



EMILIA ŁOJEK<sup>1</sup>, JOANNA STAŃCZAK<sup>2</sup>, AGNIESZKA WÓJCIK,  
BERNICE MARCOPULOS<sup>3</sup>

<sup>1</sup> University of Warsaw

<sup>2</sup> Psychological Test Laboratory of Polish Psychological Association

<sup>3</sup> James Madison University, Harrisonburg

## VALIDITY AND RELIABILITY OF THE POLISH ADAPTATION OF THE RUFF FIGURAL FLUENCY TEST

Normative studies of the Polish adaptation of The Ruff Figural Fluency Test (RFFT) were conducted on 475 men and women aged 16-79, taking into consideration such factors as gender, education, and place of residence. Clinical studies were also performed on a group of patients with left-, right-, or bilateral hemispheric brain lesions, Parkinson's disease, Huntington's disease, progressive obturational lung disease, dementia and depression. The results support the utility of the RFFT as a measure of executive functions. The validity and reliability indices of the Polish version of the test are similar to those reported by Ruff (1996). However, the sample Polish test performance differs notably from American samples performance and this difference is discussed.

*Key words:* RFFT, Polish adaptation, validity, reliability

### Introduction

#### Theoretical rationale

Fluent planning and monitoring of one's own behaviour is an indicator of the effectiveness of supervision and control of the cognitive processes (Ruff et al., 1997). Fluency can be either verbal or nonverbal. Verbal fluency is related to the processes which are responsible for verbal production and reception whereas nonverbal fluency is related to visual-spatial and psycho-motor abilities. Verbal

fluency is normally assessed by methods requiring the respondent to utter as many words as possible from a given category within a fixed time limit. The most widely-used measure of verbal fluency in the world is the Controlled Oral Word Association Test – COWAT (Borkowski et al., 1967).

Nonverbal fluency can be assessed using several different measures. For example, in the gesture tests, nonverbal fluency is measured by the number of finger arrangements in new positions (Lezak et al., 2004); whereas, in the performance test it is calculated by the number of initiated actions or activities (Christensen et al., 1999; Lancioni et al., 2001). However, the majority of clinical studies of nonverbal fluency focuses on design fluency. In such studies, the assessment procedures (which may differ depending on the specific method) require respondents either to produce completely thematically unique drawing compositions or to produce drawings based on a presented graphic design. The ability to produce original patterns on the basis of a unified graphic design is called figural fluency and can be assessed by the Ruff Figural Fluency Test (RFFT; Ruff, 1996).

Construction of the RFFT was inspired by the following tests: the *Five-point Test* (Regard et al., 1982; Ruff et al., 1987) and its prototype, the *Design Fluency Test* (Jones-Gotman & Milner, 1977). In both these tests the respondent must produce, within the given time limit, as many of the following: drawings representing real objects or abstract forms and/or figures created by joining points with the help of straight or curved lines. All of the nonverbal methods (Jones-Gotman and Milner's, 1977, and Ruff's et al., 1987), discriminate well between normal and brain-damaged individuals.

The RFFT is sometimes preferred because it discriminates better than the other fluency tests between patients with frontal and temporal lesions (Lezak et al., 2004; Ruff et al., 1994; Spreen & Strauss, 1998; Suchy et al., 2003). Studies using the RFFT have revealed major neurocognitive deficits in patients who have had traumatic brain injuries (Ruff, 1996), various degenerative diseases of the brain (Fama et al., 1998; Łojek & Bornstein, 2005) and psychiatric disorders (Ross et al., 2003; Williamson & Harrison, 2003).

### **A general overview of the RFFT**

The RFFT is an individually administered, paper-and-pencil test suitable for a wide age range (16 to 79-year-olds). The test booklet contains five white 20 x 30 cm boards showing the trial problems and test problems. Before moving to each of the test problems, the respondent is given a trial problem to familiarise him or her with the assignment that is to follow. The boards with the actual test problems contain 35 squares evenly distributed between seven rows. Each square is a separate task that contains a composition of five dots. In parts two and three of the test, these compositions are accompanied by interfering stimuli. The following interfering stimuli are used: symmetrical diamonds (part two) and asymmetrical lines connecting dots from within the square and neighbouring

squares (part three). The respondent must produce a connective or figural design by connecting at least two dots in the square with a straight line. Each new production must be different from the remaining productions. The respondent is given exactly 60 seconds to complete each task.

The intermediate score for each part of the test gives the total number of non-repetitive (original) connections and designs and the total number of repetitive designs. These scores for each part are added up to give the raw scores for the two indicators: Unique Designs (UD) and Perseverative Errors (PE). A third indicator, Error Ratio (ER), expresses the ratio of repetitions to unique designs (PE/UD). In each part of the test, the maximal UD score is 35 while the maximal PE score is 34 (there is at least one original design). The ER score can range from 0 to 34 in each part of the test. *If*  $ER = 0$ , it means that no design has been repeated; whereas, *if*  $ER = 34$ , then there is one original design that has been repeated 34 times. The higher the ER score, the more unsatisfactory the test performance. The strategies adopted to produce consecutive designs can also be analysed to obtain additional information about the respondent's performance (Ross et al., 2003; Ruff, 1996).

### **Adaptation of the RFFT to a Polish population**

Since the RFFT is a nonverbal test, the test material did not need to be modified and is identical to the original. The testing and scoring procedures are also the same as in the original. Hence, the Polish adaptation of the RFFT contains the original graphic material for all five tasks in the performance part of the test and retains aforementioned details in the research tools, such as the board measurements and the title page. The test instructions have been translated into Polish<sup>1</sup>, and great care has been taken to render a verbatim translation as much as possible, without compromising the correctness of Polish phrasing (Łojek & Stańczak, 2005).

### **Properties of the Polish adaptation of the RFFT**

Studies on the Polish adaptation of the RFFT included: elaborating norms for the Polish population, establishing the test-retest and inter-rater reliability, measuring test validity using factor analysis as well as correlations with other psychological tests, and diagnostic validity.

### **Normative study**

The normative study has been conducted by 32 qualified psychologists specially trained to do the recruitment and assessment. The examinees looked for and directly contacted persons who fulfilled demographic requirements.

---

<sup>1</sup> Since the instruction for the RFFT is short and simple the back translation procedure has not been applied.

Subjects with a positive history of psychiatric hospitalisation, drug or alcohol abuse, or neurological disorders were excluded. Finally, 475 healthy subjects (134 females and 118 males) aged 16-79 participated in the normative study. The sample is representative of the Polish population according to gender, age, education, place of residence (cities, towns and villages), and different school profiles (in the case of adolescents). The sample parameters were elaborated according to the official census published in the Statistical Yearbook for Poland (Rocznik Statystyczny, 2002) and the proportions of the various criteria in

Table 1. Sample characteristics: size and age

Group purpose and specifics	<i>N</i>	Group/Size	Age
Standardisation Norms Factor analysis	475	F 134	16-54
		F 118	55-79
		M 141	16-54
		M 82	55-79
<i>Reliability:</i>			
1. Test-retest reliability	67	25	16-54
		42	55-79
2. Inter-rater reliability	60	60	16-79
<i>Construct validity correlation with:</i>			
1. WAIS-R	444	257	16-54
		187	55-79
2. RPM-C	40	40	40-75
3. The Squares Test	32	32	40-75
4. VCAT	76	76	40-75
5. TMT Part A	433	262	16-54
		171	55-79
6. TMT Part B	404	255	16-54
		149	55-79
7. CVLT	473	274	16-54
		199	55-79
8. WCST Individuals with cortical lesions	88	F 41	17-75
		M 47	
9. UMACL	449	263	16-54
		186	55-79

F – females

M – males

the final sample are representative of the respective proportions in the Polish population.

Test-retest reliability was estimated on a random sub-sample that reflected age, level of education, gender, and place of residence of the normalisation sample. The validation data were gleaned from studies conducted on the standardization sample, incidental samples, and clinical groups. In selected cases, the sample included additional participants with various levels of education, who had been studied by students at the Faculty of Psychology at Warsaw University. These were healthy men and women without neurological or psychiatric disorders. Men and women were equally represented.

### **Clinical samples**

The utility of the RFFT for clinical diagnosis was tested in over a dozen groups of patients with cerebral dysfunction of various localisation and aetiologies. The patients were examined during their stay in medical centres (neurological units of the clinical hospitals in different Polish cities). Selection was based on the following criteria: the first occurrence of a neurological disease of the central nervous system, no psychiatric disorders or substance dependence, no episodes of loss of consciousness before the preset brain disease, aphasia or right-side paralysis. The cortical brain-lesion sample was a heterogeneous sample represented by the following variables: location of the brain damage—right-side ( $N = 75$ ), left-side ( $N = 43$ ), both hemispheres ( $N = 6$ ), frontal ( $N = 15$ ), other than frontal ( $N = 39$ ), all cerebral lobes ( $N = 18$ ). Exact information concerning localisation of the brain damage was missing for 43 patients. Sixty-seven (67) percent of the respondents were right-handed while data on handedness were missing in 27% of cases. Additional information was obtained concerning duration of illness, the aetiology of the damage (vascular, traumatic, neoplastic) and the subject's education. Patients with diagnosed dementia ( $N = 22$ ), Huntington's disease ( $N = 19$ ), and Parkinson's disease ( $N = 15$ ) were also tested. The vast majority of these patients were right-handed (96%). Other nosologic categories included chronic obturational lung disease ( $N = 31$ ) and diagnosed depression ( $N = 59$ ). This group was heterogeneous in terms of duration of illness, advancement of illness, age, education, and gender.

### **Psychometric properties of the Polish adaptation of the RFFT**

The normalisation study revealed the following patterns:

- Level of RFFT performance does not depend on gender;
- UD and ER scores change with age. Respondents under 54 years of age produce more original designs and make fewer perseverative errors than older respondents;
- Level of RFFT performance is related to the level of education and school profile. Respondents with higher education and high school pupils score highest on the test;

- Place of residence may affect fluidity in respondents under 55 years of age. Respondents who live in cities with a population of at least 100 thousand have significantly higher scores than respondents who live in rural areas, independently from educational variables.

The present findings are generally consistent with the findings reported by Ruff (1996). Both Polish and American analyses for the standardized samples showed that gender has no effect on RFFT scores but age and education do. The present study also found that RFFT score variance is related to place of residence and that residents of large cities do better on the test.

Due to the significant difference in performance between under-55s and respondents aged from 55 to 79, the remaining analyses of the basic psychometric properties of the RFFT were conducted separately for each of these two groups (275 respondents aged 16-54 and 200 respondents aged 55-79).

## Reliability

Test-retest reliability was established six months after the original testing, similarly to the Ruff's reliability studies. Test-retest reliability was estimated on a random sub-sample of the normative sample (67 subjects, aged 16-79).

Analysis of the results revealed differential absolute stability of the various RFFT scores. Test-retest statistics for the Polish normalisation sample and the UD and PE correlation coefficients (Pearson's and Spearman's correlation) are presented in Table 2. When the entire sample is considered, the correlation values are satisfactory for UD and suggest high stability of scores in the under-55 group

Table 2. Sample statistics for UD and PE and the values of Pearson's  $r$  linear correlation and  $R$ . Spearman's rank correlation coefficients in the Polish normalisation sample<sup>2</sup>

Group	Age	N	Test	Unique Designs			Perseverative Errors		
				<i>M</i>	<i>SD</i>	<i>R</i>	<i>M</i>	<i>SD</i>	<i>r/R</i>
Selected individuals from the Polish normalisation sample	16 - 54	25	1	71.76	26.19	0.83	4.92	3.98	0.35*
			2	88.80	31.86		10.56	10.58	
	55 - 79	42	1	35.29	12.13	0.50	4.24	4.82	0.35
			2	40.57	20.56		6.33	7.62	

\* correlation coefficient statistically non-significant; *M* – mean, *SD* – standard deviation  
 $r$  – Pearson linear correlation coefficient;  $R$  – Spearman rank correlation coefficient  
 Test: 1 – first administration of the RFFT; 2 – second administration of the RFFT

<sup>2</sup> Only the younger age group's UD scores had a quasi-normal distribution and therefore some of the analyses had to be done using nonparametric tests.

and much less stability in the 55-79 group. The correlations for PE and ER are low in both age groups. A similar pattern emerged in the American sample: high UD stability and lower PE stability (Ruff, 1996).

In order to examine interrater reliability, three independent judges scored 60 randomly selected test protocols. This procedure is important when dealing with tests in which there is no one correct answer and the test may be solved in many different ways. The values of the Kendall W consistency coefficients were high (over 0.94) for all the RFFT indices. The negligible effect of subjective factors on RFFT performance has also been reported by Sands (1998) and Ross et al. (2003), although Berning et al. (1998) reported slightly lower interrater reliabilities for PE (0.74) and ER (0.66).

## Validity

The procedure testing the construct and discriminant validity of the RFFT involved correlation analyses (including factor analysis) and comparison of healthy and brain-damaged individuals. Differential analysis in a clinical sample was also conducted.

## Factor analysis

In order to check the validity of the RFFT structure (i.e., to determine whether the test does indeed measure the abilities it is purported to measure), the data were factor analysed using the principal components method with Varimax rotation. The factor analysed test variables are summarised in Table 3. The factor analysis of the Polish data yielded a four-factor solution for the three factors that Ruff (1996) indicated: general intelligence, arousal, and planning. The Polish factor structure also differs from the original factor structure. Factor One in the Polish study comprises the following results: UD, all the Nonverbal Scale tests except Digit Symbol of the Wechsler Adult Intelligence Scale (WAIS-R(PL); (Brzeziński et al., 2004), Parts A and B of the Trail Making Test (TMT; Kądziaława et al., 1990) and the number of incorrect word recognitions on the California Verbal Learning Test (CVLT; Delis et al., 1987, Polish adaptation by Łojek & Stańczak, 2009)<sup>3</sup>. This factor may be described as a clear-cut visual-spatial factor, strongly saturated with attention concentration and ability to switch from one category to another and to plan by initiating a sequence of steps necessary to solve a problem. Variables relating to memory and work readiness also contribute to this factor but not so strongly (factor loadings under 0.50 or no higher than 0.30). This factor may be associated with planning and arousal. It is not exactly reflected in Ruff's (1996) analyses in which these variables were loaded with general intelligence and arousal.

---

<sup>3</sup> The Polish adaptations of the CVLT and VCAT were developed simultaneously with the Polish adaptation of the RFFT and the two tests were submitted to analogous, extensive normalisation and standardisation procedures.

Table 3. RFFT factor loadings in the Polish study

Analysed tests	Factor loadings*			
	1	2	3	4
<i>RFFT</i>				
Unique Designs	0.66	0.34		
Error Ratio				-0.60
<i>WAIS-R (PL)</i>				
Information		0.82		
Comprehension		0.78		
Vocabulary	0.37	0.77		
Similarities	0.45	0.72		
Arithmetic		0.68		
Digit Symbol		0.60		
Digit Span	0.41	0.43	0.36	
Picture Completion	0.68	0.43		
Block Design	0.67	0.40	0.31	
Object Assembly	0.66	0.31		
Picture Arrangement	0.65	0.32		
<i>Trail Making Test</i>				
Part A				
Part B				
<i>CVLT</i>				
False positives	-0.53			
Long-Delay Recall (Free + Cued)	0.40		0.84	
Short-Delay Recall (Free + Cued)	0.37		0.82	
List A Total Recall, Trials 1-5	0.45		0.79	
Recognition Hits			0.72	
Perseverations				-0.70

\* only loadings exceeding 0.30 are shown

ER in the RFFT correlates with perseverative errors in the Total Perseverative Errors score of the CVLT, suggesting that this factor reflects self-control and response self-monitoring processes. In Ruff's study (1996), error indices were part



of a factor that the author called Planning. Tests that were not included in Factor 1 made up Factor 2. Factor's 2 which structure corresponds with the general intelligence variables identified by Ruff. Factor 3 in the Polish analyses contains the remaining memory-related CVLT indicators.

### Correlation with other tests

Construct validity was assessed by correlating the same respondents' RFFT scores with level of performance on other measures of cognitive processes.

The following measures were used: the WAIS-R(PL), the CVLT, the TMT parts A and B, the Raven's Progressive Matrices Standard – Classic (RPM-C; Jaworowska & Szustrowa, 2000), the Squares Test (measuring imagining the rotation of geometric figures) included in the the APIS-Z Test Battery (a Polish original test battery measuring cristalized intelligence, Matczak, Jaworowska, Szustrowa & Ciechanowicz, 1995), the Verbal Concept Attainment Test (VCAT, Bornstein, 1983; Polish adaptation by Łojek & Stańczak), the Wisconsin Card Sorting Test (WCST; Jaworowska, 2002), and the University of Wales Institute of Science and Technology Mood Adjective Check List (UMACL, Goryńska, 2005)<sup>4</sup>.

The validation studies tested the hypothesis that level of RFFT performance would be reflected in the scores of tests involving the same cognitive processes as in the RFFT. It was predicted that RFFT scores would co-vary with performance on criterion tests, engaging such cognitive aspects as general intelligence (measured by the WAIS-R, SPM), executive functions (the WCST, VCAT, and TMT), performance speed, attention, visual-spatial abilities (the performance tests of the WAIS-R, and the Squares Test from the APIS-Z Test Battery) and working memory (the CVLT – immediate recall). The analyses supported this prediction. No matter what the respondents' age, RFFT performance involves visual-spatial abilities represented by the Squares Test ( $r = 0.42$ ), and level of RFFT performance is related to the ability to categorise material as measured by the VCAT ( $r = 0.62$  with Unique Designs and  $r = 0.42$  with Error Ratio). The positive correlation between Unique Designs and all the WAIS-R(PL) tests (from  $r = 0.33$  to  $r = 0.47$  in the younger group and from  $r = 0.37$  to  $r = 0.55$  in the older group) and with the Verbal, Nonverbal and Full scales (from  $r = 0.42$  to  $r = 0.51$  in both groups) suggest that fluidity, as measured by the RFFT, is related to general intelligence. The relation between level of RFFT performance and intellectual functioning is further supported by the low negative correlations in the younger group between the two RFFT error scores and some of the WAIS-R(PL) tests, i.e., those that engage memory and attention ( $r = -0.13$  and  $r = -0.14$  for PE and  $r = -0.14$  to  $r = -0.28$  for ER).

<sup>4</sup> The reliability coefficients for the Polish adaptations of neuropsychological tests used in the study: the WCST – interrater reliability  $r = 0.96-1.00$ ; internal consistency RPM-C  $r = 0.89-0.97$ ; UMACL  $r = 0.71-0.90$ ; the Squares Test from the APIS-Z battery  $r = 0.77-0.83$ ; WAIS-R Full Scale  $r = 0.88-0.93$ ; VCAT  $r = 0.84$ ; CVLT  $r = 0.83-0.93$ ; TMT – reliability not measured.

The relations between measures of fluidity and fluid intelligence ( $r = 0.85$ ), as measured by the RPM-C and between fluidity on the one hand and psychomotor speed and attention switching from one symbol to another needed to solve the TMT (from  $r = -0.31$  to  $r = -0.41$ ), were confirmed. The analyses also suggest that the RFFT involves similar processes to the ones that are involved in performance of the WCST, i.e., executive functions. This is suggested by the strong negative correlation between UD and the WCST error scores (from  $r = -0.25$  to  $r = -0.54$ ) and the positive correlation between UD and Number of Passed Categories on the WCST ( $r = 0.50$ ). In both age groups, the UD scores correlated positively with CVLT correct performance ( $r = 0.23$  and  $r = 0.31$ ) and negatively (albeit not always significantly) with CVLT errors (from  $r = -0.13$  to  $r = -0.25$ ), suggesting that correct RFFT performance engages working memory. Analysis of the relation between RFFT performance and mood measured by the UMACL suggests that sense of well-being and work arousal in older respondents ( $r = 0.25$ ) and anxious tension in younger respondents ( $r = 0.12$ ) may be related to RFFT performance.

### Diagnostic validity

To check the utility of the RFFT for clinical diagnosis, each clinical group was paired with a control group matched for gender, age, and education, recruited from the normalisation sample. This part of the study tested the hypothesis that brain-damage-related pathological symptoms are reflected in the level of RFFT performance. The findings are summarised in Tables 4 and 5.

In the Ruff et al. (1986) study, the RFFT UD and ER scores discriminated between healthy subjects and patients with traumatic brain damage. In the Polish study, the RFFT significantly discriminated between healthy subjects and brain-damaged subjects on one dimension only, i.e., Unique Designs (cf. Table 4). The lack of difference in Perseverance Errors was also found by Ruff et al. (1986). When analysing similar finding of Ruff et al., Lezak et al. (2004) suggest that too large within-group variance might be the cause.

As shown in Table 4, there were significant differences between the results of healthy respondents and patients with right-hemisphere lesions or lesions of non-frontal areas. There were no significant differences between patients with left-hemisphere lesions and healthy respondents. Patients with frontal-lobe brain damage had a clear tendency to produce fewer Unique Designs compared with healthy respondents and to make more Perseverative Errors ( $p = 0.06$ ), suggesting that this group has deficient executive functions.

The results presented in Table 5 also show significant differences in planning efficiency between patients with dementia and Huntington's disease on the one hand and healthy respondents on the other hand. Likewise, we can see that self-control and behaviour monitoring are reduced in patients with Parkinson's disease and patients with chronic obstructive lung disease ( $p = 0.07$ ). The RFFT

also discriminates well between patients with dementia and patients with depression. Patients with diagnosed dementia have significantly lower UD and ER

Table 4. Mean RFFT scores for brain-damaged and healthy respondents

Compared groups	N	RFFT					
		UD		PE		ER	
		M	SD	M	SD	M	SD
All CNS-damaged respondents	115	46.46	21.47	7.37	9.42	0.16	0.17
Healthy respondents	108	58.38	22.72	9.09	13.22	0.17	0.70
<i>t</i>		4.02**		-1.12		-0.32	
Group A	75	43.89	19.98	7.37	9.67	0.17	0.16
Healthy respondents	108	58.38	22.72	9.09	13.22	0.17	0.27
<i>t</i>		-4.45**		-0.96		-0.10	
Group B	34	51.18	23.32	8.00	9.47	0.16	0.18
Healthy respondents	108	58.38	22.72	9.09	3.22	0.17	0.27
<i>t</i>		-1.60		-0.45		-0.18	
Group C	6	44.33	32.83	3.00	2.83	0.55	1.20
Healthy respondents	6	58.00	25.82	11.50	11.41	0.19	0.15
<i>t</i>		-0.80		-1.77		0.73	
Group 1	18	57.39	24.67	7.17	5.85	0.14	0.16
Healthy respondents	18	59.17	19.17	10.56	13.31	0.20	0.27
<i>t</i>		-0.24		-0.99		-0.74	
Group 2	15	45.80	23.23	8.87	13.02	0.37	0.76
Healthy respondents	14	56.07	22.44	6.07	6.16	0.12	0.13
<i>t</i>		-1.21		0.73		1.22	
Group 3	39	39.69	14.98	5.38	4.52	0.14	0.11
Healthy respondents	96	58.38	22.75	9.38	13.83	0.18	0.28
<i>t</i>		-4.72**		-1.76		-0.85	
Groups A-B	109	<i>t</i> = -1.67		<i>t</i> = -0.32		<i>t</i> = 0.15	
Groups 2-3	54	<i>t</i> = 1.14		<i>t</i> = 1.47		<i>t</i> = 1.88	

Group A – respondents with right-hemisphere lesions  
 Group B – respondents with left-hemisphere lesions  
 Group C – respondents with lesions in both hemispheres  
 Group 1 – respondents with lesions in all cortical lobes  
 Group 2 – respondents with lesions in the frontal lobes  
 Group 3 – respondents with lesions in lobes other than the frontal lobes  
 UD – Unique Designs; PE – Perseverative Errors; ER – Error Ratio  
 M – mean, SD – standard deviation; *t* – Student-t statistics  
 \*\* significant at < 0.01

Table 5. Comparison of RFFT means for patients with brain deficits of various aetiologies and healthy respondents

Compared groups	N	RFFT					
		UD		PE		ER	
		M	SD	M	SD	M	SD
Group D	22	26.05	12.55	5.55	6.40	0.23	0.24
Healthy respondents	18	44.00	18.12	3.78	3.72	0.09	0.07
<i>t</i>		-3.69**	1.04	2.39*			
Group E	23	40.35	12.30	4.09	2.98	0.11	0.09
Healthy respondents	24	42.96	16.16	2.25	1.84	0.06	0.06
<i>t</i>		-0.54	2.55*	2.07*			
Group F	14	46.86	22.56	13.64	10.90	0.43	0.50
Healthy respondents	15	65.80	15.09	8.27	8.68	0.14	0.17
<i>t</i>		2.68*	-1.47	-2.11*			
Group G	31	45.55	16.20	10.90	13.14	0.26	0.32
Healthy respondents	31	47.13	17.83	6.03	7.02	0.14	0.19
<i>t</i>		-0.37	1.82	1.71			
Group H	59	53.15	20.79	6.54	11.30	0.14	0.13
Healthy respondents	56	60.57	24.70	5.82	18.12	0.21	0.36
<i>t</i>		-1.75	-1.92	-1.41			
Group H1	16	51.25	18.50	5.63	4.33	0.12	0.11
Healthy respondents	56	61.05	24.66	11.36	18.28	0.21	0.36
<i>t</i>		-1.40	-1.24	-1.01			
Group H2	13	58.69	20.09	6.00	4.71	0.11	0.11
Healthy respondents	56	61.05	24.66	11.36	18.28	0.21	0.36
<i>t</i>		-0.26	-1.04	-0.94			
Group H3	7	41.57	15.35	10.14	8.73	0.24	0.16
Healthy respondents	7	44.43	19.87	4.86	4.88	0.14	0.19
<i>t</i>		-0.30	1.40	1.12			
Groups D – H	118	$t = -5.72^{**}$	$t = -0.67$	$T = 2.14^{*}$			
Groups D – H1	75	$t = 5.01^{**}$	$t = 0.04$	$T = -1.69$			
Groups D – H2	72	$t = 5.94^{**}$	$t = 0.22$	$T = -1.60$			
Groups D – H3	66	$t = 2.71^{*}$	$t = 1.52$	$T = 0.15$			

Group D – patients with dementia; Group E – patients with Parkinson's disease;  
 Group F – patients with Huntington's disease; Group G – patients with chronic obstructive pulmonary disease;  
 Group H – patients with depression (H1 – with unipolar depression, H2 – with bipolar depression,  
 H3 – with vascular depression; data on the type or/and etiology of depression are missing for 23 respondents)

UD – Unique Designs; PE – Perseverative Errors; ER – Error Ratio

M – mean, SD – standard deviation; *t* – Student-t statistics

\*\* significant at  $< 0.01$ , \* significant at  $< 0.05$

scores. This finding may greatly contribute to the advancement of differential diagnosis. Severity of dementia assessed with the Mini Mental State Examination (MMSE; Sabe, Jason, Juejati, Leiguarda & Starkstein, 1993) correlates with RFFT scores in patients with dementia. Significant correlations were found between the MMSE on the one hand and UD ( $r = 0.76, p < .0001$ ) and ER ( $r = -0.48, p = 0.024$ ) on the other hand.

To conclude, clinical studies have demonstrated that the RFFT is a valuable assessment instrument that discriminates between healthy and brain-damaged individuals. The patients were included to the clinical groups on the basis of medical data. Patients with right-hemisphere brain lesions, diagnosed dementia (probable Alzheimer's type dementia, multiinfarct dementia and dementia of not diagnosed type, Huntington's disease, and Parkinson's disease significantly lower RFFT scores than healthy controls. A tendency towards inferior RFFT performance was found for patients with lesions of the frontal cortex and chronic obturational lung disease (in this latter group probably because of the brain hypoxia).

### **Comparison of Polish and American populations**

Comparison of the two populations was based on analysis of the UD means for the Polish population (Table 6) and the American population (Ruff, 1996). The standard deviation (SD) was the comparison criterion. We wanted to know how different the mean American and Polish scores were in terms of SD units. We assumed that the two populations performed similarly if the difference between the means for the comparison groups did not exceed  $\frac{1}{2}$  SD.

When age and education were considered, the following results were obtained:

#### *Pupils (aged 16-19)*

- Compared with their American peers the Polish subjects scored lower by over one-and-a-half SD units.

#### *Adult group aged 8-24 (no longer studying)*

- Polish respondents with secondary education (12-15 years of schooling) had scores  $\frac{3}{4}$  SD units lower than their American peers with an equivalent level of education.
- Polish respondents with vocational education (10 years of schooling) and primary education (8 years of schooling) had scores lower by  $1\frac{3}{4}$  SD and  $\frac{1}{2}$  SD respectively than the youngest American group (aged 16-24) presented by Ruff who had no more than 12 years of schooling.

#### *Adult group aged 25-54*

- Polish respondents with higher education (16-18 years of schooling) had similar results to the results of their American peers with a minimum of 16 years of schooling.
- Polish respondents with secondary education had results similar to the results of their American peers who had no more than 12 years of schooling.

Table 6. Means and standard deviations for Unique Designs and Perseverative Errors for groups selected by age, education and school profile

Group <i>N</i> = 475	<i>N</i>	Age	Education/School profile	Unique Designs		Perseverative Errors	
				<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pupils	32		High school	85.81	23.65	6.47	11.07
	34	16-19	Technical college	78.00	19.77	4.94	6.28
	10		Basic vocational school	64.90	26.86	4.60	5.19
	25		Secondary	84.20	26.01	4.96	4.42
	23	18-24	Vocational	68.26	20.10	5.43	6.74
	15		Primary	58.20	28.23	6.87	10.78
15		Higher	98.53	22.43	8.40	11.55	
Adults not currently studying	48	25-54	Secondary	77.44	23.98	6.73	6.59
	48		Vocational	70.96	18.48	11.98	16.74
	26		Primary	59.50	24.23	10.62	10.99
	12		Higher	53.17	21.12	13.92	29.12
	37	55-79	Secondary	51.57	17.15	7.49	9.11
	15		Vocational	44.47	16.63	10.20	9.66
	136		Primary	37.63	12.70	6.51	8.50

T – pupils studying at profiled high schools, vocational high schools and technical schools

*N* – sample size; *M* – arithmetical mean, *SD* – standard deviation

- Polish respondents with vocational and primary education had scores more than  $\frac{1}{2}$  and  $\frac{3}{4}$  SD lower than the scores of American peers who had no more than 12 years of schooling.

#### *Adult group aged 55-70*

- Polish respondents with higher education had scores 1SD lower than their American peers who had at least 16 years of schooling.
- Compared with Americans aged 55-70 who had no more than 12 years of schooling Polish respondents in the same age group who had secondary education had lower scores by 1 SD, vocational education –  $1\frac{1}{2}$  SD lower, and primary education –  $2\frac{1}{2}$  SD lower.

The foregoing comparison suggests that only respondents aged 25-54 with higher education have similar UD scores in the two populations. In all the remaining age and educational groups, Poles have lower scores than Americans, but we must remember that it was not always possible to match the two populations perfectly in terms of years of schooling. A review of knowledge and skills of

young adults representing 49 countries within the Programme for International Student Assessment (OECD, 2004) has shown that Polish and American students have similar performance levels for problems requiring the ability to apply basic rules and to transform the problem (broad comprehension), and neither group does generally better. Hence, the present findings are all the more bewildering.

It is also possible that lack of training is responsible for the fact that Poles, especially less educated and older Poles, produce fewer original RFFT designs than Americans. Someone who already has some experience with a situation has greater attention resources, and his or her working memory is less taxed (Sternberg, 1986). The foregoing observations lead to the conclusion that performance of even nonverbal neuropsychological tests, such as the RFFT, may be culturally determined. Hence, it would be extremely risky to adopt norms and cut-off points for pathology developed for foreign populations in domestic clinical practice (Rosselli et al., 2003).

In Ruff's (1996) study, participants were from Michigan, California and East Coast regions. Ruff does not specify whether these participants are from urban or rural areas. In our study, we found significant differences in RFFT performance between residents of large cities and villages, irrespective of their level of education. A similar pattern of neuropsychological performance also exists in the USA (e.g., Kaufman et al., 1988) and other countries (Fernandez & Marcopulos, 2008) but that effect has not been considered in the normalisation sample by Ruff in the RFFT manual.

## Concluding remarks

The purpose of this article was to present the results of the Polish RFFT normative study. We found that the psychometric parameters (reliability and validity) of the RFFT are good and similar to those obtained for the American version of the test. We have demonstrated significant correlations between the RFFT and variables representing different levels of cognitive functioning: fluid intelligence, general intelligence, and verbal memory. We have confirmed the utility of the RFFT as a measure of executive functions. Patterns of correlations between the RFFT and several measures of executive functions, i.e., the Wisconsin Card Sorting Test, the Trail Making Test, and the California Verbal Learning Test, support the validity of the RFFT as a measure of executive functions. Two RFFT indices, UD and ER, reflect the involvement of planning functions, behaviour initiating, and behaviour monitoring postulated by Ruff's (1996) figural fluidity concept. All the correlation analyses confirmed the dominant role of UD in the interpretation of RFFT scores in that this scale correlated more strongly with other variables than ER. This suggests that both indices may provide important diagnostic information, but planning and initiating capacities are more closely related to cognitive functions.



Our analyses confirmed the value of the RFFT for discrimination between healthy individuals and patients with right-hemisphere lesions and the sensitivity of the PE score to the test performance of patients with frontal lobe dysfunction. The RFFT also discriminates between patients with dementia or Huntington's disease and healthy individuals. Patients with mild to moderate dementia have significantly more trouble solving the RFFT than patients with clinically diagnosed depression.

The validity and reliability parameters of the Polish adaptation of the RFFT are generally comparable to the analogous parameters of the original. Whatever the reason for this similarity, it still needs to be explained why Polish respondents in general, and less educated and older Polish respondents in particular, produce much fewer Unique Designs than American respondents (Fernandez & Marcopulos, 2008).

## Acknowledgements

The normative studies were funded by the Psychological Test Laboratory of the Polish Psychological Association and the validation studies were funded by the Faculty of Warsaw, University of Warsaw (Grants numbers: 2000 – BW 1492/7, 2001 – BW 1536/1, BST 714/12, 2002 – BST 773/14, 2003 – BST 843/27, 2004 – BST 986/10).

## References

- Anastasi, A. & Urbina, S. (1999). *Testy psychologiczne (Psychological Tests)*. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Berning, L.C., Weed, N.C., & Aloia, M.S. (1998). Interrater reliability of the Ruff Figural Fluency Test. *Assessment*, 5 (2), 181-186.
- Borkowski, J., Benton, A., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia*, 5 (2), 135-140.
- Bornstein, R.A. (1983). Verbal Concept Attainment Test: Cross-validation and validation of a booklet form. *Journal of Clinical Psychology*, 39 (5), 743-745.
- Brzeziński, J. (1984). *Wybrane zagadnienia z psychometrii i diagnostyki psychologicznej [Selected Issues in Psychometrics and Psychological Diagnostics]*. Poznań: UAM.
- Brzeziński, J., Gaul, M., Hornowska, E., Jaworowska, A., Machowski, A., & Zakrzewska, M. (2004). *Skala Inteligencji D. Wechslera dla Dorosłych. Wersja Zrewidowana – renormalizacja WAIS-R(PL) [The Wechsler Adult Intelligence Scale – Revised. Re-Normalisation of the WAIS-R(PL)]*. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Christensen, H., MacKinnon, A.J., Lorten, A.E., Jorm, A.F., Henderson, A.S., Jacomb, P., & Rodgers, B. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: Individual differences in change scores as a function of age. *Psychological Aging*, 14 (3), 365-379.



- Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test CVLT. Research Edition. Adult Version. Manual*. San Antonio, TX: The Psychological Corporation.
- Fama, R., Sullivan, E.V., Shear, P.K., Cahn-Weiner, D.A., Yesavage, J.A., Tinklenberg, J.R., & Pfefferbaum, A. (1998). Fluency Performance Patterns in Alzheimer's Disease and Parkinson's Disease. *Clinical Neuropsychologist*, 12 (4), 487-499.
- Fernandez, A.L., & Marcopulos, B. (2008). A comparison of normative data for the Trail Making Test from several countries: Equivalence of norms and considerations for interpretation. *Scandinavian Journal of Psychology*, 49 (3), 239-246.
- Goryńska, E. (2005). *Przymiotnikowa Skala Nastroju UMACL. Podręcznik [The UMACL Manual]*. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Jaworowska, A. (2002). *Test Sortowania Kart z Wisconsin. Podręcznik [The Wisconsin Card Sorting Test Manual]*. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Jaworowska, A. & Szustrowa, T. (2000). *Test Matryc Ravena w wersji Standard. Formy: Klasyczna, Równoległa, Plus. Polskie standaryzacje. [The Raven Progressive Matrices – Standard. Classic, Parallel and Plus Forms. Polish Standardisations]*. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Jones-Gotman, M. & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal cortical lesions. *Neuropsychologia*, 15 (4-5), 653-674.
- Kądzielawa, D., Mroziak, J., Bolewska, A., & Osiejuk, E. (1990). *Podręcznik do Baterii Testów Halsteda-Reitana [The Halstead-Reitan Test Battery Manual]*. Warszawa: Wydział Psychologii UW.
- Kaufman, A.S., McLean, J.E., & Reynolds, C.R. (1988). Sex, race, residence, region and education differences on the 11 WAIS-R subtests. *Journal of Clinical Psychology*, 44 (2), 231-248.
- Lancioni, G.E., O'Reilly, M.F., Campodonico, F., & Mantini, M. (2001). Promoting performance fluency in a person with profound intellectual disability and blindness. *Behavioural and Cognitive Psychotherapy*, 29 (3), 373-377.
- Lezak, M.D., Howieson, D.B., Loring, D.W., Hannay, H.J., & Fischer, J.S. (2004). *Neuropsychological Assessment. Fourth Edition*. New York: Oxford University Press.
- Łojek, E. (2001). *Neuropsychologia osób zakażonych wirusem HIV i chorych na AIDS [The Neuropsychology of HIV-Infected Individuals and AIDS Patients]*. Warszawa: Wydawnictwo Medyk.
- Łojek, E. & Bornstein, R.A. (2005). The stability of neurocognitive patterns in HIV infected men: Classification considerations. *Journal of Clinical and Experimental Neuropsychology*, 27 (6), 665-682.

- Łojek, E. & Stańczak, J. (2005). *Test Płynności Figuralnej Ruffa. Polska standaryzacja i normalizacja. Podręcznik* [The Ruff Figural Fluency Test. Polish Standardisation and Normalisation]. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Łojek, E. & Stańczak, J. (2009). *Kalifornijski Test Ucznienia się Językowego (CVLT). Polska normalizacja* [The California Verbal Learning Test (CVLT). Polish Normalization]. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Matczak, A., Jaworowska, A., Szustrowa, T., & Ciechanowicz, A. (1995). *Bateria Testów APIS-Z. Podręcznik* [The APIS-Z Test Battery Manual]. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Nęcka, E. (2000). Procesy uwagi [Attention processes]. In J. Strelau (Ed.), *Psychologia. Podręcznik akademicki. Psychologia ogólna. Tom 2* [Psychology. An Academic Textbook. General Psychology. Volume 2] (pp. 77-96). Gdańsk: GWP.
- OECD (2004). *Learning for Tomorrow's World – First Results from PISA 2003*. Organisation for Economic Co-Operation and Development, <http://www.oecd.org/dataoecd/1/60/34002216.pdf> [15.04.2007].
- Regard, M., Strauss, E., Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills*, 55 (3), 839–844.
- Rocznik Statystyczny Rzeczypospolitej Polskiej* [Statistical Yearbook of the Republic of Poland] (2002). Warszawa: Zakład Wydawnictw Statystycznych.
- Ross, T.P., Foard, E.L., Hiott, F.B., & Vincent, A. (2003). The reliability of production strategy scores for the Ruff Figural Fluency Test. *Archives of Clinical Neuropsychology*, 18 (8), 879-891.
- Rosselli, M. & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition*, 52 (3), 326-333.
- Ruff, R.M. (1996). *Ruff Figural Fluency Test. Professional Manual*. Odessa: Psychological Assessment Resources.
- Ruff, R.M., Allen, C.C., Farrow, C.E., Niemann, H., & Wylie, T. (1994). Figural fluency: Differential impairment in patients with left versus right frontal lobe lesions. *Archives of Clinical Neuropsychology*, 9 (1), 41-55.
- Ruff, R.M., Evans, R.W., & Lawrence, F.M. (1986). Impaired verbal and figural fluency after head injury. *Archives of Clinical Neuropsychology*, 1 (2), 87-101.
- Ruff, R.M., Light, R.H., & Evans, R.W. (1987). The Ruff Figural Fluency Test: A normative study with adults. *Developmental Neuropsychology*, 3 (1), 37-51.
- Ruff, R.M., Light, R.H., Parker, S.B., & Levin, H.S. (1997). The psychological construct of word fluency. *Brain and Language*, 57 (3), 394-405.
- Sabe, L., Jason, L., Juejati, M., Leiguarda, R., & Starkstein, S. (1993). Sensitivity and specificity of the Mini-Mental State Exam in the diagnosis of dementia. *Behavioural Neurology*, 6 (4), 207-210.

- Sands, K.A. (1998). Nonverbal fluency: A neuropsychometric investigation. *Dissertation Abstracts International. Section B: The Sciences and Engineering*, 58 (8-B), 4470.
- Schretlen, D.J., Munro, C.A., Anthony, J.C., & Pearlson, G.D. (2003). Examining the range of normal intraindividual variability in neuropsychological test performance. *Journal of the International Neuropsychological Society*, 9 (6), 864-870.
- Spree, O. & Strauss, E. (1998). *A Compendium of Neuropsychological Tests. Administration, Norms, and Commentary. Second Edition*. New York: Oxford University Press.
- Sternberg, R.J., & Wagner, R.K. (Eds.) (1986). *Practical Intelligence: Nature and Origins of Competence in the Everyday World*. Cambridge: Cambridge University Press.
- Suchy, Y., Sands, K., & Chelune, G.J. (2003). Verbal and nonverbal fluency performance before and after seizure. *Journal of Clinical and Experimental Neuropsychology*, 25 (2), 190-200.
- Williamson, J.B., & Harrison, D.W. (2003). Functional cerebral asymmetry in hostility: A dual task approach with fluency and cardiovascular regulation. *Brain and Cognition*, 52 (2), 167-174.