proLékaře.cz | 1.11.2026

doi: 10.14735/amcsnn201873

Alternative forms parallel to the Czech versions of Rey Auditory Verbal Learning Test, Complex Figure Test and Verbal Fluency

Alternativní formy pro české verze Paměťového testu učení, Reyovy-Osterriethovy komplexní figury a testu Verbální fluence

Abstract

Objective: The study evaluates the equivalence of alternative forms of three frequently used neuropsychological testing methods in their Czech versions, i.e. the Rey Auditory Verbal Learning Test (RAVLT), the Complex Figure Test (CFT) and letter and categorical Verbal Fluency (VF). *Methods:* In a within-subject design, 59 healthy volunteers were assessed using the original test forms and their proposed alternative forms, first at baseline and then after three months. The performance of the original and alternative forms of the tests was subjected to the t-test and U-test for paired samples based on the normality of distribution. *Results:* There was no significant mean difference between the original and alternative forms of the RAVLT. In the CFT, the Taylor figure scores were higher than the original Rey-Osterrieth figure scores, suggesting that the figures are not equal in difficulty. In the letter VF, the newly proposed "BTL" set was found to be more difficult than the original "NKP". For the categorical VF, the category of "given names" can be considered as an alternative form to the "animals" category. *Conclusions:* The RAVLT alternative forms 1 and 2 introduced in the Neuropsychological Battery of Psychiatric Center Prague are equivalent to the original RAVLT. The "BTL" VF set and Taylor Figure were found not to be equivalent to their original test forms "NKP" and Rey-Osterrieth Figure. "Given names" may be a good alternative form to the "animals" category in VF. The study contributes with normative data for the original test forms for a specific demographic sample.

Souhrn

Cíle: Studie ověřuje ekvivalenci alternativních forem ke třem běžně používaným neuropsychologickým testovým metodám, a to Paměťovému testu učení (Rey Auditory Verbal Learning Test; RAVLT), Reyově-Osterriethově komplexní figuře (Complex Figure Test; CFT) a testu fonemická a kategoriální Verbální fluence (VF). *Metody:* Padesát devět zdravých dobrovolníků podstoupilo v rámci vnitrosubjektového designu vyšetření původními a navrhovanými alternativními formami, a to v prvním měření a následně po třech měsících. Výkony v původních a alternativních verzích byly srovnány pomocí t-testu a U-testu pro párový výběr na základě normality rozložení výkonů. *Výsledky:* Mezi výkonem v původní a v alternativních formách RAVLT nebyl nalezen signifikantní rozdíl. U CFT byly výkony vyšší v Taylorově figuře než v původní Reyově-Osterriethově figuře, což naznačuje, že figury nejsou srovnatelně náročné. U fonemické VF byla nově navržená sada "BTL" shledána jako těžší než původní "NKP". U sémantické VF vyšla kategorie "křestní jména" jako srovnatelně náročná alternativa ke kategorii "zvířata". *Závěry:* Alternativní formy 1 a 2 pro RAVLT uvedené v Neuropsychologické baterii Psychiatrického centra Praha jsou srovnatelně náročné k původní verzi RAVLT. Sada "BTL" VF a Taylorova figura nebyly srovnatelné v náročnosti vzhledem k jejich původním verzím "NKP" a Reyově-Osterriethově figuře. "Křestní jména" mohou být srovnatelnou alternativou ke kategorii "zvířata" u VF. Studie přispívá normativními daty specifické demografické skupiny pro původní testové verze.

This study is a result of research funded by the Czech Science Foundation (GAČR) (grant number 16-13093S) and project Nr. LO1611 with financial support from MEYS under the NPU I program.

The authors declare they have no potential conflicts of interest concerning drugs, products, or services used in the study.
Autoři deklarují, že v souvislosti s předmětem studie nemají žádné komerční zájmy.
The Editorial Board declares that the manuscript met the ICMJE "uniform requirements" for biomedical papers.
Redakční rada potvrzuje, že rukopis práce splnil ICMJE kritéria pro publikace zasílané do biomedicínských časopisů.

M. Paštrnák[1,2], K. Šulcová[1,2], A. Dorazilová[1], M. Rodriguez[1]

[1] National Institute of Mental Health, Klecany, Czech Republic
[2] 3rd Faculty of Medicine, Charles University in Prague, Czech Republic

✉
Mgr. Martin Paštrnák
National Institute of Mental Health
Topolová 748
250 67 Klecany
Czech Republic
e-mail: martin.pastrnak@nudz.cz

Accepted for review: 27. 6. 2017
Accepted for print: 3. 1. 2018

Key words

cognition – neuropsychological tests – task performance and analysis – visual perception – verbal learning

Klíčová slova

kognice – neuropsychologické testy – testový výkon a analýza – vizuální percepce – verbální učení

proLékaře.cz | 1.11.2026

Introduction

Cognitive impairment is present in a broad range of neurological and psychiatric conditions. The heterogeneity of cognitive impairment in terms of the affected cognitive domains and its severity is manifested not only across various disorders [1,2], but also between patients with the same disorder [3,4]. Even in the same patient, a cognitive deficit can vary in its manifestation during the course of the disorder [5,6]. Hence, for a systematic neuropsychological assessment, repeated measurement is often required, mostly in order to delineate changes in the cognitive performance as a result of the treatment, experimental intervention, or pathological process. Major limitations to many neuropsychological methods are insufficient parallel forms of the original tests, and a lack of normative data for these alternative forms. This is a major methodological issue, as repeated measurement with the same test material poses the risk of the practice effect [7].

In this study, we focus on the Czech versions of three extensively used neuropsychological tests, i.e. the Rey Auditory Verbal Learning test (RAVLT) [8–10] (Czech version: Preiss [10]), the Rey-Osterrieth Complex Figure Test (ROCFT) [11,12] (Czech version: Košč and Novák [13]) and verbal fluency (VF) [14] (Czech version: Preiss [15]); and their parallel forms (AVLT: Preiss, Rodriguez and Laing [16], CFT: Taylor [17], and VF: based on Kopeček and Kuncová [18]).

RAVLT is used to evaluate verbal learning and memory and examines the ability to learn a list of words. The test consists of a repeated presentation and free recall of a 15-word list (list A), an interference list (list B), two post-interference free recall trials (immediate, delayed) and recognition of the target words presented with distractors. The test enables the assessment of immediate word span, total acquisition, proactive and retroactive interference, delayed recall and recognition [19]. Considering the practice effect, significant improvement to almost all measures of RAVLT appeared upon retesting after one month [20,21] and six months [22]. On the other hand, the results of studies conducted so far show that the use of alternative forms parallel to the original forms in retesting showed no improvement [20,21,23]. Currently, there are two existing alternative forms parallel to the original RAVLT form in the Czech Republic [24]. However, the equivalence characteristics of these forms have not been evaluated so far and no normative data are available yet.

ROCFT is used to examine visual perception, construction, executive function and visual memory [19]. The test stimulus is a figure composed of 18 geometric segments and uses a 36-point scoring system that reflects construction and memory functions [25]. In this study, we examine the ability to copy the figure, to recall the figure after 3 and 30 min. Considering the practice effect, a different figure is recommended for retest purposes [19] as studies show that participants are able to recall the basic features of the figure even after one year [26]. Moreover, the practice effect in planning and using a specific strategy cannot be easily attenuated in a retest situation, because the assessed person is already familiar with the task demands after the first assessment. This complicates the process of issuing a relevant retest form for the ROCFT. One alternative version for the ROCFT, the Taylor figure (TCFT), has been found to be equivalently difficult in the copy trial but easier in both recall trials than the ROCFT [27,28]. Despite the easier characteristics of the TCFT, this version is often used as an alternative to the ROCFT in clinical practice in the Czech Republic. Therefore, its characteristics as an alternative form need to be examined.

The VF test examines the ability to generate a list of words and is a highly sensitive indicator for frontal brain dysfunction in terms of word recall and flexibility of the word searching process [19]. In the letter (or phonemic) VF variation, the participant has a limited time to produce as many words as possible beginning with a designated letter. In the category (or semantic) VF variation, the participant has a limited time to produce words of a given category (e.g. animals, clothing). Functional magnetic resonance imaging techniques as well as lesion studies distinguish different brain activation patterns between the two types of VF tasks. While both tasks are sensitive to frontal brain dysfunction, the letter VF is more associated with left prefrontal cortex activation [29,30] and reflects word searching strategy, psychomotor speed and speech abilities [31]. The category VF on the other hand is associated more with the temporal cortex and hippocampi activation, reflecting mainly memory for word generation and is independent of the strategy used for word generation [29,30,32]. Hence, there is a degree of specificity in distinguishing between frontal and temporal cortices and their interconnected subcortical circuit dysfunction using letters and category VF.

The current version of the letters VF used the most frequently in the Czech Republic uses the letters N, K and P based on the original characteristics of "FAS" form of VF [15]. However, the "NKP" set is considered to be problematic, with one of the main issues being the letter "N". The letter "N" enables using a strategy of creating negations of verbs using the Czech prefix "ne-" which is a different process of word generation compared to the letters "K" and "P" [33]. Another major problem is the difficulty in establishing a matching alternative form to "NKP". In relation to this, several studies suggest using different letter sets, e.g. use S instead of N in the "NKP" set (for the new set: "KPS") [33], arguing that the S letter has more similar characteristics to K and P than N; or use the "NKT" set with a good match with the "BPL" set as the alternative form [18]. While the characteristics of these new sets are promising, "NKP" is still the most frequently used letter set in clinical practice. Moreover, the practice effect in the letters VF has been found to be considerable in both foreign and Czech literature, and studies addressing this issue recommend using alternative letter sets for retest purposes [18,34,35]. Hence, a parallel letter set to "NKP" is still required.

In category VF, the most used test category in the Czech Republic is "animals" [33,36]. While normative data for categories other than animals do exist for the elderly population, equal alternative forms have not been established yet [18].

The aim of this study was to examine the equivalency of the already existing or newly proposed parallel forms and to validate their use. We hypothesized that there would be no significant difference in task performance between the original and alternative test forms in a within-subject design.

Methods

Participants

We recruited participants among university students. All of the participants signed an informed consent form. The exclusion criteria were: 1. a history of neurological and psychiatric disorders; 2. substance or alcohol abuse; 3. any medical condition affecting neurocognition; 4. age below 20 or above

proLékaře.cz | 1.11.2026

30 years; 5. psychology or psychology-related field of study; 6. exposure to any psychological tests prior to the study.

Materials

Rey Auditory Verbal Leraning Test (RAVLT)

The Czech RAVLT consists of list A (target), list B (interference) and a 50-word list (target words presented with distractors) used for the recognition task. The alternative Form 1 and Form 2 with corresponding recognition task versions for the original A and B lists used in this study were introduced in "Neuropsychologická baterie Psychiatrického centra Praha (Neuropsychological Battery of Psychiatric Center Prague)" (PCP) [24] and were developed based on the translation of the validated alternative forms of the RAVLT introduced by Crawford et al [20]. The alternative word lists were compared with the original lists for consistency in terms of word length and semantic characteristics.

Verbal Fluency (VF)

The Czech version of the letters VF uses the letter set "NKP" [15]. No previous Czech study has brought a matching parallel form to "NKP" and therefore we attempted to develop one. Firstly, we reviewed the frequency dictionary from the Czech corpus (http://www.ujc.cas.cz/phword) for letters with a similar frequency and studies addressing the issue so far [18,33,34]. The considered letters were B, D, L, M, T, R, S, and V. From these letters, we excluded letters and letter combinations found unsuitable in former studies ("VRS" according to Štorková et al [34]), letters D and M according to Kopeček and Kuncová leaving B, T and L as potential candidates [18]. Due to the fact that one study [18] showed a good match between the sets NKT and BPL, and our need was to propose an alternative form of "NKP", we switched the letters T and P. Thus, we selected the letters B, T and L as potential parallel forms for "NKP". We conducted a pilot study where 10 participants (6 females, age range 24–30 years) were asked to generate as many words as possible within 60 seconds for each letter using the standard VF instructions. The participants generated a similar sum of words for the "NKP" and "BTL" sets.

For the semantic part of VF, we utilized the most commonly used categories for categorical VF "animals" and "vegetables" (e.g. [33,36]). For alternative categories, we selected Czech "given names" and "professions". To ensure the relevance of testing of newly proposed categories, we conducted a pilot study where 10 participants (6 females, age range 24–30 years) generated a similar sum of words for both sets of categories ("animals" + "vegetables" and "given names" + "professions").

Tab. 1. Demographic variables.

Demographic variable	Males	Females
N	30	29
age: M ± SD	24.03 ± 2.08	23.9 ± 1.95
age: range min – max	20 – 28	21 – 27

Complex Figure Test (CFT)

The CTF figure used in this study [13] is the original figure introduced by Rey [11]. The alternative to the ROCFT used in the present study is the commonly used figure by Taylor with the same number of segments and an identical scoring system [9].

Procedure

The study was approved by the Ethics Committee of the National Institute of Mental Health, Klecany. All of the participants were assessed at baseline randomly by the original or the alternative test form and then after three months with the opposite form. As the RAVLT has two alternative forms, we assigned the alternative Form 1 or Form 2 to the participants randomly. The tests were individually administered by trained pregraduate psychology students under the supervision of certified clinical psychologists.

The RAVLT was administered according to the standard administration procedure from "Neuropsychologická baterie" PCP [24]. The measures of interest were: sum of recall of the 5× administered list A (RAVLT – A1–A5), recall of list B (RAVLT – B), followed by the recall of list A without prior presentation (RAVLT – A6) and delayed recall after 30 min (RAVLT – A7).

The ROCFT and TCFT were administered according the Meyers and Meyers administration procedure (copy, 3-min and 30-min delayed recall, no recognition) [25]. Both figures consisted of 18 segments and each correctly drawn and placed segment received 2 points. Participants were asked to construct a copy of the presented figure with no time restriction on a blank sheet of paper. After 3 min, the participants were asked to recall and draw the figure (immediate recall) and after 30 min the participants were asked to recall the figure once (delayed recall). For the reliability of scoring, the copy and both delayed recalls were evaluated independently by three certified psychologists. The measures of interest were: copy, recall after 3 min and recall after 30 min.

The VF test was administered according to Thurstone [14] and Preiss [15] for the letters VF and the recommended instructions from Nikolai et al. [28] for the categorical VF. In the letters VF, the participants were asked to generate as many words as quickly as possible in a given time of 60 seconds beginning with the letters N, K and P (and B, T, L for the alternative form). In the category VF, the participants had 60 seconds to produce as many words as possible from the category of "animals" and afterwards "vegetables" ("clothes" and "given names" respectively as alternatives), with all the words and names belonging to the Czech vocabulary.

Statistical analyses

All of the statistical analyses were performed using IBM SPSS Statistical software (version 23, SPSS Inc., Chicago, IL, USA). The normal distribution of the data was inspected using the Shapiro-Wilk test of normality. The t-test for paired samples was used in cases of normative data distribution and the Wilcoxon matched paired test for repeated measures was used for abnormal data distribution to compare test-retest differences, as well as group performance differences between groups defined by the order of the test-retest administration and gender. The α-level was set at 0.05.

Results

Demographic results

Fifty-nine healthy volunteers aged 20 to 30 years were recruited for the study (30 males,

proLékaře.cz | 1.11.2026

Tab. 2. RAVLT test-retest scores (retest 1: N = 33; retest 2: N = 23).

Test – score	Raw test scores mean ± SD		Form differences original – alternative 1/2	
	original form	alternative form 1	t-test /Wilcoxon	p value
RAVLT – A1–A5	61.34 ± 5.87	61.69 ± 5.71	t = –0.37	0.71
RAVLT – B	7.88 ± 2.43	7.91 ± 2.43	t = 0.18	0.86
RAVLT – A6	13.41 ± 1.43	13.22 ± 1.79	Z = –0.25	0.80
RAVLT – A7	13.44 ± 1.46	13.13 ± 1.91	Z = –0.70	0.48
Test – score	**original form**	**alternative form 2**	**t-test /Wilcoxon**	**p value**
RAVLT – A1–A5	61.76 ± 6.46	61.38 ± 6.49	t = 0.26	0.80
RAVLT – B	7.57 ± 2.54	8.48 ± 2.58	t = –1.66	0.11
RAVLT – A6	13.38 ± 1.43	13.10 ± 2.07	Z = –0.65	0.52
RAVLT – A7	13.19 ± 1.75	13.14 ± 1.93	Z = –0.06	0.95

RAVLT – Rey Auditory Verbal Learning Test

Tab. 3. Inter-raters correlations in rating of Complex Figure Tests by Spearman´s rho.

CFT score	Spearman´s rho	p value
ROCFT – copy		
rater1 – rater2	0.73	0.000***
rater1 – rater3	0.63	0.000***
rater2 – rater3	0.81	0.000***
ROCFT – 3 min		
rater1 – rater2	0.89	0.000***
rater1 – rater3	0.93	0.000***
rater2 – rater3	0.89	0.000***
ROCFT – 30 min		
rater1 – rater2	0.89	0.000***
rater1 – rater3	0.90	0.000***
rater2 – rater3	0.91	0.000***
TCFT – copy		
rater1 – rater2	0.68	0.000***
rater1 – rater3	0.57	0.000***
rater2 – rater3	0.66	0.000***
TCFT – 3 min		
rater1 – rater2	0.88	0.000***
rater1 – rater3	0.85	0.000***
rater2 – rater3	0.95	0.000***
TCFT – 30 min		
rater1 – rater2	0.91	0.000***
rater1 – rater3	0.87	0.000***
rater2 – rater3	0.88	0.000***

ROCFT – Rey-Osterrieth Complex Figure; TCFT – Taylor Complex Figure Test; CFT – Complex Figure Test
*** p < 0.001

29 females; average age 23,97 years; SD 2) (Tab. 1). All of the participants were university students or university graduates with non-psychology or psychology-related fields of study.

Original and alternative form comparison

Rey Auditory Verbal Leraning Test (RAVLT)

Two groups were defined by receiving the alternative form 1 (N = 33; 16 females and 17 males; average age 24.36 years; SD 1.97) and the alternative form 2 (N = 23; 13 females and 10 males; average age 23.48 years; SD 2.04). At baseline, alternative form 1 was administered 20× and alternative form 2 14×. Three participants were excluded due to their unavailability for the second measurement. No significant mean differences were found in the performance between male and female groups in the original and alternative forms (1, 2), both at baseline and after three months. There was no significant mean difference between groups defined by the order of administration of the original and alternative forms (original – alternative vs. alternative – original). Due to the fact that the effect of gender and the order of the test-retest administration was non-significant, we compared the test-retest differences of the mean performance of all of the participants in RAVLT – A1–A5, RAVLT – B, RAVLT – 6 and RAVLT – 7 and found no significant difference between the original and alternative forms 1 or 2 (Tab. 2).

proLékaře.cz | 1.11.2026

Tab. 4. Comparison of original and alternate test form scores (N = 57).

Test – score	Raw test scores mean ± SD		Form differences	
	original form	alternative form	t-test /Wilcoxon	p value
CFT – copy/ rater1	32.88 ± 2.54	33.57 ± 1.86	Z = –2.59	0.01*
CFT – copy/ rater2	33.10 ± 2.34	34.19 ± 1.96	Z = –3.21	0.001**
CFT – copy/ rater3	33.47 ± 2.46	33.98 ± 1.95	Z = –1.69	0.09
CFT – 3 min/ rater1	25.04 ± 5.00	28.32 ± 3.63	Z = –5.41	0.000***
CFT – 3 min/ rater2	24.63 ± 4.94	27.77 ± 4.04	t = –5.30	0.000***
CFT – 3 min/ rater3	24.61 ± 5.11	27.39 ± 4.17	Z = –4.06	0.000***
CFT – 30 min/ rater1	24.45 ± 4.90	28.28 ± 3.77	Z = –5.60	0.000***
CFT – 30 min/ rater2	24.44 ± 5.06	27.87 ± 4.29	t = –5.26	0.000***
CFT – 30 min/ rater3	24.19 ± 4.95	27.45 ± 4.31	t = –5.10	0.000***

CFT – Complex Figure Test
* p < 0.05; ** p < 0.01; *** p < 0.001

Tab. 5. Comparison of original and alternate test form for letter and category verbal fluency (N = 57).

Test – score	Raw test scores mean ± SD		Form differences	
	original form	alternative form	t-test /Wilcoxon	p value
VFT – N/B	15.29 ± 5.11	16.26 ± 3.90	t = –1.61	0.11
VFT – K/T	19.37 ± 5.37	15.95 ± 4.20	t = 5.58	0.000***
VFT – P/L	19.46 ± 4.58	15.72 ± 4.01	t = 6.62	0.000***
VFT – total NKP/BTL	54.05 ± 13.13	47.93 ± 10.07	t = 4.85	0.000***
VFT – ani/gn	26.02 ± 6.19	24.75 ± 5.55	Z = –1.42	0.16
VFT – veg/professions	14.04 ± 3.02	19.86 ± 4.71	Z = –5.95	0.000***
VFT – ani+veg/gn+ professions	40.05 ± 7.87	44.61 ± 8.65	Z = –3.42	0.001**

VFT – Verbal Fluency Test; ani – animals; veg – vegetables; gn – given names
** p < 0.01; ***; p < 0.001

Complex Figure Test (CFT)

At baseline, the ROCFT was administered 23× and TCFT was administered 34×. Considering the order of test administration, there were no significant mean differences in the copy and both recall trials between the baseline and after three months scores in both ROCFT and TCFT (e.g. TCFT administered at baseline compared to TCFT administered after three months). There were no significant differences between the groups defined by gender. The correlations of scoring of the three raters showed a strong inter-rater correlation in all three measurements (Tab. 3). Analysis of the whole group of participants revealed mean differences between the original and alternative test forms. The mean difference for the copy trial was significant in two raters; however, the differences of the raw scores were minimal. The scores of the third rater showed no significant mean difference between the copies. For recall after 3 min and recall after 30 min, the mean differences between the original ROCFT and the alternative TCFT test form were significant in all three raters (Tab. 4).

Verbal Fluency (VF)

The original "NKP" form was administered 23× and "BTL" 34× at baseline. Two participants were excluded from the sample due to their unavailability for the second measurement. The comparison of mean differences of groups defined by the order of the original-alternative form administration and by gender revealed no significant differences. The comparison of mean sum of word generation of "NKP" and of "BTL" showed a significant difference. When we compared individual letter VFs, only the "N/B" letters were equivalent to each other as they showed no significant mean difference (Tab. 5). In category VF, the mean sum of "animals/vegetables" and the mean sum of "given names/professions" also showed a significant difference. The comparison of separate categories revealed that "animals" and

proLékaře.cz | 1.11.2026

"given names" showed no significant difference in their means.

Discussion

The main aim of this study was to evaluate whether alternative forms parallel to RAVLT, ROCFT and VF produce scores equivalent to their original versions. For the RAVLT and ROCFT, we used the already existing alternative test forms and examined their characteristics as a retest option. For the "NKP" set we examined the "BTL" set as a potential parallel retest form. A secondary output of the study is an extension of the already existing population standards for the three original tests with the collected data for a 20–30-years-old university degree sample with the perspective of expanding the standards for other demographics in the future. To confirm the equivalent difficulty of the alternative forms, we used a design in which the participants were assessed twice, once with the original form and once with the alternative form, with a 3-month delay. We compared the performance of the first and the second measurement and based on the significance of their mean differences concluded whether the forms were equivalently difficult or not.

The RAVLT parallel forms 1 and 2 adopted from Crawford et al [20] by Preiss et al [24] showed a good match as an alternative task as no significant difference between the original and both the first and the second form was found. This suggests that the alternative forms 1 and 2 can be used as a valid retest option. The limitation of these results is the relatively small sample size for both separate forms, since we had to divide the whole group approximately in half. Another issue is the specificity of the relatively homogenous experimental sample as good and stable memory capacity is a common prerequisite for obtaining or wanting a university degree, which may have influenced the overall performance. This issue should be remedied in a future study on a larger and more heterogeneous sample.

The ROCFT alternative form, the TCFT, has the same scoring system and is composed of the same number of components as the original ROCFT. However, the TCFT has been considered to be easier than the ROCFT in previous research in terms of better overall performance in recall trials [24,28]. According to Hamby et al, the TCFT replicates the visuoperceptual properties of the ROCFT but not its organizational qualities [37]. Thus, the simpler structure of the TCFT yields similar copy scores but has better delayed recall performance scores [19]. Our results are consistent with these results as participants performed significantly better in the TCFT free recall after 3 and 30 min regardless of the order of administration. As for the copy, the mean difference was also statistically significant in the scores of the two raters (Tab. 4), but also marginal in terms of raw scores [19]. In clinical practice, a raw score difference of 1 point is usually considered as non-significant. One explanation for the significant mean difference in the copy trials may be a larger sample of participants (N = 57) as larger samples more likely tend to generate statistical significance even in minimal raw score differences. Our results suggest that the TCFT can be used as a parallel form for the ROCFT in the copy trial. However, performance differences in recall trials between the TCFT and ROCFT should be taken into consideration during the interpretation of the results in a retest situation. The inter-rater reliability in the present study showed a good correlation between raters, thereby confirming the usefulness of Meyers and Meyers administration and scoring [25].

In conclusion, the results suggest that caution should be applied when using the TCFT as an alternative to the ROCFT in a research context. Other studies showed a good match between the ROCFT and its parallel form – the modified Taylor figure (mTCFT) [28,38]. In this study, we did not use the mTCFT as the Czech version has not been adopted yet. The mTCFT, already evaluated abroad, may be a good alternative form for the ROCFT, and should be investigated in further research.

The commonly used VF letter set "NKP" has repeatedly shown to be problematic in terms of developing a suitable alternative form [18,29]. In the present study, our parallel set "BTL" has shown to be more difficult than "NKP" and is therefore unsuitable as an equivalent alternative form. This finding supports the claim that "NKP" is not a good baseline option for repeated measurement. We also compared individual letter pairs of "NKP" and "BTL" in terms of word generation. Only the "N/B" pair was found to be equivalent, which is congruent with the findings in the study [18]. This suggests that when obtaining the letters VF data in clinical practice and research, "NKP" is not suitable for repeated measurement designs as it still lacks a good equivalent letter set. Creating an entirely different set of letters together with matching parallel forms may resolve the issue for the letter VF task [33]. Interestingly, it is yet possible to use "NKP" in repeated administration after longer time periods because some studies suggest the practice effect in the letters VF is non-significant after one year [39].

Considering category VF, dual sets of "animals/vegetables" and "given names/professions" have been found not to be equivalent. However, when evaluated separately, the data show that there is no significant difference in means between "animals" and "given names", suggesting that the "given names" category in the category VF task is a potential alternative to "animals". While this result is promising, further research should be conducted on larger samples with more demographic heterogeneity.

One limitation of this study is the absence of a direct control of the practice effect. The sole purpose for developing alternative forms is to attenuate the practice effect. However, the use of alternative forms may produce effects of practice as well. As for the RAVLT and both semantic and phonemic VF, studies by Lemay et al [40] and Benedict and Zgaljardic [23] showed that different alternative forms produced practice effects comparable to the original test forms. Other studies showed that even repeated administration of alternative forms produced only marginal practice effects for VF and RAVLT [20,21,23,40,41]. According to Lezak [19,39], tests that have a large speed component require an unfamiliar or infrequent practice mode of response, or have a single solution – particularly if it can be easily conceptualized once it is attained – are more likely to show significant practice effects. As VF and RAVLT do not correspond to these test characteristics, the effect of practice is only marginal in RAVLT and VF when using alternative forms as was also supported by the studies mentioned above. Although the use of alternative forms may attenuate practice effects, they may still occur in novel tests, tests with high cognitive demands or those in which the participant learns to use an effective test-taking strategy or has acquired "test-wiseness" [19,30]. This may be the case of CFT and is also the main reason why we consider TCFT to have reduced criterion validity as an alternative form to ROCFT. As stated above, the best remedy in order to attenuate the practice

proLékaře.cz | 1.11.2026

effect in CFT-like tests is to use alternative forms and longer time periods between testing [41,26]. In summary, the ROCFT, RAVLT and VF all show a major reduction in the practice effect when using parallel forms [23,26,28,39,41]. Still, it is reasonable to assume that in our study the practice effect did occur to some degree and should be addressed in following studies by including repeated administration of both the original and alternative forms.

Another major limitations of this study are the sample size, its homogeneity and its narrow demographic characteristics. Psychiatric and neurological disorders in the population spread across all education levels and gender and most have a typical age of onset. For example, the mean age of onset of major depressive disorder is 20 years [44], for schizophrenia the peak age of onset is between 20 and 25 years in men and between 25 and 30 in women, for bipolar disorder the average age of onset is 18 years [42], for Parkinson's disease the average age of onset is 60, and 65 for Alzheimer's disease [43]. The usefulness of our results in clinical practice is therefore limited. Moreover, the practice effect is likely to vary in different disorders even when using alternative forms. The method to attenuate the effect of practice in a clinical setting is to compare repeated performance measurements of healthy controls and specific patient groups [44]. Following this, patient group performance change differing from the performance gain of controls (the effect of practice) after repeated administration of a neurocognitive test is considered to be related to the disorder or intervention [7,44]. For example, the practice effect in memory tests and category VF in diagnosed Alzheimer's disease is non-significant regardless of using alternative or original forms [45]. In schizophrenia on the other hand, most neurocognitive tests show the effect of practice [7]. This implies the need to examine alternative forms not just for a much broader and larger demographic sample but also for different patient groups in order to provide normative standards and the necessary diagnostic sensitivity and specificity.

While the demographic characteristics of our sample pose limits for broader conclusions, they also contribute supplement data for the existing normative standards of the original forms. Our extension is beneficial especially for the ROCFT, as normative data for this demographic group are missing. For the original letter and categorical VF and the RAVLT, the data provide extension of normative standards with the perspective of further extensions of the standards with broader demographic samples in subsequent studies.

Conclusion

In this study, we evaluated the original and alternative forms of three frequently used neuropsychological tests – the RAVLT, CFT and letter/categorical VF. The results of the study imply that the RAVLT alternative forms 1 and 2 introduced in "Neuropsychologická baterie" PCP are equivalent to the original RAVLT [24]. The study did not prove that "BTL" was a good alternative to "NKP". However, we confirmed that the letters "N" and "B" are equivalently difficult and, considering the category VF, the category of "given names" might be a good alternative form to the "animals" category. As for the ROCFT, the TCFT was found to be easier than the original form and should be used with caution as an alternative test in research. The study contributes normative data for the original test forms for a specific demographic sample.

References

1. Hill SK, Reilly JL, Harris MS et al. A comparison of neuropsychological dysfunction in first-episode psychosis patients with unipolar depression, bipolar disorder, and schizophrenia. Schizophr Res 2009; 113(2–3): 167–175. doi: 10.1016/j.schres.2009.04.020.

2. Bora E, Pantelis C. Meta-analysis of cognitive impairment in first-episode bipolar disorder: comparison with first-episode schizophrenia and healthy controls. Schizophrenia Bull 2015; 41(5): 1095–1104. doi: 10.1093/schbul/sbu198.

3. Gilbert E, Merette C, Jomphe V et al. Cluster analysis of cognitive deficits may mark heterogeneity in schizophrenia in terms of outcome and response to treatment. Eur Arch Psychiatry Clin Neurosci 2014; 264(4): 333–343. doi: 10.1007/s00406-013-0463-7.

4. Gold S, Arndt S, Nopoulos P et al. Longitudinal study of cognitive function in first-episode and recent-onset schizophrenia. Am J Psychiatry 1999; 156(9): 1342–1348. doi: 10.1176/ajp.156.9.1342.

5. Jahshan C, Heaton RK, Golshan S et al. Course of neurocognitive deficits in the prodrome and first episode of schizophrenia. Neuropsychology 2010; 24(1): 109–120. doi: 10.1037/a0016791.

6. Sponheim SR, Jung RE, Seidman LJ et al. Cognitive deficits in recent-onset and chronic schizophrenia. J Psychiatr Res 2010; 44(7): 421–428. doi: 10.1016/j.jpsychires.2009.09.010.

7. Szoke A, Trandafir A, Dupont ME et al. Longitudinal studies of cognition in schizophrenia: meta-analysis. Br J Psychiatry 2008; 192(4): 248–257. doi: 10.1192/bjp.bp.106.029009.

8. Rey A, L'examen clinique en psychologie. Paris: Presses universitaires de France 1958.

9. Taylor EM. Psychological appraisal of children with cerebral defects. Cambridge: The Commonwealth Fund by Harvard University Press 1959.

10. Preiss M. Paměťový test učení: manual pro dospělé a děti. Bratislava: Psychodiagnostika. 1999.

11. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. Arch Psychologie 1941; 28: 286–340.

12. Osterrieth PA. Le test de copie d'une figure complexe; contribution à l'étude de la perception et de la mémoire. Genève: Archives de Psychologie 1944.

13. Košč M, Novák J. Rey-Osterriethova komplexní figura. Brno: Psychodiagnostika 1997.

14. Thurstone TG, Thurstone LL. Primary mental abilities tests. Chicago: Science Research Associates 1962.

15. Preiss M. Verbální fluence, metoda vyšetření poškození mozku u dětí a dospělých. Csl Pscychol 1997; 3: 244–249.

16. Preiss M, Rodriguez M, Laing H. Neuropsychologická baterie Psychiatrického centra Praha: klinické vyšetření základních kognitivních funkcí. Praha: Psychiatrické centrum; 2002.

17. Taylor LB. Localization of cerebral lesions by psychological testing. Clin Neurosurg 1969; 16: 269–287.

18. Kopeček M. Kuncová A. Efekt nácviku generování slov a testování alternativní verze. Pilotní studie. Psychiatrie 2006; 10(4): 211–215.

19. Lezak MD. Neuropsychological assessment. New York, USA: Oxford University Press 2004.

20. Crawford JR, Stewart LE, Moore JW. Demonstration of savings on the AVLT and development of a parallel form. J Clin Exp Neuropsychol 1989; 11(6): 975–981.

21. Munjir N, Othman Z, Zakaria R et al. Equivalence and practice effect of alternate forms for Malay version of Auditory Verbal Learning Test (MAVLT). EXCLI J 2015; 14: 801–808. doi: 10.17179/excli2015-280.

22. Sithinamsuwan P, Hutchings N, Ananworanich J et al. Practice effect and normative data of an HIV-specific neuropsychological testing battery among healthy Thais. J Med Assoc Thai 2014; 97 (Suppl 2): S222–S233.

23. Benedict RH, Zgaljardic DJ. Practice effects during repeated administrations of memory tests with and without alternate forms. J Clin Exp Neuropsy 1998; 20(3): 339–352.

24. Preiss M, Bartoš A, Čermáková R et al. Neuropsychologická baterie Psychiatrického centra Praha: klinické vyšetření základních kognitivních funkcí. 3. vyd. Praha: Psychiatrické centrum 2012.

25. Meyers JE, Meyers KR. Rey Complex Figure Test and recognition trial professional manual. Florida: Psychological Assessment Resources 1995.

26. Yamashita H. One-year delayed recall performance of the Rey-Osterrieth Complex Figure in a healthy young adult sample. Appl Neuropsychol 2009; 16(2): 141–143. doi: 10.1080/09084280802623064.

27. Strauss E, Spreen O. A comparison of the Rey and Taylor figures. Arch Clin Neuropsychol 1990; 5(4): 417–420.

28. Yamashita H. Comparability of the Rey-osterrieth complex figure, the Taylor complex figure, and the modified taylor complex figure in a normal sample of Japanese speakers. Psychol Rep 2006; 99(2): 531–534.

29. Baldo JV, Schwartz S, Wilkins D et al. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. J Int Neuropsychol Soc 2006; 12(6): 896–900.

30. Birn RM, Kenworthy L, Case L et al. Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. Neuroimage 2010; 49(1): 1099–1107. doi: 10.1016/j.neuroimage.2009.07.036.

31. Dick AS, Bernal B, Tremblay P. The language connectome: new pathways, new concepts. The Neuroscientist 2014; 20(5): 453–467. doi: 10.1177/1073858413513502.

proLékaře.cz | 1.11.2026

32. Ryan L, Cox C, Hayes SM et al. Hippocampal activation during episodic and semantic memory retrieval: Comparing category production and category cued recall. Neuropsychologia 2008; 46(8): 2109–2121. doi: 10.1016/j.neuropsychologia.2008.02.030.
33. Nikolai T, Štěpánková H, Michalec J et al. Testy verbální fluence, česká normativní studie pro osoby vyššího věku. Cesk Slov Neurol N 2015; 111(3): 292–299. 10.14735/amcsnn2015292.
34. Štorková P, Preiss M, Kopeček M. Efekt nácviku testu verbální fluence a testování alternativní verze. Pilotní studie. Psychiatrie 2004; 8(3): 187–190.
35. Beglinger LJ, Gaydos B, Tangphao-Daniels O et al. Practice effects and the use of alternate forms in serial neuropsychological testing. Arch Clin Neuropsychol 2005; 20(4): 517–529.
36. Hummelova Z, Janousova E. Limits of Verbal Fluency Tests use in the differential diagnostic of neurological diseases. Cesk Slov Neurol N 2014; 77(4): 487–492.
37. Hamby SL, Wilkins JW, Barry NS. Organizational quality on the Rey-Osterrieth and Taylor Complex Figure Tests: a new scoring system. Psychological assessment 1993; 5(1): 27.
38. Hubley AM, Jassal S. Comparability of the Rey-Osterrieth and Modified Taylor Complex Figures using total scores, completion times, and construct validation. J Clin Exp Neuropsychol 2006; 28(8): 1482–1497.
39. Basso MR, Bornstein RA, Lang JM. Practice effects on commonly used measures of executive function across twelve months. Clin Neuropsychol 1999; 13(3): 283–292.
40. Lemay S, Bédard MA, Rouleau I et al. Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. Clin Neuropsychol 2004; 18(2): 284–302.
41. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. Clin Neuropsychol 2012; 26(4): 543–570.
42. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Arlington: American Psychiatric Pub 2013.
43. World Health Organization. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: World Health Organization; 1992.
44. Darby D, Maruff P, Collie A et al. Mild cognitive impairment can be detected by multiple assessments in a single day. Neurology 2002; 59(7): 1042–1046.
45. Cooper DB, Epker M, Lacritz L et al. Effects of practice on category fluency in Alzheimer's disease. Clin Neuropsychol 2001; 15(1): 125–128.

XIV. MEZINÁRODNÍ WORKSHOP FUNKČNÍ MAGNETICKÉ REZONANCE
Arcidiecézní muzeum, Olomouc
5. – 6. dubna 2018
www.fmri2018.upol.cz

XIV. OLOMOUCKÉ NEUROIMUNOLOGICKÉ SYMPOZIUM S MEZINÁRODNÍ ÚČASTÍ
Arcibiskupský palác, Olomouc
17. – 18. května 2018
www.ms2018.upol.cz

XIII. OLOMOUCKÉ CEREBROVASKULÁRNÍ SYMPOZIUM
Lékařská fakulta Univerzity Palackého, Olomouc
10. října 2018
www.cmp2018.upol.cz

PREDIAGNOSTIC, PREMOTOR AND NON-MOTOR PARKINSONISM
Umělecké centrum Univerzity Palackého, Olomouc
13. – 14. prosince 2018
www.parkinson2018.upol.cz

Organizační zajištění:
Konferenční servis Univerzity Palackého v Olomouci,
Biskupské nám. 1, 771 11 Olomouc