Movement Disorders
Vol. 9, No. 1, 1994, pp. 76–83

Unified Parkinson's Disease Rating Scale Characteristics and Structure

P. Martínez-Martín, A. Gil-Nagel, L. Morlán Gracia, J. Balseiro Gómez, J. Martínez-Sarriés, †F. Bermejo, and ‡The Cooperative Multicentric Group

Section of Neurology, Hospital Universitario de Getafe, Universidad Complutense, Getafe, Madrid; †Department of Neurology, Hospital "12 de Octubre," Universidad Complutense, Madrid; and ‡The Cooperative Multicentric Group (See below)

Summary: Our purpose was to verify some basic aspects of validation of the Unified Parkinson's Disease Rating Scale (UPDRS). One hundred and sixty-seven Parkinson's disease (PD) patients were included. Group A (n = 40) was simultaneously assessed by five raters who applied the UPDRS and other PD rating scales (PDRS). A set of timed tests, the Mini-Mental State Examination (MMSE), and the Hamilton Scale for Depression (HSD) were administered by an independent examiner. Group B (n = 127) was individually assessed through the UPDRS and the other PDRSs by one neurologist in four different hospitals. The UPDRS was administered in 16.95 ± 7.98 min. The internal consistency was high (Cronbach's alpha = 0.96). Nevertheless, the items related to depression, motivation/initiative, and tremor were scarcely consistent. The Interrater reliability was satisfactory (all the items had $k > 0.40$). There was a high correlation of the UPDRS with the Hoehn and Yahr staging (r_s = 0.71; $p < 0.001$) and some timed tests (finger tapping; arising from chair), but also with the MMSE and HSD ($r_s = 0.53$; $r_s = 0.64$; $p < 0.001$). The convergent validity with the other PDRS (Intermediate Scale and Schwab and England Scale) was very high (r_s = 0.76–0.96; $p < 0.001$). The factor analysis identified six factors that explained 59.6% of the variance. The dimension "tremor" showed a remarkable independence. The UPDRS is a multidimensional, reliable, and valid scale, with some inconveniences derived from its internal consistency, discriminant validity, and pragmatic application. **Key Words:** Unified Parkinson's Disease Rating Scale—Validation—Internal Consistency—Interrater reliability—Factor analysis.

Usually, qualitative rating scales are used to assess the functional condition of patients with Parkinson's disease (PD). Many different rating scales for PD (PDRS) have been used in clinical studies, with multiple problems arising when the interpretation or comparison of the results was attempted (1).

The need of a common and uniform method for evaluation of PD prompted the creation of the Unified Parkinson's Disease Rating Scale (UPDRS) Development Committee in 1984 (2). The UPDRS has become the most widely used PDRS. However, some of the metric characteristics that support the validity of such tools have never been checked (or published). The purpose of the present study is to verify some aspects of the validity and the dimensionality of the UPDRS.

Address correspondence and reprint requests to Dr. P. Martínez-Martín at Sección de Neurología, Hospital Universitario de Getafe, Carretera de Toledo, Km. 12.5, 28905, Getafe (Madrid), Spain.

‡M. C. Macias, C. Jimenez-Rojas, E. Marañon, Servicio de Geriatria, Hospital Universitario de Getafe, Madrid; J. M. Grau Veciana, C. Roig Arnall, O. Bruna, C. Junqué Plaja, Servicio de Neurología, Hospital Santa Cruz y San Pablo, Barcelona; S. Giménez-Roldán, Servicio de Neurología, Hospital Gregorio Marañón, Madrid; and J. A. Burguera, Servicio de Neurología, Hospital La Fe, Valencia.

PATIENTS AND METHODS

Patients

Forty PD patients from the Hospital Universitario de Getafe were included in the study as Group A. A set of timed tests and scales (see the following) was applied to this group following a set protocol. One hundred twenty-seven different PD patients were evaluated by other neurologists (the Multicentric arm; Group B) in order to obtain an independent valuation of the various PDRSs. The main characteristics of both groups of patients are shown in Table 1. There were patients in stages I to V of the Hoehn and Yahr's classification, with a higher representation of the stages II to IV in both arms of the study.

Examiners and Tests

Group A

The patients were assessed by six raters in two subsequent phases: an examiner (always the same, E.M.) recorded data about the patient (name, age, sex), evolution of disease, and treatment. Afterward, he applied the timed tests, the Mini-Mental State Examination (MMSE) (3), and the Hamilton Scale for Depression (HSD) (4). The timed tests included were (a) finger tapping. Number of complete movements in 20 seconds, each hand separately; (b) time (in seconds) spent in folding a paper sheet twice and putting it into an envelope; (c) number of times the patient gets up from a chair in 20 seconds; (d) time (in seconds) spent in walking 3 m, back and forth.

Immediately after, each patient was simultaneously evaluated by five other examiners (three neurologists plus two resident physicians in geriatrics). These examiners established by agreement the stage on the Hoehn and Yahr's classification (HYC). Afterwards, they simultaneously applied the UPDRS (Version 3.0) (2) and the Intermediate Scale for Assessment of PD (ISAPD) (5). The interview and the main aspects of the examination were conducted by one of the neurologists (in a rotational order) in a demonstrative way. Each rater had at his or her disposal the UPDRS and the ISAPD and also specifically designed forms for recording the data. In addition, the examiners asked the patient or her or his caregiver to clear doubts about the questions included in the scales. They checked any aspect of the examination that was considered ambiguous or unascertainable without a direct valuation (e.g., the rigidity).

In order to obtain as independent an assessment as possible, the examiners were not allowed to interchange opinions during the evaluation. The scores obtained by these five raters were only used for the interrater reliability analysis. Only the data obtained by one of the neurologists (A.G.-N.) were considered for the rest of the analysis.

Group B

The examiners were another six neurologists belonging to four different hospitals. They recorded, individually, some clinical data, HYC stage, and they also individually applied the UPDRS and the ISAPD to their patients. Therefore, in this group, each patient was evaluated by only one neurologist.

In both groups, a patient's self-evaluation from 0 (absolutely disabled; very poor condition) to 10 (absolutely able; very good condition) and the time spent to administer the UPDRS were also recorded.

In the present study, three components were considered in the UPDRS. The first part was integrated

TABLE 1. *Description of the series*

	Group A (n = 40)			Group B (n = 127)		
	$\bar{x}$	SD	Range	$\bar{x}$	SD	Range
Age[a]	69.57	11.12	44–88	64.97	8.91	42–84[b]
Age at onset[a]	64.80	10.58	40–80	—	—	—
Duration of disease[a]	4.96	4.36	0.5–19	6.55	4.22	0.5–20[b]
Hoehn and Yahr stage	2.75	1.08	1–5	2.62	0.76	1–4
Subjective score	4.96	2.77	0–10	5.04	2.74	0–10
UPDRS (NC)	42.96	25.97	6–96.4	32.48	15.89	0–68[b]
ISAPD (NC)	17.05	11.74	0–39	13.00	7.42	0–33[b]
Schwab and England Scale	66.17	30.10	0–10	78.93	14.66	20–100[b]

$\bar{x}$, sample mean; [b]Differences between Groups A and B are significant ($p < 0.05$).
[a] In years.
UPDRS, Unified Parkinson's Disease Rating Scale (Ref. 2); ISAPD, Intermediate Scale for Assessment of Parkinson's Disease (Ref. 1); (NC), scores of the Complications of Therapy subscale, not included.

by the items of the subscales I to III (items 1 to 31): Mentation, Behavior and Mood; Activities of Daily Living, and Motor Examination. This part is properly devoted to assess parkinsonian manifestations, and therefore, it was the object of the analysis in the present study. Each item was considered as a component of a scale devised to measure one construct: the condition of the PD patient. The sum of the scores of this part (items 1 to 31) in the "on" period was considered the "total UPDRS."

A single score (maximum 4) was generated for each item. When it was pertinent, the examined function or clinical manifestation was assessed in all body regions, and the final score was based on the average (if symmetric) or the worst function (if asymmetric).

The second part, including the items 32 to 42, is the subscale IV "Complications of therapy." The peculiarity of this section requires it to be analyzed separately, and this was not the aim of the present study.

Finally, the third part includes the HYC and the Schwab and England Activities of Daily Living Scale (SES). The HYC was taken as the "gold standard," and the SES was considered a functional scale to be compared with the UPDRS.

Analysis of the Results

The score assigned to each item (in the "on" period) and sum of the scores belonging to the first part (items 1 to 31) were used as crude data. Specific statistical measures (Cronbach's alpha, kappa of Fleiss, and the Relative Ranges Method) (6–9) were used to assess the internal consistency and the interrater reliability. Nonparametric correlation (Spearman rank correlation) was calculated to determine some specific aspects of the internal consistency and validity. Limits of statistical significance were given between 0.05 and 0.001. Additionally, a multivariate regression analysis was applied to establish the relationship between the UPDRS, the criterion (HYC), and the SES.

A factor analysis of the first part of the UPDRS was carried out in order to explore the dimensionality of the scale. Orthogonal and oblique rotations were used in the factors' extraction method.

The Modified HYC and the SES (2) were considered additional methods of PD evaluation in order to determine some aspects of validity. The application of the original versus the modified HYC was also analyzed.

RESULTS

The prevalence of the aspects evaluated by the UPDRS in the present study is shown in Table 2.

Reliability: Internal Consistency

As a comprehensive measure of this aspect, the Cronbach's alpha coefficient was calculated, with a result of 0.96. Additionally, the item total and the interitem correlations were verified through the Spearman rank correlation. In the first part of the UPDRS (items 1 to 31), 17 items showed a highly significant correlation (r_s from 0.60 to 0.81; $p < 0.001$) with the total score: 5, Speech; 9, Cutting food; 10, Dressing; 11, Hygiene; 12, Turning in bed; 15, Walking; 18, Speech; 19, Facial expression; 23, Finger taps; 24, Hand movements; 25, Alternating movements of hands; 26, Leg agility; 27, Arising from chair; 28, Posture; 29, Gait; 30, Postural stability; and 31, Body bradykinesia.

TABLE 2. *Prevalence and interrater reliability of the UPDRS items*

Item	Prevalence (%)	κ[a]	RR[b]
1. Intellectual impairment	67.5	0.79	6.5
2. Thought disorder	20.0	0.80	4.5
3. Depression	79.3	0.60	13.0
4. Motivation/initiative	58.1	0.71	10.0
5. Speech	52.5	0.69	9.5
6. Salivation	40.6	0.77	9.0
7. Swallowing	26.8	0.82	4.0
8. Handwriting	76.8	0.71	9.0
9. Cutting food	88.1	0.85	5.5
10. Dressing	87.5	0.82	5.5
11. Hygiene	86.2	0.83	4.5
12. Turning in bed	81.8	0.75	8.5
13. Falling	28.7	0.70	5.0
14. Freezing	49.3	0.71	7.5
15. Walking	91.2	0.72	8.5
16. Tremor	86.8	0.68	13.5
17. Sensory complaints	59.3	0.46	21.5
18. Speech	90.0	0.53	14.0
19. Facial expression	85.6	0.42	20.5
20. Tremor at rest	80.0	0.54	16.0
21. Action of postural tremor	56.2	0.48	17.0
22. Rigidity	67.5	0.86	4.5
23. Finger taps	80.6	0.47	15.5
24. Hand movements	69.3	0.60	11.0
25. Rapid alternating movements	77.5	0.60	14.0
26. Leg agility	85.6	0.51	15.5
27. Arising from chair	40.0	0.73	8.5
28. Posture	82.5	0.49	16.5
29. Gait	62.5	0.66	10.5
30. Postural stability	67.5	0.90	3.0
31. Body bradykinesia	86.6	0.43	19.0

[a] Kappa of Fleiss (Ref. 8).
[b] Relative ranges (Ref. 9). The interrater agreement may be considered moderate when RR is ≤15, high when RR is ≤10, and very high when RR ≤ 5.0 (theorical limits).

The other 14 items also correlated at a significant level (r_s from 0.22 to 0.50). The items 20, Tremor at rest (r_s = 0.22); 3, Depression (r_s = 0.23); and 4, Motivation/Initiative (r_s = 0.24) had the lowest correlations. A matrix of interitem correlations (Spearman rank correlation) showed that five of them displayed as nonsignificant a high proportion of their correlations: 3, depression; 4, motivation/initiative; 16, tremor; 20, tremor at rest; and 21, action or postural tremor.

Reliability: Interrater Agreement

The results of this analysis appear in Table 2. As a whole, all the items showed an adequate interobserver reliability. It was moderate for the following items: 17, sensory complaints; 19, facial expression; 21, action or postural tremor; 23, finger taps; 28, posture; and 31, body bradykinesia ($0.40 < k < 0.50$). Similar results were obtained through the Relative Ranks Methods (9). The correlation between both methods was $r = 0.95$ ($p < 0.001$). The correlations among the total scores obtained from the five examiners who rated the Group A was very high ($r_s = 0.98$; $p < 0.001$).

Validity

Pragmatic Validity

The UPDRS is easily applied in the actual practice. The average time spent in applying it (complete) was 16.95 ± 7.98 min (19.86 ± 6.49 min in Group A; 10.03 ± 9.0 min, in Group B).

Criterion-Related Validity

The correlation between the HYC (criterion) and the total score of the UPDRS was $r_s = 0.71$ ($p < 0.001$). The variation induced by taking into account the original HYC instead of the modified one (as appears in the UPDRS) (2) was negligible ($r_s = 0.69$).

The stepwise multiple regression analysis (Table 3) showed that 69% of the HYC variance was explained by five items of the UPDRS. The items that had significant predictive value were 30, postural stability; 10, dressing; 29, gait; 13, falling; and 12, turning in bed.

An additional verification of concurrent validity was established by means of the correlation among the total score of the UPDRS and the timed tests (Table 4). Correlation coefficients ranged from 0.25 (nonsignificant; Walking a distance) to 0.76 ($p < 0.001$; Arising from chair).

TABLE 3. *Stepwise multiple regression analysis*[a]

Step	Entered	Multiple R	Multiple RSQ	Change in RSQ	F to enter
1	30. Postural stability	0.76	0.58	0.58	231.49
2	10. Dressing	0.80	0.64	0.06	30.72
3	29. Gait	0.81	0.67	0.02	10.61
4	13. Falling	0.82	0.68	0.01	8.97
5	12. Turning in bed	0.83	0.69	0.009	5.30

[a] Dependent variable: Hoehn and Yahr staging.

TABLE 4. *Correlation between the UPDRS and the timed tests*

	r_s	p <
Finger tapping (left)	−0.67	0.001
Finger tapping (right)	−0.49	0.01
Finger tapping (mean)	−0.59	0.001
Arising from chair	−0.76	0.001
Walking a distance	0.25	NS
Test of the envelope	0.35	0.05

Group A, n = 40.
r_s, Spearman rank correlation coefficient; NS, not significant.

Construct Validity

The discriminant validity was determined through the correlation between the UPDRS and two scales devised to measure constructs different from PD: MMSE and HSD. Spearman rank coefficients were −0.64 and 0.53 ($p < 0.001$), respectively (Table 5).

The convergent validity was determined, establishing the correlation between the scores of the UPDRS and those obtained from other PDRS: Patient's Self-evaluation, ISAPD, and SES. Spearman

TABLE 5. *Construct validity of the UPDRS*[a]

	Group A	Group B	Both
Mini-Mental Status Exam	−0.64	—	—
Hamilton Scale for Depression	0.53	—	—
Subjective patient's score	−0.51	—	—
ISAPD	0.95	0.91	0.92
Schwab and England Scale	−0.96	−0.76	−0.81

ISAPD, Intermediate Scale for Assessment of Parkinson's Disease (Ref. 5).
[a] Spearman rank correlation coefficients $p < 0.001$ (Group A, n = 40; Group B, n = 127).

coefficients ranged from 0.51 to 0.96 ($p < 0.001$) (Table 5).

Dimensionality

Factor analysis of the UPDRS showed the existence of six factors that explained 59.61% of the variance. Table 6 shows the factor loadings and communality of the 31 included items, considered as independent variables (orthogonal analysis). The identified factors were defined as follows: Factor I, Mobility of the extremities, including bradykinesia and rigidity; Factor II, Stability, gait and general mobility; Factor III, Functional ability; Factor IV, Tremor; Factor V, Communication/expression; and Factor VI, Bradykinesia and gait, this latter with coefficients lower than 0.30. Similar results were obtained after oblique rotation (Table 7). The correlations between factors ranged from 0.28 to 0.57, but the factor "tremor" was uncorrelated to any other.

DISCUSSION

The evaluation of the PD patients is carried out by means of PDRS. These scales are the most used and pragmatic clinical tools to measure severity, in the absence of a specific parameter to be measured (e.g., density of neurons in the sustantia nigra) or uniform, comprehensive system of objective measures (1,10). However, the subjective nature of these instruments makes it necessary to verify their certainty and efficacy (reliability and validity) through a statistical process (6). Only partial aspects of the metric qualities of the UPDRS have been published (2,11,12).

The reliability of the UPDRS is remarkable. The Cronbach's alpha is the most used method to assess the internal consistency of measurement scales. The alpha obtained in the present study (0.96) is indicative of very high internal consistency, but the effects of redundancy (several items focused on the

TABLE 6. *Factor analysis (orthogonal rotation) of the UPDRS*

Items		Factors						
N.	Name	I	II	III	IV	V	VI	Communality
23.	Finger tapping	0.88						0.78
24.	Hand movements	0.83						0.73
25.	Rapid alternating mov.	0.79						0.73
26.	Leg agility	0.66	0.51					0.70
31.	Body bradykinesia	0.62	0.33	0.42			0.28	0.81
22.	Rigidity	0.55		0.34				0.52
28.	Posture	0.51	0.51	0.41				0.77
30.	Postural stability	0.35	0.81					0.83
15.	Walking	0.27	0.72	0.27				0.72
29.	Gait	0.33	0.70				0.26	0.69
12.	Turning in bed	0.36	0.55	0.32				0.68
11.	Hygiene	0.46	0.55	0.51				0.83
13.	Falling		0.54					0.48
27.	Arising from chair	0.50	0.54	0.32				0.69
14.	Freezing		0.52					0.44
9.	Cutting food	0.44	0.44	0.59				0.81
10.	Dressing	0.45	0.47	0.54			−0.25	0.82
5.	Speech	0.37	0.25	0.53		0.50		0.78
6.	Salivation			0.51				0.52
20.	Tremor at rest				0.93			0.81
16.	Tremor				0.90			0.79
18.	Speech		0.50	0.27		0.79		0.78
1.	Intellectual impairment		0.32	0.31				0.49
2.	Thought disorder		0.29	0.33				0.49
3.	Depression							0.38
4.	Motivation/initiative			0.48				0.45
21.	Action tremor	0.36			0.48			0.49
7.	Swallowing		0.28	0.45				0.51
17.	Sensory complaints	0.33						0.32
8.	Handwriting		0.45			0.33		0.50
19.	Facial expression	0.46	0.25	0.42		0.39		0.68
	Eigenvalue	5.63	5.33	3.40	2.09	1.50	0.52	
	% Variance	18.16	17.19	10.96	6.77	4.86	1.67	

Loadings less than 0.25 have been deleted.

TABLE 7. *Factor analysis (oblique rotation) of the UPDRS*

Items		Factors						
N.	Name	I	II	III	IV	V	VI	Communality
23.	Finger tapping	0.91						0.78
24.	Hand movements	0.83						0.73
25.	Rapid alternating mov.	0.79						0.73
26.	Leg agility	0.61			0.29			0.70
11.	Hygiene	0.34	0.65					0.83
10.	Dressing	0.33	0.63					0.82
9.	Cutting food	0.30	0.61					0.81
18.	Speech			1.00				0.78
5.	Speech			0.62			0.31	0.78
29.	Gait				0.72			0.69
30.	Postural stability				0.71			0.83
27.	Arising from chair	0.32			0.54			0.69
20.	Tremor at rest					0.94		0.81
16.	Tremor					0.92		0.79
15.	Walking		0.38		0.39			0.72
14.	Freezing		0.31				−0.27	0.44
17.	Sensory complaints	0.29						0.32
8.	Handwriting			0.44				0.50
19.	Facial expression	0.30		0.47				0.68
13.	Falling			0.27	0.33			0.48
21.	Action tremor	0.31				0.47		0.49
22.	Rigidity	0.49						0.52
1.	Intellectual impairment		0.36					0.49
2.	Thought disorder							0.49
3.	Depression							0.38
4.	Motivation/initiative							0.45
12.	Turning in bed	0.25	0.44				0.43	0.68
28.	Posture	0.30			0.38			0.77
6.	Salivation		0.28				0.31	0.52
7.	Swallowing		0.42					0.51
31.	Body bradykinesia	0.41		0.25	0.38		0.35	0.81
	Eigenvalue	3.95	2.28	2.25	2.24	2.08	1.00	

Loadings less than 0.25 have been deleted.

same aspect of the construct) should be kept in mind to explain this figure. Internal consistency increases with the number of items and substantially depends on the homogeneity of the summed score in relation to the construct and on the interitem correlation.

To attain some basic information about these points, the item-total correlation and the interitem matrix of correlations were obtained. Those items related to depression (item 3), motivation/initiative (item 4), and tremor (items 16, 20, and 21) appear to be poorly related to the other aspects, in agreement with previous studies (1,5,13). Their inclusion can be questioned in scales designed to evaluate the functional severity of PD. On the other hand, these manifestations of the disease are properly considered by multidimensional scales with specific parts devoted to assess every aspect of the disease, such as the UPDRS.

Concerning the interrater reliability, the UPDRS should be considered a highly reliable scale. Only six items showed kappa coefficients lower than 0.50 (Table 2), but all of them were above 0.40 (moderate interrater agreement) (14). Probably the reason for these results, which are close to those found by Fahn et al. (2), is the adequate construction of the scale (number, order, and precise definition of the scoring ranks) (1,15–17). The previous interrater agreement to determine the patient's stage on the HYC is the most plausible explanation of the exceedingly high index observed in the item 30, Postural stability.

In relation to the validity aspects, some findings deserve comment. As the authors of the UPDRS pointed out, the scale can be completed in 10–20 min (2). In the present study, the upper ends of the range were 38 and 40 min (Groups A and B, respectively). These figures could limit the use of the UPDRS in clinical settings with a heavy burden of work. Repetition of parts of the scale may be required (e.g., the motor subscale) for multiple testing. In this way, the time of application is consid-

erably reduced, but the advantages of the UPDRS with respect to other PDRS are probably reduced, too. The significant difference observed between Groups A and B may be explained by the design of the study (five simultaneously acting raters, some of them inexperienced with PDRS, in Group A).

The criterion-related validity of the UPDRS has to be considered very satisfactory. A criticism could be made on the choice of the HYC as the gold standard, due to its relative insensitivity and unreliability (1,2,5,9,18). Nevertheless, it is the most used method of establishing the severity of the disease with a simple index. In addition, other investigators have verified the validity of different PDRS using the HYC as a comparison (5,17,19,20).

The objective timed tests had a variable correlation, good to moderate, with the total score of the UPDRS, except for "Walking a distance, back and forth" (Table 4). The existence of a subgroup of patients with a predominant, early gait disorder is a possible explanation of the results obtained with this modification of the "effective method devised by Webster" (10).

The finger-tapping test with the right hand correlated at a lower level ($p < 0.01$) than that with the left hand ($p < 0.001$). It has been previously found that the correlation is better between the patient's condition and the clinical expression of the disease on the left side of the body or with the alternating movements with the left hand (21,22). This finding may be related to the asymmetric onset and evolution of PD manifestations (21) or to a lesser ability of the nondominant hand to compensate for motor deficits. We do not have a solid explanation for this finding.

The fair discriminant validity of the UPDRS with regard to the MMSE and HSD (Table 5) has to be explained by the coincidence, at least partially, of their respective constructs. Certainly, the UPDRS includes items assessing mental status and mood. More important, the combination of PD with dementia or depression is frequent and an associated PD–dementia–depression has been proposed as a possible form of PD (23–26).

The convergent validity of the UPDRS with the ISAPD (1) was excellent (Table 5). Previously named NEV (Nueva Escala de Valoración) (5,22), the ISAPD was devised from other PDRS through a selection of items by statistical methods (22). Its validation will be published in the near future. It contains 13 items scoring from 0 to 3 and a section for evaluation of dyskinesias and fluctuations.

The correlation of the UPDRS with the SES was also very high, mainly in Group A (Table 5). The difference in the results obtained by Group A and Group B probably is a consequence of bias introduced from the Multicentric multiauthored branch.

The factor analysis identified six factors, pointing out the multidmensionality of the UPDRS (Table 6). Factors I, II, and III were related to mobility of the extremities, stability/gait/general mobility, and functional ability, respectively. Factor IV is an independent one, exclusively involving tremor. Bradykinesia, speech (item 5) and facial expression weighted on four factors.

Obviously, a comprehensive assessment of PD required a multidimensional scale. In previous studies, using different PDRS (13,17), three factors accounted for ~70% of the total variance. The factors "gait and balance" and "tremor" of these studies were equivalent in components to the Factors II and IV in the present study. The design and length of the UPDRS justify a more complex dimensional structure and also that only about 60% of the variance is explained by the six factors. An oblique rotation analysis was also carried out under the theoretical assumption (with no a priori judgment) that all the items of UPDRS were related to one single underlying condition, namely PD (Table 7). This approach modified the order of the factors and some components of Factor VI. A more defined clustering of variables within each factor was obtained, although the relation between some of them (e.g., freezing and intellectual impairment) is difficult to explain. The remarkable independence of the dimension "tremor" among the manifestations of PD also appears. This fact has been pointed out in previous studies (1,20,27).

In summary, the following points may be highlighted:

1. The internal consistency of the UPDRS is high (Cronbach's alpha = 0.96), although redundancy probably inflates this index.
2. The items related to depression and tremor show a fair internal consistency.
3. In agreement with previous studies (2), the interrater reliability is satisfactory.
4. The UPDRS may be applied in a reliable way, even by nonneurologists, if clear instructions and brief training are given, as previously recommended (1,28).
5. The length of the UPDRS may be problematic in some clinical settings, threatening its pragmatic validity, as suggested before (17). The use of

subscales of the UPDRS mitigates this problem.
6. Other aspects of validity (criterion-related and convergent validity) qualify the UPDRS as an adequate instrument to assess PD.
7. The discriminant validity in relation to the MMSE and the HSD is low.
8. The UPDRS is a complex, multidimensional scale, with six identified factors that explain about 60% of the variance.

Addendum

A very interesting paper by Baas et al. (29) has appeared since the submission of our manuscript. These authors have identified eight factors (66% of the total variance) characterizing the clinical profile of Parkinson's disease by a principal component analysis applied to a battery of rating scales and objective tests.

REFERENCES

1. Martínez-Martín P. Rating scales in Parkinson's disease. In: Jankovic J, Tolosa E, eds. *Parkinson's disease and movement disorders*. 2nd ed. Baltimore: Williams and Wilkins, 1993:281–292.
2. Fahn S, Elton RL, and Members of the UPDRS Development Committee. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Calne DB, Goldstein M, eds. *Recent developments in Parkinson's disease*, vol. 2. Florham Park, NJ: Macmillan Health Care Information, 1987:153–164.
3. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
4. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.
5. Martínez-Martín P, Bermejo Pareja F. Escalas de evaluación en la enfermedad de Parkinson. In: Giménez-Roldán S, ed. *Escalas de evaluación en enfermedad de Parkinson y trastornos del movimiento*. Barcelona: Editorial MCR, 1989:5–41.
6. LaRocca NG. Statistical and methodologic considerations in scale construction. In: Munsat TL, ed. *Quantification of neurologic deficit*. Boston: Butterworths, 1989:49–67.
7. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951;16:297–334.
8. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–382.
9. Martínez-Martín P, Carrasco de la Peña JL, Ramo C, Antigüedad A, Bermejo F. Estudio sobre la fiabilidad interobservador de escalas cualitativas en enfermedad de Parkinson (I). *Arch Neurobiol* 1987;50:309–314.
10. Lang AET, Fahn S. Assessment of Parkinson's disease. In: Munsat TL, ed. *Quantification of neurologic deficit*. Boston: Butterworths, 1989:285–301.
11. Parkinson Study Group. Correlates of clinical decline in early Parkinson's disease [Abstract]. *Neurology* 1987;37 (suppl. 1):278.
12. Jankovic J, McDermott M, Carter J, et al. Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort. *Neurology* 1990;40:1529–1534.
13. Reynolds NC, Montgomery GK. Factor analysis of Parkinson's impairment: an evaluation of the final common pathway. *Arch Neurol* 1987;44:1013–1016.
14. Longstreth WT, Koepsell TD, van Belle G. Clinical neuroepidemiology. I. Diagnosis. *Arch Neurol* 1987;44:1091–1099.
15. McKenzie CR, Charlson ME. Standards for the use of ordinal scales in clinical trials. *Br Med J* 1986;292:40–43.
16. Côté R, Battista RN, Wolfson CM, Hachinski V. Stroke assessment scales: guidelines for development, validation and reliability assessment. *Can J Neurol Sci* 1988;15:261–265.
17. Henderson L, Kennard C, Crawford TJ, et al. Scales for rating motor impairment in Parkinson's disease: studies of reliability and convergent validity. *J Neurol Neurosurg Psychiatry* 1991;54:18–24.
18. Marsden CD, Schachter M. Assessment of extrapyramidal disorders. *Br J Clin Pharmacol* 1981;11:129–151.
19. Lieberman A, Dziatolowski M, Gopinathan G, Kupersmith M, Neophytides A, Korein J. Evaluation of Parkinson's disease. In: Goldstein M, Calne DB, Lieberman A, Thorner MD, eds. *Ergot compounds and brain function: neuroendocrine and neuropsychiatric aspects*. New York: Raven Press, 1980:277–286.
20. Montgomery GK, Reynolds NC, Warren RM. Qualitative assessment of Parkinson's disease: study of reliability and data reduction with an abbreviated Columbia scale. *Clin Neuropharmacol* 1985;8:83–92.
21. Korten JJ. Correlations in 80 patients with Parkinson's disease. In: Lakke JPWF, Korf J, Wesseling H, eds. *Parkinson's disease: concepts and prospects*. Amsterdam: Excerpta Medica, 1977:101–110.
22. Martínez-Martín P. Parámetros evolutivos en la Enfermedad de Parkinson [Thesis]. Madrid, Spain: Universidad Complutense, 1987. 132 p.
23. Bermejo F, Enguita MA, Del Ser T. Depresión y demencia en la enfermedad de Parkinson. *Med Clin (Barc.)* 1986;87: 90–95.
24. Sano M, Stern Y, Williams J, Coté L, Rosenstein R, Mayeux R. Coexisting dementia and depression in Parkinson's disease. *Arch Neurol* 1989;46:1284–1286.
25. Starkstein S, Preziosi TJ, Berthier ML, Bolduc PL, Mayberg HS, Robinson RG. Depression and cognitive impairment in Parkinson's disease. *Brain* 1989;112:1141–1153.
26. Starkstein SE, Bolduc PL, Mayberg HS, Preziosi TJ, Robinson RG. Cognitive impairments and depression in Parkinson's disease: a follow up study. *J Neurol Neurosurg Psychiatry* 1990;53:597–602.
27. Zetusky WJ, Jankovic J, Pirozzolo FJ. The heterogeneity of Parkinson's disease: clinical and prognostic implications. *Neurology* 1985;35:522–526.
28. Geminiani G, Cesana BM, Tamma F, et al. Interobserver reliability between neurologists in training of Parkinson's disease rating scales. A multicenter study. *Mov Disord* 1991; 6:330–5.
29. Baas H, Stecker K, Fischer PA. Value and appropriate use of rating scales and apparative measurements in quantification of disability in Parkinson's disease. *J Neural Transm* [P-D Sect] 1993;5:45–61.