_____

The effects of computerized cognitive testing on the performance of children with and without Attention Deficit/Hyperactivity Disorder symptoms

Los efectos de las pruebas cognitivas computarizadas en el rendimiento de los niños con y sin síntomas de Trastorno de Déficit de Atención/Hiperactividad

Os efeitos da testagem computadorizada na performance em testes cognitivos de crianças com e sem sintomas de Transtorno de Déficit de Atenção/Hiperatividade

Mariana Braga Fialho [1], ORCID 0000-0002-2445-5945
Elisa Rodrigues de Souza [2], ORCID 0000-0002-7783-6944
Marcela Mansur-Alves [3], ORCID 0000-0002-3961-3475

[1] *Centro Universitário Newton Paiva Belo Horizonte, Brazil*
[2] *Pontifícia Universidade Católica de Minas Gerais (PUCMinas), Brazil*
[3] *Universidade Federal de Minas Gerais (UFMG), Brazil*

***Abstract*:** The objective was to assess the impact of a test digital format on evaluating cognitive functions of children with ADHD symptoms, along with the differential impacts in comparison to the typically developed groups. The sample consisted of 99 children aged 7 to 9 years (M = 7.99, SD = 0.802), being 52 on the clinical group. All participants were assessed with paired batteries of digital and traditional tests. There were significant differences between the evaluated groups on known-compromised constructs for children with ADHD (d= -0,027 to -0,617), the highest being on the digital tasks. However, there were no significant differences in performance when comparing the results within the groups according to their computer-based and paper-and-pencil measures scores. The results suggest that the digital format does not impair the evaluation of children with ADHD symptoms, does not generate significantly different impacts between the clinical and comparison groups. Limitations and possible implications of these results will be discussed.

Keywords: cognitive evaluation; digital test; attention deficit hyperactivity disorder.

***Resumen*:** El objetivo fue evaluar el impacto del formato digital de prueba en la evaluación de las funciones cognitivas de los niños con síntomas de TDAH, junto con los impactos diferenciales en comparación con los grupos típicamente desarrollados. La muestra consistió en 99 niños de 7 a 9 años (M = 7.99, DE = 0.802), siendo 52 en el grupo clínico. Todos los participantes fueron evaluados con baterías emparejadas de pruebas digitales y tradicionales. Hubo diferencias significativas entre los grupos evaluados en construcciones comprometidas conocidas para niños con TDAH (d= -0.027 a -0.617), el más alto en las tareas digitales. Sin embargo, no hubo diferencias significativas en el rendimiento al comparar los resultados dentro de los grupos de acuerdo con sus puntajes de medición basados en computadora y en papel y lápiz. Los resultados sugieren que el formato digital no perjudica la evaluación de los niños con síntomas de TDAH, no genera impactos significativamente diferentes entre los grupos clínicos y de comparación. Se discutirán las limitaciones y posibles implicaciones de estos resultados.

Palabras clave: evaluación cognitiva; prueba digital; trastorno por déficit de atención e hiperactividad.

***Resumo*:** O objetivo do estudo foi mensurar o impacto da testagem digital para avaliar as funções cognitivas de crianças com sintomas de TDAH, assim como os impactos diferenciais em comparação com grupos com desenvolvimento típico. A amostra foi composta por 99 crianças entre 7 e 9 anos (M = 7,99, DP = 0,802), sendo 52 do grupo clínico. Todos os participantes foram avaliados com baterias emparelhadas de testes digitais e tradicionais. Houve diferenças significativas entre os grupos em construtos tipicamente apontados por serem comprometidos em crianças com TDAH (d= -0,027 a -0,617), com maiores tamanhos de efeito nas tarefas digitais. Entretanto, não houve diferenças significativas na performance ao comparar os resultados dentro dos grupos ao considerar as pontuações nos testes computadorizados e suas versões de papel. Os resultados sugerem que o formato digital não prejudica a avaliação de crianças com sintomas de TDAH, não gerando impactos significativamente diferentes entre os grupos clínico e de controle. Limitações e possíveis implicações destes resultados serão discutidas.

Palavras-chave: avaliação cognitiva; teste computadorizado; transtorno de déficit de atenção e hiperatividade.

———

Correspondence: Elisa Rodrigues de Souza, Pontifícia Universidade Católica de Minas Gerais (PUCMinas), Brazil. E-mail: elisardsouza@gmail.com

Historically, technology and modernization have vastly influenced the way we live and work. Notwithstanding, cognitive assessment measures remain similar to what they were several years ago. Although the area has advanced in the last decade, even the digitally developed tools have yet to be sufficiently explored. Among these measures, we can define computer-based/digital instruments as those that use the computer interface or other digital devices in their administration, scoring or interpretation (Kane & Parsons, 2017; Parsey & Schmitter-Edgecombe, 2013).

Computerized testing guarantees advantages, such as more motivating tasks, greater standardization in the application, automated scoring and immediate feedback. Thus, many studies are focused on analyzing the comparability between digital and traditional forms of evaluation. The development itself of computerized testing aimed for this greater efficiency (less expeditious applications, adaptive possibilities and broader reach), as well as for better psychometric properties (Csapó, Molnár & Nagy, 2014; Moncaleano & Russell, 2018; Zygouris & Tsolaki, 2015).

Conversely, possible disadvantages of said tools must also be considered, including the following: the level of the examinee's familiarity with digital devices, which may influence their performance; discrepancy between data generated on different platforms due to hardware and operating systems; possible reduction of direct contact between examinee and evaluator; higher potential for misuse, among others (Bauer et al., 2012; Lumsden, Edwards, Lawrence, Coyle & Munafò, 2016; Zygouris & Tsolaki, 2015).

Another aspect to be examined is the adaptation itself of the paper-and-pencil tests to the digital environment. New psychometric studies must be done taking into account the possibility of disparity between the two versions properties. The diverse effects caused by a digital testing environment must also be contemplated owing to distinct groups reacting differently to the platform. Individuals with ADHD (Attention Deficit / Hyperactivity Disorder), for example, do not react as healthy controls do to highly motivating tasks, which may affect the outcome on computer-based measures (Csapó et al., 2014; Kane & Parsons, 2017; Lumsden et al., 2016).

In order to better comprehend the differences, one needs to study ADHD itself. This disorder is understood as a complex disorder of neurobiological development, which symptoms may present in the preschool years and extend into adulthood. According to the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5; American Psychiatric Association [ADA], 2013), ADHD is characterized by persistent symptoms of inattention and/or hyperactivity-impulsivity. It can lead to personal, social, academic and professional losses, as along with worse performance in tasks that require attention and executive functions (EF). ADHD is a common disorder in childhood, with an estimated prevalence of 5.3 % among children and adolescents (Mahone, 2012; Willcutt, 2012).

Different models are proposed to comprehend the cognitive profile of individuals with ADHD. Among them, there is Barkley's that describes a hierarchical relationship in which a central deficit in response inhibition would lead to secondary impairments in other executive functions, such as self-regulation, working memory, discourse internalization and reconstitution of experiences. The aforementioned damages would then lead to decreased control of motor behavior, rendering the individual inept to environmental demands. Sergeant argues that these deficits may not be specific to ADHD, but they are possibly related to cognitive-energy dysfunctions. Sonuga-Barke and Halperin, however, attribute the

disorder's cognitive heterogeneity to the multiple developmental pathways, associated with inhibitory control, reward mechanisms and temporal perception, in which different deficits may be complementary (Barkley, 1997; Sergeant, 2005; Sonuga-Barke & Halperin, 2010).

A unique cognitive model that fully explains the clinical phenotype of ADHD is yet to be described. Nonetheless, the literature has consistently demonstrated the association between the disorder and deficits in various cognitive functions, with emphasis on inhibitory control, working memory, sustained attention and processing speed (Delgado et al., 2012; Messina & Tiedemann, 2009; Rueda & Muniz, 2012; Willcutt, Doyle, Nigg, Faraone & Pennington, 2005).

Despite the association between ADHD and worse performance in various cognitive functions at the individual level, these factors are not sensitive enough nor are the tests sufficient for a diagnostic, which is done using clinical assessment. However, considering ADHD's cognitive heterogeneity, it is of utmost importance to carry out a comprehensive and cognitive evaluation in order to draw up a treatment plan fit for each individual. To do so, it is necessary to identify appropriate tools for the evaluation of individuals with such disorder. Another aspect to be considered is whether the attractiveness and similarities to games of computer-based testing enables the release of dopamine, improving attention levels, thus masking attention deficit (Dovis, Van der Oord, Wiers & Prins, 2011; Lumsden et al., 2016).

Among the computer-based measures currently being used in evaluations for children with ADHD, there are CPTs (Continuous Performance Tests) and cognitive assessment batteries such as CANTAB (Cambridge Neuropsychological Test Automated Battery). Nonetheless, the comparison between digital and paper-and-pencil measures, as well as possible impacts of this format to the cognitive evaluation of people with ADHD, is still underexplored, lacking studies especially in the Brazilian scenario. The disorder is also present in a significant part of the population, causing broad damages to those individuals, and establishing, therefore, an urgent demand for the aforementioned studies (Fried, Hirshfeld-Becker, Petty, Batchelder & Biederman, 2012; Wang et al., 2011).

This study's objective was to assess the impact of a test digital format on evaluating cognitive functions of children with ADHD symptoms, along with the differential impacts in comparison to the typically developed groups within different-format tasks. In addition to that, the study also aimed to verify whether there was a significant difference between effect sizes on computer-based and alike paper-and-pencil tests on the aforementioned groups. Among the hypothesis we expect significant differences between the groups' scores. There is also the possibility that computer-based testing could be used for the clinical group, and that the impact of its platform would not be significantly different in comparison to the latter group.

_____

Method

The present study is part of a larger research project developed in the Laboratory of Evaluation and Intervention in Health (LAVIS) of the Department of Psychology of UFMG (Universidade Federal de Minas Gerais), in partnership with the Center for Development of Technologies of Inclusion (CEDETI) of the Pontifical Catholic University of Chile (PUC-Chile), to do a cross-cultural adaptation and validation of TENI (Test of Infant Neuropsychological Evaluation) for Brazil. The study was approved by UFMG's Research Ethics Committee under CAAE (51216815.9.0000.5149).

Participants

The sample of the present study was selected from 237 children in 2nd, 3rd and 4th grade from public and private schools. It was calculated considering the total population of 92,358 children enrolled in the 2nd, 3rd and 4th grades of elementary school in the city of Belo Horizonte, the capital of the state of Minas Gerais (Brazil). The confidence level was 90 % and a sampling error of 5 %, considering that approximately 5 % of the children population would present ADHD. Thus, the indicated sample was 52 children with symptomatology compatible with ADHD, which was precisely the number of the final sample obtained in the study. Among those in the clinical group, 9 (17 %) were using psychopharmaceuticals to treat ADHD. These children had their participation previously authorized by their parents, who received and filled out a socioeconomic questionnaire and a scale of perceived inattention and hyperactivity symptoms (Swanson, Nolan & Pelham Scale Version IV - SNAP IV; Mattos, Serra-Pinheiro, Rohde & Pinto, 2006).

With the provided information, 55 children with a clinical score (as defined by Costa, de Paula, Malloy-Diniz, Romano-Silva & Miranda, 2018) of 1.72 for inattention and 1.17 for hyperactivity in SNAP-IV were selected, three of them were excluded from the sample by reason of intellectual disability or intelligence quotients below 70. The 47 members of the control group were then selected so that there was no significant difference in age, sex, school type or grade to the clinical one (sample pairing). Table 1 presents descriptive and comparative data of the clinical and control groups, including their characterization.

Table 1.
Characterization of the clinical and non-clinical sample, according to age, school type, sex, maternal schooling, family income, IQ and school year (n = 99)

	Total sample	Non-clinical	Clinical	
Variable				***p***
n	99	47 (47.5%)	52 (52.5%)	_
Sex	Male 60 (60.6%)	Male 27 (27.27%)	Male 33 (33.3%)	.681[a]
Age	*M* = 7.99 (*SD* = 0.802)	*M* = 8.11 (*SD* = 0.759)	*M* = 7.88 (*SD* = 0.832)	.177[b]
School year	2nd - 36 (36.4 %) 3rd - 34 (33.3%) 4th - 29 (29.3%)	2nd - 13 (13.1%) 3rd - 19 (19.2%) 4th - 15 (15.2%)	2nd - 23 (23.2%) 3rd - 15 (15.2%) 4th - 14 (14.1%)	.183[b]
Type of school	Pub. 52 (52.5%)	Pub. 20 (20.2%)	Pub. 32 (32.3%)	.071[a]
Family income	1 to 5 wages - 59 (59.6%) > 5 wages - 37 (37.4%)	1 to 5 wages - 22 (22.2%) > 5 wages 22 (22.2%)	1 to 5 wages - 37 (37.4%) > 5 wages 15 (15.2%)	.351[a]
Degree of Mother's Instruction	a- 4 (4.21%) b- 8 (8.43%) c- 40 (42.11%) d- 43 (45.27%)	a- 1 (1.05%) b- 2 (2.11%) c- 21 (22.11%) d- 21 (22.11%)	a- 3 (3.16%) b- 6 (6.32%) c- 19 (20%) d- 22 (23.16%)	.430[b]
IQ	*M* = 100.23 (SD = 13.9)	*M* = 102.44 (SD = 14.4)	*M* = 98.23 (SD = 13.2)	.115[b]

Notes: M = Mean; SD = Standard Deviation; Pub = Participants from public schools. Mothers or caretaker's degree of education: A = illiterate to incomplete elementary school; B = complete elementary school to incomplete secondary education; C = complete high school to incomplete university; D = complete higher education. [a] = Pearsons Chi-Squared Test; [b] = Mann-Whitney U test; IQ = intelligence quotient. All *p* > .05 demonstrating that the difference between the groups were not significant.

_____

Materials & Instruments

Neuropsychological Evaluation of Children - TENI

TENI is a computer-based cognitive assessment battery for children aged 3 to 9 years developed by CEDETI of PUC-Chile. It is on tablets and has nine subtests, which evaluate various cognitive constructs ensuring a broad examination in a short period of time. All subtests were created as games with an attractive and user-friendly interface. Its Chilean version has good psychometric properties and presents high evidence of validity and reliability (Cronbach's alpha between .8 and .9). Further analysis was done by specialists, calculating the correlations as well as using the test-retest and split halves methodology to evaluate the tool's reliability (Delgado et al., 2012). The Brazilian cross-cultural adaptation and validation was done with a sample of typically developed children from Minas Gerais. The factor analysis indicated acceptable and significant results. The intercorrelation calculated between subtests evidenced a convergent and discriminant validity along with significant and strong correlation among subtests that assessed similar constructs. Moreover, tasks that evaluated theoretically distinct constructs pointed to divergence. The subtests used in the present study are described in Table 2 (Martins, Barbosa-Pereira, Valgas-Costa & Mansur-Alves, 2020).

Table 2.
Description of TENIs subtests

Subtest	Evaluated function	Short description	Scoring (Minimum/ Maximum)
The Alternative Universes	Concentrated Attention	Pairs of similar images are presented to the child. He/she should then point out the difference between them.	Correct answers (0/10).
Todd and the Earthworms	Sustained Attention	On the tablet screen the child sees a conveyor belt where apples are rolling. They are asked to touch the screen whenever an apple with a worm appears.	Touches classified as correct (0/176), omission errors (0/176) or commission errors (0/no max.). Answer time is also registered in milliseconds.
Tic-Tac	Naming Speed	A screen with drawings (ball, house, cat, apple, elephant, tree) is presented to the child, she/he is asked to name them as quick as possible	Time registered in milliseconds, the highest the value, the lowest the score.
Bzz!	Visuomotor Skills	There are bees flying randomly on the screen. The child has to touch and "smash" as many as possible within a minute.	The child has a minute to click on as many bees as possible (there always 10 on the screen) and the system registers the amount "smashed".
Bzz! Inhibition	Executive Functions	During the Bzz! Game the evaluator leaves the room for five minutes after instructing the child not to touch the screen whilst he/she is gone.	The system registers for how long the child remained without touching the screen (0 seconds/ 300 seconds).

The Mexican House (copy)	Visuospatial Skills/Planning	The child is asked to copy a drawing on the table using their fingers, the model is displayed on a separated piece of paper.	There are 8 evaluated elements on the drawing. Each one receives a score ranging from zero for wrong structure and location to four for correct structure and location. It adds up to 32 points (0/32).
The Mexican House (recall)	Memory	After copying the model, the child is asked to draw the figure from memory. The paper model is no longer available.	There are 8 evaluated elements on the drawing. Each one receives a score ranging from zero for wrong structure and location to four for correct structure and location. It adds up to 32 points (0/32).
The Mole	Executive Functions-Working Memory	On one screen you see a grid with holes in which a mole appears on. The child should observe the sequence in which it appears and touch the holes in the same order.	The system registers how many clicks the child got right (0/14).
The Farm	Executive Functions - Serialization	The child is shown a sequence of animals and a line at the end indicating that one is missing. The child must choose which completes the series correctly.	Correct answers (0/27).

Wechsler Abbreviated Intelligence Scale–WASI

WASI is a brief assessment of intelligence suitable for individual applications in clinical and research settings. The instrument is composed of four subtests (vocabulary, block design, similarities and matrix reasoning), and the IQ, however, can be calculated with only two subtests (vocabulary and matrix reasoning). The tool has a high reliability index, with Cronbach's alpha ranging from .82 to .92 on the subtests (Trentini, Yates & Heck, 2014).

Rey's Complex Figure

It evaluates the neuropsychological functions of visual perception and immediate memory. Additionally, it is widely used to investigate issues related to the actions planning and execution beyond visual memory. Its psychometric studies evidenced a high internal consistency, with Cronbach's alpha estimated at .89 for copy and .83 for recall (Oliveira & Rigoni, 2010).

Corsi Block-Tapping Test

It evaluates visuospatial short-term memory and its executive function component, when there is inversion of the items, and the task requires the use of working memory. The test consists of a tray with randomly arranged blocks which the examinee must tap on according to the sequence previously showed by the examiner, firstly in a direct order and then inversely. Although it has not been adapted for the Brazilian population, international studies have shown a linear increase in test performance conforming to age (Corsi, 1972).

Nine Hole Peg Test

It is a measure of finger dexterity, fine manual speed and visual motor coordination that can also be useful for the motor evaluation of different clinical groups (e.g., patients with cerebral palsy, Parkinson's disease and multiple sclerosis). Studies using the tool demonstrated a progressive increase in the manual dexterity alongside age (Kane & Parsons, 2017; Kellor, Frost, Silberberg, Iversen & Cummings, 1971; Willcutt, 2012).

Rapid Naming Task

Also known as Rapid Automatized Naming (RAN), it is a screening task that evaluates one of the precursors of reading. The person must sequentially, quickly and successively recall the name of distinct symbols previously introduced by the evaluator (Capellini, Smythe & Silva, 2017).

Five Digit Test (FDT)

FDT is a measure of nuclear executive functions, such as inhibitory control and cognitive flexibility, as well as simple attentional processes, including reading and counting. The stimuli used is meagerly influenced by formal schooling and social differences. The instrument has good internal consistency and a Guttman coefficient above .90 (Sedó, de Paula & Malloy-Diniz, 2015).

BPA - Psychological Battery for Attention Evaluation

It is a comprehensive evaluation used to assess one's attention and its components, i.e., sustained, divided and alternating attention. The stimuli used (both target and distractor) are abstract shapes, which minimizes the impact of schooling on performance. The General Attention level (AG) is calculated upon completion using the sum of the three subtests. The accuracy was verified using test-retest, and the correlations between applications ranged from r = .68 to .89, p < .05 (Rueda, 2013).

SNAP-IV - Swanson, Nolan, and Pelham, Version IV

SNAP-IV is a public domain questionnaire used to assess the Attention Deficit Disorder and Hyperactivity Disorders (ADHD); the instrument was based on the diagnostic criteria for ADHD established by the DSM-IV. The tool has evidence of validity and reliability with an exploratory and confirmatory factor analysis suggesting an adequate adjustment to the structure of the questionnaire factors. The coefficients of internal consistency were high, with Cronbach's alpha of .94 for inattention and .92 for hyperactivity. Other than that, using a ROC curve analysis, a cutoff score was established, providing balance between sensitivity and specificity. Using the mean score, the cutoff for inattention was 1.72 (AUC=0.877, sensitivity of 0.79 and specificity of 0.81) and 1.17 for hyperactivity (AUC=0.788, sensitivity of 0.70 and specificity of 0.73) (Costa et al., 2018; Mattos et al., 2006).

Socioeconomic Questionnaire

The questionnaire was created for the TENI validation studies, and it had to be filled out by the guardians of the participating children. The questions included topics on the family monthly income, parents' educational level, number of people dependent on said income, neighborhood safety issues, child ethnicity, psychiatric or neurological diagnoses, and medications, among other social and economic indicators.

Procedures

The study was conducted through a partnership with three schools and a reference center for inclusive education. Furthermore, the institutions signed a Consent Form and were informed of all the processes. The students were then invited to participate through the Invitation Letter sent to the guardians. An Informed Consent Form (TCLE, in Portuguese), a socioeconomic questionnaire and a copy of SNAP-IV were also sent along with the letter. Besides the consent of those responsible, the children's assent was requested during the first interview.

The evaluations were carried out by a team of research assistants and psychology professionals during the school year of 2018. The assessments were individual and occurred in two meetings with each child, lasting approximately 70 minutes each. Upon completion and correction of each form, the parents of each participant received a feedback letter describing the child's performance on the tasks. The partner institutions also received feedback through letters with the overall presentation of the participants' results.

Results

Correlations between cognitive subtests in paper-and-pencil and computer-based tools

There was a verified correlation between the participants' performance in both digital and traditional tasks that evaluated similar constructs. As shown in Table 3, the correlations were overall significant with moderate magnitude. Nonetheless, there was a low correlation between the sustained attention digital task (Toddy) and BPA's general attention score. The visuomotor skills evaluation also displayed a low correlation when analyzing the traditional and digital measures to the tasks that required manual dexterity (Nine Hole and Bzz! Visuomotor). Finally, the only correlation that was not statistically significant occurred between the inhibition items on the digital task of BZZ! (gratification delay) and on the FDT (which linked an automated response through a change in environmental demand).

Table 3.

Correlations between digital and traditional tasks

Evaluated function	Digital subtest	Comparable traditional test	r
Concentrated Attention	Alternative Universes	BPA - Concentrated Attention Points	.449**b
Sustained Attention	Toddy - Score	BPA - General Attention Points	.208*b
	Toddy – Ans. Time	BPA - General Attention Points	-.601**b
	Toddy - Omissions	BPA - General Attention Points	-.426**b
	Toddy - Commissions	BPA - General Attention Points	-.391**b
	Toddy - Adjusted Mean	BPA - General Attention Points	.43**b
Naming Speed	TIC Tac	Rapid Naming Task	.604**b
Visuospatial Skills / Hand-Eye Coordination	Bzz! (viso-motor)	Nine Hole Peg Test	.224*a
Executive Functions	Bzz! (inhibition)	FDT - Inhibition	-0.147b
Visuospatial Skills/ Planning	The Mexican House (copy)	Rey Complex Figure (copy)	.465**b
Visual Episodic Memory	The Mexican House (recall)	Rey Complex Figure (recall)	.477**b
Visuospatial Short-Term Memory	The Mole	Corsi Block-tapping test (forward)	.456**b
Executive Functions - Serialization	The Farm	Matrix Reasoning (WASI)	.582**b

Notes: [a] Pearson correlation coefficient. [b] Spearman Correlation Coefficient. * Significant correlation with $p < .05$ (two-tailed). ** Significant correlation with $p < .01$ (two-tailed).

Performance comparison between clinical and nonclinical groups in digital and traditional measures

We compared averages between computer-based and traditional tasks using both raw scores and standardized Z scores on clinical and nonclinical groups. Table 4 presents the averages and deviations (independent T-test or Mann-Whitney on variables with non-normal distribution), and the effect size for the difference.

The children in the clinical group presented significantly worse performances in all the tasks and aspects within Toddy, as well as in the Mexican's House Copy, Bzz! Inhibition and The Farm. On the traditional tests, the control group showed significantly better results on the rapid naming task and on the Nine Hole Peg Test.

Graph 1 shows the values referring to the group comparisons effect size, with the statically significant ones being in bold. The computer-based versions of the sustained attention, visuospatial ability/planning, and serialization tests, all showed significant differences in performance when comparing both groups, and their effect sizes were also superior in comparison to the traditional measures. A larger effect size was also perceived on the digital tasks that evaluated visual episodic memory and visuospatial short-term memory contrasting with their paper-and-pencil counterparts, although the mean difference was not significant. The traditional measures of manual dexterity and rapid naming presented greater effects than their digital correspondents. The same situation occurred with the paper-and-pencil concentrated attention test; however, for this test, the differences between the groups were not significant.

Table 4.

Comparison of means and size of effect of comparisons between clinical and control groups

	Clinical Group M (SD)	Non-Clinical Group M (SD)	Statistical Testing (df)	p	d
TENI - Alternative Universes					
RS	5.63 (1.95)	6.23 (1.77)	$U_{(97)} = 1023.5$	.159	-0.287
Z score	-0.151 (1.04)	0.167 (0.94)	$U_{(97)} = 1023.5$	.157	-0.287
TENI- Toddy (Score)					
RS	74.90 (8.99)	78.72 (9.32)	$U_{(95} = 827$	**.012**	**-0.528**
Z score	-0.199 (0.97)	0.212 (1.00)	$U_{(95)} = 827$	**.012**	**-0.529**
TENI - Toddy (Answer Time)					
RS	541.90 (141.98)	471.96 (64.65)	$U_{(95)} = 821.5$	**.010**	**0.537**
Z score	0.291 (1.22)	-0.310 (0.55)	$U_{(95)} = 821.5$	**.011**	**0.536**
TENI - Toddy (Default Errors)					
RS	5.16 (5.64)	3.02 (3.26)	$U_{(95)} = 885$	**.035**	**0.439**
Z score	0.218 (1.19)	-0.232 (0.69)	$U_{(95)} = 885$	**.035**	**0.439**
TENI - Toddy (Comission Errors)					
RS	64.60 (14.91)	56.57 (16.80)	$U_{(95)} = 834.5$	**.013**	**0.516**
Z score	0.239 (0.92)	-0.254 (1.03)	$U_{(95)} = 834.5$	**.014**	**0.516**
TENI - Toddy - Adjusted Mean					
Z score	-0.237 (0.81)	0.252 (0.66)	$U_{(95)} = 773$	**.004**	**-0.617**
TENI - Bzz! - Inhibition (time)					
RS	194.15 (131.52)	257.89 (98.11)	$U_{(97)} = 923.5$	**.011**	**-0.523**
Z score	-0.251 (1.09)	0.278 (0.81)	$U_{(97)} = 923.5$	**.012**	**-0.523**
TENI - Bzz! – Viso-motor Ability					
RS	72.63 (20.20)	73.13 (16.95)	$t_{(97)} = -0.131$	.896	-0.027
TENI – Mexican House (copy)					
RS	21.79 (4.27)	24.04 (5.24)	$U_{(93)} = 814$	**.019**	**-0.496**
Z score	-0.233 (0.88)	0.238 (1.07)	$U_{(93)} = 814$	**.019**	**-0.496**
TENI - Mexican House (recall)					
RS	18.86 (6.01)	20.75 (5.02)	$t_{(88)} = -1.593$	.115	-0.343
Z score	-0.149 (1.06)	0.186 (0.89)	$t_{(88)} = -1.593$	.115	- 0.334
TENI - Tic-Tac - Time					
RS	20389.23 (9953.79)	16580.89 (4450.52)	$U_{(97)} = 949$	.056	0.392
Z score	0.225 (1.24)	-0.249 (0.55)	$U_{(97)} = 949$	.056	0.392
TENI – The Mole					
RS	5.42 (1.75)	6.11 (1.79)	$U_{(97)} = 973$	.075	-0.364
Z score	-0.180 (0.97)	0.199 (0.99)	$U_{(97)} = 973$	.075	-0.364

TENI - The Farm					
RS	13.65 (8.07)	18.02 (7.56)	**$U_{(97)} = 826$**	**.005**	**-0.582**
Z score	-0.256 (0.99)	0.283 (0.934)	**$U_{(97)} = 826$**	**.005**	**-0.582**
BPA - Concentrated Attention Points					
RS	46.8 (20.0)	53.50 (25.2)	$U_{(95)} = 947.5$	.101	-0.338
Z score	-0.142 (0.88)	0.151 (1.10)	$U_{(95)} = 947.5$	.100	-0.338
BPA - General Attention Points					
RS	116.14 (62.9)	138.13 (62.2)	$U_{(94)} = 946.5$	.137	-0.308
Z score	-0.167 (0.99)	0.181 (0.98)	$U_{(94)} = 946.5$	.135	-0.308
Rapid Naming Task					
RS	43.11 (16.6)	36.45 (6.7)	**$U_{(97)} = 916.5$**	**.032**	**0.441**
Z score	0.238 (1.25)	-0.263 (0.50)	**$U_{(97)} = 916.5$**	**.032**	**0.441**
Nine Hole Peg Test					
RS	25.13 (4.5)	23.37 (3)	**$U_{(92)} = 929$**	**.040**	**0.422**
Z score	0.211 (1.14)	-0.233 (0.76)	**$U_{(92)} = 929$**	**.040**	**0.422**
FDT - Inhibition					
RS	53.20 (22.45)	53.25 (30.7)	$U_{(94)} = 1121$	.825	-0.046
Z score	0.000 (0.84)	0.001 (1.15)	$U_{(94)} = 1121$	.823	-0.046
Rey Complex Figure (copy)					
RS	13.61 (6.10)	15.19 (6.42)	$t_{(93)} = -1.228$	.222	-0.252
Z score	-0.119 (0.97)	0.132 (1.02)	$t_{(93)} = -1.228$	.222	-0.255
Rey Complex Figure (recall)					
RS	6.49 (4.7)	6.98 (4.2)	$U_{(92)} = 979.5$	.363	-0.189
Z score	-0,052 (1.05)	0.059 (0.94)	$U_{(92)} = 979.5$	.360	-0.189
Corsi Cubes – Forward					
RS	6.04 (1.6)	6.39 (1.9)	$U_{(96)} = 1144$	.709	-0.076
Z score	-0.092 (0.92)	0.104 (1.09)	$U_{(96)} = 1144$	.706	-0.076
Matrix Reasoning (WASI)					
RS	15.71 (6.4)	16.87 (6.2)	$U_{(97)} = 1089.5$	.355	-0.188
Z score	-0.087 (1.01)	0.097 (0.98)	$U_{(97)} = 1089.5$	.352	-0.188

*Note*s: RS = Raw Score; Z score = Scoring converted to z score controlling age; *M* = Mean; *SD* = Standard Deviation; *Df* = Degree of freedom; *p* = Statistical significance; *d* = Effect size; *U* = Results of the non-parametric Mann-Whitney test; *t* = Test result; *T* = independent samples. Negative *d* value indicates higher result for the nonclinical group, since the means of the clinical group were included in the first formula.

Within groups comparison of digital and traditional tests

In order to identify whether the use of digital tests generates results significantly different from those found in traditional tests for children who present symptoms of inattention and hyperactivity, the averages obtained in both forms of application were compared with the use of the Wilcoxon test (nonparametric equivalent to the dependent T-test), which is indicated for comparison in repeated measurements. Both the standardized scores (Z score) of the digital and the traditional tests were used, according to the evaluated constructs, and they resulted on positive correlations.
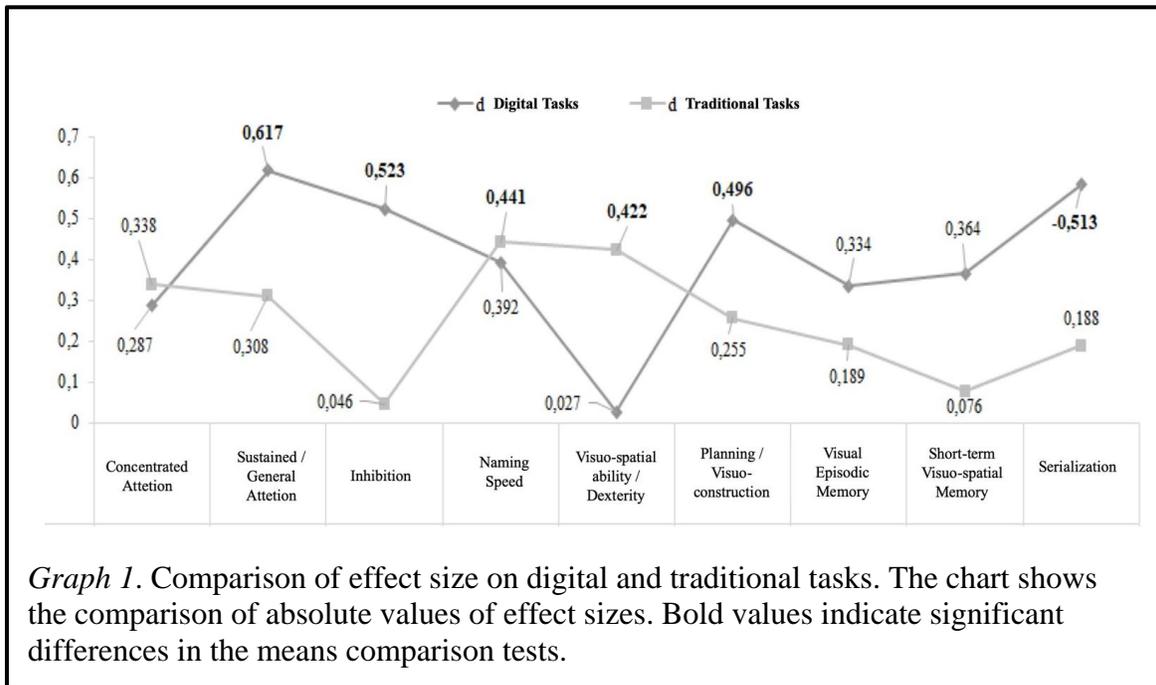
Table 5 shows that there were no significant differences in performance when comparing the results within the groups according to their computer-based and paper-and-pencil measures scores, which also resulted on small effect sizes. Yet, children in the clinical group showed better performance in traditional tasks, while the nonclinical group showed better results in digital tasks (except on the rapid naming task which has an opposite result pattern). Overall, the effects are small for both groups, being smaller on the clinic one.

Table 5.
Intragroup comparison of means in traditional and digital tests

	Non-clinical Group (*n* = 47)			Clinical Group (*n* = 52)		
	M (*SD*)	*Z*	*d*	*M* (*SD*)	*Z*	*d*
Z - Alternative Universes	0.167 (0.94)	- 0.963 *ns*	0.284	-0.151 (1.04)	-0.014 *ns*	-0.004
Z - BPA – Conc. Att. - Points	0.151 (1.10)			-0.142 (0.88)		
Z – Toddy – Adjusted Mean	0.252 (0.66)	- 0.803 *ns*	0.238	-0.237 (0.86)	-0.728 *ns*	-0.211
Z – BPA – General Attention	0.181 (0.98)			-0.167 (1.00)		
Z – Mexican House – Copy	0.238 (1.08)	- 0.397 *ns*	0.120	-0.233 (0.88)	-0.360 *ns*	-0.105
Z – Rey – Copy	0.132 (1.02)			-0.119 (0.97)		
Z – Mexican House - Recall	0.186 (0.89)	- 0.809 *ns*	0.261	-0.149 (1.07)	-0.390 *ns*	-0,113
Z – Rey – Recall	0.060 (0.94)			-0.052 (1.06)		
Z – Tic-Tac – Naming Speed	-0.249 (0.55)	- 0.011 *ns*	0.003	0.225 (1.24)	-0.055 *ns*	-0.015
Z – Rapid Naming Task	-0.263 (0.51)			0.238 (1.25)		
Z – The Mole – STM	0.200 (1.00)	- 0.486 *ns*	0.144	-0.180 (0.98)	-0.847 *ns*	-0.237
Z – Corsi Block-tapping Forward	0.104 (1.09)			-0.092 (0.92)		
Z – The Farm - Serialization	0.283 (0.93)	- 1.291 *ns*	0.383	-0.256 (1.00)	-1,175 *ns*	-0.330
Z – Matrix Reasoning	0.097 (0.98)			-0.088 (1.01)		

Notes: Results related to Wilcoxon signal testing appropriate for dependent samples. The similar digital and traditional standardized tests were grouped by transformation into Z score. *Ns* = Non-significant value with *p* > .05. Negative *d* value indicates highest result in the traditional test, since the digital tests were included in the first formula. STM = short-term memory. The digital tests entered as time 1 and the traditional ones as time 2.

Graph 1. Comparison of effect size on digital and traditional tasks. The chart shows the comparison of absolute values of effect sizes. Bold values indicate significant differences in the means comparison tests.

Discussion

The use of digital resources to aid on the process of cognitive functions evaluation is already a reality, mainly abroad, showcasing several studies that use computerized tests as tools in their investigations. Still, it is important to verify the possible impacts that digital measures can cause to the evaluation, in addition to pointing out that such impacts may differ on distinct samples, for instance, children with ADHD that may be impacted differently from those with typical development (Lumsden et al., 2016).

Thus, the present study aimed to verify the impact of computer-based tests on the evaluation of cognitive functions in children with ADHD symptoms. In order to do that, the performance in digital and traditional tasks were compared in a sample of 52 children with clinical ADHD symptoms and 47 children with fewer symptoms. As socioeconomic factors influence children's neurocognitive development, the groups were selected so as not to have significant differences regarding the type of school (public and private), family income or maternal education level. The groups also did not present significant differences in age, school year, sex and IQ (Ursache & Noble, 2015).

The study found that digital and traditional tests that evaluate the same functions generate similar results. There were significant correlations in measures that assess the same construct, such as short-term visuospatial memory, planning ability, naming speed, visual episodic memory and reasoning. The correlations ranged from low to high, indicating comparability between digital and traditional tests. Similar results are described in the systematic review by Lumsden et al. (2016), who reported that most of the cognitive tests using technology and gaming features presented correlations between computer-based and traditional tests with values ranging from $r = .45$ to $r = .60$.

In the current study, besides having significant correlations between tests that evaluated the same constructs, the correlations were higher for tasks with greater similarity in the stimuli and forms of application. Both the traditional and digital versions of the rapid naming speed, for example, required the participant to name the figures that appeared on the screen/paper as fast as possible; on account of that similarity, the correlations reached $r = .717$ ($p < .01$). Yet, the tests involving manual dexterity, in which the stimuli and application were distinct, and that scored differently (one per number of hits and the other per run time), presented low correlations such as $r = -.242$ ($p < .05$).

Whilst searching for similar computer-based and paper-and-pencil tests, there was an attempt to use comparable tools. Nevertheless, there were not many traditional instruments validated for Brazilian children within the selected age group. For instance, the digital sustained attention task (Toddy) was compared to the BPA's overall scores for a lack of better options. Although the evaluated construct is not exactly the same, in both tests, the participant should remain attentive for an extended period of time. While the digital one lasts seven uninterrupted minutes, the traditional battery consists of three tasks in a row, lasting around eight and a half minutes with brief interruptions between them. In general, the two measures presented significant and moderate correlations.

Another comparison between tests assessing distinct facets of related constructs was performed between the inhibition tasks (Bzz! and FDT Inhibition). According to Diamond (2013), inhibitory control is a function that involves the ability to do what is necessary, by controlling the attention, thoughts, emotions and behaviors to override a strong internal predisposition. One aspect of this ability is self-control, which involves restraining emotions and impulsiveness. The other, in turn, is the inhibitory control of attention, related to the interference regulation in perception, which allows to selectively respond to some stimuli and suppress others. This function works under the working memory, usually relating more to it than to other inhibitory aspects. As a result, it was not surprising to see the not significant correlations between the Bzz! and FDT tasks, since one was closely associated to the attentional aspects of inhibition, whilst the other evaluated the affective facet.

The present study also verified the performance differences between the clinical and non-clinical groups in digital and traditional tests that evaluated cognitive functions; and, in order to analyze them, we compared the means on both formats. Children with ADHD symptoms had worse mean performance on all measures, with significant differences mainly on the digital ones. Significant differences were also observed in the digital tests of sustained attention, inhibition, seriation and visuomotor/planning ability, with effect sizes between $d = 0.43$ and $d = 0.61$, effects greater than their traditional pairs. On the paper-and-pencil tools, there were significant differences on the tasks of manual dexterity and rapid naming, with effects sizes of $d = 0.42$ for the former and $d = 0.44$ for the latter.

Searching for significant differences between typically developed and ADHD children's performance is common in the literature, and the results are often variable. In a meta-analysis, Willcutt et al. (2005) considered 83 articles that investigated executive functions in patients with ADHD and showed that in only 65 % of the studies significant differences were reported. In those studies, the clinical group presented significantly worse performance than the control, with effect sizes between d = 0.43 and d = 0.69 on different functions. Among the analyzed tasks, there were CPTs, Rey's Complex Figure (Copy) and working memory tests. The authors report that studies using standardized cut-off points for group definition, rather than clinical diagnosis, also found significant differences in cognitive performance, but with smaller effect sizes (d = 0.41 ± 0.16).

The aforementioned meta-analysis was used by Fried et al. (2012) to compare the effects sizes obtained by the difference in the performance of groups with ADHD and control on CANTAB and paper-and-pencil tests. These authors mention that the use of CANTAB generated results similar to those reported by Willcutt, but with slightly lower effects, variating between d = 0.39 and d = 0.63. The present study had results compatible with those presented on both articles, for instance, the Mexican House (Copy) had an effect size of d = 0.49, while the CANTAB planning task had a d = 0.41, and the Complex Rey Figure Copy had a d = 0.43. TENI's sustained attention task had a similar effect size to other studies that involved CPTs. The meta-analytic data presented effect sizes of d = 0.51 and d = 0.64 for the errors by default and commission, respectively, while TENI's were d = 0.44 and d = 0.51. Toddy's task response time is based on continuous performance models for sustained attention, and its effect size (d = 0.54) was similar to those of CANTAB (d = 0.39), ClinicaVR: Classroom-CPT-VC (d = 0.59) and the meta-analytic data (d = 0.61) (Fried et al., 2012).

In this study, there were significant differences between groups when using the traditional rapid naming task (d = 0.44). Similarly, the study by Bidwell, Willcutt, DeFries & Pennington (2007), which had 266 children with ADHD and 332 with typical development, reported that the clinical group presented significantly worse performance on the RAN, although the effect size was d = 0.71. On the traditional dexterity test, the clinical group had worse performance (d = 0.42), which is consistent with national studies, including the one conducted by Oliveira, Cavalcante and Palmares (2018). Their research reported that 43.38 % of children with ADHD had major dexterity losses, as well as aiming and balance difficulties, which were absent on the control group.

Therefore, even though significant differences were observed on the group's performance in digital tasks regarding executive functions (planning, reasoning and inhibition), sustained attention, naming speed and visuomotor ability, these differences were not statistically evidenced throughout all the paper-and-pencil tests that composed the evaluation battery. However, Willcutt et al. (2005) show that many studies do not report statistically significant differences in ADHD and comparison groups. Moreover, Raiford, Drozdick & Zhang (2015) state that although the effect size may vary between studies that rely on different samples of participants with ADHD, the expected outcome pattern remains the same. This pattern, in which groups with symptoms of inattention and hyperactivity show worse performances in certain cognitive functions, was perceived in both digital and traditional tests, although the differences were not always significant.

According to Raiford et al. (2015) the impact of cognitive tests computerized administration for individuals with specific clinical conditions is not broadly investigated. Therefore, this study also aimed to verify whether the digital format of cognitive tests generated differential impact when comparing children of the clinical with nonclinical groups. Lumsden et al. (2016) state that many digital tests use the gaming format to evaluate and/or train cognition. Due to its attractiveness, some are target at specific clinical groups such as ADHD. Still, they also ponder the fact that by providing a structured, feedback-rich environment, it is possible that the symptoms of inattention and hyperactivity are minimized and scarcely detected. In this case, inattentive and hyperactive people would be expected to perform better on gaming-format tasks. Thus, establishing the need to further investigate whether or not tests turned into games could invalidate the evaluation, taking into consideration that if the tools mask the deficit that they intend to measure, they could generate a differential influence on groups such as those with ADHD in comparison to typical developed children.

To further understand what kind of influence the alteration of the testing environment could have on the groups' performance, the standardized scores of both types of evaluation were compared using paired sample analysis. There were no statistically significant differences between performance in digital and traditional tests, neither on the group composed of children with symptoms of inattention and hyperactivity, nor for the comparison group. It should be noted that all effect sizes were small (d between -0.004 and 0.383).

Although the differences found were not significant, there was a differential pattern between groups. In general, the clinical group presented better performance on the traditional tasks, while the nonclinical group had better averages on the digital counterparts. This pattern is consistent with the fact that only computer-based tasks generated significant differences between groups, the exception being the RAN. Hence, contrary to what was previously reported by Lumsden et al. (2016), despite the use of a digital game-like task, children with ADHD symptoms did not have their performance benefited by the platform. However, the differential pattern may be due to the lack of isomorphism between the tasks, as well as distinction of the evaluated constructs.

The differential pattern, however, may be due to the lack of isomorphism between the tasks, which may even be evaluating disparate constructs. As an example, on the inhibition tasks the constructs are different, because the FDT evaluates the attentional inhibitory control while Bzz! measures self-control. Another case was on Todd, that assessed sustained attention, a known compromised ability on ADHD individuals, while the BPA evaluated alternating and divided attention which are not as affected on this group (Frazier, Demaree & Youngstrom, 2004; Nigg, 2005).

In summary, the present study shows evidences that the digital format did not cause losses in the evaluation of children with ADHD symptoms. Both the computer-based and the traditional tests generated similar results, especially on alike tasks. As expected, children with symptoms of inattention and hyperactivity at the clinical level demonstrate significantly worse performance on digital tasks that assess constructs typically reported as deficient in ADHD. Intragroup assessments have shown that children with typical development usually perform better on digital tests, while those with ADHD symptoms demonstrate an opposite pattern. However, it should be noted that such differences do not reach the level of statistical significance. Also, the minor differences may not be due to the test format but to the

distinction between evaluated constructs. Therefore, it may be possible that the test format does not significantly impact distinct results on the studied groups.

Final considerations

This investigation contributes to the field indicating that digital tools may be useful for the evaluation of children with symptoms of ADHD. Additionally, the study also presents evidence that the differential impact was not significant between these children and the group without a clinical presentation of symptoms. Nevertheless, it is worth noting that this research had its limitations, some of which are: the use of a non-probabilistic sample, selected for convenience; not using a sample with clinical ADHD diagnostic use of only one source of information about the children's symptoms (only the parents' opinion was requested); no comorbidities were evaluated and controlled; difficulty in selecting traditional tests validated for the age group that evaluated the same constructs.

Further studies are required in order to address such limitations, with more extensive and diversified samples; along with a confirmation of the clinical diagnosis provided by specialized professionals, detailing the deficit's subtype. Another suggestion would be the comparison between tasks that evaluate other functions typically reported as deficit in ADHD, for example, working. Furthermore, comparisons could be made between TENI and different digital tests such as the Visual Attention Test (TAVIS) to verify whether there is a differential impact on clinical groups and controls in distinct measures (Duchesne & Mattos, 1997).

References

American Psychiatric Association [ADA]. (2013). *Diagnostic and statistical manual of mental disorders* (5th Ed.). Arlington, VA.

Barkley, R.A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94. doi: 10.1037/0033-2909.121.1.65

Bauer, R., Iverson, G., Cernich, A., Binder, L., Ruff, R., & Naugle, R. (2012) Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of Clinical Neuropsychology, 27*(3), 362-373. doi: 10.1093/arclin/acs027

Bidwell, L. C., Willcutt, E. G., DeFries, J. C., & Pennington, B. F. (2007). Testing for neuropsychological endophenotypes in siblings discordant for Attention Deficit and Hyperactivity Disorder. *Biological Psychiatry, 62*, 991-998. doi: 10.1016/j.biopsych.2007.04.003

Capellini, S.A., Smythe, I., & Silva, C. (2017). *Protocolo de avaliação de habilidades cognitivo-linguísticas.* Livro do profissional e do professor. Marília: Fundepe.

Corsi, P.M. (1972). *Human memory and the medial temporal region of the brain.* Montreal: McGill University.

Costa, D., de Paula, J., Malloy-Diniz, L., Romano-Silva, M., & Miranda, D. (2018). Parent SNAP-IV rating of attention-deficit/hyperactivity disorder: accuracy in a clinical

sample of ADHD, validity, and reliability in a Brazilian sample. *Jornal de Pediatria, 95*(6), 736-743. doi: 10.1016/j.jped.2018.06.014

Csapó, B., Molnár, G., & Nagy, J. (2014). Computer-based assessment of school readiness and early reasoning. *Journal of Educational Psychology, 106*(3), 639-650. doi: psycnet.apa.org/doi/10.1037/a0035756

Delgado, M. T., Uribe, P. A., Aparicio, A., Benavente, C., Thibaut, C., & Díaz, R. (2012). *Test de Evaluación Neuropsicológica Infantil - Manual Técnico-Interpretativo*. Santiago, Chile: Cedeti.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology, 6*, 135-168. doi: 10.1146/annurev-psych-113011-143750

Dovis, S., Van der Oord, S., Wiers, R., & Prins, P. (2011). Can motivation normalize working memory and task persistence in children with Attention-Deficit/Hyperactivity Disorder? The effects of money and computer-gaming. *Journal of Abnormal Child Psychology, 40*(5), 669-681. doi: 10.1007/s10802-011-9601-8.

Duchesne, M. & Mattos, P. (1997). Normatização de um teste computadorizado de atenção visual (TAVIS). *Arquivos de Neuro-Psiquiatria, 55*(1), 62-69. doi: 10.1590/S0004-282X1997000100010

Frazier, T., Demaree, H., & Youngstrom, E. (2004). Meta-Analysis of Intellectual and Neuropsychological Test Performance in Attention-Deficit/Hyperactivity Disorder. *Neuropsychology, 18*(3), 543-555. doi: 10.1037/0894-4105.18.3.543

Fried, R., Hirshfeld-Becker, D., Petty, C., Batchelder, H., & Biederman, J. (2012). How informative is the CANTAB to assess executive functioning in children with ADHD? A controlled study. *Journal of Attention Disorders, 19*(6), 468-475. doi: 10.1177%2F1087054712457038

Kane, R. & Parsons, T. (2017). Introduction to neuropsychology and technology. In: R. Kane & T. Parsons. *The role of technology in clinical neuropsychology* (1st ed.) (pp. 3-23). New York: Oxford University Press. doi: 10.1093/oso/978019023 4737.001.0001

Kellor, M., Frost, J., Silberberg, N., Iversen, I., & Cummings, R. (1971). Hand strength and dexterity. *The American Journal of Occupational Therapy, 25*(2), 77-83.

Lumsden, J., Edwards, E., Lawrence, N., Coyle, D., & Munafò, M. (2016). Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. *JMIR Serious Games, 4*(2), e11. doi: 10.2196/games. 5888

Mahone, E.M. (2012). Neuropsychiatric differences between boys and girls with ADHD. *Psychiatric Times, 29*(10), 34-43.

Martins, P. S. R., Barbosa-Pereira, D., Valgas-Costa, M. & Mansur-Alves, M. (2020). Item analysis of the Child Neuropsychological Assessment Test (TENI): Classical test theory and item response theory. *Applied Neuropsychology: Child*. doi: 10.1080/21622965.2020.1846128

Mattos, P., Serra-Pinheiro, M. A., Rohde, L. A., & Pinto, D. (2006). Apresentação de uma versão em português para uso no Brasil do instrumento MTA-SNAP-IV de avaliação de sintomas de transtorno do déficit de atenção/hiperatividade e sintomas de transtorno desafiador e de oposição. *Rev Psiquiatr Rio Gd Sul, 28*(3), 290-297. doi: 10.1590/S0101-81082006000300008

Messina, L.F. & Tiedemann, K.B. (2009). Avaliação da memória de trabalho em crianças com transtorno do déficit de atenção e hiperatividade. *Psicologia USP, 20*(2), 209-228. doi: 10.1590/S0103-65642009000200005

Moncaleano, S. & Russell, M. (2018). A historical analysis of technological advances to educational testing: a drive for efficiency and the interplay with validity. *Journal of Applied Testing Technology, 9*(1).

Nigg, J. (2005). Neuropsychologic Theory and Findings in Attention-Deficit/Hyperactivity Disorder: The State of the Field and Salient Challenges for the Coming Decade. *Biological Psychiatry, 57*(11), 1424-1435. doi: 10.1016/ j.biopsych.2004.11.011

Oliveira, C., Cavalcante, J., & Palhares, M. (2018). Características motoras de escolares com Transtorno de Déficit de Atenção com Hiperatividade. *Cadernos Brasileiros De Terapia Ocupacional, 26*(3), 590-600. doi: 10.4322/2526-8910.ctoAO1139

Oliveira, M.S. & Rigoni, M.S. (2010). *Figuras Complexas de Rey: Teste de Cópia e de Reprodução de Memória de Figuras Geométricas Complexas*. São Paulo: Casa do Psicólogo.

Parsey, C. & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *The Clinical Neuropsychologist, 27*(8), 1328-1361. doi: 10.1080/13854046.2013.834971

Raiford, S. E., Drozdick, L.W., & Zhang, O. (2015). *Q-interactive Special Group Studies: The WISC –V and Children with Autism Spectrum Disorder and Accompanying Language Impairment or Attention-Deficit/Hyperactivity Disorder. Q-interactive Technical Report 11*. Bloomington, MN: Pearson.

Rueda, F. J. (2013) *Bateria Psicológica para Avaliação da Atenção - BPA*. São Paulo: Vetor.

Rueda, F.J. & Muniz, M. (2012). Evidência de validade convergente da Bateria Psicológica para Avaliação da Atenção - BPA. *Estudos Interdisciplinares em Psicologia, 3*(2), 62-181.

Sedó, M., de Paula, J. J., & Malloy-Diniz, L. F. (2015). *O Teste dos Cinco Dígitos*. São Paulo: Hogrefe.

Sergeant, J. A. (2005). Modeling Attention-Deficit/Hyperactivity Disorder: a critical appraisal of the Cognitive-Energetic model. *Biological Psychiatry, 57*(11), 1248- 1255. doi: 10.1016/j.biopsych.2004.09.010

Sonuga-Barke, E. & Halperin, J. (2010). Developmental phenotypes and causal pathways in Attention Deficit/Hyperactivity Disorder: potential targets for early intervention? *Journal of Child Psychology and Psychiatry, 51*(4), 368-389. doi: 10.1111/j.1469-7610.2009.02195.x

Trentini, C.M., Yates, D.B., & Heck, V.S. (2014). *Escala Wechsler Abreviada de Inteligência: WASI – Manual*. São Paulo: Casa do Psicólogo.

Ursache, A. & Noble, K. (2015). Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology, 53*(1), 71-82. doi: 10.111/psyp.12547

Wang, L., Huang, Y., Chiang, Y., Hsiao, C., Shang, Z., & Chen, C. (2011). Clinical symptoms and performance on the continuous performance test in children with attention deficit hyperactivity disorder between subtypes: a natural follow-up study for 6 months. *BMC Psychiatry, 11*(65). doi: 10.1186/1471-244X-11-65

Willcutt, E. (2012). The prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: a meta-analytic review. *Neurotherapeutics, 9*(3), 490-499. doi: 10.1007/ s13311-012-0135-8

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol Psychiatry, 57*(11), 1336-1346. doi: 10.1016/j.biopsych.2005.02.006

Zygouris, S. & Tsolaki, M. (2015). Computerized cognitive testing for older adults: a review. *American Journal of Alzheimer's Disease & Other Dementias, 30*(1), 13-28. doi: 10.1177/1533317514522852

Authors' participation: a) Conception and design of the work; b) Data acquisition; c) Analysis and interpretation of data; d) Writing of the manuscript; e) Critical review of the manuscript.
M. B. F. has contributed in a, b, c, d; E. R. S. in b, d; and M. M. A. in a, c, e.

Scientific editor in charge: Dra. Cecilia Cracco.