Card Sorting Performance in Parkinson’s Disease: A Comparison Between Acquisition and Shifting Performance*

Karel P.M. van Spaendonck¹, Hans J.C. Berger¹, Martin W.I.M. Horstink², George F. Borm³, and Alexander R. Cools⁴

¹Department of Medical Psychology, University of Nijmegen, ²Department of Neurology, University of Nijmegen, ³Department of Medical Statistics, University of Nijmegen, ⁴Department of Psychoneuropharmacology, University of Nijmegen

ABSTRACT

In the present study we tested the hypothesis that learned irrelevance underlies the frequently observed poor performance of Parkinson’s disease (PD) patients on card sorting tests. If learned irrelevance accounts for the poor performance of PD patients on card sorting tests, PD patients and control subjects (CS) will not differ in the acquisition phase, during which basic concept formation is assessed, but they will differ in the subsequent shifting phases. We presented three distinct card sorting tests with an identical format to 51 PD patients and 24 normal controls. The groups did not differ with respect to intelligence, memory, or attention. PD patients showed a slightly better performance in the acquisition phase. In the first shifting phase, the performance of PD patients was significantly poorer than that of control subjects after correction for basic concept formation. In the second shifting phase this difference disappeared. We conclude that learned irrelevance does not account for the poor performance of PD patients in card sorting tests. The results are discussed in terms of self-generation of problem solving strategies.

During the last decade neuropsychological studies have provided evidence that patients with Parkinson’s disease (PD) show certain cognitive deficits, which have been grouped under headings such as an “impaired executive function” or “frontal lobe syndrome” (Brown & Marsden, 1987; Dubois, Bollier, Pillon, & Agid, 1991; Taylor, Saint-Cyr, & Lang, 1986). Impaired executive function can be assessed by a range of tests (Taylor & Saint-Cyr, 1992). Among these tests, card sorting is a prominent choice. Impaired performance of PD patients on card sorting tests, especially the Wisconsin Card Sorting Test (WCST), is a frequently recurring finding in neuropsychological studies on PD. Although this test was originally developed to assess mental flexibility in normal subjects (Grant & Berg, 1948), it later gained the reputation to detect frontal lobe dysfunction (Milner, 1963, 1964). Several forms and scoring procedures have been proposed to increase the discriminative power of the test (Heaton, 1981; Nelson, 1976). However, there is no consensus as far as the underlying deficit of poor performance is concerned. Various cognitive dysfunctions have been suggested as representing the basic impairment, such as an impairment of shifting set (Lees & Smith, 1983; Taylor et al., 1986), maintaining set (Flowers & Robertson, 1985; Taylor et al., 1986), and concept formation (Bowen, Kamienny, Burns, & Yahr, 1975; Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991).

Recently, Owen et al. (1993) have suggested that deficits in set-shifting tests, including the WCST, may originate from the disruption of two distinct forms of attentional set-shifting ability: perseveration and learned irrelevance. Perseve-

* Correspondence to: K.P.M. van Spaendonck, Department of Medical Psychology, University of Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.
Accepted for publication: June 8, 1995.
RATION occurs when a subject is unable to shift from a previously relevant dimension; it reflects an increased tendency to respond to a dimension which previous positive feedback indicated as the relevant one. Learned irrelevance occurs when a subject is unable to shift to a previously irrelevant dimension; it is due to an active inhibition of responding to a dimension which previous negative feedback indicated as an irrelevant one. In a highly sophisticated experiment, based on the intra- and extradimensional shift paradigm of Downes et al. (1989), Owen and colleagues presented two independent set-shifting tasks in which the relative contribution of perseveration and learned irrelevance was assessed in a pre- and post-shift condition to PD patients and to patients with frontal lobe damage. They found evidence that perseveration is characteristic for frontal patients, and that learned irrelevance is characteristic for PD patients.

In the present study, we tested the hypothesis whether learned irrelevance can also account for the performance of PD patients on card sorting tests. We presented three card sorting tasks in the WCST format. After the acquisition phase, during which basis concept formation was assessed, there were two unannounced shifts to a category, that had been irrelevant in the preceding phase. In a previous study (Cools, van den Bercken, Horstink, van Spaendonck, & Berger, 1984), we had already found evidence that differences in card sorting performance between PD patients and control subjects (CS) become obvious only in the second phase, after correction for baseline performance in the first phase. If learned irrelevance affects the performance of PD patients on card sorting tests, the performance of PD patients will be poorer than that of CS each time that subjects have to shift to a category which was irrelevant in the preceding phase. Unlike most studies that used the overall measures derived from the current manual (Heaton, 1981), the hypothesis of Owen and colleagues takes into account that the WCST is essentially a test to be performed in consecutive phases and, thus, may not involve identical abilities.

Therefore, the present study deals with the phase-wise analysis of three homologous card sorting tests.

METHOD

Subjects
The PD group consisted of 71 consecutive patients from the outpatients’ Department of Neurology at the University of Nijmegen. The diagnosis of idiopathic PD was based on the presence of tremor, rigidity, and bradykinesia. The symptoms were evaluated by one neurologist (M. H.) a week before psychological testing. The score of the Unified Parkinson’s Disease Rating Scale (UPDRS) (Fahn, Elton, & the members of the UPDRS Development Committee, 1987) was used as a measure of severity of PD. Patients with non-PD symptoms, EEG and CT scan abnormalities, cerebrovascular disease, and dementia according to DSM-III-R criteria had been excluded. Medication was neither changed nor stopped during the investigation. In a recent study (van Spaendonck, Berger, Horstink, Buytenhuijs, & Cools, 1993), we have found evidence that anticholinergic therapy deteriorates performance of PD patients on card sorting tests. For this reason, patients on anticholinergic therapy (n = 14) were also excluded.

The control group (CS), which consisted of 32 subjects, had no central neurological symptoms. Most control subjects were spouses of the PD patients who participated in this study. Since shifting performance can only be assessed correctly in subjects who are able to achieve the acquisition phase (cf. Cooper et al., 1991), all subjects who were unable to achieve the acquisition phase of any card sorting test were excluded. Eventually, 51 PD patients and 24 CS participated in the present study. Thirty-one PD patients were male and 20 were female. The mean age was 53.9 years (SD = 0.2). Forty-one patients were new, the others were on levodopa therapy. Eleven CS were male and 13 were female. Most of them were spouses of PD patients who participated in this study. Their mean age was 52.7 years (SD = 8.5). There were no significant differences between PD patients and CS with respect to male/female ratio, or age. All participants gave informed consent. Table 1 shows the relevant demographic data and the disease profile of the PD group.

Since comparison of consecutive phases presupposes that the subject achieves at least the first phase, subjects who were unable to achieve the acquisition were excluded from analysis. From the original pool of 71 PD patients, 6 additional patients were excluded in addition to the 14 patients on anticholinergic therapy. That represents 28% of the total subject pool. From the original group of 32 CS, 8 were excluded, which equals 25%. 

CARD SORTING
Table 1. Demographical and Clinical Characteristics.

<table>
<thead>
<tr>
<th></th>
<th>PD patients (n = 51)</th>
<th>CS (n = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>male/female</td>
<td>M (SD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M (SD)</td>
</tr>
<tr>
<td>male/female</td>
<td>31/20</td>
<td>11/13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(p = .33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(p = .33)</td>
</tr>
<tr>
<td>Age</td>
<td>53.9 (10.2)</td>
<td>52.7 (8.5)</td>
</tr>
<tr>
<td>Age at onset of disease</td>
<td>50.7 (10.2)</td>
<td></td>
</tr>
<tr>
<td>Duration of disease</td>
<td>3.3 (2.4)</td>
<td></td>
</tr>
<tr>
<td>UPDRS sum score</td>
<td>9.7 (3.5)</td>
<td></td>
</tr>
</tbody>
</table>

Tests

Reference Tests
In order to assess intelligence, memory, and attention we included a series of reference tests in our battery. We selected four subtests from the Dutch adaptation of the WAIS-R (Stinissen, Willems, Coetsier, & Hulsman, 1970) to exclude biasing differences in intelligence between both groups, and to exclude general mental deterioration: Vocabulary and Similarities, measuring verbal intelligence, Picture Completion and Block Design, measuring visuoperceptual intelligence (Cyr & Brooker, 1984). Memory was assessed by means of the Dutch adaptation (Deelman, 1972) of the Rey Auditory Verbal Learning Test (Lezak, 1983). A list of 15 semantically unrelated words was presented orally five times. The subject was asked to state the memorized words after each presentation. The number of correct words after five presentations was the score of memory performance. Attention was assessed with the Stroop Color-Word Test (Jensen & Rohwer, 1966). Only Part B (speed of color bar identification) and Part C (speed of color vs. word identification) were utilized. The score was obtained by subtracting the number of seconds that the subject required to read the simple Part B from the time (in seconds) needed to read the complex Part C.

Criterion Tests.
We presented three card sorting tests: the Wisconsin Card Sorting Test (WCST; Heaton, 1981) and a verbal and spatial variation. Assessment of card sorting by means of a number of tests with identical format, but with heterogeneous material, prevents us from basing our conclusions on the unique variance of one single test. When both variations of the WCST were developed, we adopted the traditional dichotomy of verbal and visuoperceptual intelligence in order to increase the heterogeneity of the sorting criteria. In each test the subject began to sort on the basis of trial and error according to a category that was known only to the experimenter. A false response was corrected by the experimenter. To minimize the impact of memory, the correct (or corrected) matchings were laid down just below the appropriate stimulus cards. After a fixed number of correct responses the experimenter changed the category twice without informing the subject. The first phase, in which basic concept formation is assessed, is labelled as "acquisition phase"; whereas the second and third phases are labelled as "shifting phases". Since we presented three card sorting tests, we decided not to repeat the tests. Thus, each test consisted of an acquisition phase and two shifting phases.

We presented the WCST in the original form. The simplified form of Nelson (1976), which was developed for detecting frontal lobe lesions, was not suitable for the aim of the present study, since the subjects are explicitly instructed after the acquisition that the sorting rule has been changed, which makes the task of the subject not only more simple, but also quite different (Braff et al., 1991; Goldman, Axelrod, & Tompkins, 1992).

The verbal variation of the WCST was the animals sorting test (AST). This test consists of three series of 24 cards each. On each card the name of an animal species is printed. Initially, subjects had to match 'bird versus mammal'. After seven correct responses the experimenter changed the sorting category to 'herbivore versus carnivore' and, finally, to 'native versus exotic'. If a subject still failed to make a correct assignment after 17 trials, the experimenter proceeded with the next series. The spatial sorting test (SST) consists of two stimulus cards and three series of 20 response cards, each depicting two figures that are assembled from a line and a triangle in a varying spatial relationship (see Figure 1). Three patterns out of the total of conceivable spatial relations are selected as criteria in assigning the cards to one of the two sample cards as follows: identical versus reversal, top position versus middle position, line crossing an angular point versus coinciding with the vertical side. The instructions given to the subjects are analogous to those for animal sorting.
As already indicated, the Heaton manual only provides overall measures, and these measures can hardly, or not at all, be transformed into measures that enable us to compare performance per phase. Also, the definition of perseverative errors, being the most powerful discriminative measure of the test, is not identical in the acquisition phase and in the shifting phases; nevertheless, these unequal quantities must be added to one perseverative error score in order to increase the diagnostic discriminative power of the test according to the manual (Heaton, 1981). This makes perseverative errors unsuitable for a comparison of performance in distinct phases. In previous studies (Berger et al., 1993; van Spaendonck et al., 1993), we have found that the number of trials that a subject requires before he makes the fixed number of correct assignments in succession is an appropriate score to measure performance. The total number of trials in all phases together correlated strongly with the total number of categories, being a conventional measure. In each phase, only one category can be scored, which makes this measure unsuitable to assess performance within each single phase. Therefore, we used the number of trials per phase that a subject requires before he reaches the fixed number of correct assignments successively (WCST: 10; AST: 7; SST: 7). We are fully aware that the modifications in the administration and scoring of the traditional WCST, which enabled us to assess card-sorting performance per phase, make a comparison of our results with those of other studies rather precarious.

Data analysis
Card sorting performance was assessed and standardized per phase, and the sum of scores was calculated per phase, resulting in an overall score for acquisition, the first shift, and the second shift. Finally, the difference scores between the first shift and the acquisition and between the second and the first shift were calculated. Differences between the groups were evaluated using Student’s $t$ test or, where appropriate, the chi-square test. Although the sex imbalance (F/M ratio 118% in CS vs. 65% in PD) was not significant, we also analyzed the results including a correction for cofactor sex (linear model), only to find that sex was not related to test performance, and that the results were almost identical, whether or not sex was included in the analysis. Thus, for the sake of simplicity, we choose to present the results of the $t$ test only, this being the most straightforward method. The significance level was set at .05 (two-sided).

As shown in Table 2, the groups did not differ in intelligence, memory, and attention. With respect to card sorting performance, PD patients and CS achieved almost the same number of categories. Thus, there was no difference in overall performance. The standardized sum scores of the acquisition phase for both groups did not differ significantly. Actually, PD patients performed slightly better than CS in the acquisition phase. However, the standardized sum scores of PD patients and CS in the first shifting phase was on the verge of significance ($p < .07$). The performance of PD patients was much poorer than that of CS and reached significance after correction for baseline performance in the acquisition phase. In contrast, the standardized sum scores of the groups in the second shifting phase were no longer any different. The two groups performed fairly identically in the second shifting phase after correction for baseline performance in the acquisition phase.
Table 2. Test Performance. Means and Standard Deviations in Parentheses.

<table>
<thead>
<tr>
<th></th>
<th>PD patients (n = 51)</th>
<th>CS (n = 24)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td></td>
</tr>
<tr>
<td>WAIS VIQ</td>
<td>112.1 (13.6)</td>
<td>117.1 (14.8)</td>
<td>ns</td>
</tr>
<tr>
<td>WAIS PIQ</td>
<td>114.3 (15.0)</td>
<td>117.3 (16.7)</td>
<td>ns</td>
</tr>
<tr>
<td>RAVLT</td>
<td>40.7 (9.1)</td>
<td>44.3 (10.4)</td>
<td>ns</td>
</tr>
<tr>
<td>Stroop C-B</td>
<td>43.5 (23.6)</td>
<td>40.4 (15.8)</td>
<td>ns</td>
</tr>
<tr>
<td>Sum categories</td>
<td>7.6 (1.4)</td>
<td>8.0 (1.1)</td>
<td>ns</td>
</tr>
<tr>
<td>SS acquisition</td>
<td>5.2 (5.4)</td>
<td>6.0 (6.8)</td>
<td>ns</td>
</tr>
<tr>
<td>SS first shift</td>
<td>31.7 (23.6)</td>
<td>21.6 (18.1)</td>
<td>&lt; .07</td>
</tr>
<tr>
<td>SS second shift</td>
<td>38.8 (23.9)</td>
<td>36.6 (26.5)</td>
<td>ns</td>
</tr>
<tr>
<td>CSS first shift</td>
<td>26.5 (23.0)</td>
<td>15.6 (17.6)</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>CSS second shift</td>
<td>33.7 (23.5)</td>
<td>29.6 (28.6)</td>
<td>ns</td>
</tr>
</tbody>
</table>

Note. RAVLT = Rey Auditory Verbal Learning Test
Stroop C-B = Stroop Color Word Test card C minus card B
SS = standardized sum of trials transformed into percentage of the maximum number of trials
CSS first shift = SS first shift corrected for acquisition performance
CSS second shift = SS second shift corrected for acquisition performance

DISCUSSION

In the present study we tested the hypothesis that learned irrelevance accounts for the poor performance of PD patients on card sorting tests. We compared the performance of PD patients and CS in three consecutive phases of card sorting tests. PD patients and CS did not differ with respect to age, intelligence, or attention. Data for education levels are not available, but we do not think that this invalidates our study, because almost all controls were spouses of the PD patients, and because the data concerning the level of general intelligence in Table 2 indicate that there were no significant differences in general intelligence between PD patients and CS. We presented three card sorting tests with identical instructions but heterogeneous materials, the format being that of the original form of the WCST. For reasons indicated earlier, the simplified form of the WCST (Nelson, 1976) was not suitable for the present study. Therefore, the outcome of the present study cannot be compared with that of studies using the Nelson version of the WCST.

In contrast to the studies of Taylor et al. (1986) and Cooper et al. (1991), we found no evidence that PD patients were impaired in concept formation. Actually, the performance of our PD patients in the acquisition phase was even slightly better than that of CS. There are several explanations for this discrepancy. First, we excluded patients who did not achieve the acquisition phase, whereas Taylor and colleagues, and Cooper and colleagues have provided no information in this respect. Although exclusion of those patients in our study may have led to a relatively high performance on the acquisition phase, it is unlikely that this underlies the difference. The percentage of PD patients who were excluded on this criterion (11%; n = 6) was smaller than that of CS who were excluded (25%, n = 8), suggesting that PD patients showed a slightly better concept formation.

A second explanation for the noted discrepancy may be the fact that we excluded PD patients on anticholinergic therapy. Recently, we have found that this therapy selectively impairs card sorting performance (van Spaendonck et al., 1993), a finding that was recently confirmed by Farina et al. (1994). Given the results of the
present study, we reanalysed the data of our study on anticholinergic therapy in order to establish to what extent anticholinergic therapy affected each distinct phase and discovered that 8 of the 11 PD patients on anticholinergic therapy were unable to achieve the acquisition phase. Thus, inclusion of PD patients on anticholinergic therapy can bias the performance in the acquisition phase, which is a measure of concept formation.

A third explanation for the noted discrepancy may be the fact that our PD patients had no cognitive deterioration in terms of their performance on the various subtests of the WAIS, whereas the PD patients in the study of Cooper et al. (1991) performed significantly more poorly than did CS on the Digit Span backwards and Picture Arrangement subtests of the WAIS. Although the PD patients in the Cooper et al. study did not show abnormal scores on the National Adult Reading Test (NART) or on the WAIS Vocabulary subtest and reportedly had no shortcomings in premorbid or current intelligence, the noted deficits may have contributed to their impaired performance on the acquisition phase.

With respect to cognitive shifting, most studies base their conclusions on the number of categories achieved. If PD patients achieve significantly fewer categories than CS, they are considered to be impaired in cognitive shifting (Blonder, Gur, Gur, Saykin, & Hurtig, 1986; Bowen et al., 1975; Caltagirone, Carlesimo, Nocentini, & Vicaro, 1989; Taylor et al., 1986). However, if there was no significant difference in this respect, it is concluded that PD patients were not impaired in cognitive shifting (Cooper et al., 1991; Mohr, Litvan, Williams, Fedio, & Chase, 1990). In the present study, we found that PD patients were impaired in cognitive shifting in spite of the fact that they achieved approximately the same number of categories as control subjects (7.6 vs. 8.0). Apparently, the number of categories is a score which is not subtle enough for detecting the actual impairment. The present study shows that PD patients who achieved as many categories as CS needed significantly more trials to shift from the acquisition phase to the next phase. Thus, PD patients were impaired in cognitive shifting.

This result is not biased by the fact that some PD patients on levodopa therapy participated in the present study. The studies by Downes et al. (1989), Lange et al. (1992) and Owen et al. (1993) have indicated that set-shifting in PD may be ameliorated rather than deteriorated by levodopa therapy. Thus, our PD patients showed poor cognitive shifting in spite of levodopa therapy.

Our results are not consistent with the hypothesis of Owen et al. (1993) that the poor performance of PD patients on card sorting tests is due to learned irrelevance. If PD patients suffered from learned irrelevance, they should show poor performance each time they have to shift to a previously irrelevant category. However, our PD patients only showed a poor performance in the first shifting phase. In the second shifting phase, PD patients and CS needed approximately the same number of trials to shift to a category, which was not only irrelevant in the acquisition phase, but also irrelevant in the first shifting phase, that is, to a two-fold irrelevant category. Thus, PD patients were not inhibited by learned irrelevancy in the second shifting phase.

Therefore, the question arises whether or not the differences in the first shifting phase are due to learned irrelevance. According to the concept of learned irrelevance, a subject learns which category is irrelevant when he makes a wrong assignment and gets feedback that he is incorrect. Inversely, the subject does not learn that a certain category is irrelevant when he does not make an assignment to that category. However, many subjects in the present study started with the correct category and achieved the acquisition without any failure, that is, without learning explicitly which categories were irrelevant. For example, nearly all subjects (44 PD and 21 CS) started with the correct category in the acquisition of the AST, that is, they did not learn that other categories were irrelevant. Nevertheless, PD patients required more trials than CS in the first shifting phase of the AST (6.1 vs. 3.6).

A complementary explanation is needed. Many neuropsychological studies in PD have
established that PD patients are impaired in problem solving which requires the processing of internal information. In contrast, PD patients are not impaired in problem solving that is guided by external cues. In a previous study based on this paradigm (Buytenhuijs et al., 1994), we found that PD patients organized information to be recalled to a significantly lower degree than did control subjects. Control subjects spontaneously generate categories in which this information could be stored, without any instruction to do so, a finding that is in agreement with the results of a study by Taylor, Saint-Cyr, and Lang (1990). In contrast, PD patients adhered to the externally imposed sequence on which the information to be recalled was presented. A corresponding finding was reported by Downes et al. (1993), who found that CS used significantly more semantically related word clusters in an alternating fluency task than did PD patients.

Transposing these findings to the present study, it is plausible that control subjects spontaneously group the cards to be sorted into categories already in the acquisition phase, irrespective of the relevancy of these categories. In the acquisition phase, PD patients focus on the explicit instruction and feedback of the experimenter, which explains their slightly better performance. Since PD patients and CS performed equally well in the acquisition phase, they did not differ as far as responding to feedback is concerned (i.e., grouping according to relevancy). In the first shifting phase, CS benefit from their spontaneous organization during the acquisition, because they already dispose of alternative categories. The incompatible feedback of the first shifting phase forces the PD patients to start the generation of alternative categories, which requires proportionally more trials. Owing to this generative information processing in the first shifting phase, PD patients, like control subjects, dispose of alternative categories in the second shifting phase. For this reason, the degree of complexity in the second shifting phase is similar for both groups. Accordingly, both groups do not differ in the number of trials in the second shifting phase.

In sum, the present study provides evidence that PD patients who are not impaired in concept formation suffer from an impaired ability to shift from the acquisition phase to the first shifting phase, but have an intact ability to shift from the first shifting phase to the next shifting phase in three distinct card sorting tests with identical format. We conclude that a reduced self-generation of problem solving strategies, not learned irrelevance, underlies the observed deficit. Since the patients who participated in this study were relatively young and mildly disabled, our study does not exclude that learned irrelevance may affect the card-sorting performance of PD patients in a later stage of the disease.

REFERENCES


CARD SORTING

Neurology, Neurosurgery, and Psychiatry, 52, 334-337.