

EVIDÊNCIAS DE VALIDADE DE UMA VERSÃO APLICATIVO MOBILE DO TESTE TOKEN: Comparação com o teste token computadorizado

Vinicius Marangoni Noro Veiga (IC) e Elizeu Coutinho de Macedo (Orientador)

Apoio: PIBIC CNPq

RESUMO

O interesse e a demanda por estudos envolvendo a incorporação de novas tecnologias na área da saúde vem crescendo na comunidade científica. A tecnologia oferece avaliação e intervenção automática e direcionada, com possibilidade de ser feita a distância, aumentando o acesso a diagnósticos e tratamento. O segmento que produz o maior número de publicações é o de aplicativos mobile para smartphones e tablets voltados para avaliação, intervenção e acompanhamento de tratamento. Uma das áreas que pode usufruir desse recurso é a avaliação neuropsicológica, através de aplicativos formalizados para avaliação ou intervenção. No exterior já existem publicações sobre aplicativos usados para avaliação e treinamento de habilidades de linguagem receptiva. No Brasil, entretanto, não há nenhum instrumento formal na área desenvolvido para a plataforma de dispositivos portáteis. Foi desenvolvido uma versão do teste token em aplicativo para plataforma Android, com subsequente avaliação de evidências de validade desse, através da comparação com os resultados de outra versão computadorizada do teste. Foram avaliadas 51 crianças, entre 6 e 10 anos de idade, com as duas versões. As análises estatísticas mostraram coeficientes de fidedignidade satisfatórios, além de correlações moderadas com significância entre os dois testes. As análises também mostram que os participantes foram melhor na versão aplicativo do que na versão computadorizada, e que levaram um tempo menor para concluir a tarefa no aplicativo que no computador. Não foram encontradas diferenças significativas entre o desempenho de meninos e meninas nas versões. As análises de validade se mostraram promissoras, mas exigem estudos subsequentes sobre o aplicativo.

Palavras-chave: Linguagem receptiva. Aplicativo Mobile. Teste token.

ABSTRACT

The interest and the requirement of implementation of new technologies on health has grown among scientific Community. Technology offer automated and self directed screening and intervention, widening the access to diagnosis and treatment throw distance. The sector of research with major increase in scientific publications is about the use of mobile applications for smartphones and tablets for screening, intervention and treatment follow up. Neuropsychological assessment is a discipline that can benefit from mobile applications by incorporating then to formal assessment and intervention instruments. In international jornal there are already publications about applications used to screening and training receptive

language skills. In Brazil, meanwhile, there is no formal mobile application designed to access that subject. The current study developed a mobile Android application version of the token test, and evaluated its validity evidences by comparing it with another computerized version of the test. 51 children between 6 and 10 years old were evaluated with mobile version of token test for tablets and computerized version of token test for Windows. Statistical analysis has shown good reliability coefficients for both tests and moderate correlations between them within the significance interval. Statistical analysis have shown also that participants scored more on the tablet application in comparison to the computerized one, and spented significantly less time on tablet than on computer to finish the test. There were no significant difference between boys and girls performances. The validity analysis has shown promising results, but it is still needed more studies to validate the test.

Keywords: Receptive language. Mobile application. Token test.

1. INTRODUÇÃO

O interesse pela utilização de dispositivos tecnológicos para a avaliação e intervenção em saúde mental tem crescido na literatura científica. Parte desse interesse se justifica pela tecnologia oferecer avaliação e intervenção automática e direcionada, com possibilidade de ser feita a distância, aumentando o acesso a diagnósticos e tratamento (HALL et al, 2019). Entre as categorias de abordagens tecnológicas utilizadas para esse fim, podemos citar a avaliação computadorizada, terapia virtual, jogos interventivos, intervenção baseada em realidade virtual, aplicativos, robôs, modelagem de vídeo e dispositivos e sensores de monitoramento em tempo real, sendo os aplicativos mobile para smartphones e tablets a categoria com o maior número de publicações (VALENTINE et al, 2020). Devido à crise sanitária de escala global ocorrida no início do ano de 2020, e a subsequente restrição da circulação de pessoas, a demanda por plataformas que possibilitem a avaliação e tratamento de maneira remota e com qualidade aumentou vertiginosamente (HOLMES et al, 2020). Nesse contexto são importantes novos estudos que demonstrem alternativas para esse fim.

As aplicações de base Android tem se popularizado pela sua acessibilidade e pelos poucos obstáculos para a disponibilização das aplicações na loja de aplicativos. Uma das atividades em saúde que pode se adaptar a essa plataforma é a avaliação neuropsicológica. A neuropsicologia é uma disciplina que estuda as relações entre cognição e comportamento com o sistema nervoso central, em manifestações normais ou patológicas, e para esse fim, faz uso de instrumentos de avaliação estruturados e validados, com boas propriedades psicométricas, especificamente de precisão e validade (MACEDO et al, 2007).

Uma das áreas de estudo da neuropsicologia que apresenta crescente interesse no uso de instrumentos digitais baseados em aplicativos é a área de pesquisa, avaliação e intervenção em linguagem. Podemos citar alguns estudos que desenvolveram e se utilizaram de aplicativos para realizar avaliação de linguagem receptiva em crianças (SCHAEFER et al, 2018), e outro para avaliação e treinamento de habilidades relacionadas a linguagem receptiva em crianças (FIRDOUS et al, 2019).

De acordo com a perspectiva da psicologia cognitiva, a linguagem oral é o uso de um meio organizado de combinar as palavras para fins de comunicação e possibilita a comunicação entre os indivíduos (STERNBERG, 2016). A linguagem oral, portanto, pode ser dividida entre receptiva e expressiva, sendo a linguagem receptiva referente a capacidade decodificação e compreensão da informação, e a linguagem expressiva relacionada a capacidade de produção e transmissão de informação através da linguagem. A avaliação de linguagem receptiva constitui um critério relevante no cenário da avaliação neuropsicológica

de crianças, tendo em vista que a compreensão da linguagem, falada ou escrita, é fundamental para o desenvolvimento de linguagem expressiva (CAPOVILLA et al, 1997).

Alterações ou atraso na linguagem oral podem estar relacionados a diferentes distúrbios de linguagem, como afasia, disfasia, dislexia ou outros transtornos específicos de linguagem, além de ter alta taxa de comorbidade com outros transtornos do neurodesenvolvimento, como Transtorno do Espectro Autista e Transtorno de Déficit de Atenção e Hiperatividade (ARANGO-LASPRILLA et al., 2017; VALENTINE et al, 2020).

Entre os testes neuropsicológicos para avaliação de linguagem receptiva disponíveis no Brasil, validados e normatizados, podemos citar o Teste de Vocabulário por Imagens Peabody (CAPOVILLA et al, 1997) e o Teste Token (MACEDO et al, 1998). Ambos os testes também tem versões computadorizadas validadas, para ambiente Windows (MACEDO et al, 2006; MACEDO et al, 2007).

Os testes computadorizados apresentam certas vantagens com relação às suas versões tradicionais, como medição precisa do tempo de resposta, apresentação precisa de estímulos e administração e pontuação automática do teste. Além disso, a aplicação automatizada economiza tempo e permite o aplicador focar no comportamento do colaborador. Com a diminuição do custo de aparelhos digitais e a generalização de seu uso na população, é provável que esses aparelhos se tornem parte da realidade prática do contexto de avaliação (GERMINE, REINECKE; CHAYLOR, 2019).

Não existem ainda, na realidade brasileira, testes de avaliação de linguagem digitais disponíveis para a plataforma de dispositivos portáteis como tablets e smartphones. O objetivo do presente estudo foi desenvolver e buscar evidências de validade de uma versão aplicativo mobile do Teste Token, para tablets de sistema operacional Android, a partir de comparações com outra versão computadorizada já validada.

2. REFERENCIAL TEÓRICO

O teste Token foi desenvolvido em 1962 por De Renzi e Vignolo com a intenção de detectar déficits de linguagem em adultos com afasia. Entretanto ele também tem sido utilizado com crianças e adolescentes (MALLOY-DINIZ et al, 2007). No Brasil, foram conduzidos dois estudos de validade do Teste Token para população infantil, um para crianças de 7 a 10 anos (MALLOY-DINIZ et al., 2007), e outro para crianças de 4 a 6 anos (MACEDO et al., 2007). A versão original do teste consiste em 62 comandos com aumento gradual de dificuldade, os quais usam símbolos (tokens) nas formas geométricas de quadrados e círculos, nas cores azul, branco, vermelho, verde ou amarelo (DE RENZI; VIGNOLO, 1962).

O Teste Token é amplamente usado em países da América Latina e na Espanha, e é considerado um dos 20 testes neuropsicológicos mais usados nesses países, devido sua

simplicidade, fácil aplicação e acessibilidade (ARANGO-LASPRILLA et al, 2017). Foram desenvolvidas outras versões do teste, como a versão reduzida “Shortened Version of the Token Test”, com 36 comandos, desenvolvida por (DE RENZI; FAGLIONI, 1978), e versões computadorizadas, como a de MCNEIL et al., (2015) e a de MACEDO et al, (1998).

A versão computadorizada do teste token por MACEDO et al, (1998) apresenta 40 comandos e foi utilizada como base para o desenvolvimento do Teste Token App Mobile, para tablets. Para o desenvolvimento de uma nova versão digital de um teste, válida para aplicação, desenvolvida a partir de uma versão tradicional de papel e lápis ou computadorizada, essa precisa compreender evidências de equivalência com sua versão parâmetro. Segundo o *Guidelines on Computer-Based and Internet Delivered Testing* (INTERNATIONAL TEST COMMISSION, 2005), o desenvolvimento de um teste computadorizado a partir de uma versão já existente precisa atender o parâmetro de evidência de equivalência entre as versões: fidedignidade comparável; correlação entre as versões dentro do intervalo de confiança estimado; as versões devem se correlacionar de forma similar com outros testes e critérios independentes; e produzem médias e desvios padrão equiparáveis ou calibrados apropriadamente para gerar escores equivalentes.

É do interesse da avaliação neuropsicológica a transição dos testes informatizados em computador para a plataforma mobile, uma vez que estes aparelhos são mais versáteis, oferecem vantagens em portabilidade e manuseio, têm média de preço mais acessível e capacidade de processamento razoável ou idêntica em comparação aos notebooks e desktops para executar os requisitos dos aplicativos de avaliação. Também oferecem uma forma sistemática de aquisição (compra) e atualização, a partir de uma loja de aplicativos (Play Store e Apple Store), coisa que as versões computadorizadas não oferecem, o que contribui para sua acessibilidade.

OBJETIVO

O Objetivo Geral consistiu em buscar evidências de validade do Teste Token Aplicativo Mobile por correlação com a versão computadorizada do Teste Token para ambiente Windows para avaliar linguagem receptiva em crianças de 6 a 9 anos. Já os objetivos específicos foram: i) verificar a precisão da versão para tablet do Teste Token, e da versão computadorizada para Windows; ii) comparar o desempenho na versão tablet entre as faixas etárias; iii) comparar o desempenho das crianças em função da idade com a pontuação obtida nas duas versões computadorizadas do Teste Token; iv) conduzir análises de correlação entre as duas versões computadorizadas do teste token; v) comparar o desempenho entre meninos e meninas em função da idade.

3. METODOLOGIA

3.1 PARTICIPANTES

Caracterização dos sujeitos:

Foram avaliadas 51 crianças com desenvolvimento típico, de ambos os sexos, com idades de 6 a 10 anos, dos cinco anos escolares do Ensino Fundamental I de uma escola particular da zona oeste do município de São Paulo, selecionada por critério de conveniência. A amostra foi composta por 16 crianças de 6 anos (31,4%), 12 de 7 anos (23,5%), 9 de 8 anos (17,6%), 9 de 9 anos (17,6%), e 5 de 10 anos (9,8%).

A frequência das crianças por ano escolar e por sexo, estão dispostas, respectivamente, na tabela 1 e 2.

Tabela 1 e 2 – Frequência e porcentagem por ano escolar; Frequência e porcentagem por sexo

Ano Escolar			
	Frequência	Porcentagem	Porcentagem Cumulativa
1	16	31,4	31,4
2	13	25,5	56,9
3	10	19,6	76,5
4 e 5	12	23,5	100
Total	51	100	100,00

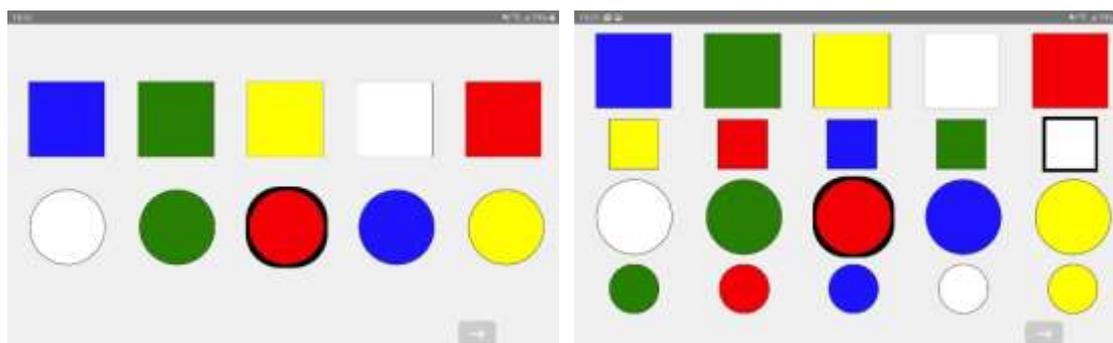
Sexo			
	Frequência	Porcentagem	Porcentagem Cumulativa
Feminino	22	43,1	43,1
Masculino	29	56,9	100
Total	51	100	

A participação de todos os sujeitos foi voluntária, mediante assinatura dos pais e do representante da instituição do Termo de Consentimento Livre e Esclarecido, e assinatura do Termo de Assentimento Livre e Esclarecido por parte do participante. Os critérios de inclusão da pesquisa foram a ausência de evidências de deficiência sensorial, intelectual ou motora, relatadas pela escola. Os critérios de exclusão foram participantes que não concluíram um dos dois testes ou que faltaram em um dos dias da coleta.

3.2. INSTRUMENTOS

3.2.1. Teste Token Apicativo Mobile

Imagens 1 e 2 – À esquerda, prancha correspondente às partes 1 e 3 do teste token App, com um estímulo selecionado, e à direita, prancha de estímulos correspondente às partes 2 e 4, com dois estímulos selecionados



É uma versão do Teste Token em aplicativo de base Android, desenvolvida a partir da versão computadorizada de Macedo et al. (1998). É composto de 40 itens, divididos em 4 partes com 10 itens cada. As partes 1 e 2 possuem um comando para cada item. Já nas partes 3 e 4 possuem 2 comandos por item, sendo que o participante deve responder o primeiro e em seguida o segundo. As quatro partes apresentam um grau crescente de dificuldade, entretanto dentro de cada parte os itens apresentam o mesmo nível de dificuldade. Na parte 1 os comandos apresentam duas informações (ex., “toque no Quadrado Verde”). Na parte 2, o comando apresenta três informações (ex., “toque no Círculo Amarelo Pequeno”). Na parte 3, o comando é composto por quatro informações (ex., “toque o Quadrado Vermelho e o Círculo Verde”). E na parte 4 são seis informações (ex., “toque o Círculo Azul Pequeno e o Quadrado Branco Grande”). Após ouvir o comando de um item, o participante seleciona as figuras que ele considera corretas. Para cada tarefa realizada corretamente, o participante recebe um ponto, e a pontuação máxima possível no teste é de 40 pontos.

Para o desenvolvimento do Teste Token Mobile foi utilizado o framework de desenvolvimento React Native, que possibilita criar código de uma só vez, tanto para o sistema Android quanto para o IOS, que são os dois grandes sistemas operacionais da maioria dos celulares no mundo todo. Isso tornou possível a criação de aplicativos para celular de forma mais rápida e fácil, sem a necessidade de ter que programar em Java para Android e em Objective-C para IOS. A versão final do aplicativo foi disponibilizada para dispositivos de base Android na Play Store sob o nome de “token teste”.

3.2.2. Teste Token Computadorizado

É uma versão computadorizada do teste token para ambiente Windows de Macedo et al. (1998), desenvolvida a partir da versão de Di Simoni (1978). É composto de 40 itens, divididos em 4 partes com 10 itens cada. Funciona de forma similar à versão de base Android, porém a forma de seleção dos estímulos é feita a partir de mouse ou touchpad.

3.3. PROCEDIMENTOS

A coleta de dados aconteceu na escola, em local com ausência de distração auditiva ou visual. Todos os participantes realizaram os testes utilizando fones de ouvido com proteção externa. Os participantes realizaram o teste token computadorizado primeiro, com seis meses de diferença para a aplicação do teste token aplicativo.

Para a análise de dados, inicialmente foi realizada análise de precisão dos escores totais do teste token aplicativo e do teste token computadorizado com o Coeficiente Alfa de Cronbach, e análise de precisão dos subscores com o Coeficiente Lambda 2 de Guttman. Em seguida foi feita a verificação da distribuição normal dos dados pelo Teste de Kolmogorov-

Smirnov. Devido a amostra não-normal, foi utilizado o Teste de Wilcoxon para comparação de desempenho e tempo entre as duas versões, Correlação de Spearman entre os escores totais e parciais entre as duas versões, Teste de Kruskal-Wallis para a medir o efeito dos anos escolares no desempenho, Teste de Man-Whitney para comparação de desempenho entre grupos etários e para comparação de desempenho entre os sexos.

4. RESULTADO E DISCUSSÃO

4.1 Dados gerais da amostra

Inicialmente, são apresentados os dados dos resultados dos 51 participantes nos dois instrumentos, disponíveis na tabela 3.

Tabela 3 – Estatísticas descritivas da amostra geral

	Mínimo	Máximo	Média	Desvio Padrão	Assimetria	
	Statistic	Statistic	Statistic	Statistic	Statistic	Erro Padrão
Comp Total	16	40	29,333	5,8126	-0,363	0,333
Comp Parte 1	4	10	9,137	1,3419	-2,069	0,333
Comp Parte 2	2	10	8,431	1,9001	-1,601	0,333
Comp Parte 3	1	10	7,843	1,9937	-0,988	0,333
Comp Parte 4	0	10	4	2,4658	0,475	0,333
App Total	18	40	32,51	4,718	-0,892	0,333
App Parte 1	7	10	9,647	0,6877	-2,094	0,333
App Parte 2	3	10	9,02	1,5164	-2,004	0,333
App Parte 3	3	10	8,824	1,4242	-2,054	0,333
App Parte 4	0	10	5,02	2,534	-0,061	0,333

Comp Parte 1 e App Parte 1 correspondem aos primeiros 10 itens do teste token computadorizado e do teste token aplicativo mobile, respectivamente. As partes 2 correspondem aos itens de número 11 ao 20, a parte 3, dos itens 21 ao 30 e a parte 4, dos itens 31 ao 40, sendo Comp Total e App Total a soma das quatro partes.

Observa-se que todos os 51 sujeitos responderam ao teste token computadorizado e ao teste token app mobile, com um intervalo de seis meses entre as duas aplicações. Entre as duas versões, podemos destacar que os participantes apresentaram maiores médias no teste token aplicativo do que na versão computadorizada, acertando, de uma maneira geral, uma média de 3 pontos a mais no escore total. As subpartes do teste (1, 2, 3 e 4), em sequência, apresentaram quantidade de acertos decrescente nas duas versões, o que é condicente ao aumento de complexidade da tarefa.

4.2 Análise de dados

Foi feita inicialmente uma análise dos resultados da versão computadorizada e da versão aplicativo através da análise da consistência interna dos escores totais com o coeficiente Alfa de Cronbach. Na análise do escore total da amostra geral na versão

computadorizada, o coeficiente foi de 0,84, já na versão aplicativo mobile, o coeficiente para o escore total da amostra geral foi de 0,81, o que indica boa precisão e consistência interna semelhante entre a consistência interna nos dois instrumentos.

Para a análise da fidedignidade das subpartes dos testes, ou seja, da parte 1, parte 2, parte 3 e parte 4 do Token Computadorizado e do Token App, nas quais o número total de itens é de 10 em cada, o coeficiente mais adequado foi o Lambda2 de Guttman. Os coeficientes de cada subparte dos dois testes podem ser visualizados na tabela 4:

Tabela 4- Lambda2 de Guttman por subpartes do token versão computadorizada e aplicativo

Comp parte 1	0,70
Comp parte 2	0,73
Comp parte 3	0,67
Comp parte 4	0,69
App parte 1	0,42 – itens sem variabilidade
App parte 2	0,72
App parte 3	0,60
App parte 4	0,71

Os coeficientes do token computadorizado parte 1, parte 2, e o token aplicativo parte 2 e parte 4 apresentaram valor de consistência boa, enquanto que as partes 3 e 4 do computadorizado, e a parte 3 do aplicativo apresentaram coeficiente moderado. A parte 1 do token mobile apresentou coeficiente abaixo do mínimo tolerável de 0,60 e falta de variabilidade entre os itens, por conta de todos os participantes acertarem os 10 itens.

Para a análise de comparação de desempenho da amostra geral entre as duas versões, o procedimento estatístico escolhido foi o teste de Wilcoxon para amostras pareadas. Os resultados do teste podem ser visualizados na tabela 5. Houve diferença significativa na parte 1, parte 3, parte 4, e no escore total, sendo que na parte 2 houve uma tendência por conta do valor próximo ao $p = 0,05$. Escores maiores foram observados na versão em aplicativo.

Tabela 5 – Teste de Wilcoxon de amostras pareadas para comparação de desempenho entre as versões aplicativo e computadorizada do teste token

		Posto Médio	Soma dos Postos	Z	p
App parte1 – Comp Parte1	Postos Positivos	9,2	92	-2,39	0,017
	Postos Negativos	16,82	286		
App parte2 – Comp parte2	Postos Positivos	16,27	179	-1,84	0,066
	Postos Negativos	17,36	382		

App parte3 – Comp parte3	Postos Positivos	12,55	125,5	-3,299	0,001
	Postos Negativos	20,79	540,5		
App parte4 – Comp parte4	Postos Positivos	18,23	237	-2,881	0,004
	Postos Negativos	23,63	709		
App_Total - Comp_Total	Postos Negativos	18,23	200,5	-3,444	0,001
	Postos Positivos	23,92	789,5		

Para a comparação do tempo médio utilizado para completar cada item de cada uma das versões do teste, o procedimento estatístico adotado foi o teste de Wilcoxon para amostras pareadas. Na tabela 6 de estatísticas descritivas, é possível ver que o tempo médio dispendido por item na versão computadorizada é significativamente maior que na versão mobile, assim como a variação média do tempo por item no teste computadorizado também é maior que na versão tablet. Na tabela 7 pode-se observar que houve diferença significativa entre o tempo médio por item empregado em cada versão, com $p < 0,001$. Isso indica, a princípio, que o manuseio pela plataforma tablet foi mais fácil.

Tabela 6 - Estatísticas descritivas da comparação de tempo total nas duas versões

	Média	Desvio padrão	Mínimo	Máximo
Tempo_Comp	8124,846	3899,266	3842,600	23793,830
Tempo_App	4625,702	1211,376	2733,4	7690,6

Tempo computado em milissegundos (1s = 1000)

Tabela 7 - Ranqueamento do teste de Wilcoxon de amostras pareadas para comparação das medidas de tempo em ambas as versões

		N	Posto Médio	Soma dos Postos	Z	p
Tempo_App – Tempo_Comp	Postos Negativos	48	27,02	1297	-5,943	0,0001
	Postos Positivos	3	9,67	29		

Para a análise de correlação entre o desempenho da amostra geral no teste token computadorizado e no teste token aplicativo, o ρ de Spearman se mostrou como alternativa adequada ao r de Pearson. De uma maneira geral, foram encontradas correlações moderadas e baixas entre os dois instrumentos. Os dados podem ser visualizados na tabela 8.

As correlações entre os escores totais e entre a quarta parte dos testes foram moderadas, com coeficiente de 0,456 e $p=0,001$, e de 0,559 e $p<0,001$, respectivamente. As correlações entre a quarta parte do token app com escore total do computadorizado, e a quarta parte do computadorizado com o escore total do app foram moderadas, ambas com $p<0,001$. A correlação da terceira parte do computador com o escore total do app também foi moderada, com significância de $p=0,002$, e a correlação da parte 3 computadorizada teve correlação moderada com a parte 4 do App, com $p<0,001$. Os demais coeficientes de correlação foram baixos, embora alguns deles apresentem bons índices de significância.

Tabela 8 – Correlação de Spearman entre os escores totais e parciais das duas versões

			Comp_Total	Comp Parte 1	Comp Parte 2	Comp Parte 3	Comp Parte4
Spearman's rho	App_Total	Coeficiente de Correlação	0,456	0,135	0,317	0,430	0,502
		P	0,001	0,346	0,023	0,002	0,000
	App Parte1	Coeficiente de Correlação	0,108	0,001	0,017	0,121	0,16
		P	0,451	0,994	0,904	0,399	0,264
	App Parte2	Coeficiente de Correlação	0,316	0,138	0,153	0,272	0,325
		P	0,024	0,334	0,283	0,053	0,02
	App Parte3	Coeficiente de Correlação	0,197	0,265	0,151	0,297	0,097
		P	0,166	0,06	0,292	0,034	0,5
	App Parte4	Coeficiente de Correlação	0,490	0,062	0,351	0,454	0,559
		P	0,000	0,667	0,011	0,001	0,000

Para a análise de variância do efeito dos anos escolares sobre o escore no Teste Token, em ambas as versões, foi escolhido o teste Kruskal Wallis. Na tabela 9 estão expostas as estatísticas descritivas. Nela é possível observar que da terceira para a quarta parte do teste há em média uma queda de 4 pontos no número de acertos. Essa queda é homogênea e acontece em ambas as versões. O teste revelou efeito significativo dos anos escolares na versão computadorizada $p=0,033$ e tendência na versão aplicativo $p= 0,072$.

Nas tabelas 10 e 11 são apresentados os postos médios de acertos para cada ano escolar, ou seja, o escore correspondente a mediana de acertos de cada ano, assim como a significância da diferença de pontuação entre os anos. As medidas com diferença significativa foram o escore total do teste token computadorizado (Comp Total), com $p=0,033$, e a parte 3 do computadorizado (Comp Parte 3), com $p=0,043$, enquanto que no aplicativo o escore total e a parte 4 apresentaram tendência somente, com $p=0,07$ e $p=0,07$, respectivamente. Nos escore total do App e nas partes Comp 4 e App 4, é possível observar aumento progressivo nos postos médios com relação aos anos, embora não ocorra diferença significativa.

Tabela 9 – Estatísticas descritivas do teste Kruskal-Wallis para efeito de anos escolares

	Ano Escolar	Comp Total	Comp parte1	Comp parte2	Comp parte3	comp parte4	App total	App parte1	App parte2	App parte3	App parte4	
Média	1	26,8	8,75	8,19	7,06	3,13	30,3	9,63	8,81	8,00	3,88	
	2	28,4	9,31	8,08	7,46	3,54	32,9	9,69	9,31	9,23	4,69	
	3	30,9	9,00	9,20	8,50	4,20	32,9	9,70	8,70	8,80	5,70	
	4	29,9	9,57	7,86	8,29	4,14	32,6	9,29	8,71	9,43	5,14	
	5	36,0	9,60	9,40	9,40	7,40	37,6	10,0	10,0	9,60	8,00	
Desvio Padrão	1	3,87	1,58	2,00	1,62	1,92	5,07	0,703	1,14	2,02	2,28	
	2	6,03	1,66	1,84	2,44	2,01	2,56	0,468	1,00	0,706	1,88	
	3	5,07	0,788	1,19	1,94	1,86	4,65	0,651	1,88	0,997	2,81	
	4	6,09	0,507	2,41	1,42	2,48	5,33	1,06	2,43	0,507	2,29	
	5	5,40	0,828	1,24	0,828	3,18	1,92	0,00	0,00	0,507	1,60	
Mínimo	1	18	5	2	5	1	18	8	7	3	0	
	2	16	4	3	1	0	28	9	7	8	2	
	3	19	8	7	4	0	25	8	5	7	1	
	4	21	9	4	6	1	22	7	3	9	2	
	5	26	8	7	8	2	34	10	10	9	5	
Máximo	1	33	10	10	10	7	37	10	10	10	7	
	2	35	10	10	10	6	39	10	10	10	9	
	3	36	10	10	10	7	40	10	10	10	10	
	4	36	10	10	10	7	37	10	10	10	8	
	5	40	10	10	10	10	39	10	10	10	9	
Percentil	25th	1	24,8	8,00	7,75	6,00	1,75	28,0	9,75	8,50	6,75	2,75
		2	26,0	10,0	8,00	6,00	3,00	32,0	9,00	9,00	9,00	4,00
		3	29,0	8,00	9,00	7,00	3,00	29,0	10,0	7,00	8,00	3,00
		4	22,0	9,00	5,00	7,00	1,00	29,0	9,00	9,00	9,00	3,00
		5	36,0	10,0	10,0	9,00	6,00	38,0	10,0	10,0	9,00	8,00
	50th	1	27,0	9,50	9,00	6,50	3,00	31,5	10,0	9,00	9,00	4,00
		2	30,0	10,0	9,00	8,00	3,00	33,0	10,0	10,0	9,00	4,00
		3	32,5	9,00	10,0	9,50	4,50	34,0	10,0	10,0	9,00	6,50
		4	31,0	10,0	9,00	8,00	4,00	36,0	10,0	10,0	9,00	6,00
		5	38,0	10,0	10,0	10,0	9,00	38,0	10,0	10,0	10,0	9,00
	75th	1	29,0	10,0	9,25	8,25	4,00	34,3	10,0	10,0	9,25	6,00
		2	33,0	10,0	9,00	9,00	5,00	34,0	10,0	10,0	10,0	6,00
		3	34,0	10,0	10,0	10,0	5,00	36,0	10,0	10,0	10,0	8,00
		4	36,0	10,0	10,0	10,0	7,00	37,0	10,0	10,0	10,0	7,00
		5	40,0	10,0	10,0	10,0	10,0	39,0	10,0	10,0	10,0	9,00

Tabela 10 e 11 – Ranqueamento do teste kruskal-wallis para anos escolares em ambas versões

	Ano	N	Posto Médio	Qui-quadrado	p		Ano	N	Posto Médio	Qui-quadrado	p
Comp_Total	1	16	18,5	8,75	0,033	App_Total	1	16	19,69	7,001	0,072
	2	13	24,27				2	13	25,04		
	3	10	30,6				3	10	27,1		
	4	12	34,04				4	12	34,54		
	Total	51					Total	51			
Comp1	1	16	22,84	4,958	0,175	App1	1	16	25,88	0,168	0,983
	2	13	30,54				2	13	25,27		
	3	10	20,65				3	10	27,2		
	4	12	29,75				4	12	25,96		
	Total	51					Total	51			
Comp2	1	16	23,38	4,252	0,235	App2	1	16	21	4,221	0,239
	2	13	21,62				2	13	27,88		
	3	10	32,4				3	10	25,4		

	4	12	28,92				4	12	31,13						
	Total	51					Total	51							
Comp3	1	16	18,78	8,162	0,043	App3	1	16	20,19	6,563	0,087	App4	1	16	19,84
	2	13	24,27				2	13	28,73				2	13	23,58
	3	10	31,85				3	10	23,35				3	10	29,9
	4	12	32,63				4	12	33				4	12	33,58
	Total	51					Total	51					Total	51	
Comp4	1	16	20,69	5,318	0,15	App4	1	16	19,84	7,014	0,071	App4	1	16	19,84
	2	13	24,04				2	13	23,58				2	13	23,58
	3	10	28,7				3	10	29,9				3	10	29,9
	4	12	32,96				4	12	33,58				4	12	33,58
	Total	51					Total	51					Total	51	

Durante a análise do efeito de idade nos escores, foi observado o agrupamento de faixas etárias em relação ao desempenho. Dessa forma, escolhemos agrupar os dados da amostra em dois grupos etários diferentes, sendo o Grupo 1 composto por crianças das faixas etárias de 6 e 7 anos de idade, e o Grupo 2 composto por crianças de 8 a 10 anos de idade. Para a análise foi utilizado o teste de Man-Whitney para duas amostras independentes. O teste revelou efeito significativo na versão computadorizada $p=0,005$ e na versão aplicativo para tablet $p=0,04$. As estatísticas descritivas estão expostas na Tabela 12.

Tabela 12 – Estatísticas descritivas do teste de Man-Whitney para grupos etários

	Média		Desvio Padrão		Mínimo		Máximo		percentil		25th		50th		75th	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Grupos Etários	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Comp parte1	8,96	9,35	1,67	0,775	4	8	10	10	8,00	9,00	10,0	10,0	10,0	10,0	10,0	10,0
Comp parte2	8,07	8,87	1,94	1,79	2	4	10	10	7,75	8,00	9,00	10,0	9,00	10,0	10,0	
Comp parte3	7,18	8,65	2,06	1,61	1	4	10	10	6,00	8,00	7,50	9,00	9,00	10,0	10,0	
Comp parte4	3,32	4,83	2,02	2,74	0	0	7	10	2,00	3,00	3,00	5,00	5,00	6,50	6,50	
Comp Total	27,4	31,7	5,06	5,86	16	19	35	40	25,8	28,5	27,0	33,0	31,0	36,0	36,0	
App parte1	9,68	9,61	0,612	0,783	8	7	10	10	9,75	9,50	10,0	10,0	10,0	10,0	10,0	
App parte2	9,00	9,04	1,12	1,92	7	3	10	10	8,75	9,00	9,00	10,0	10,0	10,0	10,0	
App parte3	8,54	9,17	1,73	0,834	3	7	10	10	8,00	9,00	9,00	9,00	10,0	10,0	10,0	
App parte4	4,25	5,96	2,20	2,64	0	1	9	10	3,00	3,50	4,00	7,00	6,00	8,00	8,00	
App Total	31,5	33,8	4,45	4,81	18	22	39	40	30,3	31,0	32,5	35,0	34,0	37,5	37,5	

Pode-se observar na coluna dos percentis que, da terceira para a quarta parte do teste, os escores caem em média 4 pontos. A queda é homogênea e em média os participantes pontuaram um ponto a mais na versão tablet.

Nos resultados do teste de Mann-Whitney, na tabela 13, pode-se observar que houve diferença significativa entre os postos médios de acertos do Grupo 1 e do Grupo 2 na parte 2, 3, 4, escore total e tempo total do teste token computadorizado. O Grupo 2 pontuou em média 10 pontos a mais que o Grupo 1 nessas mesmas partes do teste, sendo que no escore total na versão computadorizada, a diferença foi de 12 pontos. Sobre o tempo total para realização do token computadorizado, o Grupo 1 levou em média o dobro do tempo que o Grupo 2 levou para concluir o teste.

Para o teste em versão aplicativo mobile, houve diferença significativa na quarta parte, nos escores totais e no tempo total do teste. O Grupo 2 acertou cerca de 10 pontos a mais que o Grupo 1 na parte 4, e 8 pontos a mais no escore total. O Grupo 1 precisou de 10 UT(Unidades de Tempo) a mais que o Grupo 2 para finalizar o teste. Nas partes 2 e 3 do token app, embora o Grupo 2 tenha acertado em média 5 e 4 pontos a mais, respectivamente, que o Grupo 1, o teste de significância não acusou coeficiente menor que 0,05 para essas medidas.

Tabela 13 – Ranqueamento dos escores e escores parciais dos grupos etários 1 e 2

Idade		N	Posto Médio	Soma dos Postos	Mann-Whitney U	p
Comp1	Grupo 1	28	25,91	725,50	319,500	0,958
	Grupo 2	23	26,11	600,50		
	Total	51				
Comp2	Grupo 1	28	21,89	613,00	207,000	0,024
	Grupo 2	23	31,00	713,00		
	Total	51				
Comp3	Grupo 1	28	20,80	582,50	176,500	0,005
	Grupo 2	23	32,33	743,50		
	Total	51				
Comp4	Grupo 1	28	22,29	624,00	218,000	0,047
	Grupo 2	23	30,52	702,00		
	Total	51				
Comp_Total	Grupo 1	28	20,70	579,50	173,500	0,005
	Grupo 2	23	32,46	746,50		
	Total	51				
Tempo_Comp	Grupo 1	28	34,43	964,00	86,000	0,0001
	Grupo 2	23	15,74	362,00		
	Total	51				
App1	Grupo 1	28	26,20	733,50	316,500	0,891
	Grupo 2	23	25,76	592,50		
	Total	51				
App2	Grupo 1	28	23,61	661,00	255,000	0,161
	Grupo 2	23	28,91	665,00		

	Total	51				
App3	Grupo 1	28	24,04	673,00	267,000	0,271
	Grupo 2	23	28,39	653,00		
	Total	51				
App4	Grupo 1	28	21,57	604,00	198,000	0,018
	Grupo 2	23	31,39	722,00		
	Total	51				
App_Total	Grupo 1	28	22,14	620,00	214,000	0,040
	Grupo 2	23	30,70	706,00		
	Total	51				
Tempo_App	Grupo 1	28	30,86	864,00	186,000	0,010
	Grupo 2	23	20,09	462,00		
	Total	51				

Realizamos uma comparação entre meninos e meninas em todas as medidas da versão tablet e da versão computadorizada, e não houve diferença significativa no desempenho entre os sexos em ambas as versões, como descrito na Tabela 14. O que sugere que, a princípio, não existe diferença no desempenho de meninos e meninas no teste token e não são necessárias tabelas normativas diferentes para os dois sexos.

Tabela 14 – Comparação de desempenho de meninos e meninas em todas as medidas

	Sexo	N	Postos Médios	Soma dos postos	Mann-Whitney U	p
Comp_Total	Feminino	22	27	594	297	0,675
	Masculino	29	25,24	732		
	Total	51				
Comp1	Feminino	22	24,59	541	288	0,511
	Masculino	29	27,07	785		
	Total	51				
Comp2	Feminino	22	29,48	648,5	242,5	0,131
	Masculino	29	23,36	677,5		
	Total	51				
Comp3	Feminino	22	28,59	629	262	0,27
	Masculino	29	24,03	697		
	Total	51				
Comp4	Feminino	22	25,05	551	298	0,687
	Masculino	29	26,72	775		
	Total	51				
App_Total	Feminino	22	26,23	577	314	0,924
	Masculino	29	25,83	749		
	Total	51				
App1	Feminino	22	24,48	538,5	285,5	0,403
	Masculino	29	27,16	787,5		
	Total	51				
App2	Feminino	22	27,82	612	279	0,401
	Masculino	29	24,62	714		
	Total	51				
App3	Feminino	22	26,52	583,5	307,5	0,817
	Masculino	29	25,6	742,5		
	Total	51				

App4	Feminino	22	25,34	557,5	304,5	0,781
	Masculino	29	26,5	768,5		
	Total	51				

5. DISCUSSÃO

Os objetivos propostos por esse trabalho envolveram a verificação da precisão das duas versões do teste token, comparação do desempenho nos dois testes em função de faixa etária, ano escolar e sexo, além de correlação entre o desempenho nas duas versões e de comparação do tempo dispendido na tarefa em cada uma das plataformas. De maneira geral, tanto o objetivo geral quanto os objetivos específicos foram respondidos.

As análises de precisão apresentaram bons coeficientes, tanto para os escores totais quanto para os parciais nos dois instrumentos. Apenas a Parte 1 do teste token app apresentou coeficiente de fidedignidade abaixo do mínimo tolerável pois houve ausência de variabilidade nas respostas, ou seja, todos os participantes acertaram os 10 itens da primeira parte do aplicativo. Esse resultado indica que a parte 1 do Token App foi fácil de ser resolvida, o que pode ter relação com a aplicação do teste token app ter ocorrido depois da aplicação do token computadorizado, facilitando a resolução da tarefa.

A comparação de desempenho entre as duas versões mostrou diferença significativa entre os resultados dos escores totais, das partes 1, 3 e 4, e tendência para a parte 2 com valor próximo de $p=0,05$. Na análise de correlação do desempenho dos participantes nas duas plataformas, foram observadas correlações baixas e moderadas, com bons índices de significância. Segundo HUTZ et al. (2015), são esperadas correlações altas entre testes que avaliam o mesmo construto. Nesse sentido, eram esperadas correlações maiores entre o desempenho nos dois instrumentos, que podem ter sido afetadas pela não randomização da amostra.

Como as crianças foram sistematicamente melhores na versão aplicativo, acertando em média 1 ponto a mais do que na versão computadorizada, esse foi o principal fator que favoreceu a divergência no desempenho entre as plataformas, citada nas análises acima. Isso reflete que a versão tablet foi mais fácil de ser realizada. Para justificar esse efeito, é possível levantar algumas hipóteses. A primeira é que pode ter ocorrido um efeito de maturação entre a aplicação do token computadorizado e a do token app. Todos os 51 participantes responderam o aplicativo depois de terem sido avaliados no token computadorizado, com um intervalo de tempo de seis meses entre os dois, pois este havia sido coletado em um estudo prévio. A não randomização dessa amostra é um fator limitante desse estudo, e sugerimos para estudos subsequentes randomizem a amostra através do método das duas metades.

A comparação do tempo gasto para realizar os itens da tarefa mostra que a aplicação do teste token aplicativo mobile é significativamente mais rápida do que a aplicação computadorizada, com $p < 0,001$. Tal relação não era esperada, pois em estudo prévio, MACEDO et al. (2007) mostraram não haver diferenças significativas entre o tempo de administração do teste token computadorizado e o teste token tradicional em papel.

Uma hipótese a ser considerada é a possibilidade das crianças estarem mais familiarizadas com dispositivos touchscreen. No mesmo estudo, MACEDO et al. (2007) discute a possibilidade de os escores reduzidos na versão computadorizada em comparação com a versão tradicional do teste token serem justificados pelas demandas de manuseio da interface mouse e teclado no computador. Como no aplicativo mobile retiramos essas variáveis e inserimos a interface touchscreen, uma de nossas hipóteses era que o teste em formato aplicativo se assemelharia mais a versão tradicional do que a versão computadorizada. O fato das crianças atingirem escores maiores em menor tempo na versão aplicativo pode ser uma evidência nesse sentido.

Para verificar essa associação, seria interessante um estudo posterior comparando o teste token para tablets com a versão tradicional de papel do teste, a dizer, a versão reduzida de DE RENZI e FLAGLIONI (1978). A comparação do teste token app com outro instrumento que acesse um constructo similar, como vocabulário receptivo (ex. Teste de Vocabulário por Imagens Pebody) pode ser uma alternativa para buscar evidências de validade convergente.

Na comparação de desempenho em função dos anos escolares, foi observado efeito positivo do tempo de escolarização no desempenho do teste token computadorizado, e no teste token aplicativo mobile foi observado apenas uma tendência, com significância próxima de $p = 0,05$. Em estudo de CAPOVILLA et al. (1997), foi encontrado efeito positivo significativo da escolarização para habilidades de vocabulário receptivo e expressivo, o que corrobora com o efeito da escolarização no token computadorizado, embora fosse interessante um estudo posterior com amostra maior para discriminar melhor esse efeito no Token App.

Na comparação do desempenho entre grupos etários, foi observado efeito positivo nos escores com o aumento da idade cronológica. Esse resultado foi observado através do agrupamento das faixas etárias em dois grupos distintos. Como os grupos refletem dois períodos do desenvolvimento, com idades cronologicamente próximas, é possível observar uma relação positiva no desempenho em função do aumento da idade, já que os grupos apresentaram diferença significativa nas pontuações dos escores totais e parciais, sendo que o grupo 2 acertou em média 10 itens a mais que o grupo 1. Tal efeito está de acordo com o resultado de outros estudos que encontraram progressão no número de acertos no teste token

em função da idade em crianças, como MALLOY-DINIZ et al. (2007) e ARANGO-LASPRILLA et al. (2017).

Na comparação do desempenho entre os sexos, não foram encontradas diferenças significativas entre as pontuações totais e parciais de meninos e meninas. Esse resultado está de acordo com o resultado de outros estudos como o de MACEDO et al. (2007), MALLOY-DINIZ et al. (2007) e ARANGO-LASPRILLA et al. (2017).

Entre as limitações do estudo, podemos colocar a amostra reduzida, em especial nas faixas etárias mais velhas. Outra limitação do estudo foi que os dados não apresentaram distribuição normal, o que nos forçou realizar análises estatísticas não-paramétricas, que tem um poder inferencial muito inferior aos métodos estatísticos paramétricos. Ademais, podemos colocar como limitação todos os participantes serem de escola privada, e seria interessante em estudo posterior abranger crianças do ensino público também.

6. CONSIDERAÇÕES FINAIS

Através das análises descritivas e inferenciais dos dados, foi observado aumento do número de acertos associado ao aumento da idade cronológica e ano escolar. Segundo Mecca (2010), testes que se propõem a analisar períodos do desenvolvimento podem considerar o aumento gradual da pontuação em função da idade e da escolarização como evidências de validade.

Foram observadas também evidências de validade por critério a partir da correlação e da comparação intraparticipantes das medidas de desempenho entre as versões: embora as correlações tenham sido moderadas, os testes de significância corroboram que essas associações não foram por acaso, o que pode indicar que a magnitude da correlação pode ter sido afetada pela não randomização da amostra. Outro ponto é que, segundo MECCA (2010), testes que se propõem a avaliar períodos do desenvolvimento são muito sensíveis ao espaço de tempo entre aplicações, uma vez que as crianças estão em um processo contínuo de desenvolvimento. Como houve seis meses de intervalo entre a coleta dos testes, tal fato pode ter impactado nos resultados, favorecendo o desempenho no aplicativo.

No que tange à Fidedignidade, podemos dizer que o instrumento apresentou bons índices de consistência interna, tanto para o escore total quanto para os escores parciais. Os bons índices de precisão do token app, e sua semelhança com os coeficientes do teste token computadorizado podem ser considerados outra fonte de evidências de validade.

De maneira geral, foi possível levantar evidências de validade e de fidedignidade para o Teste Token Aplicativo Mobile para uma população entre 6 e 10 anos de idade do ensino

particular da cidade de São Paulo. Também foi possível formular caminhos para novos estudos de validação do instrumento. Segundo a International Test Commission, a busca por evidências de validade e de precisão são um requisito fundamental para atestar a qualidade de um instrumento, a fim que esse possa ser oferecido para uso na população posteriormente.

7. REFERÊNCIAS

ARANGO-LASPRILLA, J. C. *et al.* Shortened Version of the Token Test: Normative data for Spanish-speaking pediatric population. **NeuroRehabilitation**, v. 41, n. 3, p. 649–659, 2017. Disponível em: <<https://doi.org/10.3233/nre-172244>>.

CAPOVILLA, F. *et al.* Versão brasileira do teste de vocabulário por imagens peabody: dados preliminares. **Distúrbios da Comunicação**, 1997.

DE RENZI, E.; VIGNOLO, L. A. The token test: A sensitive test to detect receptive disturbances in aphasics. **Brain**, v. 85, n. 4, p. 665–678, 1962.

DE RENZI, E.; FAGLIONI, P. Normative data and screening power of a shortened version of the Token test. **Cortex**, 14: 41-9, 1978.

INTERNATIONAL TEST COMMISSION. Guidelines on computer-based and internet delivered testing. 2005. Disponível em: <<https://www.intestcom.org>>.

FIRDOUS, S.; WAHID, M.; UD DIN, A.; BAKHT, K.; KHAN, M. Y. A.; BATOOL, R.; NOREEN, M. Android based receptive language tracking tool for toddlers. **International Journal of Advanced Computer Science and Applications**, v. 10, n. 3, p. 589–595, 2019. Disponível em: <<https://doi.org/10.14569/IJACSA.2019.0100375>>.

GERMINE, L.; REINECKE, K.; CHAYTOR, N. S. Digital neuropsychology: challenges and opportunities at the intersection of science and software. **Clinical Neuropsychologist**, v. 0, n. 0, p. 1–16, 2019. Disponível em: <<https://doi.org/10.1080/13854046.2018.1535662>>.

HALL, C. L. *et al.* Investigating a therapist-guided, parent-assisted remote digital behavioural intervention for tics in children and adolescents - “Online Remote Behavioural Intervention for Tics” (ORBIT) trial: Protocol of an internal pilot study and single-blind randomised controlled trial. **BMJ Open**, v. 9, n. 1, p. 1–12, 2019.

HOLMES, E. A. *et al.* Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. **The lancet. Psychiatry**, v. 0366, n. 20, p. 1–14, 2020.

HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. *Psicometria*. Porto Alegre: editora Artmed, 2015.

MACEDO, E. C.; CAPOVILLA, F. C.; CHARIN, S.; DUDUCHI, M. Versões computadorizadas de testes neuropsicológicos: Teste Boston para Diagnóstico Diferencial das Afasias (Boston-

COMP), Teste Boston de Nomeação (TBN-Comp) e Teste Token para Crianças (Token-Comp). In: CAPOVILLA, F. C.; M. J. GONÇALVE, M. J., MACEDO, E. C. (Orgs.). **Tecnologia em (re)habilitação cognitiva: uma perspectiva multidisciplinar**. São Paulo: Edunisc, 1998.

MACEDO, E. C. et al. Avaliando linguagem receptiva via teste de vocabulário por imagens peabody: versão tradicional versus computadorizada. **Psicologia: Teoria e Prática**, v. 8, n. 2, p. 40–50, 2006.

MACEDO, E.; FIRMO, L.; DUDUCHI, M.; CAPOVILLA, F. Avaliando linguagem receptiva via Teste Token: versão tradicional versus computadorizada. **Avaliação Psicológica**, v. 6, n. 1, p. 61–68, 2007.

MALLOY-DINIZ, L. F.; BENTES, R. C.; FIGUEIREDO, P. M.; BRANDÃO-BRETAS, D.; DA COSTA-ABRANTES, S.; PARIZZI, A. M.; BORGES-LEITE, W.; SALGADO, J. V. Normalización de una batería de tests para evaluar las habilidades de comprensión del lenguaje, fluidez verbal y denominación en niños brasileños de 7 a 10 años: Resultados preliminares. **Revista de Neurologia**, v. 44, n. 5, p. 275–280, 2007.

MCNEIL, M. R. et al. Reliability and validity of the computerized Revised Token Test: Comparison of reading and listening versions in persons with and without aphasia. **Journal of Speech, Language, and Hearing Research**. v. 24, n. May, p. 272–281, 2015.

MECCA, T. P. Tradução, adaptação, fidedignidade e evidências de validade da bateria de visualização e raciocínio da “leiter international performance scale-revised”. 2010. **Tese** (Mestrado em Distúrbios do Desenvolvimento) -Universidade Presbiteriana Mackenzie, São Paulo, 2010.

SCHAEFER, B.; EHLERT, H.; KEMP, L.; HOESL, K.; SCHRADER, V.; WARNECKE, C.; HERRMANN, F. Stern, gwiazda or star: Screening receptive vocabulary skills across languages in monolingual and bilingual German–Polish or German–Turkish children using a tablet application. **Child Language Teaching and Therapy**, 2018.

STERNBERG, J.; STERNBERG, K. *Psicologia cognitiva*. São Paulo: **Editora Cengage Learning**, 2016.

VALENTINE, Althea Z.; BROWN, Beverley J.; GROOM, Madeleine J.; YOUNG, Emma; HOLLIS, Chris; HALL, Charlotte L. A systematic review evaluating the implementation of technologies to assess, monitor and treat neurodevelopmental disorders: A map of the current evidence. **Clinical Psychology Review**, [S. l.], v. 80, p. 101870, 2020. DOI: 10.1016/j.cpr.2020.101870. Disponível em: <<https://doi.org/10.1016/j.cpr.2020.101870>>.

Contatos: viniciusnveiga@gmail.com; elizeu.macedo@uol.com.br.