neuropsychological followup would seem essential to assess the safety of any procedure that involves ablating part of an elderly patient's brain.

This section reviews the reports of neuropsychological followup of pallidotomy that were published up to the end of 1997. The majority of these articles were published in abstract form, and describe the pre- and postoperative assessment of patients using a large battery of neuropsychological tests. There are methodological and practical limitations to some of the published research that must be considered before the results are reported, however.

6.1.1 Limitations of the review

(a) Publication in abstract form

10 of the 15 studies reviewed are reported in abstract or letter form (see Table 6.1). These studies have not been subject to peer review, and so their results must be treated with caution. The limitations of the abstract format mean that the names of the neuropsychological tests used are often not specified, and results are sometimes reported by cognitive domain (e.g. "no deficits were seen in visuospatial ability"). Methodological details are often omitted, such as the use of alternate forms to minimise practice effects in the postoperative testing session

Authors	Format	Patient numbers	Lesion	Controls
Iacono et al. '94	Abstract	15	? Unilateral	10 healthy
Cole et al. '95	Abstract	14	Unilateral	None
Baron et al. '96	Paper	15	Unilateral	None
Ghika et al. '96	Abstract	4	Bilateral	None
Green et al. '96	Abstract	13	Unilateral	None
Bowers et al. '97	Abstract	21	Unilateral	None
Lang et al. '97	Letter	>40	? - Unilateral	None
Manning et al, '97	Abstract	10	Unilateral	None
Masterman et al. '97	Abstract	24	Unilateral	None
Narabayashi et al. '97	Paper	12	Unilateral plus thalamotomy	Thalamotomy patients
Soukup et al. '97	Paper	14	6 left 8 right	None
Stebbins et al. '97	Abstract	9	Unilateral	None
Troster et al. '97	Paper	9	6 Left, 3 Right - stimulation	None
Uitti et al. '97	Paper	9	9 left	None
von Falkenhayn et al. '97	Abstract	10	? - Stimulation	None

Table 6.1: This table shows all of the studies that are reviewed in this section. A "?" in the lesion column indicates that the type of lesion is not explicitly stated. It is assumed that these lesions are unilateral.

(b) Variability in measurement

There is little agreement in the studies reviewed as to what constitutes "significant" postoperative improvement or deterioration. Most studies test the difference between pre- and postoperative group mean test scores, but levels of α vary. Some studies adopt a conservative α of $p < 0.01$ to allow for multiple comparisons in a battery of tests (e.g. Uitti et al. 1997), yet Stebbins et al. (1997) report a $p < 0.1$ reduction in digit span as a "decline in performance". Unless stated otherwise, this review will only report changes of $p < 0.05$.

The use of group means to assess performance can conceal clinically significant information. If pallidotomy leaves 10% of patients with significant deficits of executive function, this may not alter the group mean performance sufficiently to register a difference at $p < 0.05$. Tröster et al. (1997) maximise the information provided in their table of results by listing the number of patients who improve or deteriorate on a test by more than one or two S.D.s, as well as group means. This format still omits information; for example, it is impossible to determine if the three patients showing a > 1 S.D. decline in phonemic fluency are the same as those three that show a similar deficit of delayed free recall. A minority of deteriorated patients may, or may not, account for deficits seen across a number of cognitive domains. Future outcome studies must address individual patient outcomes, and the constellation of deficits in individuals.

(c) Variability of surgical procedure

This review includes assessments of both unilateral and bilateral pallidotomies, and studies often treat left and right hemisphere pallidotomies as a single group (for an exception, see Lang et al. 1997). Electrophysiological recording has been used to identify the target site in some studies, but not in others. There is insufficient data in the published studies to separate the effects of these factors. The results of all of these types of pallidotomy are reported together, and variability will be added to the data as a consequence.

(d) Methodological issues surrounding repeated testing

Pre- and postoperative testing requires that patients perform exactly the same neuropsychological test twice. Individuals' scores differ between 'test' and 'retest' phases in some tests due to practice effects. For example, once an individual has completed the WCST, s/he will be aware of the test's requirement to shift rule, which is unanticipated in the first presentation. Lezak (1995) describes the WCST as a "one shot" test (p. 623), yet three studies use the WCST pre- and postoperatively to assess changes in executive/attentional function (Soukup et al. 1997, Tröster et al. 1997, Uitti et al. 1997). None of these studies use controls to rule out practice/retest effects on the WCST, and only Soukup et al. (1997) concede that the lack of a control group represents a "significant limitation" to their study. In the absence of a control group, provision of alternate forms and consideration of the test-retest reliability of the instruments used is critical.

None of the 15 studies reviewed make and reference to the test-retest reliability of the components of their neuropsychological batteries. Some of the instruments used

will doubtless have a high reliability, but the fact that this is never referred to implies indifference. Two studies (Bowers et al. 1997, Soukup et al. 1997) specify that alternate forms of a test were used, but the remaining studies either do not use alternate forms or do not mention it. For example, Uitti et al. (1997) use components of the Wechsler Memory Scale - Revised (WMS-R) pre- and postoperatively; the WMS-R has no alternate form, so at 'retest', patients are relearning items that were novel at 'test'. Where alternate forms are available, there may still be a "test sophistication effect" (Soukup et al. 1997) that leads to improved scores at the second administration of the test. The provision of an appropriate parkinsonian control group, tested at similar intervals to the surgical patients, can effectively control for this. Only one out of the 15 studies reviewed here used a control group, and it consisted of age-matched healthy volunteers.

In summary, many important methodological details are not provided in the studies reviewed, and where these details are available, the methods used are often not ideal. As a consequence, only limited conclusions can be drawn from this review. The review can inform us about which domains of cognition and which neuropsychological tests are most sensitive to pallidotomy. It cannot provide clinically important information such as the percentage rate of significant negative or positive neuropsychological outcomes. This information is critical if the full potential risks and benefits of pallidotomy are to be assessed. Fully informed consent cannot be given by patients in the absence of this data. The methodological limitations of the studies reviewed also mean that basic science questions about the function of the globus pallidus and basal ganglia-thalamocortical loops cannot be conclusively answered. The ideal neuropsychological followup of pallidotomy would distinguish between bilateral, and right and left hemisphere unilateral lesions. A best medical care control group would be used, and there would be long term follwup, both to differentiate transient and chronic deficits and to assess the rate of progression of cognitive deficits in the surgical group. Future studies should identify any features in the preoperative neuropsychological profile that are risk factors for cognitive deterioration, and functional imaging techniques should be used to identify the neural substrates of cognitive deficits or improvements.

6.1.2 Review

The results of the studies reviewed are summarised by cognitive domain (measures of overall competence, memory, attentional/executive function, language,

visuospatial/visuoconstructive ability). A category is included for mood/neuropsychiatric change, as the onset or relief of (for example) depression will influence scores on many neuropsychological tests. The categories used are inclusive; as most neuropsychological tests require multiple cognitive abilities, some tests will appear in more than one category. For example, verbal fluency tests will appear in both the 'language' and 'executive function' sections.

6.1.2.1 Measures of overall competence

Some studies have reported the effects of pallidotomy on global measures of cognitive ability, such as IQ and dementia rating scales. Predominantly, no change is reported. Narabayashi et al. (1997) report no postoperative change in WAIS score, and Uitti et al. (1997) find no significant alteration in WAIS verbal IQ. Manning et al. (1997) are presumably referring to a full scale IQ test when they report unchanged "overall intelligence", just as they are probably referring to the MMSE or Mattis Dementia Rating Scale (DRS) when they find unchanged "dementia". Three further studies find unchanged total DRS scores (Baron et al. 1996, Tröster et al. 1997, Uitti et al. 1997), though change on one DRS subscale was found by Tröster et al. (1997), and is reported in the 'visuospatial' section below.

6.1.2.2 Memory

A range of instruments have been used to test immediate and delayed memory for verbal and pictorial material after pallidotomy. The most commonly used format for verbal material is that of the Rey Auditory-Verbal Learning Test (RAVLT), which has been developed into the California Verbal Learning Test (CVLT). This procedure requires the subject to learn a list of words over several presentations, and can yield scores of immediate and delayed recall, delayed recognition and interference. None of these measures changed after pallidotomy in the studies reported by Baron et al. (1997), Tröster et al. (1997), Uitti et al. (1997) (all using the CVLT) or Soukup et al. (1997) (using the RAVLT). Baron et al. (1996) found a decrease in delayed recognition memory scores between their postoperative and 1 year assessment, but disease progression may account for this deterioration. The final two studies that use the CVLT report contradictory results. Bowers et al.'s (1997) study is unique in finding improvements in delayed cued recall and recognition memory after pallidotomy; it is also the only study to report the use of alternate CVLT forms in the postoperative phase. In contrast, Lang et al. (1997) found an unspecified ("serial list learning") deficit in

patients with left pallidotomy that resolved after 6 months. The WMS-R logical memory subtest was also used to index verbal memory, but it showed no change after pallidotomy (Uitti et al. 1997, Tröster et al. 1997).

Digit span is a special case of verbal memory, due to the high attentional demands of the task. Uitti et al. (1997) found unchanged WAIS digit span scores, but Baron et al. (1996) found significant improvement in backward digit span between pre- and postoperative assessments. One year after surgery, Baron et al. (1997) found that this improvement had disappeared. In the absence of a control group, it is impossible to know whether to describe this as a 'transient improvement' that fades, or an improvement that is subsequently affected by normal disease progression.

The principal paradigm for testing memory for pictorial material has been to test recall of a complex figure, such as the Rey figure. The results arising from the use of this test are again inconsistent. Soukup et al. (1997) report no change in recall after pallidotomy, yet Lang et al. (1997) report a deficit in patients with right pallidotomy that resolves after six months. Most convincingly, Iacono et al. (1994) report that patients improve more than controls at the second (post-surgery) administration of the test. This is the only report of an improvement in pictorial memory, however; the WMS-R visual reproduction and figural memory subscales have been found to be unchanged after pallidotomy (Tröster et al. 1997, Uitti et al. 1997), and recognition memory for faces (Warrington Recognition Memory Test, Soukup et al. 1997) is also unaffected.

Procedural memory has only been tested in one study. Von Falkenhayn et al. (1997) used a four choice reaction time task, in which a ten-item sequence was embedded. When patients were not undergoing globus pallidus stimulation, normal implicit learning was seen in the form of decreased reaction time with increased exposure to the embedded sequence. With the stimulator switched on, however, reaction times became generally faster, but there was no implicit learning of the sequence. This is an interesting result, particularly as recent studies suggest the substrate of procedural learning to be the cerebellum, not the basal ganglia (Pascual-Leone et al. 1993). Despite this interesting result, however, there is no consistent, replicated evidence for improvement or deterioration of any aspect of memory after pallidotomy.

6.1.2.3 Attention and executive function

Tests such as the WCST and the Stroop test are often described as both tests of selective attention and executive function. A number of tests described in this section will fall

into both categories. This section of the review will proceed from the tests that are most 'attentional' and least 'executive' to executive tasks with minimal attentional involvement.

Soukup et al. (1997) administered the Bell test to detect problems with visual scanning or neglect, and the WAIS-R picture completion subtest which can be seen as testing attention to visual detail. No change was seen on either test after pallidotomy. Part A of the Trail making test also tests visual scanning, as subjects are required to connect numbered dots on a sheet in order; part B of the test introduces a 'divided attention' component as subjects are required to connect dots on the basis of two sequences that must both be kept in mind. Neither Bowers et al. (1997) or Uitti et al. (1997) found any postoperative change in either the 'attentional' part A or more 'executive' part B of this test.

The Stroop test and WCST are both sensitive to the effects of frontal damage, and both require subjects to pay selective attention to one aspect of a complex stimulus. The shifting component of the WCST also requires executive control, and executive function is needed to overcome attentional interference in the Stroop test. No changes have been reported in Stroop performance (Baron et al. 1996, Bowers et al. 1997, Tröster et al. 1997, Uitti et al. 1997) or on the WCST (Soukup et al. 1997, Tröster et al. 1997, Uitti et al. 1997) after pallidotomy. Masterman et al. (1997) also report no change in "set-shifting", which may refer to the WCST.

Verbal fluency tests have been used as indices of executive/frontal function, divergent thinking and verbal intelligence. There is certainly an executive component to the search for new words under specific constraints. Both letter and category (phonemic and semantic) fluency have been used in patients undergoing pallidotomy. Baron et al. (1996), Bowers et al. (1997) and Soukup et al. (1997) found letter fluency to be unchanged after pallidotomy. In contrast, Uitti et al. (1997) found a postoperative deficit in letter fluency, and Lang et al. (1997) found a deficit in left pallidotomy patients that resolved after six months. Category fluency appears to be even more sensitive to pallidotomy. Uitti et al. (1997) and Masterman et al. (1997) found significant postoperative decreases in category fluency, and Lang et al. (1997) found chronic category fluency deficits in left pallidotomy patients that were still present one year after surgery. Bowers et al. (1997) may have found intact category fluency, but it is unclear from their abstract whether unchanged "verbal fluency" includes a category fluency test. A full discussion of these verbal fluency results will be undertaken in the "Language" section below.

Finally, the cognitive estimates test was used by Tröster et al. (1997), and no postoperative change in score was found. If appears that executive function is not affected in the majority of patients undergoing pallidotomy. The presence of decreases in verbal fluency in the context of generally spared executive function may imply that language problems are the primary cause of this deficit. A minority of patients may still undergo changes in executive function, however; Lang et al. (1997) observed changes of a "frontal/executive" nature in 20% of their patients, which emphasises the need to report more than just group mean performance in pallidotomy followup studies.

6.1.2.4 Language

The only tests related to language that are affected by pallidotomy are the verbal fluency tests described above. Performance on the Boston Naming Test is unchanged after surgery (Baron et al. 1996, Soukup et al. 1997, Tröster et al. 1997, Uitti et al. 1997) as is vocabulary knowledge and performance on the Token test of the Multilingual Aphasia Examination (Uitti et al. 1997). Lang et al. (1997) note that 69% of patients with left hemisphere pallidotomies reported word finding complaints on questionnaires, however. It is unclear whether these patients were specifically asked about problems that had arisen after surgery, or simply about their current status (which may have been impaired preoperatively). It seems likely that surgery caused these problems, however, as only 25% of patients with right pallidotomies reported such complaints. It is possible that there are differences between the processes required to carry out an internally cued search for a word and those required to search for a word prompted by an external cue (as in the Boston Naming Test). If 'internal search' processes were more vulnerable to pallidotomy, this could account for problems in word-finding and verbal fluency in the presence of intact confrontation naming.

6.1.2.5 Visuospatial and visuoconstructive function

No consistent changes are seen in visuospatial function after pallidotomy. Block design (Soukup et al. 1997, Uitti et al. 1997) and judgement of line orientation (Baron et al. 1996, Soukup et al. 1997, Uitti et al. 1997) are unchanged after surgery. The Hooper Visual Organisation Test has been used in two studies, showing no postoperative change (Baron et al. 1996, Tröster et al. 1997). Bower et al. (1997) make a more general reference to unchanged "visuospatial" abilities. Tröster et al. (1997) are unique in finding a postoperative deficit on the construction subscale of the DRS, which was not seen in the other studies that used the DRS (see Section 1).

6.1.2.6 Mood / neuropsychiatric change

Depression scales are frequently used to determine whether the functional improvements seen after pallidotomy are accompanied by improved mood. Most studies find that levels of depression are unchanged. The GDS and Beck Depression Inventory (BDI) have registered no significant postoperative change in four studies (Cole et al. 1995, Baron et al. 1996, Tröster et al. 1997, Uitti et al. 1997). Two studies also used the Hamilton Depression Rating Scale, finding levels of depression unchanged after surgery (Cole et al. 1995, Baron et al. 1996). Narabayashi et al. (1997) did see an improvement in the depression score derived from the MMPI, and Tröster et al. (1997) saw a nonsignificant trend ($p < 0.06$) to improvement on the BDI, but the overall picture is of unchanged levels of depression. This can be used to argue, following Cole et al. 1995, that depression seen in patients with PD is not a reaction to physical disability, as it does not improve as activities of daily life improve after surgery. Parkinsonian depression may stem more from neurochemical changes than from an individual's appraisal of his or her situation.

Anxiety was shown to decrease after surgery in one study (Tröster et al. 1997, using the Beck Anxiety Inventory), but not in two others that used different instruments (State Trait Anxiety Inventory, Bowers et al. 1997; Spielberger State Anxiety Inventory, Baron et al. 1996). A drop in anxiety would be predicted intuitively; presumably, preoperative anxiety about surgery is high. Other reports of neuropsychiatric change give disparate findings. Narabayashi et al. (1997) found reductions in the MMPI indices of hypochondriasis and hysteria, whereas Masterman et al. (1997) found worsening scores on their Neuropsychiatric Inventory in seven out of nine patients who had preoperative problems. Abulia and obsessive-compulsive symptoms have been found in nonparkinsonian patients with bilateral globus pallidus lesions (Bhatia and Marsden 1994, Lopez-Villegas et al. 1997). These symptoms were also found in two of the four patients undergoing bilateral pallidotomy in Ghika et al.'s (1996) study. Unilateral pallidotomy may be safer; Masterman et al. (1997) found no group mean change in scores on the Starkstein Apathy Scale in patients with a single lesion. There are other reports of poor neuropsychiatric outcome after pallidotomy (such as Lang et al.'s (1997) 20% with "frontal/executive" changes), but no attempt has been made to determine an overall percentage risk for such changes. Neuroimaging studies will determine whether such outcomes are due to extrapallidal damage consequent to surgery.

6.1.2.7 Summary

Pallidotomy appears to have no serious adverse neuropsychological consequences for most patients. Visuospatial function is completely spared, and most tests of language and executive/attentional function show no change after the operation. There are some contradictory findings on memory function, but the reports of positive and negative changes would appear to 'cancel out' in the presence of a majority of studies finding no change. Verbal fluency is the only test that records deficits after surgery with some consistency, and this may be due to impairment of processes that mediate internally cued search for words. It remains possible, however, that a number of subtle deficits are being masked by practice effects. For example, a surgical patient's score might remain unchanged at 'retest', suggesting no change in cognitive function, but in fact s/he might not be showing the improvement due to practice seen in an unoperated patient.

6.2 Introduction to Study 6

Ventrolateral pallidotomy is gaining increasing acceptance as a treatment for severe Parkinson's disease (Favre et al., 1996 - see Section 1.7.3.5). Pallidotomy has been shown consistently to abolish or reduce levodopa-induced dyskinesias (Samuel et al. 1998), and some studies have reported that it also improves akinesia and bradykinesia (Baron et al. 1996, Lozano et al. 1995). However, there is relatively little consensus on the cognitive consequences of this surgery due to a relative neglect of pre- and postoperative neuropsychological assessment.

Studies of cognitive change after pallidotomy have generally used similar designs. A large battery of neuropsychological tests has been administered before and after surgery, and any postoperative change in test score has been attributed to the effects of the pallidal lesion (e.g. Soukup et al. 1997, Tröster et al. 1997, Uitti et al. 1997). The most consistent changes have been found using verbal fluency tasks, with Lang et al. (1997) and Uitti et al. (1997) finding postoperative decreases in letter fluency and three groups finding decreases in category fluency (Lang et al. 1997, Masterman et al. 1997, Uitti et al. 1997). A number of other neuropsychological changes have been reported, but there is little consistency between the results of different studies and the changes are usually described as clinically minor. However, the power of past studies accurately to detect cognitive changes has been limited by failure to account for practice or test sophistication effects. Patients perform the same test before and after surgery, yet only one study (Iacono et al. 1994) has tested control subjects to account for the postoperative improvement in performance that might be anticipated due to practice. Even this study used healthy subjects as controls rather than patients with PD, which does not account for the fact that patients with PD may respond differently to practice to healthy subjects. It is clearly premature to describe postoperative improvements in test scores as 'cognitive improvements' caused by the surgery when there is no way to assess the positive effects of practice on performance. Failure to control for practice effects might also lead to a subtle postoperative cognitive deficit being 'missed'. A postoperative cognitive deficit would worsen test scores, but this reduction might be canceled out by improvements in scores due to practice. As a consequence, no overall change in test score would be seen after surgery, despite a postoperative decline in cognitive function.

The present study differs from previous research in two ways. Only those domains of cognitive function that appeared likely to be affected by surgery were tested.

This contrasts with previous studies' attempts to provide a comprehensive assessment of all domains of cognitive function. A control group of patients with Parkinson's disease who were not undergoing surgery were also tested, in order to account for practice effects. Tests were chosen for this study principally on the basis of their sensitivity to frontal lobe function, particularly to the function of the dorsolateral prefrontal cortex (DLPFC). Several lines of evidence suggest that pallidotomy might alter cognitive abilities associated with DLPFC function. The lateral dorsomedial segment of the globus pallidus (ldm-GPi) projects to the DLPFC via the thalamus, and receives projections from the DLPFC via the head of the caudate nucleus. This anatomical circuit is known as the 'dorsolateral prefrontal' or 'complex' loop (Alexander et al., 1986), and some investigators have suggested that disruption of this loop is the primary source of the cognitive deficits of Parkinson's disease (Taylor and Saint-Cyr, 1995). There is depletion of dopamine in the caudate nucleus of patients with Parkinson's disease (Kish et al., 1988) which would disrupt the dorsolateral prefrontal loop, and alterations in DLPFC function might explain why the cognitive deficits of Parkinson's disease resemble those seen after damage to the frontal lobes (Taylor et al. 1986). In theory, pallidotomy should not affect the function of the dorsolateral prefrontal loop, as the ventrolateral globus pallidus (vl-GPi) is targeted, which is within a separate circuit known as the 'motor loop'. However, variations in lesion size or placement might lead to the surgery having an effect on the ldm-GPi or fibres that arise from it. There are two possible consequences of such damage. Outflow from the ldm-GPi could simply be interrupted, leading to 'frontal'-style deficits of executive function such as those sometimes seen in nonparkinsonian patients with selective globus pallidus lesions (Lopez-Villegas et al. 1996). Alternatively, the damage could have a positive effect on the dorsolateral prefrontal loop. Pallidotomy is thought to alleviate motor symptoms by decreasing pallidal inhibition of ventrolateral thalamic nuclei within the 'motor' loop, allowing restoration of activity between these thalamic nuclei and supplementary motor cortex. It is possible that disruption of the ldm-GPi projections to the thalamus might have a similar disinhibitory effect, allowing increased activity in the ventral anterior and mediodorsal thalamic nuclei, and their projection sites in the DLPFC. Data from functional neuroimaging studies appear to support this latter possibility. Ceballos-Baumann et al. (1994) reported restored bloodflow to the DLPFC in a single case after pallidotomy, a finding that was replicated by Eidelberg et al. (1996) and Samuel et al. (1997).

The anatomical and neuroimaging data described above strongly support the inclusion of tests that are sensitive to frontal function in the assessment of pallidotomy. The first part of the test battery used in this study consists of three tests that have been shown to be sensitive to both frontal lobe function and to the progression of cognitive deficits in Parkinson's disease. As pallidotomy has been described as 'reversing' or 'reducing' parkinsonian motor symptoms (Laitinen et al. 1992, Dogali et al. 1995, Lozano et al. 1995), tests have been included that are sensitive to the progression of cognitive deficits in Parkinson's disease in order to detect any similar 'reversal' of cognitive symptoms. Two tests were taken from the CANTAB battery (Fray et al. 1996). The Stockings of Cambridge test is a variant of the widely-used Tower of London task , and has been shown to be sensitive to frontal lobe function by both lesion studies (Owen et al., 1990) and neuroimaging studies (Baker et al., 1996). Certain measures derived from this test can also discriminate between the cognitive deficits suffered by patients with moderate and severe Parkinson's disease (Owen et al., 1992). Performance of difficult problems on this test has recently been shown to be associated with increased cerebral bloodflow in the right GPi (Owen et al., 1998) in healthy subjects. As the patients in this study underwent lesions in the right GPi, it was felt to be important to examine the consequences of these lesions for performance of the Stockings of Cambridge task. The spatial working memory test from the CANTAB battery was also used, which has been shown to be sensitive to damage to the frontal lobes in humans (Owen et al., 1990) and also to disease progression in Parkinson's disease (Owen et al., 1992). The third test has been developed in our laboratory, and is a test of attentional set-shifting that includes reversal, intra- and extradimensional shifts (Slamecka et al. 1968). An analogous test from the CANTAB battery has been shown to be sensitive to frontal lobe damage in humans and monkeys (Owen et al. 1991, Dias et al. 1996).

Two tests were selected for the second part of the battery on the basis on past research into pallidotomy. Relatively consistent deficits are found in verbal fluency after pallidotomy, and letter, category and alternating fluency tests were included in an attempt to replicate this. In contrast, variable results have been found using verbal learning tasks such as the Auditory-Verbal Learning Test (AVLT) and the California Verbal Learning Test (CVLT). Lang et al. (1997) found a transient deficit on the CVLT after left pallidotomy, whereas Bowers et al. (1997) found improvements in the delayed cued recognition and recall component of the test. The AVLT was included (Rey 1964,

as described in Lezak 1995) in the second part of the test battery due to its apparent sensitivity to both positive and negative changes after pallidotomy.

6.3 Method

6.3.1 Subjects

6.3.1.1 Surgical patients

A consecutive series of three patients were tested, all of whom had been referred to undergo right pallidotomy at the Department of Neurosurgery at Dundee Royal Infirmary. One patient did not give consent to be tested postoperatively and was therefore not included. The two remaining patients, GM and JG, both suffered from advanced Parkinson's disease complicated by fluctuating medication response (both Hoehn-Yahr stages IV to V in 'off' phase). They both experienced rigid akinesia in their off phases and severe drug-induced dyskinesias in their on phases. Both GM and JG had been diagnosed with idiopathic Parkinson's disease by a consultant neurologist and they had both initially shown a strong clinical response to levodopa.

GM: GM was a 64-year old retired carpenter, who had been diagnosed with Parkinson's disease 15 years previously. He experienced severe bilateral drug-induced dyskinesias and was referred for right pallidotomy because he was left-handed. GM was treated only with Sinemet, as he had been unable to tolerate any other antiparkinsonian agents.

JG: JG was 49 years old and had previously worked as an accountant. She had been diagnosed with Parkinson's disease 14 years previously and had severe bilateral dyskinesias. Although right-handed, her negative parkinsonian symptoms and her dyskinesias were slightly worse on her left side and as a consequence she was referred for right pallidotomy. JG was being treated with Sinemet, Madopar CR, Pergolide and Tolcapone. The Geriatric Depression Scale (GDS) indicated that she was mildly depressed (22/30, cutoff for mild depression 15/30).

6.3.1.2 Control patients

Nine patients with mild to moderate Parkinson's disease were tested as controls, all of whom were being treated as outpatients at Dundee Royal Infirmary. The ideal control group for patients undergoing pallidotomy would be patients on a 'waiting list' for this surgery, as such patients would presumably suffer from Parkinson's disease of a similar severity to the surgical patients. However, access to such a 'waiting list' group was not available. All of the control subjects had been diagnosed with idiopathic Parkinson's disease by a consultant neurologist and had no history of any other neurological disorder. All but one of these patients was taking levodopa and additionally two were taking benzhexol and two selegiline.

All surgical and control patients gave their consent to take part in the study. Subjects completed screening tests for dementia and depression (MMSE and GDS) and also completed the NART to obtain an estimate of premorbid verbal IQ. The result of these tests are in Table 6.2.

	GM	**JG**	**Controls**
Age	64	49	69.1
MMSE	24	30	27.4
GDS	-	22	13.4
NART VIQ	97	113	109.9
Hoehn-Yahr 'on'	3	3	1.7

Table 6.2: Subject characteristics

6.3.2 Methods and Procedure

6.3.2.1 Location and timing of assessment:

It was difficult to obtain the surgical patients' best possible neuropsychological performance due to their severe and fluctuating levels of motor disability. An attempt was made to test these patients in their best possible 'on' state, that is, when they were most mobile but least troubled by dyskinesias. The most appropriate time of day for testing was established by discussion with the patient and by monitoring their movement status over the course of one day. Testing was carried out in two sessions before and two sessions after the surgery. Preoperative ('preop') testing was performed after the implantation of the Bennett's sphere (see 'Surgical methods' section below) and the postoperative ('postop') sessions were carried out between the second and the sixth day after surgery. All testing sessions were carried out at the same time of day and at the same time after the patient's last levodopa dose. Patients were tested in the quietest available room in the Neurosurgery department. Despite efforts to test patients when their dyskinesias were minimal, both patients had moderate to severe dyskinesias at the time of preop testing (score of 3 on the Modified Dyskinesia Scale (v. 2.0); Goetz et al., 1994).

Controls were tested in a quiet office either in Dundee Royal Infirmary or at the School of Psychology at St. Andrews. Two testing sessions were carried out, separated by an interval of around one week. For convenience, these two sessions are referred to as the 'preop' and 'postop' testing sessions although the control patients did not undergo an operation.

6.3.2.2 Surgical procedure

Surgery was carried out by Mr. T.R.K. Varma (neurosurgeon) and Dr. A. Forster (neurophysiologist) at Dundee Royal Infirmary. A Cosman-Roberts-Wells stereotaxic frame was used (Radionics Inc., Burlington, MA, USA) which was coregistered with a preoperative MRI image. The target of the surgery was the ventroposterolateral pallidum. Lesioning and intraoperative stimulation was carried out with an electrode barrel containing five electrodes (School of Psychology, University of St. Andrews). Intraoperative stimulation was carried out at low (5 Hz) and high (200 Hz) frequencies to elicit motor and visual responses (respectively) that might indicate proximity to the optic tract or internal capsule. Several lesions were created within the pallidum; the exact location and number was determined by response to intraoperative stimulation.

6.3.2.3 Assessment of motor function

The CAPIT committee (Langston et al., 1992) have provided guidelines for the comprehensive assessment of the clinical status of patients with Parkinson's disease before and after a surgical operation. It would have been desirable to carry out the CAPIT protocol for the surgical patients in this study, but this was impossible due to limited access to resources and patients. The UPDRS was completed for both patients in their 'best possible on' phase, but it became apparent that scores on this scale were meaningless for patients with severe choreic dyskinesias. For example, it is impossible to assess resting tremor or rigidity in a patient who is never at rest. UPDRS scores are therefore not reported.

Patients were videotaped pre- and postoperatively whilst in their 'best possible on' state. Patients were recorded whilst completing a protocol devised by Pollak (Dr. P. Pollak, Grenoble, 1992), which included components of the UPDRS and of Fahn's rating scale for tremor. However, the principal motor assessment carried out on these patients was a rating of dyskinesia severity. Patients' dyskinesias were rated pre- and postoperatively for limbs ipsi- and contralateral to the site of the surgery by a non-blind observer. The Modified Dyskinesia Scale version 2.0 (MDS 2.0, Goetz et al., 1994) was used, which rates a patient's worst observable dyskinesia on a five-point scale from 0 ('absent') to 4 ('violent dyskinesias, incompatible with any normal motor task').

6.3.3 Neuropsychological tests

Subjects completed five tests at the preop and postop stages. Three of these tests were computer based and required subjects to respond by pressing a touchscreen. The two

remaining tests required only verbal responses. These tests were not administered in a fixed order for the surgical patients due to their motor disabilities; at some points, surgical patients were unable to respond accurately to the touchscreen but could complete tests requiring verbal responses. The order of testing for surgical patients was therefore determined by their motor status, although priority was given to the computer-based tests where possible. The tests were given to the controls in a fixed order.

The three computer-based tests were the 'Stockings of Cambridge', an Attentional Set-Shifting task and a test of Spatial Working Memory. All of these tests were run on a portable computer (Datalux Databrick) using a touch-sensitive screen (MicroTouch TruePoint) to register responses. The Stockings of Cambridge test and the Spatial Working Memory test were taken from the Working Memory and Planning battery of the Cambridge Neuropsychological Test Automated Battery (CANTAB, Owen et al. 1990). The Attentional Set-shifting task was developed specifically for this study. The Stockings of Cambridge and the Spatial Working Memory tests were repeated in an identical format at the postop stage. Different stimuli and dimensions were used in the postop stage of the Attentional Set-shifting task (see below).

6.3.3.1 Stockings of Cambridge

This test has been described at length by Owen et al. (1992) and is only summarised here. Subjects were presented with two displays, one at the top and one at the bottom of the screen. These displays each consisted of three containers that resembled socks or stockings that had three coloured balls placed at various different positions within them (Figure 6.1). The positioning of the balls in the stockings differed between the top and bottom displays; the goal of the initial 'problem-solving' stage was to move the balls within the bottom display until their position in the bottom stockings was identical to the position of the balls in the top stockings. Subjects could pick a ball up by pressing it once with their finger and could put it down in a particular location by pressing their finger in that location. Certain moves were 'illegal' and were not registered by the computer. Generally, the illegal moves resembled those that could not be made with a genuine arrangement of 'balls' and 'stockings'. Subjects were not permitted to place two balls in the same position and could not move a ball that was underneath another in the same stocking until the ball above it had been moved. Balls had to be placed in the lowest possible position in a stocking and could not 'float' above a free space. The difficulty of this task was manipulated by changing the number of moves that was required to match the bottom display to the top display. The problem became more difficult as a greater number of moves were required. Subjects were required to

complete problems at four levels of difficulty, requiring either two, three, four or five moves to solve. The problems that required two or three moves to solve were regarded by most subjects as easy, as they required only direct and evident moves to match the bottom and top arrangements. The four- and five-move problems were more difficult as they required planning and the formation of subgoals for completion. Subjects were encouraged to 'think through' the entire solution to the problem before starting to move the balls. The time taken between the appearance of the pattern and the initiation of the first response was measured as it is thought to represent the time required mentally to 'solve' the problem and initiate a response.

The second phase of this test did not require subjects to solve a problem and subjects were therefore not required to engage in planning or setting subgoals. The format and task requirements of this stage were the same as for the first 'problem-solving' stage. Only one move was ever required to match the bottom display to the top display. However, as soon as this move was made a single ball was moved on the top display such that another move was required. This stage of the task required only a simple match-to-sample response and subjects were instructed to make each move as quickly as possible. This stage of the test was a 'yoked control' condition, as each set of moves that was required was identical to a set of moves that the subject had made at the problem-solving stage. Movement initiation and completion times were recorded at this stage of the test. These measures were thought to represent the time required to complete a movement without the 'planning' requirements of the problem-solving stage of the test. The time required to initiate the first movement at each new problem in this yoked control stage was subtracted from the initiation time of the corresponding problem in the problem-solving phase. By subtracting the absolute time required to make the movement from the times recorded at the problem-solving phase, it was hoped to gain a measure of 'initial thinking' time that would be independent of any generalised slowing of movement such as the bradykinesia seen in PD. In total, subjects completed two 'two-move' and two 'three-move' problems and four 'four-move' and four 'five-move' problems.

A number of performance measures can be taken from the Stockings of Cambridge, of which three were calculated for each level of problem difficulty. The mean moves taken above the minimum number of moves possible was calculated as a measure of accuracy. Two 'thinking time' measures were also taken. The first was the 'intial thinking time' referred to above. The second related to the time required to make further moves after the first and was termed 'subsequent thinking time'. Subsequent

thinking time was calculated by subtracting the average time taken to select a subsequent ball at the 'yoked control' stage from the equivalent time at the 'problem-solving' stage.

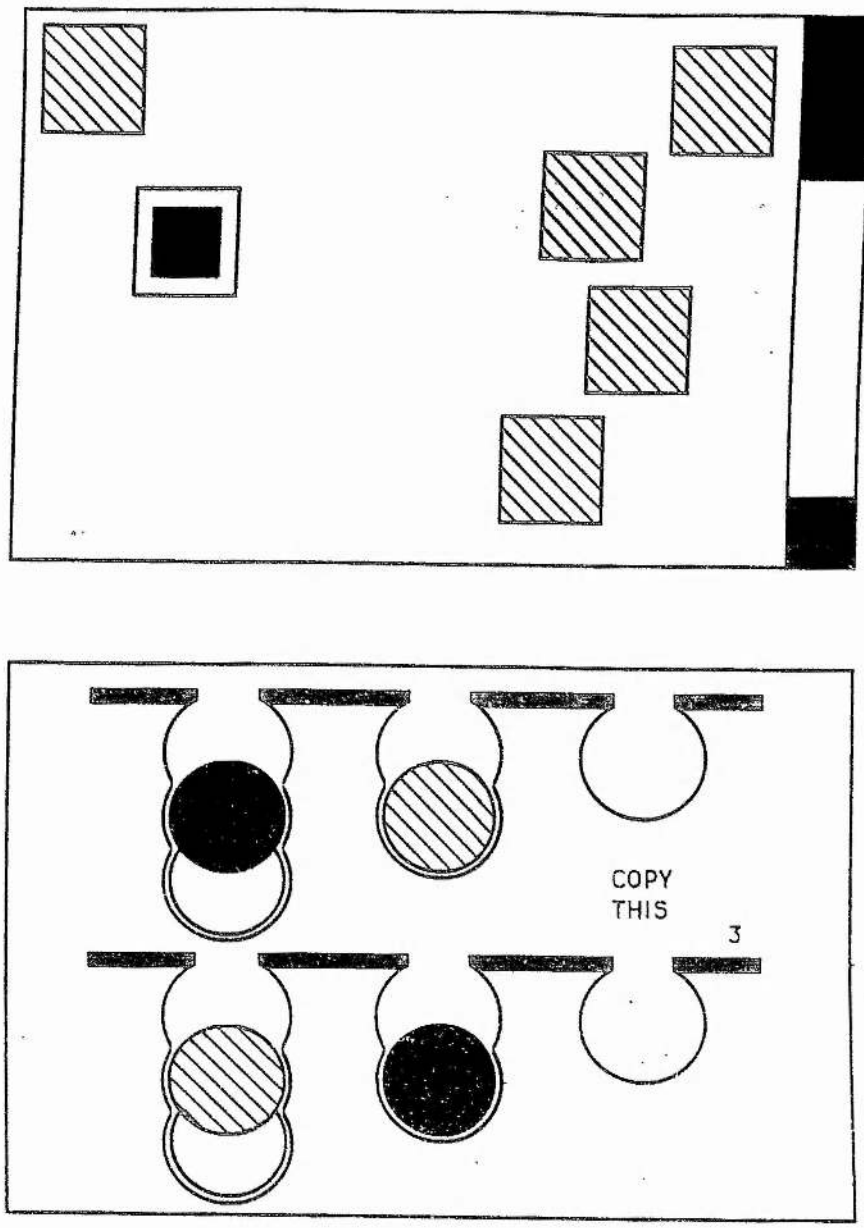


Figure 6.1: Screen shots from the Spatial Working Memory test (top panel) and the Stockings of Cambridge test (bottom panel) from the CANTAB battery. The top panel shows the 'boxes' that subjects must work through for a six-move Spatial Working Memory task (diagonal stripes), one of which contains a counter (filled square). The bottom panel shows a 3-move problem from the Stockings of Cambridge, in which the bottom set of balls must be rearranged to match the top arrangement.

6.3.3.2 Attentional Set-shifting

This task was very similar in appearance and response requirements to Study 1. It assessed subjects' ability to complete ID, ED and reversal (RV) shifts. The task was similar to that used in Studies 1 to 4 in that it was a two-choice visual discrimination task involving two perceptual dimensions. The only differences in stimulus display, response requirements and feedback to Study 1 was that stimulus display time was limited to two seconds.

Subjects completed this test in two blocks, each block including an acquisition phase, followed by an ID, ED and RV shift presented in a different order in each block. The ID and ED shifts were of a 'standard' type, involving 'total change' designs (see Section Q). The RV shift involved no change of exemplars; subjects were required to cease responding to one exemplar of the relevant dimension and to acquire a response to the other exemplar of the relevant dimension. Each shift occurred after subjects achieved a criterion of six consecutive correct responses at a previous shift. If subjects had not achieved this criterion after 30 discrimination trials at the previous shift, they were 'informed' of the correct rule by the computer in the same fashion as for Study 1. The shift the occurred when subjects were stably responding to the rule (subjects to the same conditions as Study 1).

The stimuli in the preop stage of this test were created from the dimensions shape and colour; the postop stage used shape and pattern. The dependent measure was errors to criterion; as subjects completed each type of shift twice at preop and postop stages, the mean error score for each type of shift was calculated and used in statistical testing.

6.3.3.3 Spatial Working Memory

This task required subjects to search through a series of boxes in order to find a 'counter' that was concealed in one of them. Subjects were presented with an array of boxes on the touch screen such as that illustrated in Figure 6.1 and were told to find a counter. Subjects searched through the boxes in any order that they chose, 'opening' each box by pressing it. Once a subject found a blue counter, the subject was required to place it in a 'bar' on the right-hand side of the screen and to search for more counters until the bar was full, when the trial terminated. Once a subject found a counter, they were instructed to search again, but it was emphasised that a counter would never be found in the same box twice. Thus, this task was a self-ordered search task in which

subjects were required to 'keep track' of where they had previously found counters in order to avoid searching these boxes again.

There were four levels of difficulty of this task, corresponding to the number of boxes displayed on the screen (three, four, six or eight). As the number of boxes became larger, it became harder to keep track of which boxes had previously been opened. At each level of difficulty, subjects completed four trials from which three dependent measures were calculated. The first, termed 'between search errors', represented the number of times in a trial that a subject opened a box in which a counter had already been found. The second measure was an index of strategy use. Subjects sometimes select a beneficial strategy when performing this task (Owen et al. 1990) that consists of searching through the boxes in a fixed sequence within a particular trial. Subjects repeat this sequence after finding each counter, reporting that it is easier to 'keep track' of their progress using this method. When subjects used this strategy they started each new search sequence by opening the same box; thus, the strategy score was a count of the number of search sequences started with the same box for the six and eight-move problems. A higher strategy score represented more consistent strategy use. The third dependent measure used was termed 'within-search' errors. This was an index of the number of times that a subject opened the same box twice within a particular search sequence; that is, during the set of moves made to find a single counter. Within-search errors are relatively rare as most subjects are able to keep track of the boxes that they have recently pressed. As a consequence, this dependent measure is regarded as less sensitive than 'between-search' errors.

The final two tasks were widely-used neuropsychological tests that did not require the use of a computer. In both cases, different test materials were used for preop and postop phases.

6.3.3.4 Verbal Fluency

The verbal fluency tests used in this study were administered in the standard format of asking subjects to generate as many words as possible that conform to a particular rule in one minute. Three types of fluency were assessed. Subjects completed three phonemic ('letter') fluency trials, one semantic ('category') fluency trial and one alternating fluency trial. The alternating fluency task was taken from a study by Downes et al. (1993); subjects were required to alternate between producing a word from a phonemic rule and a word from a semantic rule. When these tests are administered in their traditional format, subjects are instructed not to generate proper nouns. This instruction was omitted in this study, following Downes et al. (1993), as the use of

clusters of proper names can be considered as evidence of efficient strategy use and can be scored as such. The main dependent variable used was the total number of words (excluding repetitions and 'errors' - that is, words that breach the rule) generated in one minute. The mean score for the three phonemic trials was used.

The phonemic rules were taken from Benton and Hamsher's (1989) Multilingual Aphasia Battery, as described by Lezak (1995). The following rules were used, in the given order. Preop: phonemic, 'P', 'R', 'W'; semantic, 'towns and cities'; alternating, 'B' then 'animals'. For the postop stage: phonemic, 'C', 'F', 'L'; semantic 'countries'; alternating, 'T' then 'fruit and vegetables'.

6.3.3.5 Auditory-Verbal Learning Test (AVLT)

This classic test assesses subjects' ability to recall items from a list of 15 words, both when the list is initially presented and after repeated presentation. The AVLT was administered in the manner described by Lezak (1995), using parallel forms at the postop stage. The list (list A) is read out loud to subjects at a rate of one words per second and subjects are instructed to recall as many words from it as possible in any order. After the first presentation, list A is presented a further four times with identical instructions to subjects. Subjects usually recall more words from the list after each presentation. After five presentations of list A, subjects are presented with a list of 15 new words and asked to recall these (list B). There are two further recall stages; immediately after recalling list B, subjects are asked to recall items from list A and delayed recall for list A is tested again after thirty minutes. The dependent measures were immediate recall (items recalled at the first presentation of list A), learning (change in score between the first and fifth presentations of list A) and delayed recall (recall of list A after a 30-minute delay).

6.3.4 Data analysis

The purpose of this study was to detect any changes in cognitive performance that might occur as a consequence of pallidotomy. However, test scores may change at a 'retest' session due to factors such as practice, so control patients were tested in order to account for this. In order to identify changes in performance in the surgical patients that were greater than those seen for the control patients, 'difference scores' were calculated for all subjects by subtracting postop test scores from preop test scores. The difference scores for the surgical patients were then compared to those of the control patient group using a one-sample t-test. For the sake of simplicity, the significant results of these analyses will be referred to as 'postoperative improvements / decrements in

performance' in the surgical patients; in fact, of course, they refer to 'postoperative improvements / decrements compared to the changes seen in control patients'. Alpha was set at 0.01 for all comparisons between surgical patients and controls to account for multiple comparisons. All surgical patients' decrements or increments in performance of more than one standard deviation have also been labelled on the graphs. Positive values (e.g. '+4.0 SD') indicate improvements in performance, negative values indicate decrements. Although difference scores were used for statistical testing, absolute test scores have been plotted on the Figures. Depiction of absolute values allows identification of subjects' proximity to 'floors' and 'ceilings' in performance and allows comparison of surgical patients' absolute levels of performance to those of controls.

6.3.4.1 Missing data

A number of factors placed limitations on the collection of data from surgical patients in this study. The amount of time available for testing was limited by patients' short stay in hospital, which was further complicated by the fact that they were untestable for much of the day due to motor disability. Testing time was also limited by patients' medical care and the normal activity of the neurosurgery ward - that is, the constraints of ward rounds, physiotherapy and the availability of quiet testing space. As a consequence of these factors, two pieces of data are missing for patient GM - the GDS and the spatial working memory test. Also, the 'delayed recall' section of the AVLT was not recorded for JG.

6.4 Results

6.4.1 Clinical status

Surgery led to an almost complete abolition of dyskinesias contralateral to the lesion in both patients and also to a minor improvement in ipsilateral dyskinesias. Changes in dyskinesia severity as measured by the MDS 2.0 (Goetz et al. 1994) can be seen in Table 6.3. However, there appeared to be no improvement in negative parkinsonian symptoms. This was reflected by the fact that the surgical patients' level of medication could not be reduced postoperatively; indeed, Sinemet and Pergolide doses were increased for GM and JG respectively, as they could tolerate higher doses of these drugs without exacerbation of their dyskinesias. GM experienced postoperative dysarthria that had not resolved when he was discharged one week after surgery. This was the only complication seen. It should be noted that the surgical patients in this study were affected by dyskinesias that were at least as severe as those seen in other studies. For example, Lozano et al.'s (1995) patients were of a similar disease duration to those in this study (13.8 years compared to 14.5 years) but their mean MDS 2.0 scores were around 2. Other studies have used patients with MDS 2.0 scores of around 3 (Kishore et al., 1997; Samuel et al., 1998); JG and GM's preoperative left-side dyskinesias were clearly disabling, and were rated 4. The relative severity of JG and GM's dyskinesias compared to those of patients previously studied may account for why previous studies have not mentioned the practical problems of carrying out neuropsychological tests with subjects with severe dyskinesias.

		Preop	**Postop**
GM	*Contra*	4	1
	Ipsi	4	3
JG	*Contra*	4	1
	Ipsi	3	3

Table 6.3: Severity of the worst observable dyskinesias, before and after surgery, measured by the MDS 2.0; 0 = absent, 4 = violent, incompatible with any motor task.

6.4.2 Stockings of Cambridge

Considering each dependent variable in isolation for the Stockings of Cambridge task can be uninformative, as they are interdependent. For example, if a subject showed

lengthened subsequent thinking times after surgery, this might appear to be evidence of a postoperative deficit in performance. However, if the same subject showed reduced initial thinking times postoperatively it would be clear that the subject had merely initiated solutions more rapidly and had been forced to 'make up' the consequent lost thinking time after pressing the first ball. Therefore, the results are summarised for all dependent variables at the end of this section, though they are first described individually.

The severity of GM's dyskinesias interfered with his performance of this task to the extent that doubt must be cast on the validity of his thinking time data (both initial and subsequent). GM was severely dyskinetic when he completed the test, particularly at the preop stage. He was frequently forced to pause whilst completing a problem and substantial time sometimes elapsed as he attempted and failed to select or place a ball. Although the 'yoked control' condition in this test is intended to control for differences in movement status, it appears to be better suited to controlling for a generalised speeding or slowing of performance rather than intermittent breaks in performance. GM's timing data is reported, as he did complete the test at preop and postop stages, but should be treated with caution. However, there is no reason to doubt the validity of his 'mean moves above minimum' data.

6.4.2.1 Mean moves above minimum

Data for all subjects for this variable can be seen in Figure 6.2. It appears that the Stockings of Cambridge test is not greatly affected by practice effects, as controls show little improvement when testing is repeated (postop stage). This is particularly noteworthy as this test was repeated in identical format after an interval of only one week. The controls appear to be performing at 'ceiling' for the 2- and 3-move problems and this is also true for surgical patients at the 2-move problems. Thus, only 3-, 4- and 5-move problems will be considered in detail.

Inspection of Figure 6.2 suggests that results for the two surgical patients change in opposite directions after surgery - GM's performance appears to worsen and JG's to improve. GM showed a significant decrement in performance at 3- and 4-move problems ($t(8) = 20.45$ and 4.34 respectively, $p < 0.01$), though not for 5-move problems. However, the absence of a statistically significant deficit for 5-move problems may be an artefact; GM was performing at 'floor' for the 5-move problems at the postop stage (seven moves above minimum is the maximum possible). In contrast, JG showed significant improvement for the 3- and 4-move problems ($t(8) = 6.82$ and 8.16 respectively, $p < 0.001$) though not for the 5-move problems.

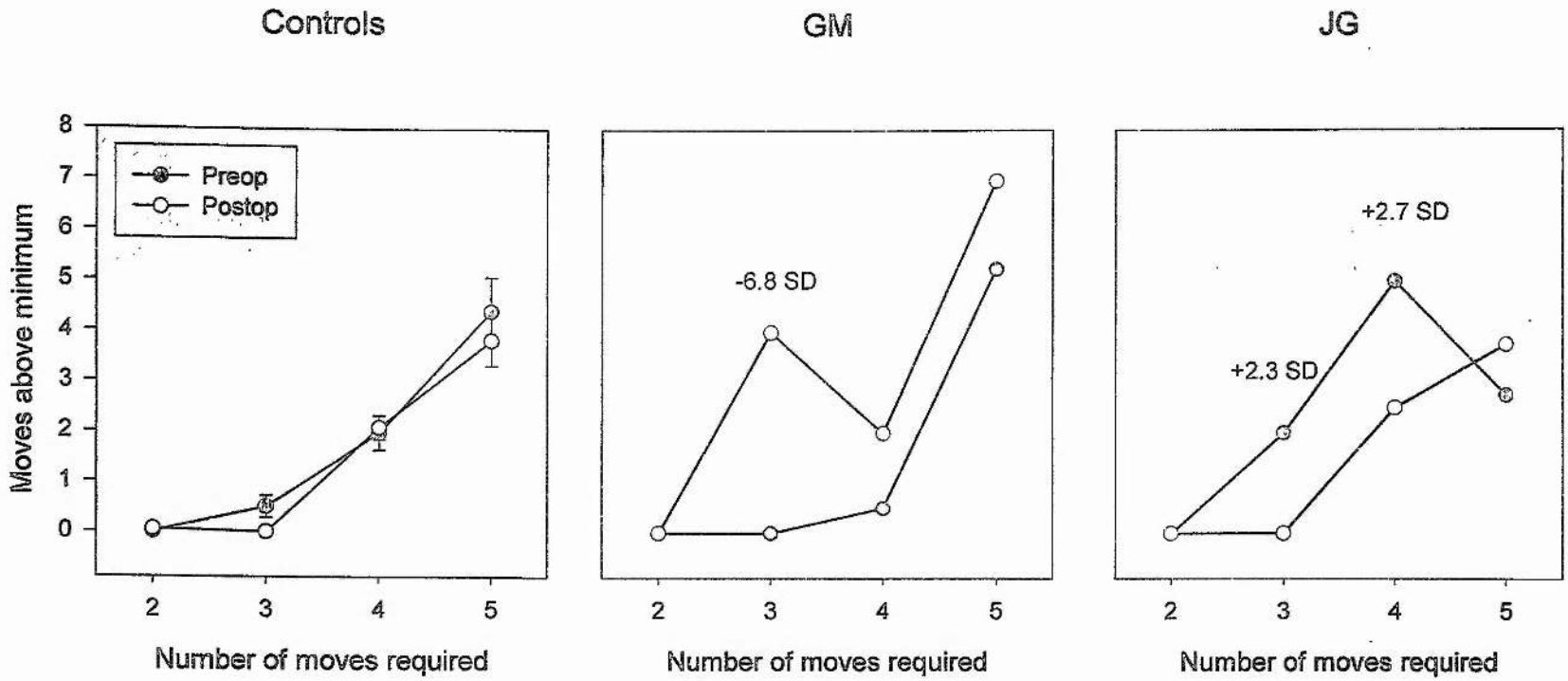


Figure 6.2: Mean moves above the minimum possible for surgical and control patients at the preop and postop stages. Error bars represent standard error of the mean.

6.4.2.2 Initial thinking times

These data are depicted in Figure 6.3. The overall pattern of data for controls is similar to that seen in other studies of performance of the Stocking of Cambridge task by patients with PD (Owen et al., 1992); an increase in thinking times between 2- and 3-move problems, with 3-, 4-, and 5-move problems differing only minimally. An inspection of Figure 6.3 suggests that controls' thinking times reduce with practice, being lower at the postop stage. However, this reduction did not reach significance (main effect of practice $F(1, 6) = 4.77, p = 0.07$). Both controls and surgical patients were at a 'ceiling' of performance for the 2-move problems, as seen for the 'moves above minimum' data, so only the 3-, 4- and 5-move problems will be considered. The pattern of GM's initial thinking times across the three levels of difficulty is not easy to interpret. His times at the postop stage were almost identical to those of controls, but at the preop stage he showed lengthened thinking times for 3- and 5-move problems compared to controls. This led to him showing significant postoperative improvements in performance at the 3- and 5-move problems ($t(8) = 4.98, p < 0.01$ and $t(6) = 5.94, p < 0.001$ respectively), but not for the 4-move problems ($t(8) = 1.6, p > 0.1$). A fuller analysis of his performance is possible when all performance measures are considered together in the 'Summary' section below.

Visual inspection of JG's performance suggests that her initial thinking times became shorter postoperatively for both 4- and 5-move problems. In fact, only the change for the 5-move problems attained statistical significance ($t(6) = 8.17, p < 0.001$); the change for the 4-move problems was a nonsignificant trend ($t(8) = 2.62, p < 0.05, > 0.01$). For the 3-move problems, JG showed a trend towards not making the performance improvements shown by controls at the postop stage ($t(8) = 2.67, p < 0.05, > 0.01$). However, her preop initial thinking times for 3-move problems were at a similar level to that reached by the controls at the postop stage. Thus, there may have been little scope for JG to improve her performance after surgery. In summary, JG showed a greater decrease in initial thinking times than controls at the postop stage for the more difficult problems.

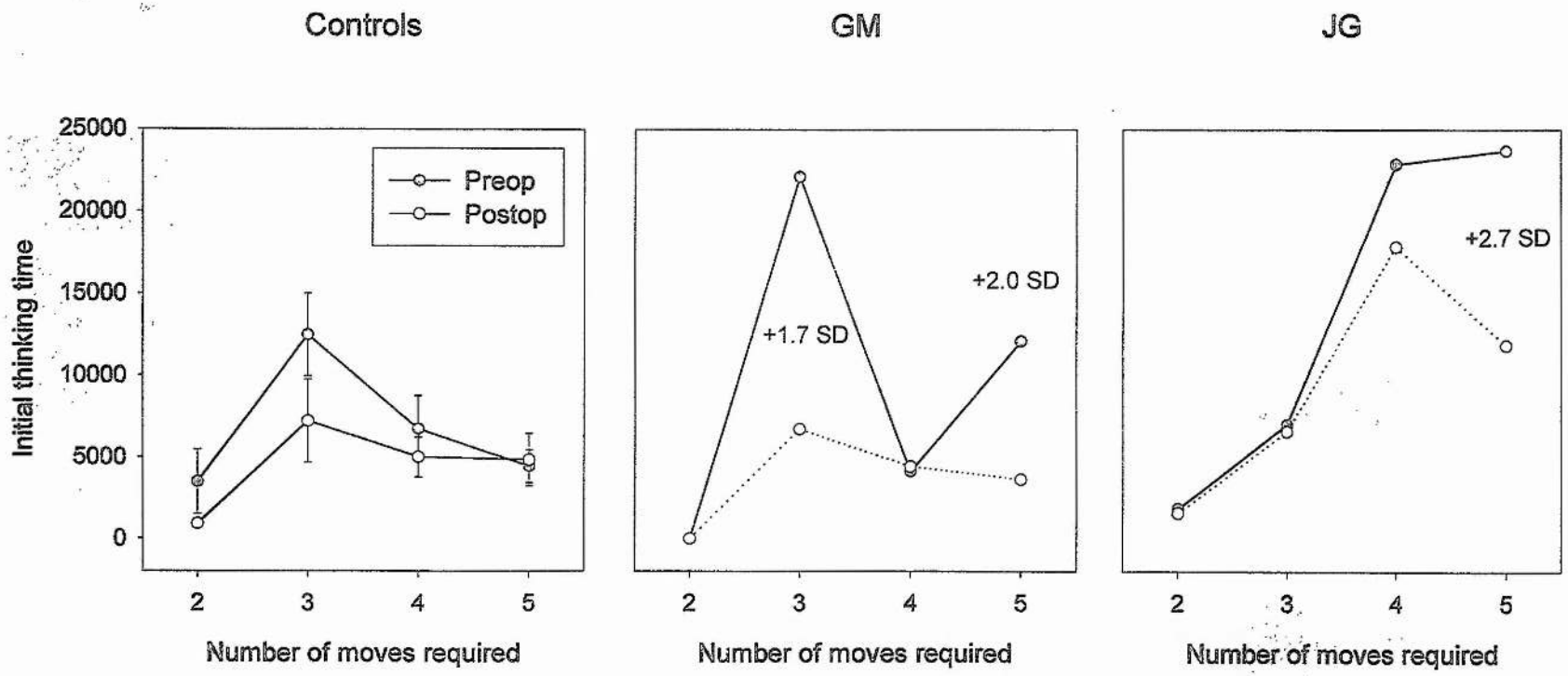


Figure 6.3: Mean initial thinking times for the surgical and control patients at preop and postop stages for all levels of difficulty. Error bars represent the standard error of the mean.

6.4.2.3 Subsequent thinking times

Subjects' subsequent thinking times can be seen in Figure 6.4. As for the 'moves above minimum' measure, controls' performance does not appear to consistently improve with practice. GM's results do not show a simple improvement or decrement of performance a the postop stage. There were three significant changes of subsequent thinking time; an increase for 2-move problems and a decrease for 3- and 4-move problems (all $p < 0.001$). There was also a trend towards lengthened subsequent thinking times for the 5-move problems ($t(6) = 3.54, p < 0.05, > 0.01$). This pattern of results is difficult to interpret and would resemble the pattern of controls' results more closely if GM's performance had been less extreme at the 2-move postop and the 3-move preop stages. These isolated slowings of subsequent thinking times may have been a consequence of GM's motor status.

JG's results are easier to interpret and give an overall impression of shortened subsequent thinking times for the postop stage. JG's change in performance at the 2- and 3-move problems did not differ from that of controls. Her subsequent thinking times were significantly shortened at postop for the 4-move problems ($t(7) = 6.10, p < 0.001$) and there was a similar trend towards shortening times for the 5-move problems ($t(6) = 2.65, p < 0.05, > 0.01$).

6.4.2.4 Summary

The effects of pallidotomy on the performance of the Stockings of Cambridge test can only be judged by assessing all of the outcomes variables together. The significant changes in GM's and JG's performance relative to controls on all three dependent variables are shown on Table 6.4 for 3-, 4- and 5-move problems. A "+" represents a significant improvement in performance after surgery (decreased moves above minimum or thinking times) and a "-" represents a decrease in performance. JG's results clearly indicate an overall improvement in performance of the Stockings of Cambridge test after surgery. At each level of problem difficulty there was a significant improvement in performance for at least one dependent variable. Also, the table does not show the trends towards improved performance seen for initial thinking times at the 3- and 4-move problems and for subsequent thinking times at the 5-move problems (all $p < 0.05, > 0.01$); these are further evidence for JG's improved performance. GM's results are more equivocal. The 'moves above minimum' data suggest decreased problem solving accuracy after surgery. However, the improvements in initial and subsequent thinking times suggest that there is no global decrement of performance. Although GM's 'thinking time' data may not be completely valid due to his motor

disorder, the overall impression remains that the principal change in GM's performance was a change in the tradeoff between speed and accuracy; that is, his performance was faster but less accurate at the postop stage. Thus, there is no evidence of an overall change in GM's performance of the Stockings of Cambridge test after surgery.

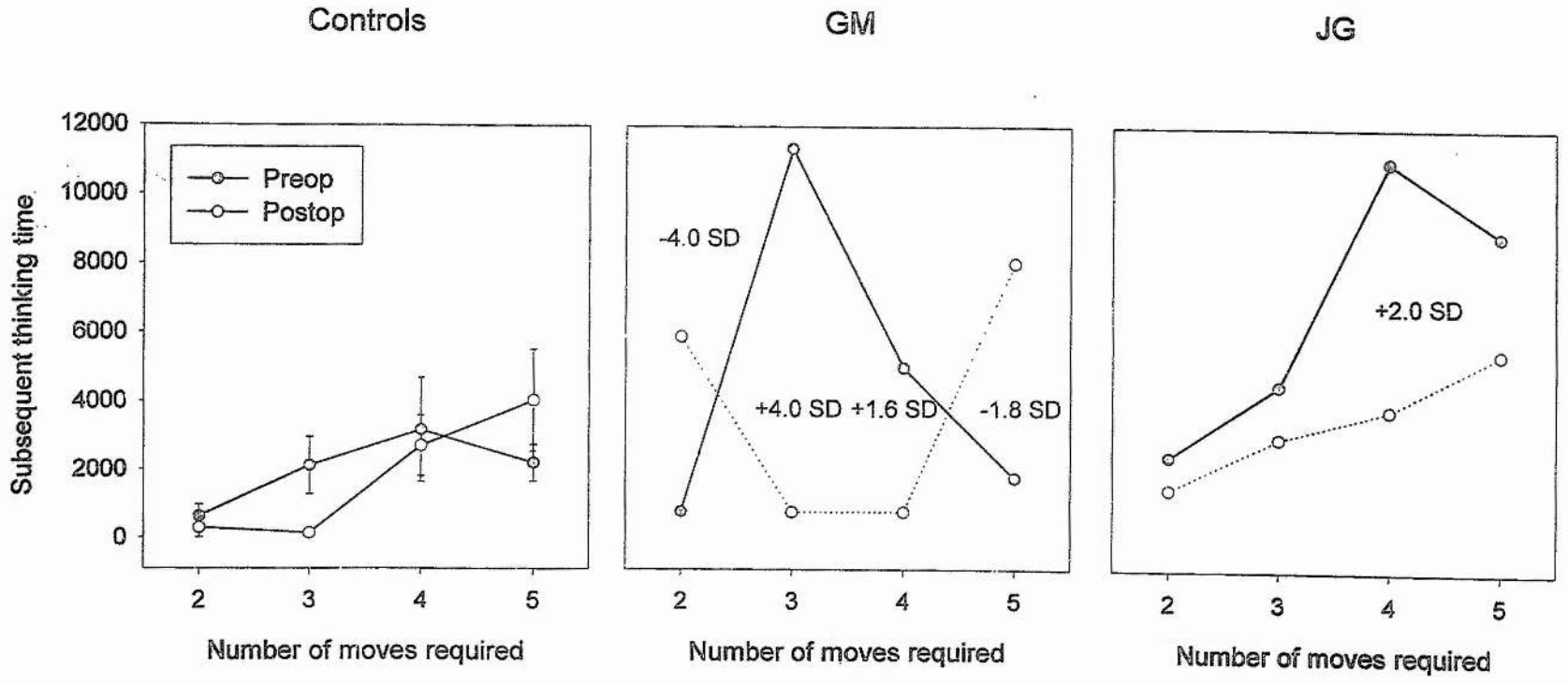


Figure 6.4: Mean subsequent thinking times for surgical and control patients at the preop and postop stages at all levels of diffficulty. Error bars represent the standard error of the mean.

		3	4	5
GM	*Moves above minimum*	-	-	
	Initial thinking time	+		+
	Subsequent thinking time	+	+	
JG	*Moves above minimum*	+	+	
	Initial thinking time			+
	Subsequent thinking time		+	

Table 6.4: Significant ($p < 0.01$) improvements (+) and decrements (-) in performance after surgery compared to controls for the three dependent variables of the Stockings of Cambridge test.

6.4.3 Attentional Set-Shifting

The absolute number of errors committed by control subjects on this test was not high, as can be seen in Figure 6.5. However, the controls were not at the 'ceiling' of performance that would be indicated by scores of one errors to criterion (as seen for JG). Controls showed the theoretically significant ID-ED difference, making more errors at the ED shift than at the ID shift ($Z = 2.19$, $p < 0.05$). Practice had no significant effect on performance of any of the shifts (all $p > 0.1$).

As described in Section 6.3.3.2, mean scores were calculated for each type of shift at the preop and postop stages. Using these scores, it can be shown that GM showed significant improvements in performance for the ID and ED shifts after surgery ($t(8) = 4.17$ and 5.84 respectively, $p < 0.01$). He made numerous errors at the preop stage but followed this with perfect performance at the postop stage. However, Figure 6.5 plots the number of errors generated at each individual shift at the preop and postop stages. It can be seen from this graph that GM's ID and ED shift performance at the preop stage was not universally poor; for both ID and ED shifts, GM made many errors at one shift but performed the other shift at a level comparable to controls. Thus, it appears that GM was capable of producing normal ID and ED shift performance at the preop stage, despite the impression of a preop deficit that was created by the method of 'averaging' shift performance.

JG's performance did not change significantly for any shift after surgery. However, her performance at the RV and ID shifts was at 'ceiling' before and after surgery and her ED shift performance was close to 'ceiling'. As a consequence, the attentional set-shifting test could only have revealed postoperative deficits for JG, as performance improvements were not possible.

In summary, there is no evidence that pallidotomy impairs attentional set-shifting by inducing 'frontal'-style cognitive deficits. The improvement of GM's performance suggests that pallidotomy may alleviate preoperative deficits of attentional set-shifting.

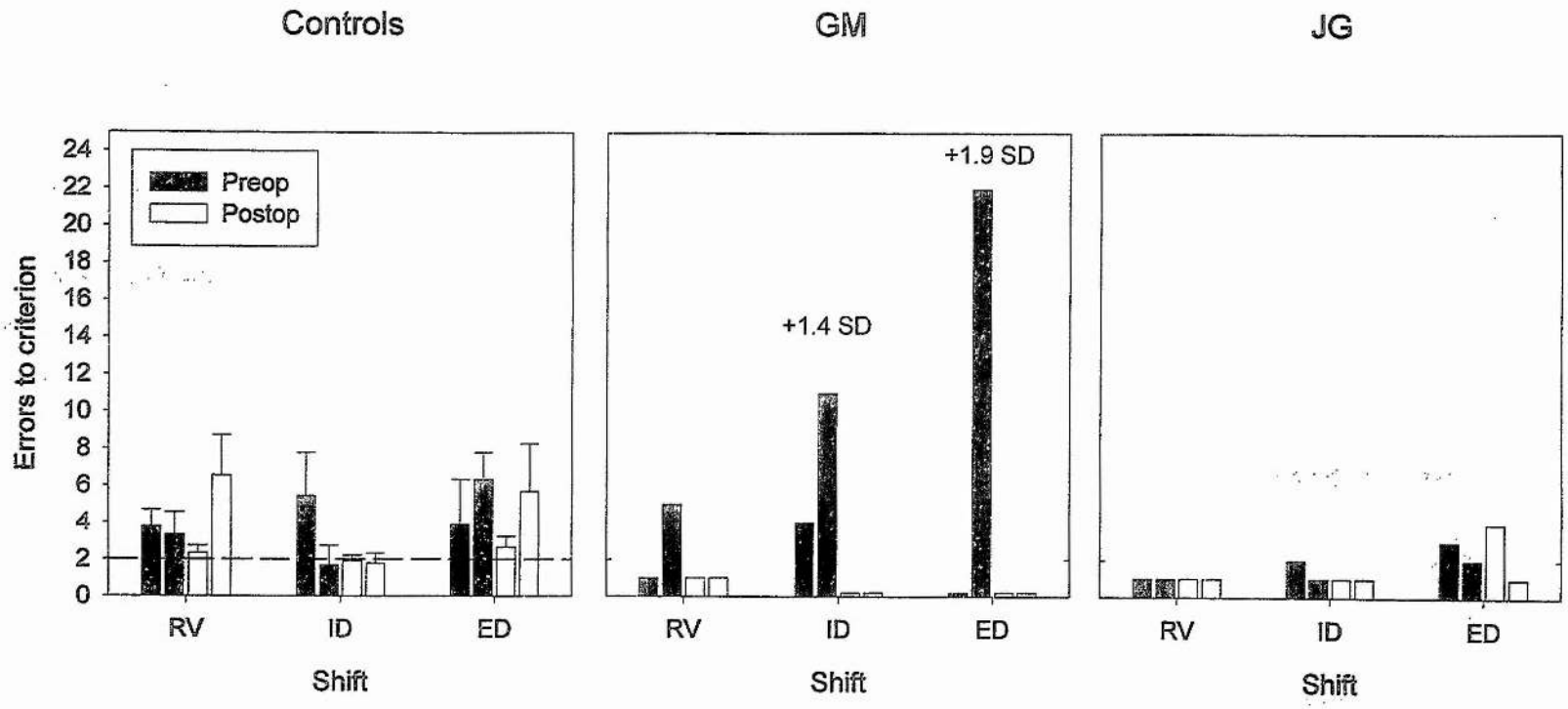


Figure 6.5: Errors to a criterion of six consecutive correct responses for all shifts.
Error bars represent the standard error of the mean.
Scores below the horizontal gridline represent 'perfect' performance.

6.4.4 Spatial Working Memory

The number of between-search errors made by controls can be seen in Figure 6.6 (unbroken lines); as anticipated, the number of errors increased with the number of boxes involved in the problem. Controls did not commit any errors for the 3-box problems and for the remaining levels of difficulty there appears to be no consistent effect of practice. Inspection of JG's data reveals increased levels of between-search errors after surgery, despite good performance at the preop stage. JG showed a significant postoperative drop in performance for the 4- and 6-box problems ($t(7) = 6.43$ and 8.37 respectively, $p < 0.001$) and a trend towards a decrement in performance for the 8-box problems ($t(7) = 3.26, p < 0.05, > 0.01$). A postoperative drop in JG's strategy score almost reached significance ($t(7) = 3.35, p < 0.05, > 0.01$), raising the possibility that surgery imposed a 'frontal-style' cognitive deficit that inhibited strategy use and therefore increased between-search errors. JG also showed a marked increase in within-search errors after surgery, despite normal preop performance (Figure 6.6, broken lines). Statistical comparison of this change with that of controls was not possible due to the low number of errors made by controls.

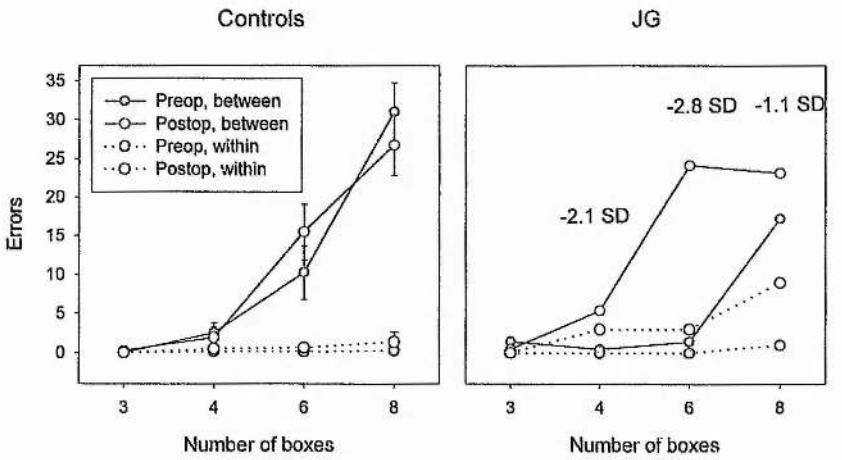


Figure 6.6: Mean between-search errors (unbroken lines) and within-search errors (broken lines) for control patients and JG for all levels of difficulty. Error bars represent the standard error of the mean.

6.4.5 Verbal Fluency

The results of the verbal fluency tests are not shown in a figure as surgical patients showed no overall change on these measures after surgery. Controls performed no better at the task at the postop stage than at the preop stage, indicating that the 'rules' used at the two stages were of equivalent difficulty. GM showed no significant change

semantic fluency ($t(8) = 7.09, p < 0.001$), but no change in phonemic or alternating fluency. However, any deficit of semantic fluency shown by JG cannot have been of great magnitude as it did not affect scores on the alternating fluency test. JG's unchanged alternating fluency scores imply that she generated words to a semantic rule at the same rate before and after surgery when performing the alternating fluency task, as half of the words generated on an alternating fluency test are created according to a semantic rule. Thus, there is no compelling evidence for the presence of a semantic fluency deficit in JG or for any other fluency deficit in either JG or GM.

6.4.6 Auditory-Verbal Learning Test

Recall scores for the five learning trials of list A are plotted in Figure 6.7. As anticipated, the control subjects recalled more words with each repetition of the list. The parallel forms used were clearly well matched for difficulty as control subjects' performance at the preop and postop stages was very similar. Neither of the surgical patients appear to have been experiencing gross memory problems at the preop stage and the overall effects of surgery on performance were minimal. The only significant change seen for the surgical patients was an increase in the 'learning' score for GM after surgery. However, this change appears to be an artefact of the method used to calculate the 'learning' score. GM's learning score at the preop stage was zero, as his recall at trials one and five was equal. However, inspection of Figure 6.7 indicates that GM's performance improved by four words between trials one and four, but this gain was lost at trial five. GM was clearly capable of learning words with repeated list presentation at the preop stage and if a learning score is calculated by subtracting trial one recall from trial four recall there is no longer any postoperative improvement of GM's learning score. Thus, there is no evidence for any change in immediate or delayed memory or learning after pallidotomy.

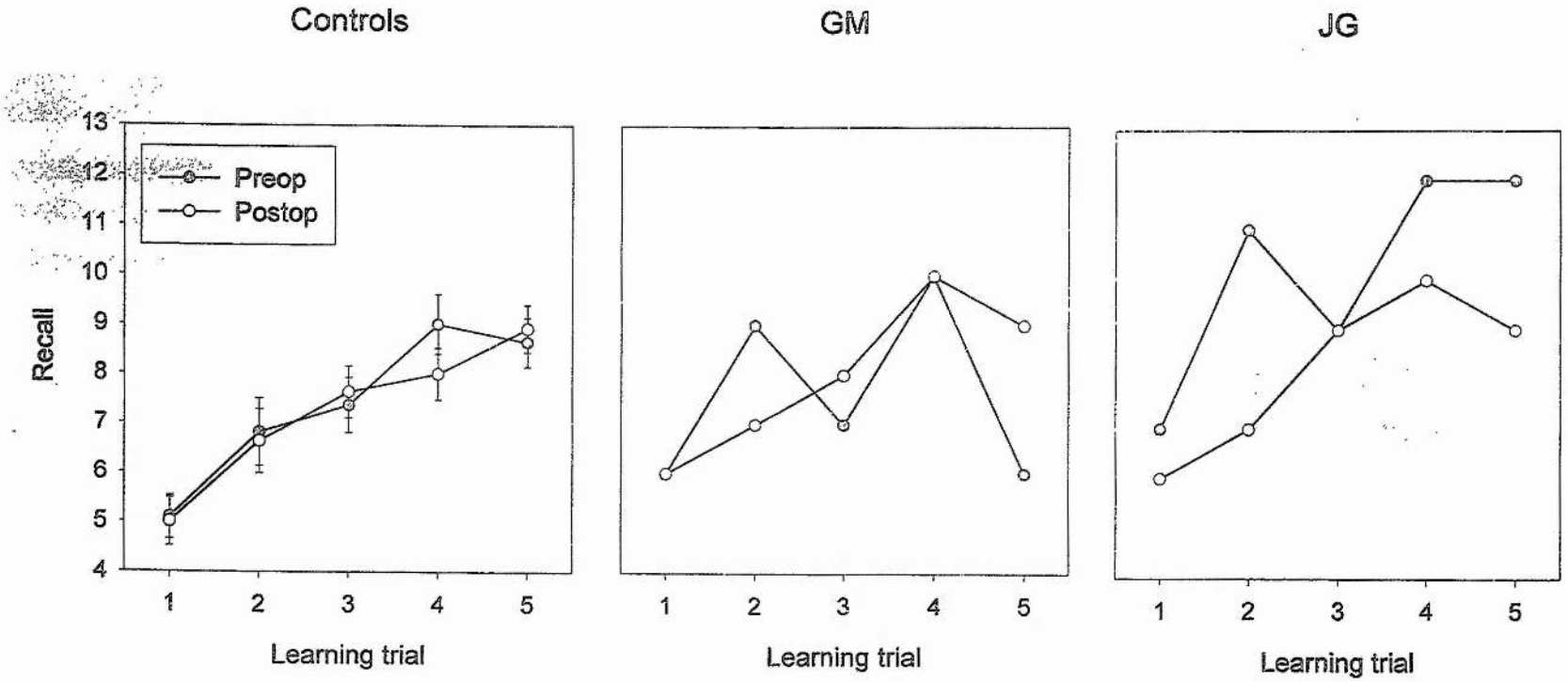


Figure 6.7: Number of words recalled for the first five learning trials of list A for surgical patients and controls at the preop and postop stages. Error bars represent the standard error of the mean.

6.4.7 Summary of results

Assessments of generally 'improved' or 'impaired' performance after surgery for JG and GM can be seen in Table 6.5. Neither GM nor JG showed any reliable changes in verbal fluency or in verbal memory as measured by the AVLT. However, the results of the Stockings of Cambridge and Spatial Working Memory tasks were less consistent. GM's performance on the Stockings of Cambridge test is best described as unchanged, as the postoperative alterations in his performance are consistent with faster but less accurate performance rather than showing a global improvement or decline in performance. In contrast, JG's performance of the Stockings of Cambridge test improved postoperatively. JG showed a clear decrement in performance of the Spatial Working Memory test that was apparent in both 'error' measures of this task. Pallidotomy caused no deficits in attentional set-shifting. However, there were apparent improvements in GM's postoperative performance of the attentional set-shifting task that will be discussed in the following section.

	Stockings of Cambridge	Attentional Set-Shifting	Spatial Working Memory	Verbal Fluency	AVLT
GM	=	↑		=	=
JG	↑	=	↓	=	=

Table 6.5: General changes in test performance after surgery for JG and GM. GM did not perform the spatial working memory test.

6.5 Discussion

Two patients with advanced PD underwent neuropsychological testing before and after unilateral pallidotomy in an attempt to detect any positive or negative consequences of the surgery for cognitive function. The surgical patients completed tests that had previously been shown to be sensitive to the effects of pallidotomy as well as tests that were known to be sensitive to frontal lobe function. Unoperated patients with PD completed the same tests twice at similar intervals to the surgical patients in order to control for any effects of practice on test scores. Surgery had only a moderate effect on the surgical patients' motor symptoms, abolishing dyskinesias contralateral to the lesion. The effects of surgery on cognitive function were also moderate. There was evidence of an improvement in the performance of the Attentional Set-Shifting task in GM and the Stockings of Cambridge task in JG; JG also showed a significant decline in the performance of the Spatial Working Memory task after surgery. Verbal memory and verbal fluency were unchanged in both patients after pallidotomy.

6.5.1 Magnitude of changes seen

The results of this study are in agreement with the majority of studies published to date, which have found that pallidotomy has a minimal effect on cognitive function (see Section 6.1). The one cognitive deficit seen after surgery in this study was small; JG's spatial working memory performance at the postop stage was within normal limits for her age group on most measures and was only a 'deficit' relative to her superior preop performance. Also, there is no data on the chronicity of the changes seen. Lang et al. (1997) have shown that some postoperative deficits are transient and it is possible that JG's deficit resolved soon after surgery. Neither GM nor JG reported that they experienced any improvement in cognitive function after surgery. One month after the surgery, JG commented that she had noted a loss of mental 'sharpness' when performing cognitively demanding tasks but she reported no problems with spatial memory. Any changes in cognitive function experienced by the patients were clearly minimal in terms of quality of life compared to the positive or negative motor consequences of surgery. Despite reporting a loss of cognitive function, JG was hoping to undergo a further pallidotomy in her left hemisphere and GM found his dysarthria so aversive that he reported regretting having undergone the surgery. Although the cognitive changes seen may have been minor in clinical terms, they are theoretically interesting as they represent selective changes in a context of unaltered overall function.

Examination of dissociations in performance between different cognitive tasks may yield information about the cognitive abilities required to perform these tasks.

6.5.2 Results for patient GM

Neither of the two changes that were seen in GM's neuropsychological profile constituted convincing evidence of a postoperative alteration in cognitive function. As noted above (Section 6.4.2.4), GM's changes in scores on the Stockings of Cambridge task appear to have been due to faster but less accurate performance. GM also showed an apparent improvement in ED and ID shifting. GM's mean performance at these shifts at the preop stage was poor but he made no errors at these shifts postoperatively. Although this would appear to be good evidence of an improvement in cognitive function, it was noted in Section 6.4.3 that GM's high scores for the ID and ED shifts at the preop stage were created by averaging one high-error and one low-error shift. Thus, GM's ID and ED shift deficit at the preop stage did not result in a consistent slowing of ID and ED shifts. It can be argued that when a subject can make rapid ID and ED shifts both before and after surgery, no true 'improvement' has occurred as a consequence of surgery. However, it may be incorrect to charaterise the shifting deficits shown by brain-injured populations as slowings that consistently retard the performance of every shift. As noted in Section 3.1.2.6, the distribution of error scores for attentional shifts is often bimodal, reflecting subjects' tendency either to make the shift quickly or to spend some time searching for a solution. It may be that the nature of a shifting deficit in an individual subject is to reduce the proportion of shifts that a subject solves quickly and thereby to increase the proportion of high error solutions rather than to abolish all low-error solutions. If this were the case, group deficits would be seen if groups of such subjects were tested, but the characterisation of an individual's performance in terms of one or two shifts could be misleading. For example, if GM's attentional set-shifting abilities had been judged by his performance on a single ID and ED shift (as in the widely-used CANTAB battery) the results of this study would have changed radically depending on whether he achieved a high-error or a low-error performance at the preop stage. If this account of shifting deficits were true, a multiple-shift test would be required to accurately assess an individual's shifting ability. There is precedent for this; for example, the WCST uses six ED-type shifts. As a consequence of these reseservations, it cannot be confidently concluded that GM's attentional set-shifting performance improved postoperatively.

6.5.3 Results for patient JG

The changes in JG's cognitive performance after surgery were less ambiguous than those of GM. JG showed a deficit in the performance of the Spatial Working Memory test whilst her verbal fluency, verbal memory and attentional set-shifting abilities were unimpaired. Most significantly, her performance on the Stocking of Cambridge appeared to improve after surgery. The interpretation of two of these results is uncertain, and requires further comment. Whilst the evidence for JG's improvement on the Stockings of Cambridge task is compelling, it must be noted that she was younger and of a higher IQ than controls; it is possible that she could show a greater ability to benefit from practice than controls and that this could account for her postoperative improvements. However, there is no literature relating to practice effects on the Stockings of Cambridge task in late middle age, so this issue cannot be resolved. JG's relative youth and high IQ make her postoperative deficits on the Spatial Working Memory task even more noteworthy, particularly as these deficits extended to problems in recalling the location of very recent responses. JG also performed well on the test of attentional set-shifting that was used in this study. As noted in Section 6.4.3, her 'optimal' preop performance made it impossible for her to show any postoperative improvement in performance. Thus, the only conclusion that can be drawn from JG's attentional set-shifting results is that pallidotomy caused no clear deficit in her performance of this test.

JG's high level of performance had not been anticipated when the attentional set-shifting test was designed. It had been assumed that surgical patients would perform less well at attentional set-shifting than controls due to the surgical patients' more advanced disease status (Owen et al., 1992). However, patients are often selected for surgery on account of their relatively intact cognitive status. This raises two issues; first, tests for the assessment of pallidotomy outcome must not be calibrated on the assumption that surgical patients' cognitive performance will be typical of other patients with severe PD. Second, the assessment of cognitive improvements after pallidotomy is impossible if patients do not have preoperative cognitive deficits. The practice of excluding patients with dementia from pallidotomy is widespread and the consequence of testing relatively 'intact' patients may be to underestimate pallidotomy's effectiveness in ameliorating cognitive deficits.

The differential effect of surgery on the performance of the Spatial Working Memory task and the Stockings of Cambridge task constitutes a dissociation; the two tasks clearly do not involve identical cognitive mechanisms. This dissociation is valid

whether JG's postoperative improvement on the Stockings of Cambridge task is a genuine restitution of cognitive processes or an effect of practice. The fact that there is impairment of performance at the Spatial Working Memory task whilst there is not even a trend towards an impairment of performance of the Stockings of Cambridge test is good evidence for the involvement of different cognitive mechanisms in the two tasks.

The dissociation between performance of the Spatial Working Memory task and the Stockings of Cambridge task was not unanticipated. It had been shown that some cognitive abilities are involved in the performance of both of these tasks, but also that they require different cognitive aptitudes. Robbins (1996) has demonstrated the existence of common abilities between the two tasks using factor analysis; there is a correlation between the proportion of perfect solutions achieved on the Stockings of Cambridge task and both the strategy and between-search error scores of the Spatial Working Memory task in healthy volunteers. Studies using positron emission tomography have also shown that the two tests share anatomical substrates. Baker et al. (1996) showed that attempting to solve more difficult Stockings of Cambridge problems was associated with increased rCBF bilaterally in the DLPFC (areas 9 and 46). Owen et al. (1996b) have found activation of similar areas in subjects performing a task similar to the Spatial Working Memory task in this study. The cognitive abilities that are shared by these tasks appear to be the 'strategic' and 'executive' functions (Robbins 1996, Owen 1997) that can benefit spatial working memory performance and that are essential for Stockings of Cambridge performance. An impairment of these 'executive' processes could not account for JG's results, as it would disrupt performance of both the Spatial Working Memory and the Stockings of Cambridge task.

The Spatial Working Memory task has also been shown to involve more simple processes of memory retrieval that are less strongly associated with performance of the Stockings of Cambridge test. The influence of these less 'executive' processes on Spatial Working Memory performance has been demonstrated by the study of a variety of brain-injured populations. Increased between-search errors on this test have been found to occur in the absence of a deficit on the strategy measure in patients with PD (Owen et al. 1992) and with temporal lobe damage (Owen et al. 1996c). Between-search errors also correlate significantly with performance of the Spatial Span test of the CANTAB battery (Robbins, 1996). This is a simple test that requires subjects to reproduce a spatial sequence of increasing length. It places strong demands on recall processes and involves little or no strategy use. It is possible that the retrieval processes associated with the Spatial Span test were impaired in JG by surgery. These processes

are likely to play a stronger role in Spatial Working Memory performance than in Stockings of Cambridge performance and impairment of this mechanism could therefore account for the selective spatial working memory deficit.

If JG were impaired in the performance of the Spatial Span test, this would be evidence of an impairment of 'memory retrieval' processes. However, JG's spatial span was tested one month after surgery and was found to be normal. This does not rule out the possibility of an impairment in memory retrieval, as JG's Spatial Working Memory performance was also within normal limits at the postop stage; both her spatial working memory and spatial span performance may have declined from high preop baselines. However, the most important piece of evidence relating to a possible deficit in memory retrieval processes is anatomical. The 'executive' and 'retrieval' processes described above have been associated with different areas of lateral frontal cortex by Petrides (1996) and Owen (1997). The 'executive' processes that are involved in the active monitoring and organisation of material are thought to be associated with the DLPFC (areas 9 and 46). In contrast, the retrieval of material from memory is associated with areas in the ventrolateral prefrontal cortex (VLPFC, areas 47, 12 and 45). Impairment of 'retrieval' processes in JG would therefore be associated with disrupted information processing in VLPFC. This is unlikely to have occurred as a consequence of pallidotomy, as VLPFC (areas 47, 12, 45) do not receive projections from the basal ganglia.

Consideration of JG's performance in the context of basal ganglia-frontal lobe connections indicates that her impairment on the Spatial Working Memory task did not arise from either of the 'executive' or 'retrieval' processes proposed by Petrides (1996) and Owen (1997). A third process appears to exist that is involved in the performance of the Spatial Working Memory task but not the Stockings of Cambridge task. The existence of this third process has already been indicated by a study of PD. Owen et al. (1992) found that their patients with mild, medicated PD showed a deficit on the Spatial Working Memory task in the presence of normal levels of strategy use and normal performance of the Spatial Span test. However, characterisation of the nature of this process is beyond the scope of this study.

6.5.4 Summary

Pallidotomy had little overall effect on cognitive performance and caused only a moderate improvement of motor function. Improvements in postoperative performance were seen on the Stockings of Cambridge test for JG, but it is possible that these were

due to enhanced practice effects that were a consequence of her relative youth and high IQ. JG also showed a postoperative deficit in performance of the Spatial Working Memory test. JG's performance on the Stockings of Cambridge test and the Spatial Working Memory test allowed identification of a cognitive process that is involved in the Spatial Working Memory test but is neither of the 'lateral frontal' working memory processes proposed by Petrides (1996) and Owen (1997).

7.0 General Discussion

7.0 General discussion

This thesis reports the results of studies of attentional and motor set in patients with PD and healthy individuals. The final study (Study 6, Section 6.0) also reports the cognitive consequences of a neurosurgical procedure that is designed to benefit motor function. The following sections will briefly summarise the results of these studies, whilst aiming at minimal repetition of previous discussion sessions. The domains of attentional and motor set will be considered separately; each summary of results will be followed by a discussion of issues arising that are relevant for future studies. The consequences of pallidotomy will be considered in a further section and the final section of this discussion will place the studies in this thesis in a broader context and raise salient questions for future research.

7.1 Attentional set

7.1.1 Summary of results

7.1.1.1 Characteristics of the ED shift deficit in patients with PD

Studies 1 to 4 in this thesis has investigated attentional set in patients with PD. A deficit in the performance of ED shifts has been taken as the purest expression of the deficits of attentional set seen in PD. Study 1 allows some conclusions to be drawn about the general characteristics of ED shift performance in patients with PD. This study showed that an ED shift deficit was present in patients with PD both when the newly relevant dimension after the shift was novel (P and LI+ shifts) and when it had been experienced before (LI shift, see Table 2.2). Thus, the deficit seen in patients with PD related solely to the requirement to shift dimensions, irrespective of patients previous experience of those dimensions. In accordance with previous studies (e.g. Downes et al., 1989) ED shifting was found to impaired, and ID shifting spared, in patients with PD. There was also an increase in errors after criterion in patients, which caused an apparent contradiction; an ED shift deficit in the presence of spared ID shifting can only be a consequence of exaggerated, inflexible set, yet such an exaggeration of set should minimise errors after criterion. This can be explained by proposing that patients form an excessively strong set before an ED shift - that is, they focus excessively on the relevant dimension. This set retards ED shifting and then disturbs discrimination performance after the shift by periodically intruding and causing 'loss of set' to the newly-relevant dimension. Flowers and Robertson (1985) originally proposed that patients with PD

were diminished in their ability to inhibit a competing attentional 'set'; this appears to be true of the patients in Study 1.

7.1.1.2 Evaluation of theories of the cause of the ED shift deficit in patients with PD

Study 1 tested and rejected the hypothesis that exaggerated learned irrelevance is the cause of the ED shift deficit in patients with PD. The learned irrelevance (LI) shift that had been devised by Owen et al. (1993b) was shown to be sensitive to mechanisms that were not learned irrelevance. A number of other possible explanations of the ED shift deficit in patients with PD were also ruled out. It was shown that the ED shift deficit cannot arise from a basic instability of set or a failure to maintain set (see Section 1.6.1). The presence of weak or unstable set before a shift would in fact make ED shifting easier, as the relative difficulty of ED shifts compared to ID shifts is caused in part by the requirement to abandon the set that was present before the shift. If this set were weak, it would be discarded easily and the ED shift would be performed rapidly. The ED shift deficit of patients with PD indicates that preshift set is exaggerated rather than diminished, as discussed above.

Brown and Marsden's (1990) theory of attentional resource depletion provides another potential explanation for the ED shift deficit of patients with PD. Brown and Marsden suggested that patients with PD are impaired in all cognitive tasks that exceed their depleted attentional resources. Thus, it is possible that the cognitive demands of ED shifting exceed the resources of patients with PD, whereas the demands of ID shifting do not. This idea is plausible, as ED shifts are clearly harder for healthy subjects than ID shifts. However, an explanation in terms of Brown and Marsden's theory cannot account for the increase in errors committed after criterion, which occur at an attentionally undemanding stage of the test. In contrast, these errors would be predicted if patients could not inhibit a strong competing set, as suggested above.

Thus, Study 1 showed that the ED shift deficit of patients with PD was due to an exaggeration of set; this exaggeration was caused neither by learned irrelevance nor a 'depletion of resources'. When considering the implications of Study 1, it became clear that the cognitive processes underlying ED shifting were poorly understood. Study 1 had cast doubt on Owen et al.'s (1993b) account of ED shifting, which had been the most contemporary and plausible account available. Therefore, Study 1 was followed by a series of investigations of the mechanisms underlying ED shifting in healthy subjects (Studies 2 to 4).

7.1.1.3 Mechanisms underlying ED shifting in healthy subjects

It was noted in Study 1 that the proposed mechanism of 'learned irrelevance' closely resembles the well-established effect of latent inhibition. Also, the 'test' phase of human latent inhibition paradigms closely resembles an ED shift. As a consequence, it was concluded that learned irrelevance/latent inhibition might influence ED shift difficulty despite the conclusions of Study 1. Studies 2 and 3 used 'preexposure' designs that were adapted from latent inhibition experiments to test the hypothesis that preexposure of the irrelevant dimension would retard ED shifting in a fashion similar to latent inhibition. Preexposure had no effect on ED shifting. As learned irrelevance is defined as a loss of associability that a stimulus (or dimension) incurs when it is presented uncorrelated with reward, this result shows that learned irrelevance cannot be involved in determining the difficulty of an ED shift. This conclusion is reinforced by evidence that a low level of preexposure of both dimensions often facilitates ED shifting ('overlearning' - see Wolff, 1967 and Section 3.2.4.3).

Having rejected the role of learned irrelevance in ED shifting, an entirely different account of the ED shifting deficit of patients with PD was proposed in Study 4. This account centred around an analysis of the process of hypothesis-testing after an ED shift. It was noted that subjects can gather information relating to both dimensions after receiving feedback from a response; that is, the negative feedback that is received after pressing a green square yields information about both the dimensions 'colour' and 'shape'. A study by Channon et al. (1993) reported that patients with PD were less able than controls to use feedback to gather information about the dimensions that they were not currently attending to and testing. Channon et al. (1993) did not demonstrate this effect using an ED shifting task, but it was noted in Section 3.3.1 that such a deficit would selectively interfere with ED shifts, but not ID shifts, in an attentional set-shifting task. Therefore, the deficits seen in Channon et al.'s study could be the source of the ED shift deficits of patients with PD. Study 4 tested the hypothesis that a restriction of the ability to gain information from an untested dimension would lead to an impairment only of the ED shift. Young healthy subjects' ability to gather information from an untested dimension was reduced by an experimental manipulation involving the use of masking and restriction of stimulus display time. This manipulation had the predicted effect; it altered ID and ED shift performance differentially rather than enhancing the difficulty of both shifts. Thus, it appears possible that the ED shift deficit of patients with PD could be accounted for by the deficit found in Channon et

al.'s (1993) study. However, it was also noted that the absolute level of errors found in Study 4 could not have been generated by an inability to learn from the untested dimension. The source of these errors has not been identified, but subjects' subjective reports suggest that inattention to the newly-relevant dimension is involved (see Section 3.3.4.4).

7.1.2 Issues arising from these results

7.1.2.1 The nature of stimuli and stimulus presentation

It is clear that the number of errors generated at an ED shift can vary substantially between studies, even when those studies use structurally identical shifts and the same dependent measures. For example, the absolute level of errors seen in Study 1 was different to that seen in Owen et al.'s (1993) study, which used identical shifts and demographically similar subjects. Stimulus differences appear to account for these varying error levels; different studies use stimuli and dimensions that vary in salience and discriminability, and these factors affect the ease of a shift. The influence of stimulus characteristics on shift difficulty has been show most clearly in the analysis of the ID shifts used in Study 4. Subjects completed four sequential ID shifts at the beginning of the test; all of these shifts involved the same relevant and irrelevant dimensions, and experimental conditions were the same throughout. Despite the fact that the shifts only differed in the specific stimuli used, the shifts differed significantly in difficulty (see Section 3.3.4.1). Errors were higher when subjects were required to shift to stimuli that were less discriminable. This finding has clear implications for the design of studies that involve the comparison of ID and ED shift performance. If ID and ED shifts use stimuli that are not comparable in difficulty, then manipulations that affect only one type of shift must be interpreted with caution.

Use of relatively simple stimuli has meant that subjects often perform at ceiling in studies of attentional set-shifting. In past studies using the CANTAB battery, both controls and patients are often at ceiling in the performance of ID shifts. This is not usually problematic, as most CANTAB studies have involved demonstrating deficits of ED shift performance in relation to ID shift performance, making ceiling effects largely irrelevant. However, there is a danger of missing a subtle patient deficit when both patients and controls are performing at ceiling at an ED shift. For example, a subtle deficit could have been missed in patients with frontal lobe damage at the LI shift in Owen et al.'s (1993) study, as both the patients and their control group were close to

'ceiling'. However, comparing control groups that are at ceiling with patient groups that are not at ceiling creates more serious problems for data analysis.

7.1.2.2 Analysis of learning-to-criterion data from ED shift paradigms

Conventional parametric analysis of error data from ED shifting tasks is often inappropriate for two reasons. First, the amount of variance in the data of groups that are at a 'ceiling' or away from a 'ceiling' in performance can differ greatly. Second, the distribution of data is often bimodal, or at least positively skewed. There is little consensus in the past literature on ED shifting as to what is the most appropriate method of analysis for errors / trials to criterion data. Investigators have used ANOVAs, either on square-root transformed data (median errors, Roberts et al., 1988; mean trials, Downes et al., 1989) or untransformed data (means, Elliott et al., 1995; Owen et al., 1993b). With the benefit of hindsight, it may be less than optimal to use parametric tests to compare groups of subjects between which the standard errors differ by a factor of five, as in Owen et al.'s (1993b) study (LI shift, MED PD patients versus controls). Such criticisms are not serious, as the CANTAB literature has usually dealt with quite large, unambiguous effects. Nonetheless, this should not deter the search for more appropriate methods of analysis. This thesis is indebted to Thornton et al.'s (1996) suggestion of data analysis via the construction of survival or hazard tables. Whilst the Life Tables analysis used in Studies 2 to 4 is more simple than the methods used in Thornton et al.'s (1996) paper, it can account for the artificial scores of subjects who never reach criterion as well as the bimodal distribution of data. Trials to criterion data from ED shifts conform to all of the assumptions of Life Tables analysis. Thus, the various survival / hazard tables approaches to data analysis (e.g. Life Tables, Kaplan-Meier Survival Analysis, Cox's Proportional Hazards Regression Analysis) may be optimal for analysing data from ED shift tasks.

7.1.2.3 ED shifts as a measure of attentional set-shifting ability

Data from GM in Study 6 raised doubts as to whether the performance of a single ED shift can accurately reflect a subject's attentional set-shifting abilities. The bimodal distribution of ED shift data meant that GM's error scores on his two preop ED shifts varied between very high and very low. It was unclear which shift reflected his ED shifting abilities. The distribution of ED shift error data means that testing attentional set-shifting ability with a single shift is similar to testing memory by requiring recall of a single item; the resultant score is effectively one or zero. This approach may be valid

in the context of a group design. A higher percentage of patients with amnesia than controls would forget a single item and this would reflect a memory deficit. Indeed, investigators have chosen to report CANTAB ED shift performance in these terms; Lange et al. (1992) reported that only 50% of patients successfully complete the ED shift compared to 74% of controls. However, the recall of a single item could never be a good indicator of an individual's memory performance, and neither can a single ED shift reflect an individual's attentional set-shifting ability. Thus, any test involving only a single ED and ID shift (such as the CANTAB attentional set-shifting test) can be of no use in a single case design or in a clinical setting. It is to be hoped that tests incorporating multiple shifts will be developed in the future.

When more 'parametric' measures of attentional set-shifting ability become available, it will be possible to describe an individual's performance with respect to that of a control group. Future studies will be able to report more statistics about the distribution of deficits in a patient group, such as the percentage of patients scoring more than one standard deviation below the control mean. It is no longer sufficient to conclude that patients are 'impaired' or 'relatively impaired' (Downes et al., 1989) at ED shifting on the grounds that the mean error score of a group of patients is significantly higher than that of controls. The population of patients with PD is likely to be as heterogeneous with respect to cognitive symptoms as it is to motor symptoms. Therefore, it is quite likely that some patients will show intact ED shift performance, just as some patients show no tremor. Indeed, Downes et al. (1989) reported that over 60% of their medicated patients with PD successfully completed an ED shift. There is a need for multiple single-case designs in this type of research, where both group means and individual performance are reported (e.g. Della Sala et al., 1995; Logie et al., 1996). It could be the case that some patients with PD in fact show 'loss of set' and superior ED shift performance, or that there is a sharp disjunction between patients with intact ED shifting and those with a deficit. Until multiple single-case designs are adopted, none of these possibilities can be detected.

A number of methodological issues have been raised in this section. Although these issues are quite diverse, it is interesting to note that they are identical to the list of problems with latent inhibition studies reported by Lubow (1997)[1].

[1] "Methodological issues in human LI research include: dichotomous data, ceiling effects, inappropriate data analysis, individual versus group data, and the need for within-subjects designs and parametric studies." Lubow (1997), p. 75

7.2 Motor set

7.2.1 Summary of results

7.2.1.1 Under what conditions do patients with PD show an RT deficit?

The quantitative analysis carried out in Section 4.0 allowed the assessment of general hypotheses about the conditions under which patients with PD show an RT deficit. It was clear from this analysis that patients with PD consistently show slowed simple RTs. However, the presence of deficits on choice RT tasks was less consistent. The quantitative analysis found that the absolute speed at which controls could perform a choice RT task was an important variable for determining whether a PD deficit was present; deficits were more likely in tasks at which controls could achieve 'fast' (that is, short) RTs, but were less likely at 'slow' tasks. This finding was true in a range of tasks in which control RTs varied between around 300 ms and 550 ms. In line with this prediction, an RT deficit was found in Study 5, in which controls had relatively 'fast' RTs of between 320 and 390 ms. The review in Section 4.0 also found that patients were more likely to show an RT deficit when off medication. However, it appears that to be genuinely 'off' medication, prolonged medication withdrawal is required - skipping a single dose may not suffice (see Section 4.3.2). Medication effects may account for spared patient performance on 'slow' choice RT tasks , as levodopa appears to improve RTs more on such 'slow' tasks than on 'fast' tasks (see Section 4.3.3.1).

7.2.1.2 Conclusions about motor set in PD

This section will consider what impact PD has on temporal and spatial set in turn, followed by a consideration of a third type of set that has been termed 'attentional focussing'. 'Temporal set' and 'spatial set' are defined here, as in Section 5.0, as a readiness to respond at a particular time and to a particular location. Study 5 showed that the distinction between temporal and spatial set was functionally valid. There was a dissociation between the responses of young and elderly healthy subjects to manipulations of spatial and temporal probability. Healthy elderly subjects did not show the response to spatial probability that was seen in young subjects; however, healthy elderly subjects showed a response to temporal probability that was as large as that of young controls, if not larger. Thus, whilst the RTs of young subjects may be determined by both the temporal and spatial probability of stimulus occurrence, their response to these two types of probability is mediated by functionally and neurally separate systems.

Temporal set appears to be intact in PD. Study 5 investigated the temporal 'set' that mediates RT speeding with lengthening foreperiod known as 'delay-dependent speeding'. It was shown that this effect was due to subjects' response to information about the temporal probability of stimulus occurrence. Patients with PD show normal RT variation with manipulations of temporal probability. However, they also show a consistent RT deficit that was unrelated to the changes in RT caused by manipulations of temporal probability. Thus, deficits of temporal set cannot be the cause of the RT deficits of patients with PD. It can also be concluded that the more 'cognitive' components of temporal set are intact. In order for patients with PD to show normal RT variation with temporal probability change, they must be able to estimate time delays accurately and show successful implicit learning of the relationship of these delays to temporal probability. These abilities are unaffected by PD.

Spatial set has been investigated in two sections of this thesis. First, the 'Motor preprogramming' section of the quantitative review (Section 4.2) investigated the type of spatial set that can rapidly be formed when a subject is cued about the direction of an upcoming movement. Second, in Study 5 spatial probability was manipulated such that a response was more likely - but not certain - to be required at one location than another. This was designed to create a 'set' that would allow more rapid responses to the more probable location. A quantitative analysis and literature review (Sections 4.2.3.2 and 4.2.2) have shown that the first type of set is intact in PD; patients with PD benefit as much as controls from being cued about the direction of a forthcoming movement. However, no conclusions could be drawn about spatial set from Study 5, as neither patients nor their age-matched controls showed changes in RT as a result of the manipulation of spatial probability that was used. The failure of this manipulation to affect RT in elderly subjects was itself of interest, as it contrasted with young subjects' intact response to spatial probability and contributed to the dissociation described above.

The final aspect of 'motor set' that must be considered has not been well characterised, but its existence can be inferred from its effect on simple RT tasks. As noted above, patients with PD show a deficit on simple RT tasks that is more consistent than that seen in choice RT tasks. Goodrich et al. (1989) have proposed that this simple RT deficit derives from a failure to engage in 'attentional focussing' by patients with PD (for a full discussion of this hypothesis see Sections 1.5 and 4.4.2). 'Attentional focussing' is Goodrich et al.'s term for a type of motor readiness that can be achieved by healthy subjects at simple, but not choice RT tasks; as such, it can be described as a

type of 'set'. Goodrich et al. (1989, 1990) have supported their hypothesis with evidence from studies of patients with PD and healthy controls. The only other well-articulated theory that attempts to account for the consistency of the simple RT deficit in patients with PD is that of Hallett (1990) which is described in Section 4.2.2. Hallett's hypothesis of a 'lengthened motor initiation stage' could account for the PD simple RT deficit, but there is no evidence to support this assertion and this hypothesis cannot account for Goodrich et al.'s (1989, 1990) data.

The type if motor set identified by Goodrich et al. cannot be identical to either temporal or spatial set, as these are both intact in PD. Thus, a third type of set must be proposed. It appears that the phrase 'motor set' that was introduced in Section 1.5 includes at least three separate psychological processes, each of which may be differentially vulnerable to brain damage.

7.2.2 Issues arising from these results

7.2.2.1 Understanding of set

The results described above permit the continuation of Gibson's (1941) project to define 'set' precisely and to delimit different type of set. Three different types of set have been identified, only one of which is impaired by PD. It has also been shown that some types of set are only present under certain conditions. Healthy elderly controls were no quicker to make a response that they knew to be twice as probable as an alternative (a 2:1 bias - Study 5). Thus, no set had been formed to take advantage of the bias in probability. It is interesting to note that groups of younger healthy subjects (aged around 30 and 50) have also been shown to display no RT speeding towards a more probable target in a task that involved a 3:1 bias (Georgiou et al., 1996). The absence of a spatial set in Study 5 is in sharp contrast to the findings of other paradigms that succeed in forming spatial set in healthy elderly subjects. Healthy elderly subjects are able to use cues to speed their responses in studies of motor preprogramming (see Section 4.2.3.2). They also show speeding of RTs through implicit learning of sequences in serial RT tasks (e.g. Pascual-Leone et al., 1993). The failure of probability manipulations to alter RT must be accounted for; two related factors, response certainty and response probability, may account for this.

The most obvious difference between the 'probability' manipulation in Study 5 and more effective manipulations is that the probability manipulation does not allow certainty about the location of the next response. Precueing provides such certainty (in most designs) and if an embedded sequence is successfully learned, this also allows

certainty. Thus, it is possible either that organisms cannot form a set without certainty about the location of the next response, or that they judge it (implicitly or explicitly) not to be beneficial to form a set under such conditions. An alternative explanation may be that certainty is not required, but that a more substantial manipulation of probability is required to form a set that will alter RT. It has been noted above that a 2:1 bias of spatial probability had no effect on RT; interestingly, this level of temporal probability also had no effect on RT. Table 5.2 shows that for the Rising block on the more probable side, foreperiods of 1000 and 500 ms are associated with probabilities of 0.66 and 0.33, respectively (a 2:1 bias). An inspection of the relevant data points in Figures 5.1 and 5.2 shows that there is very little RT difference between these points, despite a 2:1 bias of temporal probability. Indeed, the only major RT differences seen in these graphs are RT slowings associated with low probabilities, such as at the 250 ms foreperiod in the Rising block and the 1000 ms foreperiod in the Falling block. Thus, it appears that a motor set may manifest as a slowing of response to unanticipated, low probability times or locations rather than an incremental speeding of response with increasingly high probability.

The results reported above have allowed the identification of multiple types of set involved in RT tasks. However, it appears that the conditions required to establish a set may be quite specific, and intuitive ideas may be a poor guide to creating manipulations that elicit set. For example, all elderly subjects became aware of the spatial probability bias soon after starting the experiment, which led to the expectation that they would show corresponding changes in RT; they did not show such RT changes.

7.2.2.2 Implications of the RT deficit of patients with PD

It is important to note that the mere presence of a simple or choice RT deficit in patients with PD signifies very little. The use of tasks involving response time measures has been extremely productive in investigating specific hypotheses about PD, such as the suggestion that patients are unable to use 'advance information' to speed their responses. However, it is unclear how to interpret the mere presence of slowed RTs in patients with PD. RTs are correlated with the motor symptoms of PD, but less well so than movement times (Zappia et al., 1994); RTs are clearly inferior to both clinical rating scales and movement times as an indicator of disease status. Although investigators have (perhaps somewhat lazily) associated RT deficits with 'akinesia' and deficits of response initiation, no attempt has been made to correlate RT with symptoms

such as gait freezing, in which patients' ability to initiate movements is completely absent. RTs must necessarily index some form of response initiation, but the type of response required in RT tasks - usually an externally cued, unimanual ballistic movement - is the type of response that is probably least impaired by PD. It has yet to be established that an RT deficit in patients with PD signifies anything other than the presence of brain damage (as in head injury - Miller, 1970; or Alzheimer's disease - Gordon and Carson 1990). Indeed, Section 4.0 has shown that the RT deficit of patients with PD may obey quite mundane laws relating to the overall 'speed' of an RT task. The mere presence of an RT deficit in patients with PD tells us nothing about PD.

7.3 The effects of pallidotomy on cognitive function

7.3.1 Summary of results

Both Study 6 and the review of pallidotomy studies in Section 6.1 indicate that pallidotomy has no major detrimental effects on cognitive function. JG showed a significant drop in performance of a spatial working memory task after surgery, but did not report that this created any problems in her daily life. The review in Section 6.1 concluded that deficits of verbal fluency were a relatively consistent consequence of surgery; however, neither JG nor GM showed a postoperative decline in verbal fluency. This may have been due to the location of the surgery. It was noted in Section 6.1.2.4 that verbal fluency deficits were more common after left hemisphere surgery, whereas both JG and GM underwent right pallidotomy.

Study 6 also demonstrated that performance of the CANTAB Spatial Working Memory task can involve at least three separate cognitive mechanisms. The dissociations in JG's performance and a consideration of basal ganglia connectivity showed that JG's deficit on the Spatial Working Memory task could not have involved either the 'strategic' or the 'memory recall' processes that are thought to be involved in the performance of the task (Owen, 1997). This is not an entirely novel finding, as patients with PD have been shown to have a spatial working memory deficit in the presence of intact strategy use and spatial span (recall) performance (Owen et al., 1992). However, the results from JG demonstrate the existence of a third process quite clearly, as this process is selectively impaired as a consequence of quite limited brain damage.

7.3.2 Issues arising from these results

Many of the methodological issues surrounding the followup studies of pallidotomy have already been discussed in Section 6.1.1. In followup studies of pallidotomy there is once again a need for multiple single-case designs. Reporting unchanged mean group performance after surgery can conceal the presence of consistent postoperative improvements and decrements in cognitive performance in individual patients. Although pallidotomy appears to be a neuropsychologically 'safe' procedure, serious adverse consequences have been reported (e.g. Ghika et al., 1996; Lang et al., 1997). Future large-scale studies should attempt to establish the percentage occurrence of these complications, their chronicity and their associated risk factors.

The surprising failure of most previous studies to account for test-retest effects was also discussed in Section 6.1.1. The need for control subjects in followup studies of

pallidotomy is evident, and the repeated use of neuropsychological tests without alternate forms or controls appears to be a gross error of method. A final issue that arises from reviewing past studies of pallidotomy is the apparent ignorance of most researchers of the past literature on the topic. None of the studies reviewed in Section 6.1 cited any of the psychological followup studies of pallidotomy that were carried out in the 1960s (e.g. McFie, 1960; Riklan et al., 1960) despite their obvious relevance and publication in prominent journals. It is clear that the neurosurgical techniques used 40 years ago differ from contemporary methods and that these older neuropsychological studies sometimes used tests that are now seen as antique. However, Riklan et al. (1960) used the WAIS and carried out a long term followup of pallidotomy patients; this must be relevant, and ignorance of such studies may have allowed the 'rebirth' of pallidotomy to commence without appropriate neuropsychological evaluation (Laitinen et al., 1992).

7.4 Future directions for research

The research reported in this thesis has shown that understanding of a deficit such as the WCST deficit seen in patients with PD can be refined and clarified. Increasingly well-specified hypotheses can be generated and tested to explain such a deficit. A substantial amount of time and money has been spent in establishing and investigating the ED shift deficit seen in patients with PD (Downes et al., 1989; Lange et al., 1992; Owen et al., 1992, 1993; Robbins et al., 1994). However, it has not been established that the deficit of attentional set-shifting revealed by ED shifting tasks has any negative impact on the lives of patients with PD. Results from JG have shown that an individual can undergo a change in cognitive performance that is statistically significant but unnoticeable. Also, the cause of the cognitive deficits of PD has not been clearly established. It has been assumed that the cognitive difficulties seen in PD are a direct consequence of the neuropathology of PD; however, there are other possible causal mechanisms.

The ability to attend selectively to one aspect of a compound stimulus is thought to be an aspect of 'executive' function. Tests that are thought to tap executive function have been shown to be sensitive to sleep disturbance in healthy subjects (Fluck et al., 1998) and to motor activity in patients with PD (Brown and Marsden, 1991). Sleep disturbance is prevalent in PD (Pearce, 1992) and motor activity is a requirement of most cognitive tasks. Thus, it is possible that problems such as the ED shifting deficit are not a direct consequence of neural damage, but may instead be caused indirectly by one of the non-cognitive symptoms of PD. If this were the case, then the logic of the many studies that use PD as a model of basal ganglia dysfunction or dopamine depletion would be invalid. Also, studies attempting to attribute cognitive deficits to particular brain regions would be misconceived. Further research is needed to establish whether the cognitive deficits seen in PD are directly or indirectly caused by the neuropathology of PD.

The ED shift deficit is a persistent and refractory symptom (see Section 1.6.4) that is common in patients with PD. Thus, it can be argued that that this deficit is an important symptom that should be investigated. However, there is a contrast between the results of neuropsychological tests and the subjective reports of patients. In the author's experience, individuals with early PD deny the presence of cognitive deficits whilst freely admitting to depression and other embarrassing medical conditions. Also, neither medical texts (e.g. Pearce, 1992) nor information from charities (e.g. National Parkinson Foundation, 1996; Parkinson's Disease Society, 1998) list cognitive

impairment as a presenting symptom of early PD. Thus, a high percentage of patients with early PD have ED shift deficits, but neither these patients nor observers notice any cognitive impairment, let alone report it as a troubling symptom. Deficits of attentional set-shifting only become evident when neuropsychological tests are carried out. I wish to argue that it is essential to establish what impact ED shift deficits have on activities of everyday life and the quality of life of patients with PD. Currently, the ED shift deficit seen in PD can be compared to the anosmia seen in PD. Anosmia can be present from an early stage of PD and it persists and is refractory to medical treatment (Doty et al., 1988). As with ED shift deficits, anosmia is present in some of the supposedly 'healthy' general population and can be seen in other neurological disorders. There is currently no evidence that either anosmia or attentional set-shifting deficits are in any way a serious handicap for patients with PD, in contrast to their motor symptoms. A similar argument has already been advanced about the RT deficit seen in patients with PD (see Section 7.2.2.2), as this deficit is not the best index of disease status in PD and has never been explicitly associated with movement initiation impairments.

Future research relating to deficits such as the ED shift deficit in PD can be justified in one of two ways. First, it can be argued that research such as that reported in this thesis is aimed at understanding the symptoms of PD in order to improve the lives of patients with PD (at some point, however distant). If this is the case, it is essential to establish that deficits such as the ED shift deficit are significant to patients with PD. If such deficits have little impact, then money and time should be spent on investigating the other disabling symptoms of PD. Alternatively, it can be argued that PD represents an instance of brain damage and as such the performance of patients with PD on neuropsychological tests may reveal dissociations that can inform models of normal cognitive function. An example of this approach is the demonstration of a double dissociation between the performance of patients with amnesia and patients with PD on different memory tasks (Knowlton et al., 1996). This thesis has found two such dissociations - between temporal and spatial set (Section 5.4) and between different processes involved in spatial working memory (Section 6.5.3). However, if the search for dissociations is the justification for future research, it must be accepted that patients with PD are only used because of their brain damage, not because PD is of interest *per se*. Therefore, funding should not be sought on the basis of the potential benefits of research for patients with PD. It should be noted that the two dissociations reported in this thesis arose from the consequences of aging and a pallidal lesion rather than PD.

In summary, research in this thesis has investigated the source of attentional and motor deficits in PD. In the future, focus must shift to the impact of these deficits. Once the impact of these deficits is understood, it will be clear that further research on these topics is justified.

References

Agid, Y., Ruberg, M., Raisman, R., Hirsch, E., & Javoy-Agid, F. (1990). The biochemistry of Parkinson's disease. In G. M. Stern (Ed.), *Parkinson's disease* . London: Chapman & Hall Medical.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organisation of functionally segregated circuits linking basal ganglia and cortex. *Ann. Rev. Neurosci., 9*, 357-381.

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *TINS, 13*(7), 266-271.

Baddeley, A. (1986). *Working Memory*. Oxford: Oxford University Press.

Baker, S. C., Rogers, R. D., A.M., O., Frith, C. D., Dolan, R., Frackowiak, R. S. J., & Robbins, T. W. (1996). Neural systems involved in planning: a PET study of the Tower of London task. *Neuropsychologia, 34*(4), 515-526.

Baron, S., Vitek, J. L., Bakay, R. A. E., Green, J., Kaneoke, Y., Hashimoto, T., Turner, R. S., Woodard, J. L., Cole, S. A., McDonald, W. M., & DeLong, M. R. (1996). Treatment of advanced Parkinson's disease by posterior GPi pallidotomy: 1-year results of a pilot study. *Annals of Neurology, 40*, 355-366.

Battig, K., Rosvold, H. E., & Mishkin, M. (1962). Comparison of the effects of frontal and caudate lesions on discrimination learning in monkeys. *Journal of Comparative and Physiological Psychology, 55*, 400-404.

Ben-Shlomo, Y. (1996). How far are we in understanding the causes of Parkinson's disease? *Journal of Neurology, Neurosurgery and Psychiatry, 61*, 4-16.

Benecke, R., Rothwell, J. C., Dick, J. P. R., Day, B. L., & Marsden, C. D. (1986). Performance of simultaneous movements in patients with Parkinson's disease. *Brain, 109*, 739-757.

Benecke, R., Rothwell, J. C., Dick, J. P. R., Day, B. L., & Marsden, C. D. (1987a). Disturbance of sequential movements in patients with Parkinson's disease. *Brain, 110*, 361-379.

Benecke, R., Rothwell, J. C., Dick, J. P. R., Day, B. L., & Marsden, C. D. (1987b). Simple and complex movements off and on treatment in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 50*, 296-303.

Bennett, C. H., Maldonado, A., & Mackintosh, N. J. (1995a). Learned irrelevance is not the sum of exposure to CS and US. *Quarterly Journal of Experimental Psychology, 48B*, 117-128.

Bennett, K. M. B., Waterman, C., Scarpa, M., & Castiello, U. (1995b). Covert visuospatial attentional mechanisms in Parkinson's disease. *Brain, 118*, 153-166.

Bhatia, K. P., & Marsden, C. D. (1994). The behvioural and motor consequences of focal lesions of the basal ganglia in man. *Brain, 117*, 859-876.

Bloxham, C. A., Mindel, T. A., & Frith, C. D. (1984). Initiation and execution of predictable and unpredictable movements in Parkinson's disease. *Brain, 107*, 371-384.

Bloxham, C. A., Dick, D. J., & Moore, M. (1987). Reaction times and attention in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 50*, 1178-1183.

Bonardi, C., & Hall, G. (1996). Learned irrelevance: no more than the sum of CS and US preexposure effects? *Journal of Experimental Psychology: Animal Behaviour Processes, 22*(2), 183-191.

Bowen, F., Kamienny, R. S., Burns, M. M., & Yahr, M. D. (1975). Parkinsonism: effects of levodopa on concept formation. *Neurology, 25*, 701-704.

Bowers, D., Crucian, G. P., Friedman, W. A., Triggs, W. J., Behrman, A., & Greer, M. (1997). Memory and neuropsychological status following computerized image guided stereotactic pallidotomy in Parkinson's disease. *Neurology, 48*, A233.

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*(1), 51-66.

Brown, R. G., & Marsden, C. D. (1984). How common is dementia in Parkinson's disease? *Lancet, 8414*(2), 1262-1265.

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

Brown, R. G., & Marsden, C. D. (1990). Cognitive function in Parkinson's disease: from description to theory. *Trends in the Neurosciences, 13*(1), 21-29.

Brown, R. G., & Marsden, C. D. (1991). Dual task performance and processing resources in normal subjects and patients with Parkinson's disease. *Brain, 114*, 215-231.

Brown, V. J., & Robbins, T. W. (1991). Simple and choice reaction time performance following unilateral striatal dopamine depletion in the rat: impaired motor readiness but preserved response preparation. *Brain, 114*, 513-525.

Brown, R. G., Jahanshahi, M., & Marsden, C. D. (1993a). Response choice in Parkinson's disease: the effects of uncertainty and stimulus-response compatability. *Brain, 116*, 869-885.

Brown, V. J., Schwarz, U., Bowman, E. M., Fuhr, P., Robinson, D. L., & Hallett, M. (1993b). Dopamine dependant reaction time deficits in patients with Parkinson's disease are task specific. *Neuropsychologia, 31*(5), 459-469.

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*(3), 335-342.

Canavan, A., Passingham, R. E., Marsden, C. D., Quinn, N., Wyke, M., & Polkey, C. E. (1989). The performance on learning tasks of patients in the early stages of Parkinson's disease. *Neuropsychologia, 27*, 141-156.

Candy, J. M., Perry, R. H., Perry, E. K., Irving, D., Blessed, G., Fairbairn, A. F., & Tomlinson, B. E. (1983). Pathological changes in the nucleus of Meynert in Alzheimer's and Parkinson's disease. *Journal of the Neurological Sciences, 59*, 277-289.

Caparros-Lefebvre, D., Ruchoux, M. M., Blond, S., Petit, H., & Percheron, G. (1994). Long-term thalamic-stimulation in Parkinson's disease - postmortem anatomoclinical study. *Neurology, 44*(10), 1856-1860.

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 and Psychiatry, 58*, 51-55.

Carli, M., Evenden, J. L., & Robbins, T. W. (1985). Depletion of unilateral striatal dopamine impairs initiation of contralateral actions and not sensory attention. *Nature, 313*, 679-682.

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377*(7), 59-62.

Ceballos-Baumann, A. O., Obeso, J. A., Vitek, J. L., DeLong, M. R., Bakay, R., Linazasoro, G., & Brooks, D. J. (1994). Restoration of thalamocortical activity after posteroventral pallidotomy in Parkinson's disease. *The Lancet, 344*, 814.

Channon, S., Jones, M.-C., & Stephenson, S. (1993). Cognitive strategies and hypothesis testing during discrimination learning in Parkinson's disease. *Neuropsychologia, 31*(1), 75-82.

Chesselet, M. F., & Delfs, J. M. (1996). Basal ganglia and movement disorders: an update. *Trends in the Neurosciences, 19*(10), 417-422.

Cole, S. A., Woodard, J. L., DeLong, M. R., Vitek, J. L., Bakay, R., Barron, M., Green, J., Kaneoke, Y., Hashimoto, T., & Turner, R. (1995). Pallidotomy improves motor function in patients with Parkinson's disease, but does not improve depression. *Psychosomatic Medicine, 57*, 72.

Cools, A. R., van den Bercken, J. H. L., Horstink, M. W. I. M., van Spaendonck, K. P. M., & Berger, H. J. C. (1984). Cognitive and motor shifting aptitude disorder in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 47*, 443-453.

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

Cooper, J. A., & Sagar, H. J. (1993). Encoding deficits in untreated Parkinson's disease. *Cortex, 29*, 251-265.

Cooper, J. A., Sagar, H. J., Tidswell, P., & Jordan, N. (1994). Slowed central processing in simple and go/no-go reaction time tasks in Parkinson's disease. *Brain, 117*, 517-529.

Crawford, T., Goodrich, S., Henderson, L., & Kennard, C. (1989). Predictive responses in Parkinson's disease: manual keypresses and saccadic eye meovements to regular stimulus events. *Journal of Neurology, Neurosurgery and Psychiatry, 52*, 1033-1042.

Cronin-Golomb, A., Corkin, S., & Growdon, J. H. (1994). Impaired problem-solving in Parkinson's disease: impact of a set-shifting deficit. *Neuropsychologia, 32*(5), 579-593.

Dalrymple-Alford, J. C., Kalders, A. S., Jones, R. D., & Watson, R. W. (1994). A central executive deficit in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 57*, 360-367.

Darvesh, S., & Freedman, M. (1996). Subcortical dementia: a neurobehavioral approach. *Brain and Cognition, 31*, 230-249.

Daum, I., & Quinn, N. (1991). Reaction times and visuospatial processing in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 13*(6), 972-982.

Davis, K. D., Taub, E., Houle, S., A.E., L., Dostrovsky, J. O., Tasker, R. R., & Lozano, A. M. (1997). Globus pallidus stimulation activates the cortical motor system during alleviation of parkinsonian symptoms. *Nature Medicine, 3*(6).

De la Casa, L. G., Ruiz, G., & Lubow, R. E. (1993). Latent inhibition and recall/recognition of irrelevant stimuli as a function of preexposure duration in high and low psychotic-prone normal subjects. *British Journal of Psychology, 84*, 119-132.

Deiber, M. P., Pollack, P., Passingham, R., Landais, P., Gervason, C., Cinotti, L., Friston, K., Frackowiak, R., Mauguiere, F., & Benabid, A. L. (1993). Thalamic-stimulation and suppression of parkinsonian tremor - evidence of a cerebellar dectivation using postron emission tomography. *Brain, 116*(1), 267-279.

Della Sala, S., Muggia, S., Spinnler, H., & Zuffi, M. (1995). Cognitive modelling of face processing: evidence from Alzheimer patients. *Neuropsychologia, 33*(6), 675-687.

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *TINS, 13*(7), 281-285.

Department of Health, U. K. (1997). *The new N.H.S. - modern, dependable*. London: The Stationery Office Limited.

DeVietti, T. L., Bauste, R. L., Nutt, G., Barrett, O. V., Daly, K., & Petree, A. D. (1987). Latent inhibition: a trace conditioning phenomenon? *Learning and Motivation, 18*, 185-201.

DeZubicaray, G., & Ashton, R. (1996). Nelson (1976) modified card sorting test - a review. *Clinical Neuropsychologist, 10*(3), 245-254.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*, 69-72.

Dogali, M., Fazzini, E., Kolodny, E., Eidelberg, D., Sterio, D., Devinsky, O., & Beric, A. (1995). Stereotactic ventral pallidotomy for Parkinson's disease. *Neurology, 45*, 753-761.

Donders, F. C. (1969). On the speed of mental processes. In W. Koster (Ed.), *Attention and Performance II* (Vol. 30, pp. 412-431).

Doty, R. L., Deems, D. A., & Steller, S. (1988). Olfactory dysfunction in parkinsonism: a general deficit, unrelated to neurologic signs, disease state or disease duration. *Neurology, 38*, 1237-1244.

Downes, J. J., Roberts, A. C., Sahakian, B. J., Evenden, J. L., R.G., M., & Robbins, T. W. (1989). Impaired extra-dimensional shift performance in medicated and unmedicated

Parkinson's disease: evidence for a specific attentional dysfunction. *Neuropsychologia, 27*(11/12), 1329-1343.

Downes, J. J., Sharp, H. M., Costall, B. M., Sagar, H. J., & Howe, J. (1993). Alternating fluency in Parkinson's disease. *Brain, 116*, 887-902.

Dubois, B., & Pillon, B. (1995). Do cognitive changes of Parkinson's disease result from dopamine depletion. *Journal of Neural Transmission, 45*, 27-34.

Eidelberg, D., Moeller, J. R., Ishikawa, T., Dhawan, V., Spetsieris, P., Silbersweig, D., Stern, E., Woods, R. P., Fazzini, E., Dogali, M., & Beric, A. (1996). Regional metabolic correlates of surgical outcome following unilateral pallidotomy for Parkinson's disease. *Annals of Neurology, 39*, 450-459.

Eimas, P. D. (1966). Effects of overtraining and age in intradimensional and extradimensional shifts in children. *Journal of Experimental Child Psychology, 3*, 348-355.

Elliot, R., McKenna, P. J., Robbins, T. W., & Sahakian, B. J. (1995). Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. *Psychological Medicine, 25*, 619-630.

Evarts, E. V., Teräväinen, H., & Calne, D. B. (1981). Reaction time in Parkinson's disease. *Brain, 104*, 167-186.

Favre, J., Taha, J. M., Nguyen, T. T., Gildenberg, P. L., & Burchiel, K. J. (1996). Pallidotomy: a survey of current practice in North America. *Neurosurgery, 39*(4), 883-892.

Feldon, J., & Weiner, I. (1991). The latent inhibition model of schizophrenic attention disorder: haloperidol and sulpiride enhance rat's ability to ignore irrelevant stimuli. *Biological Psychiatry, 29*, 635-646.

Ferraro, R., Balota, D., & Connor, L. (1993). Implicit memory and the 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.

Fimm, B., Bartl, G., Zimmermann, P., & Wallesch, C.-W. (1994). Different mechanisms underly shifting set on external and internal cues in Parkinson's disease. *Brain and Cognition, 25*, 287-304.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*, 381-391.

Flowers, K. A., & Robertson, C. (1985). The effect of Parkinson's disease on the ability to maintain a mental set. *Journal of Neurology, Neurosurgery and Psychiatry, 48*, 517-529.

Flowers, K. A., & Robertson, C. (1995). Perceptual abnormalities in Parkinson's disease: top-down or bottom-up processes? *Perception, 24*, 1201-1221.

Fluck, E., File, S.E., Springett, J., Kopelman, M.D., Rees, J. and Orgill, J. (1998) Does the sedation resulting from sleep deprivation and lorazepam cause similar cognitive deficits? *Pharmacology, Biochemistry and Behavior, 59(4)*, 909-915.

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

Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry, 11*, 329-336.

Frith, C. D., & Done, D. J. (1986). Routes to action in reaction time tasks. *Psychological Research, 48*, 169-177.

Georgiou, N., Bradshaw, J. L., Phillips, J. G., & Chiu, E. (1996). The effect of Huntington's disease and Gilles de la Tourette's syndrome on the ability to hold and shift attention. *Neuropsychologia, 34*(9), 843-851.

Ghika, J., Wiegner, A. W., Fang, J. J., Davies, L., Young, R. R., & Growdon, J. H. (1993). Portable system for quantifying motor abnormalities in Parkinson's disease. *IEEE Transactions on Biomedical Engineering, 40*(3), 276-283.

Ghika, J., Favre, J., Fankhauser, H., & Regli, F. (1996). Neurological and neuropsychological complications of bilateral contemporaneous pallidotomy in Parkinson's disease. *Neurology, 46*, A417.

Gibson, J. J. (1941). A critical review of the concept of set in contemporary experimental psychology. *Psychological Bulletin, 38*(9), 781-817.

Girotti, F., Carella, F., Grassi, M. P., Soliveri, P., Marano, R., & Caraceni, T. (1986). Motor and cognitive performances of Parkinsonian patients in the on and off phases of the disease. *Journal of Neurology, Neurosurgery and Psychiatry, 49*, 657-660.

Goetz, C. G., DeLong, M. R., Penn, R. D., & Bakay, R. A. E. (1993). Neurosurgical horizons in Parkinson's disease. *Neurology, 43*, 1-7.

Goetz, C. G., Stebbins, G. T., Shale, H. M., Lang, A. E., Chernik, D. A., Chmura, T. A., Ahlskog, J. E., & Dorflinger, E. E. (1994). Utility of an objective dyskinesia rating scale for Parkinson's disease: inter- and intrarater reliability assessment. *Movement Disorders, 9*(4), 390-394.

Goodrich, S., Henderson, L., & Kennard, C. (1989). On the existence of an attention-demanding process peculiar to simple reaction time: converging evidence from Parkinson's disease. *Cognitive Neuropsychology, 6*(3), 309-331.

Goodrich, S., Henderson, L., Allchin, N., & Jeevaratnam, A. (1990). On the peculiarity of simple reaction time. *Quarterly Journal of Experimental Psychology, 42A*, 763-775.

Gordon, B., & Carson, K. (1990). The basis for choice reaction time slowing in Alzheimer's disease. *Brain and Cognition, 13*, 148-166.

Gorrell, J. M., Ordidge, R. J., Brown, G. G., Deniau, J.-C., Buderer, N. M., & Helpern, J. A. (1995). Increased iron-related MRI contrast in the substantia nigra in Parkinson's disease. *Neurology, 45*, 1138-1143.

Gotham, A. M., Brown, R. G., & Marsden, C. D. (1988). "Frontal" cognitive function in patients with Parkinson's disease "on" and "off" levodopa. *Brain, 111*, 299-321.

Grafton, S. T., Waters, C., Sutton, J., Lew, M. F., & Couldwell, W. (1995). Pallidotomy increases activity of motor association cortex in Parkinson's disease: a positron emission tomographic study. *Annals of Neurology, 37*, 776-783.

Graham, S., & McLaren, I. P. L. (1998). Retardation of human discrimination learning as a consequence of pre-exposure: latent inhibition or negative priming? *Quarterly Journal of Experimental Psychology, 51B*(2), 155-172.

Grant, D., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*, 404-411.

Graybiel, A. M. (1995). Building action repertoires: memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology, 5*, 733-741.

Green, J., Vitek, J. L., Baron, M., Bakay, R. A. E., & DeLong, M. R. (1996). Neuropsychological sequelae of pallidotomy for treatment of Parkinson's disease: one-year pilot findings. *Neurology, 46*, A200.

Greve, K. W., Brooks, J., Crouch, J. A., Williams, M. C., & Rice, W. J. (1997). Factorial structure of the Wisconsin Card Sorting Test. *British Journal of Clinical Psychology, 36*, 283-285.

Hallett, M., & Khoshbin, S. (1980). A physiological mechanism of bradykinesia. *Brain, 103*, 301-314.

Hallett, M. (1990). Clinical neurophysiology of akinesia. *Revue Neurologique, 146*(10), 585-590.

Harrington, D. L., & Haaland, K. Y. (1991). Sequencing in Parkinson's disease - abnormalities in programming and controlling movement. *Brain, 114*, 99-115.

Harrison, J., Henderson, L., & Kennard, C. (1995). Abnormal refractoriness in patients with Parkinson's disease after brief withdrawal of levodopa treatment. *Journal of Neurology, Neurosurgery and Psychiatry, 59*, 499-506.

Hartikainen, P., Helkala, E.-L., Soininen, H., & Riekinnen, 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).

Heilman, K. M., Bowers, D., Watson, R. T., & Greer, M. (1976). Reaction times in Parkinson's disease. *Archives of Neurology, 33*, 139-140.

Henik, A., Singh, J., Beckley, D. J., & Rafal, R. D. (1993). Disinhibition of automatic word reading in Parkinson's disease. *Cortex, 29*, 589-599.

Hietanen, M., & Teräväinen, H. (1986). Cognitive performance in early Parkinson's disease. *Acta Neurologica Scandinavica, 73*, 151-159.

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: onset, progression and mortality. *Neurology, 17*(427-442).

Holthoff-Detto, V. J., Kessler, J., Herholz, K., Bonner, H., Pietrzyk, U., Wurker, M., Ghaemi, M., Wienhard, K., Wagner, R., & Heiss, W.-D. (1997). Functional effects of striatal dysfunction in Parkinson's disease. *Archives of Neurology, 54*, 145-150.

Hughes, A. J., Daniel, S. E., Blankson, S., & Lees, A. J. (1993). A clinicopathologic study of 100 cases of Parkinson's disease. *Archives of Neurology, 50*, 140-148.

Iacono, R. P., Meltzer, C., & Schoonenber, T. (1994). Cognitive improvements follwing stereotactic pallidotomy. *Annals of Neurology, 36*(2), 270.

Iacono, R. P., Lonser, R. R., Ulloth, J. E., & Shima, F. (1995). Postero-ventral pallidotomy in Parkinson's disease. *Journal of Clinical Neuroscience, 2*(2), 140-145.

Jacobs, D. M., Marder, K., Cöté, L. J., Sano, M., Stern, Y., & Mayeux, R. (1995). Neuropsychological characteristics of preclinical dementia in Parkinson's disease. *Neurology, 45*(1691-1696).

Jahanshahi, M., Brown, R. G., & Marsden, C. D. (1992a). Simple and choice reaction time and the use of advance information for motor preparation in Parkinson's disease. *Brain, 115*, 539-564.

Jahanshahi, M., Brown, R. G., & Marsden, C. D. (1992b). The effect of withdrawal of dopaminergic medication on simple and choice reaction time and the use of advance information in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 55*, 1168-1176.

Jahanshahi, M., Jenkins, I. H., Brown, R. G., Marsden, C. D., Passingham, R. E., & Brooks, D. J. (1995). Self initiated versus externally triggered movements 1: An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. *Brain, 118*, 913-933.

Jones, D. L., Phillips, J. G., Bradshaw, J. L., Iansek, R., & Bradshaw, J. A. (1993). Coding of movement direction and amplitude in Parkinson's disease: are they differentially impaired (or unimportant?). *Journal of Neurology, Neurosurgery and Psychiatry, 56*, 419-422.

Jordan, N., Sagar, H. J., & Cooper, J. A. (1992). Cognitive components of reaction time in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 55*, 658-664.

Kish, J. K., Shannak, K., & Hornykiewicz, O. (1988). Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. *New England Journal of Medicine, 318*, 876-880.

Kishore, A., Turnbull, I. M., Snow, B. J., de la Fuente-Fernandez, R., Schulzer, M., Mak, E., Yardley, S., & Calne, D. B. (1997). Efficacy, stability and predictors of outcome of pallidotomy for Parkinson's disease: six-month follow-up and additional 1-year observations. *Brain, 120*, 729-737.

Klockgether, T., & Dichgans, J. (1994). Visual control of arm movement in Parkinson's disease. *Movement Disorders, 9*(1), 48-56.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399-1402.

Kramer, J. H., & Duffy, J. M. (1996). Aphasia, apraxia and agnosia in the diagnosis of dementia. *Dementia, 7*(1), 23-26.

Kulisevsky, J., Avila, A., Barbanoj, M., Antonijoan, R., Berthier, M. L., & Gironell, A. (1996). Acute effects of levodopa on neuropsychological performance in stable and fluctuating Parkinson's disease patients at different levodopa plasma levels. *Brain, 119*, 2121-2132.

Labutta, R. J., Miles, R. B., Sanes, J. N., & Hallett, M. (1994). Motor program memory storage in Parkinson's disease patients tested with a delayed response task. *Movement disorders, 9*(2), 218-222.

Laitinen, L. V., Bergenheim, A. T., & Hariz, M. I. (1992). Ventroposterolateral pallidotomy can abolish all parkinsonian symptoms. *Stereotactic and Functional Neurosurgery, 58*(14-21).

Lang, A. E., Lozano, A., Tasker, R., Duff, J., Saint-Cyr, J., & Trepanier, L. (1997). Neuropsychological and behavioral changes and weight gain after medial pallidotomy. *Annals of Neurology, 41*(6), 834-835.

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 test sensitive to frontal lobe dysfunction. *Psychopharmacology, 107*, 394-404.

Lange, K. W., Sahakian, B. J., Quinn, N. P., Marsden, C. D., & Robbins, T. W. (1995). Comparison of executive and visuospatial memory function in Huntingdon's disease and dementia of the Alzheimer type matched for degree of dementia. *Journal of Neurology, Neurosurgery and Psychiatry, 58*, 598-606.

Langston, J. W., Widner, H., Goetz, C. G., Brooks, D., Fahn, S., Freeman, T., & Watts, R. (1992). Core assessment program for intracerebral transplantations (CAPIT). *Movement Disorders, 7*(1), 2-13.

Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. *Brain, 106*, 257-270.

Levine, M. (1970). Human discrimination learning: the subset-sampling assumption. *Psychological Bulletin, 74*(4), 397-404.

Levy, R., Friedman, H. R., Davachi, L., & Goldman-Rakic, P. S. (1997a). Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *Journal of Neuroscience, 17*(10), 3870-3882.

Levy, R., Hazrati, L.-N., Herrero, M.-T., Vila, M., Hassani, O.-K., Mouroux, M., Ruberg, M., Aseni, H., Agid, Y., Feger, J., Obeso, J. A., Parent, A., & Hirsch, E. C. (1997b). Re-evaluation of the functional anatomy of the basal ganglia in normal and parkinsonian states. *Neuroscience, 76*(2), 335-343.

Lezak, M. D. (1995). *Neuropsychological Assessment.* (3 ed.).

Lichter, D. G., Corbett, A. J., Fitzgibbon, G. M., Davidson, O. R., Ayton Hope, J. K., Goddard, G. V., Sharples, K. J., & Pollock, M. (1988). Cognitive and motor dysfunction in Parkinson's disease. *Archives of Neurology, 45*, 854-860.

Limousin, P., Pollack, P., Benazzouz, A., Hoffman, D., Le Bas, J.-F., Brousolle, E., Perret, J. E., & Benabid, A.-L. (1995). Effect on parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. *Lancet, 345*, 91-95.

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: the case of verbal short term memory. *Memory and Cognition, 24*(3), 305-321.

Lopez-Villegas, D., Kulisevaky, J., Deus, J., Junque, C., Pujol, J., Guardia, E., & Grau, J. M. (1996). Neuropsychological alterations in patients with computed tomography-detected basal ganglia calcification. *Archives of Neurology, 53*, 251-256.

Lozano, A. M., Lang, A. E., Galvez-Jimenez, N., Miyasaki, J., Duff, J., Hutchinson, W. D., & Dostrovsky, J. O. (1995). Effect of GPi pallidotomy on motor function in Parkinson's disease. *Lancet, 346*, 1383-1387.

Lubow, R. E., & Gewirtz, J. C. (1995). Latent inhibition in humans: data, theory and implications for schizophrenia. *Psychological Bulletin, 117*(1), 87-103.

Lubow, R. E. (1997). Latent inhibition as a measure of learned inattention: some problems and solutions. *Behavioural Brain Research, 88*, 75-83.

Luce, R. D. (1986). *Response times: their role in inferring elementary mental organisation*. Oxford: Oxford University Press.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.

Maddox, T. W., Filoteo, J. V., Delis, D. C., & Slamon, D. P. (1996). Visual selective attention deficits in patients with Parkinson's disease: a quantitative model-based approach. *Neuropsychology, 10*, 197-218.

Mahieux, F., Fenelon, G., Flahault, A., Manifacier, M. J., Michelet, D., & Boller, F. (1998). Neuropsychological prediction of dementia in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 64*(2), 178-183.

Malapani, C., Pillon, B., Dubois, B., & Agid, Y. (1994). Impaired simultaneous cognitive task performance in Parkinson's disease: a dopamine-related dysfunction. *Neurology, 44*, 319-326.

Manning, C. A., Bennett, J. P., Wikniss, S. M., Jones, M. G., & Laws, E. R. (1997). Comprehensive neuropsychological assessment of neuropsychological functioning pre- and post-unilateral posteroventrolateral pallidotomy. *Neurology, 48*, A252.

Marsden, C. D. (1990). Neurophysiology. In G. M. Stern (Ed.), *Parkinson's disease* . London: Chapman & Hall Medical.

Marsden, C. D. (1992). Dopamine and basal ganglia disorders in humans. *Seminars in the Neurosciences, 4*, 171-178.

Martin, J. P. (1967). *The Basal Ganglia and Posture*. London: Pitman Medical.

Masterman, D. L., Bronstein, J., DeSalles, A. F., Foti, D. F., & Cummings, J. L. (1997). Cognitive and behavioral performance follwing unilateral ventroposterior pallidotomy. *Neurology, 48*, A252.

Mayeux, R., Stern, Y., Rosen, J., & Levethal, J. (1981). Depression, intellectual impairment and Parkinson's diseasee. *Neurology, 31*, 645-650.

Mayeux, R., Stern, Y., Sano, M., Cote, L., & Williams, J. B. W. (1987). Clinical and biochemical correlates of bradyphrenia in Parkinson's disease. *Neurology, 37*, 1130-1134.

McDonald, C., Brown, G. G., & Gorell, J. M. (1996). Impaired set-shifting in Parkinson's disease: new evidence from a lexical decision task. *Journal of Clinical and Experimental Neuropsychology, 18*(6), 793-809.

McFie, J. (1960). Psychological effects of stereotaxic operations for the relief of parkinsonian symptoms. *Journal of Mental Science, 106*, 1512-1517.

Meco, G., Gasparini, M., & Doricchi, F. (1996). Attentional functions in multiple systems atrophy and Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 60*(4), 393-398.

Miller, E. (1970). Simple and choice reaction time following severe head injury. *Cortex, 6*, 121-127.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9*, 90-100.

Mink, J. W., & Thach, W. T. (1993). Basal ganglia intrinsic circuits and their role in behavior. *Current Opinion in Neurobiology, 3*, 950-957.

Montgomery, E. B., & Nuessen, J. (1990). The movement speed/accuracy operator in Parkinson's disease. *Neurology, 40*, 269-272.

Morrish, P. K., Sawle, G. V., & Brooks, D. J. (1995). Clinical and [^{18}F]dopa PET findings in early Parksinon's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 59*, 597-600.

Mutch, W. J., Dingwall-Fordyce, I., Downie, A. W., Paterson, J. G., & Roy, S. K. (1986). Parkinson's disease in a Scottish city. *British Medical Journal, 292*, 534-536.

Nakashima, K., Shimoyama, R., & Takahashi, K. (1993). Paired choice reaction tasks in patients with Parkinson's disease. *Acta Neurologica Scandinavica, 87*, 178-183.

Narabayashi, H., Miyashita, N., Hattori, Y., Saito, K., & Endo, K. (1997). Posteroventral pallidotomy: its effect on motor symptoms and scores of MMPI test in patients with Parkinson's disease. *Parkinsonism and related disorders, 3*(1), 7-20.

National Parkinson Foundation, I. (1996). What the patient should know . Available: http://www.parkinson.org/pdedu.html.

Näätänen, R., & Merisalo, A. (1977). Expectancy and preparation in simple reaction time. In S. Dornic (Ed.), *Attention and Performance IV* (pp. 115-138). Hillsdale, N.J.: Lawrence Erlbaum.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal deficits. *Cortex, 12*, 313-324.

Nelson, H. E. (1982). *National Adult Reading Test (NART): Test Manual*. Windsor: NFER-Nelson.

Nissen, M. J., & Bullermer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology, 19*, 1-32.

Nissenbaum, H., Quinn, N. P., Brown, R. G., Toone, B. G., Gotham, A.-M., & Marsden, C. D. (1987). Mood swings associated with the 'on-off' phenomenon in Parkinson's disease. *Psychological Medicine, 17*, 899-904.

Norman, D. A., & Shallice, T. (1980). *Attention to Action: Willed and Automatic Control of Behavior* (CHIP report 99): University of California.

Norusis, M. J. (1994). SPSS for Windows (Version 6.1). Chicago: SPSS Inc.

Owen, A. M., Downes, J. D., Sahakian, B. J., Polkey, C. E., & T.W., R. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28*, 1021-1034.

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia, 29*(10), 993-1006.

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., Beksinska, M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Sahakian, B. J., & Robbins, T. W. (1993a). Visuospatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia, 31*(7), 627-644.

Owen, A. M., Roberts, A. C., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1993b). Contrasting mechanisms of impaired attentional set-shifting in patients wiht frontal lobe damage or Parkinson's disease. *Brain, 116*, 1159-1175.

Owen, A. M., Sahakian, B. J., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1995). Dopamine-dependent frontostriatial planning deficits in early Parkinson's disease. *Neuropsychology, 9*, 126-140.

Owen, A. M., Doyon, J., Petrides, M., & Evans, A. C. (1996a). Planning and spatial working memory: a positron emission tomography study in humans. *European Journal of Neuroscience, 8*, 353-364.

Owen, A. M., Evans, A. C., & Petrides, M. (1996b). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. *Cerabral Cortex, 6*, 31-38.

Owen, A. M., Morris, R. G., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1996c). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excision, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain, 119*, 1597-1615.

Owen, A. M., Iddon, J. L., Hodges, J. R., Summers, B. A., & Robbins, T. W. (1997). Spatial and non-spatial working memory at different stages of Parkinson's disease. *Neuropsychologia, 35*(4), 519-532.

Owen, A. M. (1997). The functional organisation of working memory processes within human lateral frontal cortex: the contribution of functional neuroimaging. *European Journal of Neuroscience, 9*, 1329-1339.

Owen, A. M., Doyon, J., Dagher, A., Sadikot, A., & Evans, A. C. (1998). Abnormal basal ganglia outflow in Parkinson's disease identified with PET: implications for higher cortical functions. *Brain, 121*, 949-965.

Paddinson, R. M., & Griffith, R. P. (1974). Occurrence of Parkinson's disease in black patients at Charity Hospital in New Orleans. *Neurology, 24*, 688-691.

Palazzini, E., Soliveri, P., Fillipini, G., Fetoni, V., Zappacosta, B., Scigliano, G., Monza, D., Caraceni, T., & Girotti, F. (1995). Progression of motor and cognitive impairment in Parkinson's disease. *Journal of Neurology, 242*(8), 535-540.

Paolo, A. M., Troster, A. I., Axelrod, B. N., & Koller, W. C. (1995). Construct-validity of the WCST in normal elderly and persons with Parkinson's disease. *Archives of Clinical Neuropsychology, 10*(5), 463-473.

Parent, A. (1990). Extrinsic connections of the basal ganglia. *Trends in the Neurosciences, 13*(7), 254-258.

Parkinson, J. (1817/1997). *An essay on the shaking palsy*. London: Emap healthcare ltd.

Parkinson's Disease Society, U. K. (1998). Fact Sheet One - all about the Parkinson's Disease Society . Available: http://glaxocentre.merseyside.org/parks.html.

Partiot, A., Verin, M., Pillon, B., Teixeira-Ferreira, C., Agid, Y., & Dubois, B. (1996). Delayed response tasks in basal ganglia lesions in man: further evidence for a fronto-striatal cooperation in behavioural adaptation. *Neuropsychologia, 34*(7), 709-721.

Pascual-Leone, A., Grafman, J., Clark, K., Stewart, M., Massaquoi, S., Lou, J.-S., & Hallett, M. (1993). Procedural learning in Parkinson's disease and cerebellar degeneration. *Annals of Neurology, 34*, 594-602.

Pastor, M. A., Artieda, J., Jahanshahi, M., & Obeso, J. A. (1992). Time estimation and reproduction is abnormal in Parkinson's disease. *Brain, 115*, 211-225.

Pate, D. S., & Margolin, D. I. (1994). Cognitive slowing in Parkinson's and Alzheimer's patients distinguishing bradyphrenia from dementia. *Neurology, 44*(4), 669-674.

Pearce, J. M. S. (1992). *Parkinson's Disease and its Management*. Oxford: Oxford University Press.

Petrides, M. (1996). Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical Transactions of the Royal Society: Series B, 351*, 1455-1462.

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*(1), 77-85.

Pillon, B., Ertle, S., Deweer, B., Bonnet, A.-M., Vidailhet, M., & Dubios, B. (1997). Memory for spatial location in 'de novo' parkinsonian patients. *Neuropsychologia, 35*(3), 221-228.

Pillon, B., Deweer, B., Vidailhet, M., Bonnet, A.-M., Hahn-Barma, V., & Dubois, B. (1998). Is impaired memory for spatial location in Parkinson's disease domain specific or dependent on 'strategic' processes? *Neuropsychologia, 36*(1), 1-9.

Playford, E. D., Jenkins, I. H., Passingham, R. E., Nutt, J., Frackowiak, R. S. J., & Brooks, D. J. (1992). Impaired mesial frontal and putamen activation in Parkinson's disease: a positron emission tomography study. *Annals of Neurology, 32*(2), 151-161.

Posner, M. I. (1978). *Chronometric Explorations of Mind: the Third Paul M. Fitts Lecture*. Hillsdale, New Jersey: Erlbaum Associates.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Pullman, S. L., Watts, R. L., Juncos, J. L., Chase, T. N., & Sanes, J. N. (1988). Dopminergic effects on simple and choice reaction time performance in Parkinson's disease. *Neurology, 38*, 249-254.

Pullman, S. L., Watts, R. L., Juncos, J. L., & Sanes, J. N. (1990). Movement amplitude choice reaction time performance in Parkinson's disease may be independant of dopaminergic status. *Journal of Neurology, Neurosurgery and Psychiatry, 53*, 279-283.

Quinn, N. P. (1995). Drug treatment of Parkinson's disease. *British Medical Journal, 310*, 575-579.

Rafal, R. D., Posner, M. I., Walker, J. A., & Friedrich, F. J. (1984). Cognition and the basal ganglia: separating mental and motor components of performance in Parkinson's disease. *Brain, 107*, 1083-1094.

Rafal, R. D., Inhoff, A. W., Friedman, J. H., & Bernstein, E. (1987). Programming and execution of sequential movements in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 50*, 1267-1273.

Reber, A. S. (1985). *The Penguin Dictionary of Psychology*. London: Penguin Books.

Reid, W. G. J., Broe, G. A., Hely, M. A., Morris, J. G. L., Williamson, P. M., O'Sullivan, D. J., Rail, D., Genge, S., & Moss, N. G. (1989). The neuropsychology of de novo patients with Parkinson's disease: the effects of age of onset. *International Journal of Neuroscience, 48*, 205-217.

Reid, W. G. J., Helu, M. A., Morris, J. G. L., Broe, G. A., Adena, M., Sullivan, D. J. O., & Willaimson, P. M. (1996). A longitudinal study of Parkinson's disease - clinical and neuropsychological correlates of dementia. *Journal of Clinical Neuroscience, 3*(4), 327-333.

Revonsuo, A., Portin, R., Koivikko, L., Rinne, J. O., & Rinne, U. K. (1993). Slowing of information processing in Parkinson's disease. *Brain and Cognition, 21*, 87-110.

Richards, M., Cote, L. J., & Stern, Y. (1993). Executive functions in Parkinson's disease - set-shifting or set-maintenance. *Journal of Clinical and Experimental Neuropsychology, 15*(2), 266-279.

Riklan, M., Diller, L., Weiner, H., & Cooper, I. S. (1960). Psychological studies of the effect of chemosurgery of the basal ganglia in parkinsonism I: Intellectual functioning. *Archives of General Psychiatry, 2*, 32-42.

Robbins, T. W., & Brown, V. J. (1990). The role of the striatum in the mental chronometry of action: a theoretical review. *Reviews in the Neurosciences, 2*(4), 181-213.

Robbins, T. W., James, M., Lange, K. W., Owen, A. M., Quinn, N. P., & Marsden, C. D. (1992). Cognitive performance in multiple system atrophy. *Brain, 115*, 271-291.

Robbins, T. W., James, J., 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 systems atrophy in tests sensitive to frontal lobe dysfunction. *Journal of Neurology, Neurosurgery and Psychiatry, 57*, 79-88.

Robbins, T. W. (1996). Dissociating executive functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society - Series B, 351*, 1463-1471.

Roberts, A..C., Robbins T.W., & Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Quarterly Journal of Experimental Psychology, 40B*(4), 321-341.

Roberts, A. C., DeSalvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., & Robbins, T. W. (1994). 6-hydroxydopamine lesions of the prefrontal cortex in monkeys

enhance performance on an analog of the Wisconsin Card Sort test: possible interactions with subcortical dopamine. *Journal of Neuroscience, 14*(5), 2531-2544.

Robertson, C., & Flowers, K. A. (1990). Motor set in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 53*, 583-592.

Rosenbaum, D. A. (1980). Human movement initiation: specification of arm, direction and extent. *Journal of Experimental Psychology: General, 109*, 444-474.

Samuel, M., Ceballos-Baumann, A. O., Turjanski, N., Boecker, H., Gorospe, A., Linazasoro, G., Holmes, A. P., DeLong, M. R., Vitek, J. L., Thomas, D. G. T., Quinn, N. P., Obeso, J. A., & Brooks, D. J. (1997). Pallidotomy in Parkinson's disease increases supplementary motor area and prefrontal activation during performance of volitional movements: An H215o PET study. *Brain, 120*, 1301-1313.

Samuel, M., Caputo, E., Brooks, D. J., Schrag, A., Scaravilli, T., Branston, N. M., Rothwell, J. C., Marsden, C. D., Thomas, D. G. T., Lees, A. J., & Quinn, N. P. (1998). A study of medial pallidotomy for Parkinson's disease: clinical outcome, MRI location and complications. *Brain, 121*, 59-75.

Sawle, G. V. (1995). Living neurochemistry: the dopaminergic system. *Seminars in the Neurosciences, 7*, 173-177.

Schnider, A., Gutbrod, K., & Hess, C. W. (1995). Motion imagery in Parkinson's disease. *Brain, 118*(2), 485-493.

Schugens, M. M., Daum, I., Richter, S., Scholz, E., & Canavan, A. G. M. (1993). Proximal and distal reaction times are not differentially affected in Parkinson's disease. *Movement Disorders, 8*(3), 367-370.

Schwab, R. S. (1972). Akinesia paradoxica. *Electroencephalography and Clinical Neurphysiology, Supplement 31*, 87-92.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, B 298*, 199-209.

Sharpe, M. H. (1990). Distractibility in early Parkinson's disease. *Cortex, 26*, 239-246.

Sheridan, M. R., Flowers, K. A., & Hurrell, J. (1987). Programming and execution of movement in Parkinson's disease. *Brain, 110*, 1247-1271.

Slamecka, N. J. (1968). A methodological analysis of shift paradigms in human discrimination learning. *Psychological Bulletin, 69*, 423-438.

Soukup, V. M., Ingram, F., Schiess, M. C., Bonnen, J. G., Nauta, H. J. W., & Calverley, J. R. (1997). Cognitive sequelae of unilateral pallidotomy. *Archives of Neurology, 54*, 947-950.

Sprengelmeyer, R., Canavan, A. G. M., Lange, H. W., & Homberg, V. (1995). Associative learning in degenerative neostriatial disorders: contrasts in explicit and implicit remembering between Parkinson's and Huntingdon's diseases. *Movement Disorders, 10*(1), 51-65.

Stam, C. J., Visser, S. L., Op de Coul, A. A. W., De Sonneville, M. J., Schellens, R. L. L. A., Brunia, C. H. M., de Smet, J. S., & Gielen, G. (1993). Disturbed frontal regulation of attention in Parkinson's disease. *Brain, 116*, 1139-1158.

Starkstein, S. E., Esteguy, E., Berthier, M. L., Garcia, H., & Leiguarda, R. (1989a). Evoked potentials, reaction time, and cognitive performance in on and off phases of Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 52*, 338-340.

Starkstein, S. E., Preziosi, T. J., Berthier, M. L., Bolduc, P. L., Mayberg, H. S., & Robinson, R. G. (1989b). Depression and cognitive impairment in Parkinson's disease. *Brain, 112*, 1141-1153.

Starkstein, S. E., Mayberg, H. S., Leiguarda, R., Preziosi, T. S., & Robinson, R. G. (1992). A prospective, longitudinal study of depression, cognitive decline and physical impairments in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 55*, 377-382.

Stebbins, G. T., Gabrieli, J. D. E., Goetz, C. G., Shannon, K. M., Penn, R. D., & Masciari, F. (1997). Impaired fronto-striatal memory functioning following pallidotomy in advanced Parkinson's disease. *Neurology, 48*, A252.

Stelmach, G. E., Worringham, C. J., & Strand, E. A. (1986). Movement preparation in Parkinson's disease - the use of advance information. *Brain, 109*, 1179-1194.

Stern, Y., Mayeux, R., & Côté, L. (1984). Reaction time and vigilance in Parkinson's disease. *Archives of Neurology, 41*, 1086-1089.

Talland, G. A. (1962). Cognitive functions in Parkinson's disease. *Journal of Nervous and Mental Disease, 135*, 196-205.

Talland, G. A. (1963). Manual skill in Parkinson's disease. *Geriatrics, 18*, 613-620.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1986). Frontal lobe dysfunction in Parkinson's disease: the cortical focus of neostriatal outflow. *Brain, 109*, 845-883.

Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1987). Parkinson's disease: cognitive changes in relation to treatment response. *Brain, 110*, 35-51.

Taylor, A. E., & Saint-Cyr, J. A. (1995). The neuropsychology of Parkinson's disease. *Brain and Cognition, 28*, 281-296.

Thompson, P. D., Dick, J. P. R., Day, B. L., Rothwell, J. C., Berardelli, A., Kachi, T., & Marsden, C. D. (1986). Electrophysiology of the corticomotoneurone pathways in patients with movement disorders. *Movement Disorders, 1*, 14-23.

Thornton, J. C., Dawe, S., Lee, C., Capstick, C., Corr, P. J., Cotter, P., Frangou, S., Gray, N. S., Russell, M. A. H., & Gray, J. A. (1996). Effects of nicotine and amphetamine on latent inhibition in human subjects. *Psychopharmacology, 127*, 164-173.

Tröster, A. I., Stalp, L. D., Paolo, A. M., Fields, J. A., & Koller, W. C. (1995). Neuropsychological impairment in Parkinson's disease with and without depression. *Archives of Neurology, 52*, 1164-1169.

Tröster, A. I., Fields, J. A., Wilkinson, S. B., Pahwa, R., Miyawaki, E., Lyons, K. E., & Koller, W. C. (1997). Unilateral pallidal stimulation for Parkinson's disease: neurobehavioral functioning before and three months after electrode implantation. *Neurology, 49*, 1078-1083.

Uitti, R. J., Wharen, R. E., Turk, M. F., Lucas, J. A., Finton, M. J., Graff-Radford, N. R., Boylan, K. B., Goerss, S. J., Kall, B. A., Adler, C. H., Caviness, J. N., & Atkinson, E. J. (1997). Unilateral pallidotomy for Parkinson's disease: comparison of outcome in younger versus elderly patients. *Neurology, 49*, 1072-1077.

Van Spaendonck, K. P. M., Berger, H. J. C., Horstink, M. W. I., Buytenhuijs, E. L., & Cools, A. L. (1993). Impaired cognitive shifting in parkinsonian patients on anticholinergic therapy. *Neuropsychologia, 31*(4), 407-411.

Van Spaendonck, K. P. M., Berger, H. J. C., Horstink, M. W. I. M., Borm, G. F., & Cools, A. R. (1995). Card sorting performance in Parkinson's disease: a comparison between acquisition and shifting performance. *Journal of Clinical and Experimental Neuropsychology, 17*(6), 918-925.

Van Spaendonck, K. P. M., Berger, H. J. C., Horstink, M. W. I. M., Buytenhuis, E. L., & Cools, A. R. (1996). Executive functions and disease characteristics in Parkinson's disease. *Neuropsychologia, 34*(7), 617-626.

Velasco, F., & Velasco, M. (1973). A quantitative evaluation of the effects of L-Dopa on Parkinson's disease. *Neuropharmacology, 12*, 89-99.

von Falkenhayn, I., Ceballos-Baumann, A. O., Moringlane, J. R., Alesch, F., Pinter, M., Barcia, J. L., Tormos, J. M., & Pascual-Leone, A. (1997). Effects of subthalamic nucleus (STN) and globus pallidus internus (GPI) stimulation on procedural learning. *Neurology, 48*, A118.

Webster, D. D. (1968). Critical analysis of the disability in Parkinson's disease. *Mod Treat, 5*, 257-282.

Whitney, L., & White, K. G. (1993). Dimensional shift and the transfer of attention. *Quarterly Journal of Experimental Psychology, 46B*, 225-252.

Williams, J. H., Wellman, N. A., Rawlins, J. N. P., Geaney, D. P., Cowen, P. J., & Feldon, J. (1994). Haloperidol increases latent inhibition in high schizotypal subjects. *Schizophrenia Research, 11*(2), 162.

Willingham, D. B., Koroshetz, W. J., Treadwell, J. R., & Bennett, J. P. (1995). Comparison of Huntington's and Parkinson's disease patients use of advance information. *Neuropsychology, 9*(1), 39-46.

Wolff, J. L. (1967). Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin, 68*(6), 369-408.

Worringham, C. J., & Stelmach, G. E. (1990). Practice effects on the preprogramming of discrete movements in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 53*, 702-704.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research, 17*, 37-49.

Yokochi, F., Nakamura, R., & Narabayashi, H. (1985). Reaction time of patients with Parkinson's disease, with reference to asymmetry of neurological signs. *Journal of Neurology, Neurosurgery and Psychiatry, 48*, 702-705.

Zappia, M., Montesanti, R., Colao, R., & Quattrone, A. (1994). Usefulness of movement time in the assessment of Parkinson's disease. *Journal of Neurology, 241*, 543-550.

Zimmermann, P., Sprengelmeyer, R., Fimm, B., & Wallesch, C.-W. (1992). Cognitive slowing in decision tasks in early and advanced Parkinson's disease. *Brain and Cognition, 18*, 60-69.

Appendix 1: Program used for dimensional shifting task in Study 1

LI.BAS
C:\USR\WINDOWS\SYSTEM\GRID.VBX
C:\USR\WINDOWS\SYSTEM\MSOLE2.VBX
C:\USR\WINDOWS\SYSTEM\ANIBUTON.VBX
C:\USR\WINDOWS\SYSTEM\CMDIALOG.VBX
C:\USR\WINDOWS\SYSTEM\CRYSTAL.VBX
C:\USR\WINDOWS\SYSTEM\GAUGE.VBX
C:\USR\WINDOWS\SYSTEM\GRAPH.VBX
C:\USR\WINDOWS\SYSTEM\KEYSTAT.VBX
C:\USR\WINDOWS\SYSTEM\MSCOMM.VBX
C:\USR\WINDOWS\SYSTEM\MSMASKED.VBX
C:\USR\WINDOWS\SYSTEM\MSOUTLIN.VBX
C:\USR\WINDOWS\SYSTEM\PICCLIP.VBX
C:\USR\WINDOWS\SYSTEM\SPIN.VBX
C:\USR\WINDOWS\SYSTEM\THREED.VBX
LI1.FRM
LI2.FRM
LI3.FRM
HOWDONE.FRM
END.FRM
ProjWinSize=80,387,248,183
ProjWinShow=2
IconForm="Form1"
Title="LI"
ExeName="LI.EXE"

LI.BAS:

```
Option Explicit
   Declare Function timegettime Lib "mmsystem" () As Long
   Global nType As Integer
   Global sName As String
   Global nStage As Integer
   Global nDelay As Integer
   Global nOnetime As Long
   Global nTwotime As Long
   Global nThrtime As Long
   Global nTrials As Integer
   Global sSide As String
   Global nCorrect As Integer
   Global nBlock As Integer
   Dim nRt As Long
   Dim nMt As Long
   Global nStagecount As Integer
   Dim nSet As Integer
   Global nPiccy As Integer
   Dim sWrong As String
   Dim nLoop As Integer
   Dim nCrit(100) As Integer
   Dim sRight As String
   Global nExtra As Integer
   Global nSlide As Integer
   Global nForm As Integer

:
Sub ch1 ()
   ' colour spots
   ' colour (green) correct
   Randomize
   nPiccy = Rnd(1)
   nSet = Rnd(1)
   If nSet = 1 Then
      sWrong = "yelstri.bmp"
      sRight = "bluspo.bmp"
   Else
      sWrong = "yelspo.bmp"
      sRight = "blustri.bmp"
   End If
   If nPiccy = 1 Then
      form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
      form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
```

```
    Else
        form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
        form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
    End If
End Sub
:
Sub ch2 ()
    ' number and spots
    ' number (ONE) correct
    Randomize
    nPiccy = Rnd(1)
    nSet = Rnd(1)

    If nSet = 1 Then
        sWrong = "onesta.bmp"
        sRight = "twohash.bmp"
    Else
        sWrong = "onehash.bmp"
        sRight = "twosta.bmp"
    End If
    If nPiccy = 1 Then
        form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
        form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
    Else
        form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
        form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
    End If
End Sub
:
Sub charm ()

    ' here check the stage
    ' call the appropriate piccyloading sub
    Select Case nStage
        Case 1
            Call novel
        Case 2
            Call ids
        Case 3
            Call ch1
        Case 4
            Call ch2
        Case 5
            Call theend
    End Select
End Sub
:
Sub ctr (frmForm As Form)
    frmForm.Left = 0
    frmForm.Top = 0
    frmForm.Width = screen.Width
    frmForm.Height = screen.Height
    ' centring form1 - any form, actually.
End Sub
:
Sub feedback ()
    form2!Picture1.Visible = False
    form2!Picture2.Visible = False
    ' find out whether it's right or not
    ' write to file
    ' nWait = nOnetime - nUno
    nRt = nTwotime - nOnetime
    nMt = nThrtime - nTwotime
    Write #1, nTrials, nCorrect, nDelay, sSide, nRt, nMt' , nWait
    Unload form2
    Load form2
    ' display the appropriate piccy
    ' then probably check sliding criterion
    nCrit(nStagecount) = nCorrect
    If nStagecount = 30 Then
        Call slider1
    End If
```

```
  If nStagecount = (30 + nExtra) Then
    Call slider2
  End If
  If nStagecount = 30 + (nExtra * 2) Then
    Call slider3
  End If
End Sub
:
Sub ids ()
  ' as novel, but changed exemplars
  ' standard grcir rdsqr stuff
  Randomize
  nPiccy = Rnd(1)
  nSet = Rnd(1)
  If nSet = 1 Then
    sWrong = "rdcir.bmp"
    sRight = "grsqr.bmp"
  Else
    sWrong = "grcir.bmp"
    sRight = "rdsqr.bmp"
  End If
  If nPiccy = 1 Then
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
  Else
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
  End If
End Sub
:
Sub main ()
  nTrials = 0
  nStagecount = 0
  nExtra = Int((12 - 8 + 1) * Rnd(1) + 8)
  Load form3
  Load form1
End Sub
:
Sub novel ()
  ' intial discrimination
  Randomize
  nPiccy = Rnd(1)
  nSet = Rnd(1)
  If nSet = 1 Then
    sWrong = "blathi.bmp"
    sRight = "pinktri.bmp"
  Else
    sWrong = "pinkthi.bmp"
    sRight = "blatri.bmp"
  End If
  If nPiccy = 1 Then
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
  Else
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
  End If
End Sub
:
Sub premreload ()
  If nForm = 2 Then
    Unload form2
    Load form2
  Else
    Load form2
  End If
End Sub
:
Sub reload ()
  Unload form2
  Load form2
  ' yep, it's there all right
```

```
End Sub
:
Sub shift ()
   ' i'm putting this is as it is quite clear that i need to
   ' do lots when shifting. here it is;
    nExtra = Int((12 - 8 + 1) * Rnd(1) + 8) ' defines post-crit interval
   ' clear the array for the criteria - don't know if it's necessary
   For nLoop = 1 To 70
      nCrit(nLoop) = 0
   Next nLoop
   ' set the counter for this array to 1.
   nStagecount = 1
   ' ? put something in the results file about having shifted
   Write #1, "stage shift from", nStage, "to", nStage + 1
   ' advance nStage by 1
   nStage = nStage + 1
   If nStage = 9 Then
      Call theend
   End If
   nSlide = 0
   ' that's it, really. still the same response form, but just different
   ' stimuli coming up.
End Sub
:
Sub slider1 ()
   ' here goes with the sliding criterion
   ' 8 out of 10 (this week.)
   nSlide = 0
   ' look at last 10 responses, add up number correct
   For nLoop = (nStagecount - 9) To nStagecount
      nSlide = nSlide + nCrit(nLoop)
   Next nLoop
   Load form4
End Sub
:
Sub slider2 ()
  nSlide = 0

   ' look at last 10 responses, add up number correct
   For nLoop = (nStagecount - 9) To nStagecount
      nSlide = nSlide + nCrit(nLoop)
   Next nLoop
   If nSlide >= 8 Then
      Call shift
   Else
      Load form4
   End If
End Sub
:
Sub slider3 ()
   nSlide = 0
   ' look at last 10 responses, add up number correct
   For nLoop = (nStagecount - 9) To nStagecount
      nSlide = nSlide + nCrit(nLoop)
   Next nLoop
   If nSlide >= 8 Then
      Call shift
   Else
      Call theend
   End If
End Sub
:
Sub str1 ()
   ' spots colour
   ' spots correct
   Randomize
   nPiccy = Rnd(1)
   nSet = Rnd(1)
   If nSet = 1 Then
      sWrong = "yelstri.bmp"
      sRight = "bluspo.bmp"
   Else
```

```
    sWrong = "blustri.bmp"
    sRight = "yelspo.bmp"
  End If
  If nPiccy = 1 Then
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
  Else
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
  End If
End Sub
:
Sub str2 ()
  ' colour number
  ' green correct
  Randomize
  nPiccy = Rnd(1)
  nSet = Rnd(1)
  If nSet = 1 Then
    sWrong = "onebr.bmp"
    sRight = "twowhi.bmp"
  Else
    sWrong = "twobr.bmp"
    sRight = "onewhi.bmp"
  End If
  If nPiccy = 1 Then
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
  Else
    form2.Picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sWrong)
    form2.Picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sRight)
  End If
End Sub
:
Sub strict ()
  ' here check the stage
  ' call the appropriate piccyloading sub
  Select Case nStage
    Case 1
      Call novel
    Case 2
      Call ids
    Case 3
      Call str1
    Case 4
      Call str2
    Case 5
      Call theend
  End Select
End Sub
:
Sub theend ()
  MsgBox "Well done! You have completed this stage of the test."
  Close
  End
End Sub
:
```

LI1.FRM

```
Option Explicit

Sub Command1_Click ()
   nType = 1
  If text1.Text = "" Then
    MsgBox "You must enter some initials!"
  Else
    sName = text1.Text & "ls"
    Open "c:\usr\d\vbasic\results\" & sName For Append As #1
    Write #1, sName, "strict version", "block #", nBlock, Date
    Unload form1
    Set form1 = Nothing
    Load form2
```

```
    End If
End Sub
:
Sub Command2_Click ()
    nType = 2
    If text1.Text = "" Then
        MsgBox "You must enter some initials!"
        text1.SetFocus
    Else
        sName = text1.Text & "lc"
        Open "c:\usr\d\vbasic\results\" & sName For Append As #1
        Write #1, sName, "charm version", "block #", nBlock, Date
        Unload form1
        Set form1 = Nothing
        Load form2
    End If
End Sub
:
Sub Form_Load ()
    ctr form1
    form1.Show
End Sub
:
Sub Option1_Click ()
    nStage = 1
    nBlock = 1
End Sub
:
Sub Option2_Click ()
    nBlock = 2
    nStage = 2
End Sub
:
Sub Option3_Click ()
    nBlock = 3
    nStage = 2
End Sub
:
Sub Option4_Click ()
    nBlock = 4
    nStage = 2
End Sub
:
```

LI2.FRM

```
Option Explicit
    Dim nZeit As Integer
    Dim nWaiting As Integer
    Dim nKey As Integer
    Dim nLoop As Integer
    Dim nClick As Integer
:
Sub Form_DblClick ()
    ' End
End Sub
:
Sub Form_KeyDown (KeyCode As Integer, Shift As Integer)
    timer1.Enabled = True
    form2.Visible = False
    nWaiting = 1
End Sub
:
Sub Form_KeyUp (KeyCode As Integer, Shift As Integer)
    If nWaiting = 1 Then
        Write #1, "prem response"
        For nLoop = 1 To 50
            Beep
        Next
        MsgBox "Whoops! You released the key too early"
        Call reload
    End If
End Sub
```

```
:
Sub Form_Load ()
   nKey = 0
   nClick = 0
   If nTrials <> 0 Then
      If nCorrect = 1 Then
         form2.Picture3.Picture = LoadPicture("c:\usr\d\vbasic\graphics\yep.bmp")
      Else
         form2.Picture3.Picture = LoadPicture("c:\usr\d\vbasic\graphics\nope.bmp")
      End If
   End If
   ctr form2
   form2.Show
   form2!Picture3.Visible = True
   nWaiting = 0
End Sub
:
Sub Picture1_KeyUp (KeyCode As Integer, Shift As Integer)
   nTwotime = timegettime()
   picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\blank.bmp")
   picture2.Picture = LoadPicture("c:\usr\d\vbasic\graphics\blank.bmp")
End Sub
:
Sub Picture1_MouseDown (Button As Integer, Shift As Integer, X As Single, Y As Single)
   nThrtime = timegettime()
   nTrials = nTrials + 1
   nStagecount = nStagecount + 1
   sSide = "R"
   If nPiccy = 1 Then
      nCorrect = 1
   Else
      nCorrect = 0
   End If
 Call feedback
End Sub
:
Sub Picture2_MouseDown (Button As Integer, Shift As Integer, X As Single, Y As Single)
   nThrtime = timegettime()
   nTrials = nTrials + 1
   nStagecount = nStagecount + 1
   sSide = "L"
   If nPiccy = 1 Then
      nCorrect = 0
   Else
      nCorrect = 1
   End If
   Call feedback
End Sub
:
Sub Picture3_DblClick ()
   ' End
   nClick = nClick + 1
   If nClick = 2 Then
      Close
      End
   End If
End Sub
:
Sub Picture3_KeyDown (KeyCode As Integer, Shift As Integer)
   If nKey = 0 Then
      Picture3.Visible = False
      nWaiting = 1
      If nType = 1 Then
         Call strict
      Else
         Call charm
      End If
      Randomize
      timer1.Enabled = False
      nDelay = Int((4 - 1 + 1) * Rnd(1) + 1)
      Select Case nDelay
         Case 1
```

```
            nZeit = 100
          Case 2
            nZeit = 300
          Case 3
            nZeit = 500
          Case 4
            nZeit = 700
              ' these are the timer settings that are required to get the correct delays (as in text)
      End Select
      timer1.Interval = nZeit
      timer1.Enabled = True
      nKey = 1
    End If
End Sub
:
Sub Timer1_Timer ()
    timer1.Enabled = False
    picture1.Visible = True
    picture2.Visible = True
    form2.Visible = True
    nOnetime = timegettime()
    nWaiting = 0
End Sub
:
```

LI3.FRM

```
Option Explicit
Sub Form_KeyUp (KeyCode As Integer, Shift As Integer)
    MsgBox "Whoops! You released the key too early"
    Load form2
End Sub
:
Sub Form_Load ()
    ctr form3
    form3.Show
End Sub
:
```

HOWDONE.FRM

```
Option Explicit
    Dim sPiccy As String
:
Sub Command1_MouseDown (Button As Integer, Shift As Integer, X As Single, Y As Single)
    Unload form4
End Sub
:
Sub doneright ()
    ' find the stage, infer the exemplar
    If nType = 1 Then ' strict
      Select Case nStage
        Case 1
          sPiccy = "corrtri.bmp"
        Case 2
          sPiccy = "corrsq.bmp"
        Case 3
          sPiccy = "corrspo.bmp"
        Case 4
          sPiccy = "corrwhi.bmp"
      End Select
    Else ' charm
      Select Case nStage
        Case 1
          sPiccy = "corrtri.bmp"
        Case 2
          sPiccy = "corrsq.bmp"
        Case 3
          sPiccy = "corrblu.bmp"
        Case 4
          sPiccy = "corrtwo.bmp"
      End Select
```

```
   End If
   Write #1, "well done"
End Sub
:
Sub donewrong ()
   ' find the stage, infer the exemplar
   If nType = 1 Then ' strict
      Select Case nStage
         Case 1
            sPiccy = "wrotri.bmp"
         Case 2
            sPiccy = "wrosq.bmp"
         Case 3
            sPiccy = "wrospo.bmp"
         Case 4
            sPiccy = "wrowhi.bmp"
      End Select
   Else ' charm
      Select Case nStage
         Case 1
            sPiccy = "wrotri.bmp"
         Case 2
            sPiccy = "wrosq.bmp"
         Case 3
            sPiccy = "wroblu.bmp"
         Case 4
            sPiccy = "wrotwo.bmp"
      End Select
   End If
   Write #1, "not yet"
End Sub
:
Sub Form_Load ()
   ctr form4
   ' indeedy
   If nSlide >= 8 Then
      Call doneright
   Else
      Call donewrong
   End If
   picture1.Picture = LoadPicture("c:\usr\d\vbasic\graphics\" & sPiccy)
   form4.Show
End Sub
:
```

END.FRM

```
Option Explicit
Sub Command1_MouseDown (Button As Integer, Shift As Integer, X As Single, Y As Single)
   Close
   End
End Sub
:
Sub Form_DblClick ()
   End
End Sub
:
Sub Form_Load ()
   ctr form5
   form5.Show
 End Sub
```

Appendix 2

Studies included in advance information analysis

References	Task	Response	Inclusion	Source of values
Bloxham et al. (1984)	2 CRT and cued 2 CRT	Keyrelease	All values reported	From table 4
Girotti et al. (1986)	3 CRT and cued 3 CRT	Keyrelease	All values reported	From table 2
Stelmach et al. (1986)	8 CRT, four levels of cueing	Keyrelease	Uncertainty levels 3 (uncued) and zero (fully cued)	Est'd from fig 4
Lichter et al. (1988)	2 CRT and cued 2 CRT	Keyrelease	All values reported	From table 3
Pullman et al. (1990)	2 CRT and cued 2 CRT	Wrist flexion	High (optimal) levodopa dose	Est'd from fig 1
Jahanshahi et. al. (1992a)	4 CRT, 3 levels of cueing, SRT	Keyrelease	Fully cued and uncued 4 CRT	Est'd from fig 6
Sprengelmeyer et al. (1995)	2 CRT and cued 2 CRT	Keyrelease	Block 1 "simple" (cued 2 CRT), block 1 "choice" (2 CRT)	From table 5
Willingham et al. (1995)	2 CRT, two types of cued CRT	Keypress	"Warningonly" and "cuedwatch" conditions	Est'd fig 1

Appendix 3

Studies included in medication analysis

References	Task	Response	Manipulation	Inclusion	Source of values
Bloxham et al. (1987)	SRT	Keypress	Withdrawal, mean 9.5 hours	All "without secondary task" values	Est'd from fig 2
Pullman et al. (1988)	SRT, 2 CRT	Wrist flexion / extension	levodopa infusion at "high", "medium" and "low"levels	"High" (optimal) and "low" levels	Est'd from fig 3
Brown et al. (1993b)	2 CRT	Keyrelease	Withdrawal for at least 24 hours	All values	Est'd from fig 3
Malapani et al. (1994)	Go/nogo CRT	Keypress	Withdrawal for at least 18 hours	Visual and auditory baseline conditions	From table 2
Labutta et al. (1994)	3 CRT	Whole arm reach	"Overnight" withdrawal	All delays, collapsed across "direction"	Est'd from fig 1
Harrison et al. (1995)	SRT, 2 CRT	Keypress	Withdrawal for "about 12 hours"	Averages given for SRT and CRT (across delay)	From table 2

Appendix 4

2-CRT studies (uncued)

Reference	Response	Imperative signal	Inclusion	Source
Talland et al.. (1963)	Keypress	Light	"Disjunctive" condition	From table 1
Mayeux et al. (1987)	Finger lift	Light or tone	"RT choice"	From table 1
Goodrich et al. (1989)	Finger lift	Tactile	CRT, secondary task absent	From table 3
Worringham & Stelmach (1990)	Keyrelease	Light (LED)	Block 1 only	Est'd from fig 1
Zimmerman et al. (1992)	Key release	Screen	Tasks 2 and 3	From table 2
Bradhsaw et al. (1993)	Key press	Vibrotactile	Expt 1 and 2, "old" control groups	Est'd from figs 1 and 2
Brown, Jahanshahi & Marsden, (1993)	Key release	Light	"2-CRT, compatible"	Est'd from fig 2
Revonsuo et al. (1993)	Key press	Screen	2-CRT, "preserved" patient group	From table 5
Schugens et al. (1993)	Key release	Screen	Choice conditions for finger and hand	Est'd from fig 1
Fimm et al.. (1994)	Key release	Screen	CRT	From table 2
Pate & Margolin (1994)	Key press	Screen	"CRT-1"	From table 2
Bennett et al. (1995b)	Key press	Screen	Experiment 4	From table 2
Schnider et al. (1995)	Key press	Screen	CRT "best run"	From table 2

Appendix 5

Simple reaction times

Reference	Response	Imperative signal	Inclusion	Source of values
Talland et al. (1963)	Keypress	Light	"Simple" and "alerted all RYG", PD and control	From table 1
Heilman et al. (1976)	Keypress	Light (quick-peak incandescant)	All values	From table 1
Bloxham et al. (1987)	Button press (with thumb)	Tone	Without secondary task, on medication	Est'd from fig 2
Mayeux et al. (1987)	Finger lift	Light or tone	Simple RT; "PD without bradyphrenia" group	From table 1
Pullman et al. (1988)	Wrist flexion	Screen	Simple RT, high dose "on"	Est'd from fig 3
Crawford et al. (1989)	Keypress	Screen	Mean SRT	From table 2
Goodrich et al. (1989)	Finger lift	Tactile	SRT left and right, secondary task absent	From table 3
Reid et al. (1989)	Key press	Light	Early- and late-onset groups	From table 2
Daum and Quinn (1991)	Key release	Light	SRT	Est'd from fig 1
Jahanshahi et al. (1992a)	Key release	Screen	Random block SRT, all delays	Est'd from fig 5
Jordan et al. (1992)	Keypress	Screen	SRT, PD de novo and treated	Est'd from fig 1
Zimmerman et al. (1992)	Key release	Screen	Task 1, early and advanced PD	From table 2
Brown et al. (1993a)	Key release	Screen	Mean SRT across delays	Est'd from fig 1

Ghika et al. (1993)	Key release	Tone	Age-matched controls, left and right	From table 2
Nakashima et al. (1993)	Keypress	Tone	SRT, calculated by adding EMG and MT	Est'd from figs 1 and 2
Revonsuo et al. (1993)	Keypress	Screen	"Preserved" PD group	From text on p97
Schugens et al. (1993)	Key release	Screen	"Finger" and "hand" conditions	Est'd from fig 1
Cooper et al. (1994)	Keypress	Screen	SRT	Est'd from fig 1
Fimm etal. (1994)	Key release	Screen	Untreated and treated PD	From table 2
Klockgether et al. (1994)	Initiation of arm movement	Light (laser dot)	SRT	From text on p50
Pate & Margolin (1994)	Keypress	Screen	nd-PD and vm-PD groups	From table 2
Bennett et al. (1995b)	Keypress	Screen	Experiments 2 and 3	From table 2
Gorrell et al. (1995)	Key release	Screen	SRT	Text on p1141
Harrison et al. (1995)	Keypress	Light	SRT, mean of both PD groups	Table 2
Schnider et al. (1995)	Keypress	Screen	Run 1 and run 2	From table 2
Willingham et al. (1995)	Keypress	Screen	"Single" task, all delays	Est'd from fig 1