

RESEARCH

Open Access



# Evidences of validity and reliability of the Luria-Nebraska Test for Children

Ricardo Franco de Lima<sup>1\*</sup>, Mateus Silvestrin<sup>1</sup>, Fernanda de Lima<sup>1</sup>, Patrícia Abreu Pinheiro Crenitte<sup>2</sup> and Sylvia Maria Ciasca<sup>1</sup>

## Abstract

This paper aimed to verify evidences of validity and reliability of Luria-Nebraska Test for Children (TLN-C, in Portuguese). Three hundred eighty-seven students aged 6–13 years old, with learning difficulties, comprised the study. They were assessed with the Wechsler Intelligence Scale for Children (WISC-III) and TLN-C; and effect of age differences, as well as accuracy rating by internal consistency were investigated. Age effects were found for all subtests and in the general score, except for receptive speech subtest, even when total IQ effect was controlled. Reliability analysis had satisfactory results (0.79). The TLN-C showed evidences of validity and reliability. Receptive speech subtest requires revision.

**Keywords:** Neuropsychology, Psychometrics, Neuropsychological assessment, Validation Studies

## Background

Neuropsychological testing, in its clinical or research application, may be composed of single tests that assess certain aspects of cognitive functions or of sets of tests, known as batteries. Some test batteries are designed to investigate several facets of a single cognitive function (e.g. types of memory), while others are dedicated to a broader and deeper investigation of cognitive abilities. Between the specificity of single tests and the depth of batteries, one may find a third class of psychometric instruments, composed of instruments designed to combine relatively quick application time and cover many cognitive functions: the screening tests. To accomplish such a task, screening tests usually refrain from making a complete performance profile of the functions they assess and focus on detecting signs of deficits (Ustároz, 2007).

Independently of the type of test, it is of the utmost importance that instruments used in neuropsychological assessment exhibit adequate psychometric properties, ensuring that their application and interpretation are reliable (AERA, APA, & NCME, 1999). In Brazil, validation studies for neuropsychological instruments for children and teenagers are specially needed because only a few of such tests are available. The “Evaluation System

for Psychological Tests” (SATEPSI in Portuguese), a platform for controlling research on Brazilian validation studies and disclosure of which tests can be used in psychological day-to-day practice, showed, in November 2013, that only approximately 15 cognitive tests for such population filled the requirements at that moment (SATEPSI, 2013).

Therefore, Brazilian researchers have invested both in adapting international instruments, e.g. NEPSY (A developmental neuropsychological assessment) (Argollo et al., 2009), and developing neuropsychological batteries, e.g. the “Neuropsychological assessment battery for children”—NEUPSILIN (Salles et al., 2011). These batteries aim to measure several cognitive domains, like attention, memory, language and executive functions. For validation of both of them, studies comparing performance of control and clinical populations were conducted (Pawlowski et al. 2013; Gonçalves et al., 2013), as well as studies comparing populations of different age ranges (Zibeti et al., 2010) and studies investigating correlations with previously validated instruments for the Brazilian population (Yates et al., 2013).

The present paper is in line with these Brazilian efforts, in this case, to provide an instrument of cognitive screening with neuropsychological emphasis, for children. This instrument is the second Brazilian adaptation of Luria-Nebraska Neuropsychological Battery—Children’s Revision (LNNB-

\* Correspondence: ricardolima01@yahoo.com.br

<sup>1</sup>Universidade Estadual de Campinas, Campinas, Brazil

Full list of author information is available at the end of the article

CR) (Golden, 1987). The original battery, Luria-Nebraska Neuropsychological Battery (LNNB), measures 25 cognitive functions (from motor skills to intellectual processes) clustered in the following scales: clinical (11 functions), summary (3 functions), factual (11 functions) and optional (two specific measures of language skills) (Golden, 1987).

Previous studies with the original battery obtained several evidences of validity. Snow (1985) performed factor analysis from data of 100 children with learning disabilities, revealing three factors: language-overall intelligence, reading-writing, sensory-motor. In clinical samples the LNNB-CR identified performance differences between control and learning disabled subjects (Lewis et al. 1993), and between children and adolescents with reading disorders. The authors point out the LNNB-CR as an instrument potentially valid for investigation of neuropsychological alterations in children with learning issues (Myers et al. 1989).

Correlations between LNNB scores and instruments of intelligence assessment have been observed, like the Woodcock-Johnson Tests of Achievement-Revised (Hooper, 1995) and also between LNNB language and arithmetic measures and the Wechsler Intelligence Scale for Children—WISC (Pfeiffer et al. 1987). Boyd and Hooper (1993), exploring predictive models of intellectual performance using LNNB scores, verified that including age as a factor in the multivariate regression analysis was essential for a reasonable output. Thus, LNNB performance was able to predict performance on the WISC version used at that time (the revised version, WISC-R). Reliability studies of the original battery included the test-retest procedure on a psychiatric sample, where results were reproduced successfully with 8 months between two testing (Plaisted & Golden, 1982), and internal consistency analysis by Cronbach's Alpha, with results varying from adequate and high levels of internal consistency among the scales (coefficients ranging from 0.72 to 0.96) (Teichner et al. 1999).

Brazilian adaptations of the LNNB different versions have been proposed. Romanelli et al. (1999) presented the procedure for adaptation and standardization of Luria-Christensen version of the battery. A more recent pilot study (Crenitte et al., 2011) of an adaptation of the battery for children obtained preliminary standardization data and indicated the necessity of refining some subtests. Ciasca (1994) adapted the LNNB-CR to Portuguese, naming the instrument "Luria-Nebraska Battery for Children" (in Portuguese, *Bateria Luria Nebraska para Crianças—BLN-C*), focusing on the clinical scale, resulting, therefore, in a briefer test. The second Brazilian adaptation was proposed by Lima et al. (2005) and kept the first adaptation focus, making changes in scoring and subtests. It is named Luria-Nebraska Test for Children (in Portuguese, "*Teste Luria-Nebraska para Crianças*")—TLN-C) and investigates 10 functions in

children from 6 to 12 years of age, as follows: motor, rhythm, tactile, visual, receptive language, expressive language, writing, reading, arithmetic and immediate memory. The previously referred pilot study by Crenitte et al. (2011) used this version of the battery.

Advancing the studies of this Brazilian adaptation of the LNNB, the present study aimed to obtain validity evidences for TLN-C by investigation of its relations with external criteria (age) and external variable (intellectual quotient—IQ), investigation of its predictive capabilities regarding IQ and verification of its reliability by internal consistency analysis.

## Method

### Participants

Initially, 576 students aged 6–16 years old ( $M = 9.86$ ;  $SD = 2.06$ ) participated in the study. These children and adolescents had learning difficulties complaints (reading, writing, and arithmetic) and were referred to be assessed by an interdisciplinary team of a neurology outpatient clinic. The following exclusion criteria were adopted: intellectual quotient (IQ) below 80, that is, having intellectual classification from borderline to intellectually deficient in the Wechsler Intelligence Scale for Children—Third Edition (WISC-III) (Wechsler, 1991; Figueiredo, 2002), presenting uncorrected hearing or visual deficits, and presenting neurological or psychiatric disorders. Following these parameters, 189 students were excluded from the sample, remaining 387 participants with learning difficulties (reading, writing, and arithmetic), from both genders, aged 6–13 years old ( $M = 9.43$ ;  $SD = 1.87$ ) who attended grades 1st–9th of public schools. In this study scholar repetition rates were not considered. Table 1 shows the distribution of age groups as a function of grades and intellectual level (Full scale intellectual quotient—FSIQ, WISC-III).

### Materials

Wechsler Intelligence Scale for Children—Third Edition, WISC-III (Wechsler, 1991; Figueiredo, 2002). Scale adapted and standardized for Brazilian population, of individual administration, aimed to evaluate cognitive/intellectual capabilities of children and adolescents aged 6–16 years old. It encompasses twelve subtests with specific materials, measuring several cognitive functions. WISC-III was adopted in this study for selection of participants, and the full-scale IQ was used as a covariant in the analysis of age effects on TLN-C performance.

Teste Luria-Nebraska para Crianças—TLN-C (Lima et al. 2005). A screening test for deficits in cognitive functions, of individual administration, and composed of 120 items distributed in 10 subtests, which are: motor skill (0–15 points), rhythm (0–10 points), tactile skill (0–19 points), visual skill (0–12 points), receptive speech

**Table 1** Distribution of ages groups by gender, grades and intellectual level

Variables	6	7	8	9	10	11	12	13	f (%)
Gender									
Male	15 (68)	28 (74)	46 (68)	63 (73)	36 (64)	32 (58)	28 (74)	22 (92)	270 (70)
Female	7 (32)	10 (26)	22 (32)	23 (27)	20 (36)	23 (42)	10 (26)	2 (8)	117 (30)
Grades									
1st-4th	22 (100)	38 (100)	68 (100)	78 (91)	32 (57)	19 (35)	4 (11)	4 (17)	265 (68)
5-9th	0	0	0	8 (9)	24 (43)	36 (65)	34 (89)	20 (83)	122 (32)
IQ									
Low average	7 (32)	9 (24)	23 (34)	19 (22)	19 (34)	21 (38)	11 (29)	8 (33)	117 (30)
Average	13 (59)	17 (45)	38 (57)	37 (43)	29 (52)	25 (45)	18 (47)	12 (50)	190 (49)
High average	0	7 (18)	3 (4)	12 (14)	2 (3)	6 (11)	7 (18)	1 (4)	38 (10)
Superior	0	4 (11)	3 (4)	10 (12)	5 (9)	2 (4)	2 (6)	2 (8)	28 (07)
Very Superior	2 (9)	1 (3)	0	8 (9)	1 (2)	1 (2)	0	1 (4)	14 (04)
Total	22	38	68	86	56	55	38	24	387 (100)

Note: IQ Intellectual Quotient

(0–6 points), expressive speech (0–7 points), writing (0–15), reading (0–9 points), mathematical reasoning (0–15 points) and immediate memory (0–12 points). The child in assessment is asked to produce verbal or motor responses and, in some cases, these responses involve the manipulation of specific subtest materials. In addition to verbal instructions, some subtests use stimuli cards to elicit responses to each item. Each item is scored according to response efficiency: 0 = unable to execute task; 0.5 = troublesome task execution; 1.0 = task easily executed. Items that have binary answers (e.g. “color identification” on the visual skill subtest) are scored only with 0 or 1. Raw scores are obtained in each subtest by the sum of the points scored in its items and, by the sum of subtest scores the total score of the test is obtained.

#### Procedures and statistics

After approval by the Research Ethics Committee (Protocol n. 476.243), data from the ambulatory assessment database in the period from 2005 to 2012 were pooled. Inclusion and exclusion criteria were considered and, afterwards, the data was organized according to the research objectives. Data was analyzed using IBM SPSS Statistics 20.0 for Windows®. For inferential analysis, parametric tests were used based on the results of the Kolmogorov-Smirnov test. The statistical analysis were divided into the following steps, according to the parameters for evidence of validity (AERA, APA, & NCME, 1999): (i) To investigate validity evidences for TLN-C in its relations with external criteria, comparison of means between ages was done using analysis of covariance (Ancova), in which age was the factor and intellectual quotient (WISC-III) was the covariant. This analysis was complemented by Tukey HSD post hoc test to determine which ages were different. Effect sizes were

considered calculating the partial squared eta ( $\eta_p^2$ ); (ii) In order to investigate validity evidences for TLN-C in its relations with external variable, Pearson’s correlation was carried between TLN-C subtests and FSIQ (WISC-III). Subsequently, stepwise regression analysis was carried out to verify possible effects of TLN-C’s subtests and total score as predictors of FSIQ on WISC-III. Significance level of  $p \leq .05$  was adopted for all analyses; (iii) The TLN-C’s internal consistency was analyzed using Cronbach’s alpha ( $\alpha$ ).

#### Results

Table 2 presents descriptive analysis of TLN-C scores as a function of age and in the total sample. There was a tendency of progression of scores as age increased. This progression pattern was interrupted between some subtests and ages: Motor Skill (10 and 11 years old); Rhythm (12 and 13 years old); Tactile Skill (7 and 8; 12 and 13 years old); Visual Skill (7 and 8; 9 and 10 years old); Receptive Speech (8 and 9; 11 and 12 years old); Expressive Speech (10 and 11; 12 and 13 years old); Reading (10 and 11; 12 and 13 years old); Mathematical Reasoning (10 and 11 years old) and Immediate Memory (12 and 13 years old).

Covariance analysis (Ancova) using age as factor and IQ as covariant (Table 3), aiming to minimize possible effects of IQ over the performance in TLN-C, revealed significant effect of age on all subtests, except on “receptive speech”. Tukey HSD analysis indicated significant differences among age groups. A systematic progression of the means was present, mainly, on the total score. There was little or no progression in subtest means from 9 to 10 years onward. In the total score, from 11 years onward there was no change among age groups.

**Table 2** Descriptive statistics of TLN-C subtests according age and sample total

Subtests		Age (years)								Min.	Máx.	Total (n = 387)
		6 (n = 22)	7 (n = 38)	8 (n = 68)	9 (n = 86)	10 (n = 56)	11 (n = 55)	12 (n = 38)	13 (n = 24)			
MS	M	11.45	11.68	12.75	13.06	13.54	13.47	14.03	14.12	4	15	13.07
	SD	2.54	1.86	1.64	1.55	1.33	1.50	1.13	.95			1.73
RY	M	5.91	7.34	7.43	7.88	8.52	8.93	9.05	8.83	0	10	8.05
	SD	3.16	2.63	2.21	2.30	1.80	1.61	1.56	1.86			2.27
TS	M	12.05	13.61	13.54	14.72	14.64	15.42	16.61	15.38	7	19	14.57
	SD	2.52	2.69	2.55	2.64	2.79	2.78	2.87	3.21			2.91
VS	M	10.95	11.61	11.43	11.59	11.57	11.69	11.87	11.88	5	12	11.58
	SD	1.68	.68	.92	.86	.74	.69	.41	.45			0.85
RS	M	5.73	5.76	5.82	5.64	5.89	5.93	5.76	5.83	0	6	5.79
	SD	.55	.54	.79	.98	.37	.26	.75	.82			0.71
ES	M	6.36	6.87	6.87	6.91	7.00	6.91	6.95	6.92	3	7	6.88
	SD	1.14	.34	.38	.39	.00	.29	.23	.28			0.43
WR	M	7.91	9.34	9.68	11.88	13.07	13.11	13.47	14.08	2	15	11.66
	SD	3.71	3.29	3.44	2.83	1.92	2.02	1.87	1.35			3.24
RE	M	3.50	5.66	5.82	7.40	8.27	8.18	8.58	8.38	0	9	7.14
	SD	2.74	2.98	2.63	2.18	1.27	1.55	1.20	.92			2.49
MR	M	9.50	11.16	11.44	12.71	13.48	13.35	14.00	14.29	0	15	12.58
	SD	3.52	2.18	2.26	2.18	1.61	2.43	1.90	1.16			2.50
IM	M	8.23	9.29	9.78	9.80	10.09	10.62	10.89	10.75	5	12	9.98
	SD	1.90	1.43	1.37	1.24	1.24	.97	1.01	.99			1.41
Total	M	81.59	92.58	94.46	101.60	106.14	107.67	111.29	110.38	68	120	101.34
	SD	11.55	12.88	9.92	10.27	7.20	7.37	6.58	7.74			12.17

Note: *M* Mean, *SD* Standard Deviation, *n* number, *Min.* Minimum, *Max.* Maximum, *MS* motor skill, *RY* rhythm, *TS* tactile skill, *VS* visual skill, *RS* receptive speech, *ES* expressive speech, *WR* writing, *RE* reading, *MR* mathematical reasoning, *IM* immediate memory

Positive and significant correlations between full-scale IQ on WISC-III and all of TLN-C subtests and scores were found (Table 4). Effect size was low for Visual Skill, Receptive Speech, Expressive Speech subtests, moderate for Motor Skill, Rhythm, Tactile Skill, Writing, Reading

and Immediate Memory, and high for Mathematical Reasoning and Total. Positive and significant correlations were also obtained among all TLN-C subtests. Effect sizes ranged from low to high. High effect sizes were associated with Writing, Reading and Mathematical

**Table 3** Comparison of different ages in TLN-C subtests controlling effect of total IQ by Ancova

TLN-C	F	$p^a$	$\eta_p^2$	Tukey HSD post hoc test
MS	13.030	<.001*	.216	6,7 < 8,9 < 10,11,12,13
RY	10.648	<.001*	.184	6 < 7,8,9,10,11,12,13/ 7,8,9 < 10,11,12,13
TS	11.647	<.001*	.198	6 < 8,9,10,11,12,13/ 7, 8 < 10,11,12,13/ 9 < 11,12/ 10 < 12
VS	3.757	<.001*	.080	6 < 11,12,13
RS	1.130	.342	.023	ns
ES	5.349	<.001*	.102	6 < 7,8,9,10,11,12,13
WR	32.384	<.001*	.407	6 < 8,9,10,11,12,13/ 7,8 < 9,10,11,12,13/ 9 < 10,11,12,13
RE	27.931	<.001*	.372	6 < 7,8,9,10,11,12,13/ 7,8 < 9,10,11,12,13/ 9 < 10,11,12,13
MR	23.795	<.001*	.335	6 < 7,8,9,10,11,12,13/ 7,8 < 9,10,11,12,13/ 9 < 10,11,12,13
IM	19.854	<.001*	.296	6 < 7,8,9,10,11,12,13/ 7 < 8,10,11,12, 13/ 8 < 11,12,13/ 9 < 10, 11,12,13/ 10 < 11,12,13
Total	58.008	<.001*	.551	6 < 7 < 8 < 10 < 11,12,13

Note: *MS* motor skill, *RY* rhythm, *TS* tactile skill, *VS* visual skill, *RS* receptive speech, *ES* expressive speech, *WR* writing, *RE* reading, *MR* mathematical reasoning, *IM* immediate memory

<sup>a</sup>Analysis of Covariance (Ancova); F: Fisher's F;  $\eta_p^2$ : Partial eta squared; \* $p < .01$ ; ns: not significant

**Table 4** Pearson’s correlation matrix between total IQ and TLN-C subtests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) IQ	1											
(2) MS	.352**	1										
(3) RY	.397**	.314**	1									
(4) TS	.383**	.279**	.291**	1								
(5) VS	.246**	.102*	.167**	.173**	1							
(6) RS	.134**	.251**	.212**	.128**	.140**	1						
(7) ES	.164**	.299**	.210**	.137**	.270**	.241**	1					
(8) WR	.495**	.232**	.386**	.368**	.323**	.099*	.143**	1				
(9) RE	.452**	.228**	.317**	.319**	.326**	.112**	.217**	.795**	1			
(10) MR	.564**	.252**	.350**	.294**	.360**	.114**	.206**	.587**	.604**	1		
(11) IM	.434**	.193**	.323**	.302**	.238**	.171**	.233**	.376**	.337**	.350**	1	
(12) Total	.641**	.485**	.619**	.609**	.440**	.282**	.339**	.824**	.788**	.750**	.553**	1

Note: MS motor skill, RY rhythm, TS tactile skill, VS visual skill, RS receptive speech, ES expressive speech, WR writing, RE reading, MR mathematical reasoning, IM immediate memory

\* $p < .05$

\*\* $p < .01$

Reasoning subtests, and between total score and Rhythm, Tactile Skill, Writing, Reading and Mathematical Reasoning subtests.

A stepwise regression analysis was conducted. The Durbin-Watson’s result was  $d = 1.89$ , collinearity evaluation yielded tolerance values from .22 to .40 and VIF values from 1.0 to 4.5. The analysis provided three models organized by order (Table 5), in which figured TLN-C’s total score, Mathematical Reasoning and Reading subtests. The total score has the best predictive value about full-scale IQ on WISC-III. TLN-C and WISC-III variance could be explained in 20 % ( $R^2a = .197$ ). In such model, the TLN-C total score predicts full-scale IQ results in 45 % ( $\beta = .45$ ).

Reliability analysis by internal consistency was done using Cronbach’s alpha coefficient on full sample. The obtained value was .79, which is considered satisfactory. Regarding item-total correlation, Receptive Speech subtest is the one with least contribution to internal consistency. Subtests with the highest indexes were Writing, Reading and Mathematical Reasoning. However, no important

**Table 5** Stepwise regression analysis models for TLN-C

Models	R <sup>2</sup> a	Standard Error	F (Anova)	p	β	VIF
1. Total	.197	18.10	141.40	<.001	.45	1.0
2. Total, MR	.218	17.86	81.05	<.001	.26	2.5
3. Total, MR, RE	.231	17.71	58.34	<.05	.43	4.5
					.25	2.5
					-.21	3.1

Note: Total TLN-C’s total score, MR mathematical reasoning, RE reading, R<sup>2</sup>a. adjusted variance; F: Fisher’s F; β: Beta, standardized coefficient; VIF variance inflation factor

coefficient improvement was observed with the exclusion of any subtest (Table 6).

**Discussion**

This study aimed to: (i) obtain validity and reliability evidences for the Luria-Nebraska Test for Children from relations with external criteria (age), (ii) identify scores that predict IQ, and (iii) verify internal consistency.

The age effect analysis was performed controlling possible effects of full-scale IQ. Results showed that TLN-C’s scores increase with age. There was a systematic progression of the means, especially on the total score. This is an important type of validity evidence in neuropsychological screening tests, since sensitivity to detect changes along the development is one of the main

**Table 6** Internal consistency reliability (Alfa de Cronbach)

TLN-C	Scale mean of item deleted	Scale variance if item deleted	Correted item-total correlation	Cronbach’s Alpha if item deleted
MS	88.24	128.710	.397	.776
RY	93.25	119.086	.466	.768
TS	86.74	111.370	.446	.777
VS	89.72	140.285	.313	.787
RS	95.51	144.230	.151	.794
ES	94.42	144.695	.236	.793
WR	89.65	90.333	.756	.721
RE	94.16	103.758	.736	.727
MR	88.73	109.298	.608	.747
IM	91.32	129.473	.494	.769

Note: MS motor skill, RY rhythm, TS tactile skill, VS visual skill, RS receptive speech, ES expressive speech, WR writing, RE reading, MR mathematical reasoning, IM immediate memory

parameters that allow the establishment of normative data (Pasquali, 2010).

Along the development from preschool age to adolescence there is acquisition and refinement of cognitive functions. This result is supported by the maturation of the nervous system (especially the myelination and optimization of neural networks by synaptic pruning) and environmental stimulation that usually puts the child before many cognitive challenges, mainly in school activities (Osborn & Pereira, 2012).

The detection of differences in almost all TLN-C's subtests points that it has effectively measured both perceptual-motor and abstract functions, successfully differentiating development levels. This differentiation is carried out by detecting the maturation level of basic perceptual-motor functions and development level of academic skills. These two axes present on TLN-C, the first one with little influence from formal education and the second one directly linked to it, help to understand the increasing differences found on the performances until 10 years, the relative separation between ranges 6–10 and 11–13 years old, and the systematic differences in total score. It is especially relevant that differences among ages were present controlling IQ influence (except for one subtest), which confirms that they are related to age.

The verification of age effects is common in cognitive test validation, since cognitive functions can develop with aging and experience. This external variable is so relevant in this kind of assessment that, after the normalization process, it is common for normative tables of reference for result interpretation to be organized by age ranges. A recent example is the validation e normatization of the newest Brazilian adaptation of the WISC (Rueda et al. 2013).

The Receptive Speech subtest was the only one insensible to detect any changes with age. This subtest measures a basic cognitive skill, in the sense that it is a prerequisite for children to be able to comprehend what is demanded of them whenever they receive a verbal instruction. Even in this case, gains in this ability are expected along children development as they manage, increasingly, to: (i) comprehend more elaborate verbal sentences; (ii) retain more content as their immediate memory improves; and (iii) organize them with their working memory (Carneiro, 2008; Dias & Landeira-Fernandez, 2011). Therefore, the absence of differences on this score points to the need of task reformulation, so it may entail more levels of complexity.

Another observed result was the small changes in subtest means on the range from 9 years on, and in the total score from 11 years on. These results provide evidence about subtest difficulty and its adequacy to the age range the test is designed for. In a screening test it's especially

important to include simple items, enabling the detection of subtle deficits, and to avoid to include overly demanding items. The absence of differences found between some age ranges may point to the need of inclusion of harder items in several subtests, so they may become more sensible to performance differences on the range from 9 to 13 years.

Furthermore, there was no ceiling effect and an interruption on the progression of means was found in some subtests. A ceiling effect is expected for some of TLN-C subtests because of their content (e.g. the notion of left and right, present on the Tactile Skill subtest, depend on age, and skills such as reading and mathematical reasoning depend on years of instruction) and task difficulty, which is not scalar, so that even the most difficult of them is not challenging.

In most cases, this data behavior may be explained by the sample of the study being composed of children with learning difficulties. In previous studies, the LNNB proved to be sensible in detecting performance differences between subjects with and without learning disabilities (Lewis et al., 1993; Myers et al., 1989). In this sense, the variations found may be related to the sensibility of the test to detect deficits in this population; however, comparative studies are needed to test such hypothesis. This kind of study may also help to clarify whether the similar performance of higher and lower ages in some subtests is due to a real lack of discrepancy on these functions during the developmental period covered by the test, or whether older children with learning difficulties show a performance similar to younger children due to deficits in cognitive functions. Moreover, the interruption of progression of scores occurred only in a few subtests and were insufficient to establish a new pattern.

The Pearson analysis showed that all subtests and the total score of TLN-C correlated with WISC-III's full-scale IQ. Both total scores are measures that reflect the performance on a heterogeneous set of cognitive functions. The adequate functioning of part of the functions assessed by TLN-C may be considered prerequisites for an individual to produce adequate answers on the WISC-III (exceptions being Reading, Writing and Mathematical Reasoning). For instance, a minimum of motor skill is needed in the performance tests, both these and the verbal tests have oral instructions, requiring the use of receptive speech, and the response to the second group of tasks demands the use of expressive speech.

These relations reflect the theoretical principles that neuropsychological functioning and intellectual ability are closely related and affect each other (Ardila & Bernal, 2007). In a study with the original battery for children, Gilger and Geary (1985) detected a good capability of the LNNB-CR to trace neuropsychological

deficits in expressive and receptive language functions, which were in accordance with discrepant results between verbal and performance scales in the WISC-R. More recent studies, with another largely used neuropsychological battery, the Halsted-Reitan Neuropsychological Battery, are also grounded on relations between intelligence and neuropsychological functions. A study with children presenting learning disabilities showed distinct result profiles in this battery in children from the various inferior ranges of the WISC-R (Davis et al. 2001).

Significant correlations were found among all subtests of TLN-C, showing cohesion of the test as a whole. The magnitudes of the correlations show patterns well-related to theoretical foundations. Subtests from the axis of academic skills had moderate to high magnitudes. Correlations between items with small theoretical relation, like Rhythm and Visual Skill, had low magnitudes. A finding that reinforces the cohesion of the test as a whole is that, generally, the highest correlation magnitudes happened between subtests and the total score. The obtained correlations between TLN-C and the WISC suggest validity evidences from relations with external variables, in this case, with a previously standardized instrument. Furthermore, the correlation among subtests of TLN-C suggests cohesion throughout its scores.

The regression analysis results reinforce the importance of the total score, adding to its property of reflecting the internal coherence of TLN-C, the property of contributing to intellectual performance in this sample. The results suggest that the total score of TLN-C explains better the IQ. This characteristic is in accordance with the fact that both the total score of TLN-C and of the WISC are heterogeneous and correlated measures, as discussed previously (Pfeiffer et al., 1987; Boyd & Hooper, 1993). The fact that models considering specific subtests along with the total score were less effective predictors also agrees with what we presented above about the support neuropsychological functions provide to intellectual performance.

Boyd and Hooper (1993), in a study of multivariate regression models involving age and the performance on the original battery for adults found the verbal IQ and, more markedly, the full-scale IQ, to have predictive capabilities. From their results, they suggested that the LNNB is as good as abbreviated forms of the WISC to predict intellectual performance.

The group of evidences about the total score of TLN-C, gathered in the present study, contributes with validity evidences of the instrument as a whole. However, as Pawlowski et al. (2007) point out, in an instrument of fast application that involves the assessment of several theoretical constructs (neuropsychological functions, in this case), it is also important to gather evidences about individual subtest validity, the way they are internally related and the way they relate to the total score. A step in

this direction was made in this work by the correlation analysis among subtests, and it may be complemented by other procedures, always respecting the characteristics of TLN-C, as follows: factorial analysis, relations with instruments or their parts that assess constructs similar to one or some subtests of TLN-C, and relations of the test with other external criteria apart from intelligence. It is also important to collect comparative data between control and criterion groups, since the sample of children presented herein shows learning difficulties.

Referring to the precision or reliability of TLN-C, the Cronbach's alpha coefficient showed a satisfactory result (.79). According to the Resolution 002/2003 of the Brazilian Federal Council of Psychology (CFP, 2003), the minimal acceptable value for this index is .60. Freire and Almeida (2001) suggested value intervals for classification: .80-.90, very good; .70-.80, respectable; .65-.70, acceptable; .60-.65, undesirable; below .60, unacceptable. It is also relevant to point out the coherence shown by the fact that subtests Writing, Reading, Immediate Memory presented more links with most of the other test items, since they represent complex cognitive functions that are supported by many simpler functions assessed by other subtests. The low contribution of Receptive Speech to internal consistency comes alongside the other findings about this subtest, which indicates psychometric inadequacy in its present configuration.

Moreover, in spite of being commonly applied (Ladesma et al. 2002), it should be noted that Cronbach's alpha may not be the best procedure to evaluate the reliability of batteries or screening instruments. Such instruments usually involve an important diversity of functions, which constructs are not immediately related, despite the correlations found in our results indicating that there is at least a global coherence among the subtests of the instrument evaluated here.

A closely related theoretical problem was found on the validation process of the NEUPSILIN, and the authors propose some alternatives to Cronbach's alpha that may be useful in complementing the reliability evidences of TLN-C (Pawlowski et al., 2007). Alternatives proposed by the authors are the agreement among judge scores and the test-retest procedure, which has already been used in the validation of the original battery, with results of 75 % mean steadiness between results (Plaisted & Golden, 1982).

The present study is part of a large project that has aimed to provide the TLN-C for clinical use. Notwithstanding the relevance of this study, there were limitations that should be addressed in subsequent research: (i) conduct performance comparison in TLN-C by genders and clinical subgroups; (ii) analyze correlations between the subtests of WISC and TLN-C. Moreover, studies are needed to investigate other type of the validity, as well as the normalization of the instrument.

## Conclusion

The study results provided several validity evidences for TLN-C: (i) accordance with external criteria, mainly with development, as showed by effects of age on the performance; (ii) accordance with external variable, as showed in significant correlations with a standardized test of intellectual assessment (IQ—WISC-III); (iii) predictive, expressed on the verification that the total score serves as predictor of full-scale IQ on the WISC-III; (iv) reliability, with a satisfactory alpha coefficient.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

RFL, conception and design, acquisition of data, analysis and interpretation of data, writing the article and revising it critically; MS, acquisition of data, interpretation of data, writing and revising it critically; FL, acquisition of data, critical review; PAPC, critical review; SMC, conception, critical review and orientation. All authors read and approved the final manuscript.

## Author details

<sup>1</sup>Universidade Estadual de Campinas, Campinas, Brazil. <sup>2</sup>Universidade de São Paulo, Bauru, Brazil.

Received: 15 March 2016 Accepted: 7 April 2016

Published online: 18 April 2016

## References

- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing. Washington, DC: AERA; 1999.
- Ardila A, Bernal B. What can be localized in the brain? Toward a "factor" theory on brain organization of cognition. *Int J Neurosci*. 2007;117:935–69. doi:10.1080/00207450600912222.
- Argollo N, Bueno OFA, Shayer B, Godinho K, Abreu K, Durán P, et al. Adaptação transcultural da Bateria NEPSY – avaliação neuropsicológica do desenvolvimento: estudo piloto. *Avaliação Psicol*. 2009;8(1):59–75.
- Boyd TA, Hooper SR. WISC-R IQ estimates from the Luria-Nebraska Neuropsychological Battery. *Percept Mot Skills*. 1993;77(2):683–8. doi:10.2466/pms.1993.77.2.683.
- Carneiro MP. Desenvolvimento da memória na criança: o que muda com a idade? *Psicologia*. 2008;21(1):51–9. doi:10.1590/S0102-79722008000100007.
- Ciasca SM. Distúrbios e dificuldades de aprendizagem em crianças: análise do diagnóstico interdisciplinar. Campinas: Tese de Doutorado, Faculdade de Ciências Médicas, Universidade Estadual de Campinas; 1994.
- Conselho Federal de Psicologia CFP (2003). Resolução N° 002/2003. Disponível em <http://migre.me/tvMue>. Acessado em 04/03/2014.
- Crenitte PAP, Batista ADS, Silva L, De Lima RF, Ciasca SM. Estudo piloto da adaptação da bateria neuropsicológica Luria-Nebraska para crianças (LNNB-C). *Revista Psicopedagogia*. 2011;28(86):117–25.
- Davis B, Krug D, Dean RS. Neuropsychological clusters within intelligence levels for learning disabled children. *Int J Neurosci*. 2001;106(3-4):239–51.
- Dias LBT, Landeira-Fernandez J. Neuropsicologia do desenvolvimento da memória: da pré-escola ao período escolar. *Revista Neuropsicol Latinoamericana*. 2011;3(1):19–26. doi:10.5579/rnl.2011.0061.
- Figueiredo VLM. WISC-III: Escala de Inteligência Wechsler para crianças - manual. Adaptação e padronização de uma amostra brasileira. São Paulo: Casa do Psicólogo; 2002.
- Freire T, Almeida LS. Escalas de Avaliação: construção e validação. In: Fernandes EM, Almeida LS, editors. *Métodos e Técnicas de Avaliação*. Minho: Universidade do Minho, Centro de Estudos em Educação e Psicologia; 2001.
- Gilger JW, Geary DC. Performance on the Luria-Nebraska Neuropsychological Test Battery-Children's Revision: a comparison of children with and without significant WISC-R VIQ-PIQ discrepancies. *J Clin Psychol*. 1985;41(6):806–11. doi:10.1002/1097-4679(198511)41:6<806::AID-JCLP2270410614>3.0.CO;2-N.
- Golden CJ. *Luria-Nebraska Neuropsychological Battery: Children's Revision*. Los Angeles: Western Psychological Services; 1987.
- Gonçalves HA, Mohr RM, Moraes AL, Siqueira LS, Prando ML, Fonseca RP. Componentes atencionais e de funções executivas em meninos com TDAH: dados de uma bateria neuropsicológica flexível. *J Bras Psiquiatr*. 2013;62(1):13–21. doi:10.1590/S0047-20852013000100003.
- Hooper SR. Relationship between the Luria-Nebraska Neuropsychological Battery and Woodcock-Johnson Tests of Achievement-Revised in children with psychiatric impairment. *Percept Mot Skills*. 1995;80(3):1353–4. doi:10.2466/pms.1995.80.3c.1353.
- Ladesma R, Adánez GM, Mora PV. Análise de consistência interna mediante Alfa de Cronbach: um programa baseado em gráficos dinâmicos. *Psico-USF*. 2002;7(2):143–52.
- Lewis RD, Hutchens TA, Garland B. Cross-validation neuropsychological of the discriminative battery for learning effectiveness of the Luria-Nebraska disabled adolescents. *Arch Clin Neuropsychol*. 1993;8(5):437–47. doi:10.1016/0887-6177(93)90007-N.
- Lima RF, Mello R, Massoni I, Riechi T, Ciasca SM. Teste Luria-Nebraska para Crianças -TLN-C. Campinas: Manual técnico, Faculdade de Ciências Médicas, Universidade Estadual de Campinas; 2005.
- Myers Jr D, Sweet JJ, Deysach R, Myers FC. Utility of the Luria-Nebraska Neuropsychological Battery-Children's Revision in the evaluation of reading disabled children. *Arch Clin Neuropsychol*. 1989;4(3):201–15. doi:10.1093/arcin/4.3.201.
- Osborn E, Pereira LD. Neuropsychological aspects of 10-year-old children. *Einstein (São Paulo, Brazil)*. 2012;10(4):433–8. doi:10.1590/S1679-45082012000400007.
- Pasquali L. *Instrumentação psicológica: fundamentos e práticas*. Porto Alegre: Artes Médicas; 2010.
- Pawlowski J, Rodrigues JC, Martins SCO, Brondani R, Chaves MLF, Fonseca RP, et al. Evaluación neuropsicológica breve de adultos post-accidente cerebrovascular em el hemisferio izquierdo. *Avances Psicol Latinoamericana*. 2013;31(1):33–45.
- Pawlowski J, Trentini CM, Bandeira DR. Discutindo procedimentos psicométricos a partir da análise de um instrumento de avaliação neuropsicológica breve. *Psico-USF*. 2007;2(12):211–9. doi:10.1590/S1413-82712007000200009.
- Pfeiffer SJ, Naglieri JA, Tingstrom DH. Comparison of the Luria-Nebraska Neuropsychological Battery-Children's Revision and the WISC-R with learning disabled children. *Percept Mot Skills*. 1987;65(3):911–6. doi:10.2466/pms.1987.65.3.911.
- Plaisted JR, Golden CJ. Test-retest reliability of the clinical, factor, and localization scales of the Luria-Nebraska Neuropsychological Battery. *Int J Neurosci*. 1982;17(3):163–7.
- Romanelli EJ, Riechi TJS, Ambrósio CR, Gardens GS, Mitzuck MT, Oliveira MAF, et al. Análise do processo de adaptação e padronização da Bateria Neuropsicológica Luria-Christensen para a população brasileira. *InterAÇÃO*. 1999;3:61–78.
- Rueda FJM, Noronha APP, Sisto FF, Santos AAA, Castro NR. *Escala Wechsler de Inteligência para Crianças – 4ª Edição (WISC-IV): Manual Técnico*. São Paulo: Casa do Psicólogo; 2013.
- Salles JF, Fonseca RP, Cruz-Rodrigues C, Mello CB, Barbosa T, Miranda, MC. Development of the Child Brief Neuropsychological Assessment Battery NEUPSILIN-INF. *Psico-USF*. 2011;16(3):297-305. <https://dx.doi.org/10.1590/S1413-82712011000300006>.
- SATEPSI - Sistema de Avaliação de Testes Psicológicos. *Lista dos Testes Aprovados*. Disponível em <http://satepsi.cfp.org.br/listaTeste.cfm>. Acessado em 15 de novembro de 2013.
- Snow J. Factor structure of the Luria-Nebraska neuropsychological battery-children's revision with learning-disabled children. *J Sch Psychol*. 1985;23(3):271–5. doi:10.1016/0022-4405(85)90019-6.
- Teichner G, Golden CJ, Bradley JDD, Crum TA. Internal consistency and discriminant validity of the Luria Nebraska Neuropsychological Battery - III. *Int J Neurosci*. 1999;98(1-2):141–52. doi:10.3109/00207459908994797.
- Ustároz JT. La evaluación neuropsicológica. *Intervención Psicosocial*. 2007;16(2):189–211.
- Wechsler D. *Wechsler Intelligence Scale for Children—Third Edition (WISC-III)*. San Antonio: The Psychological Corporation; 1991.
- Yates DB, Zibetti MR, Pawlowski J, Salles JF, Parente MAMP, Argimon IL, et al. WCST and Neupsilin: Relationships among executive functions, attention, memory and language. *Psicologia*. 2013;26(3):506–15. doi:10.1590/S0102-79722013000300010.
- Zibetti MR, Gindri G, Pawlowski J, Salles JF, Parente MAMP, Bandeira DR, et al. Estudo comparativo de funções neuropsicológicas entre grupos etários de 21 a 90 anos. *Revista Neuropsicol Latinoamericana*. 2010;2(1):55–67.