










Reprodutibilidade de medidas do controle inibitório, memória de trabalho e flexibilidade cognitiva em mulheres idosas

Reproducibility of inhibitory control, working memory, and cognitive flexibility measures in older women

Alan Pantoja-Cardoso¹ , José Carlos Aragão-Santos¹ , Marcos Raphael Pereira Monteiro¹ , Poliana de Jesus Santos¹ , Ana Carolina Dos-Santos¹ , Heloiana Faro² , Juan Ramon Heredia-Elvar³ , Leonardo de Sousa Fortes² , Marzo Edir Da Silva-Grigoletto¹ 

1. Universidade Federal de Sergipe, São Cristóvão, SE, Brasil

2. Universidade Federal de Paraíba, Brasil

3. Universidad Alfonso X El Sabio, Madrid, Espanha

RESUMO

Introdução: A Função Executiva é expressa nas atividades do dia a dia por meio do controle inibitório, memória de trabalho e da flexibilidade cognitiva. Apesar da importância de avaliar essas medidas, existem divergências sobre a reprodutibilidade dos testes. **Objetivo:** Testar a reprodutibilidade do *Stroop Color-Word Test*, Teste dos Cubos de Corsi e Teste de Trilhas em mulheres idosas. **Métodos:** Trinta e cinco mulheres idosas realizaram o *Stroop Color-Word Test* (Controle Inibitório), Teste dos Cubos de Corsi (Memória de Trabalho) e Teste de Trilhas (Flexibilidade Cognitiva) com uma semana entre o teste e reteste. A reprodutibilidade dos testes foi determinada pelo coeficiente de correlação intraclass, coeficiente de variação, erro padrão da medida e inspeção visual dos gráficos de Bland-Altman. **Resultados:** O *Stroop Color-Word Test* apresentou valores satisfatórios quanto à reprodutibilidade apenas para as medidas congruentes e incongruentes, com valores excelentes de coeficiente de correlação intraclass. O Teste dos Cubos de Corsi apresentou valores reprodutíveis com coeficiente de correlação intraclass moderado e bom para a sequência e escore composto, respectivamente. O Teste de Trilhas apresentou valores reprodutíveis para as partes A, B e a razão (B/A), com coeficientes de correlação intraclass entre moderado e bom. A inspeção visual nos gráficos de Bland-Altman demonstrou baixo viés em todas as variáveis. **Conclusão:** Os resultados do *Stroop Color-Word Test*, para ensaios congruentes e incongruentes, a sequência e o escore composto do Teste dos Cubos de Corsi, assim como a parte A, B e a razão (B/A) do Teste de Trilhas são medidas reprodutíveis para mulheres idosas.

Palavras-chave: confiabilidade do teste-reteste; função executiva; pessoas idosas; testes neuropsicológicos.

ABSTRACT

Introduction: Executive Function is expressed in day-to-day activities through inhibitory control, working memory, and cognitive flexibility. Despite the importance of evaluating these measures, there are disagreements about the reproducibility of the tests. **Objective:** To test the reproducibility of the *Stroop Color-Word Test*, Corsi Block-Tapping Test, and Trail Making Test in older women. **Methods:** Thirty-five older women performed the *Stroop Color-Word Test* (Inhibitory Control), Corsi Block-Tapping Test (Working Memory), and Trail Making Test (Cognitive Flexibility) within one week between the test and retest. The reproducibility of the tests was determined by the intraclass correlation coefficient, coefficient of variation, standard error of measurement, and visual inspection of the Bland-Altman graphs. **Results:** The *Stroop Color-Word Test* showed satisfactory reproducibility values only for congruent and incongruent measures, with excellent intraclass correlation coefficient values. Corsi Block-Tapping Test showed reproducible values with a moderate and good intraclass correlation coefficient for the sequence and composite score, respectively. The Trail Making Test showed reproducible values for parts A, B, and the ratio (B/A), with intraclass correlation coefficients between moderate and good. Visual inspection of the Bland-Altman plots showed low bias in all variables. **Conclusion:** The results of the *Stroop Color-Word Test*, for congruent and incongruent trials, the sequence and the composite score of the Corsi Block-Tapping Test, as well as the part A, B, and the ratio (B/A) of the Trail Making Test, are reproducible measurements for older women.

Keywords: test-retest reliability; executive function; old people; neuropsychological tests.

Recebido em: 2 de fevereiro de 2023; Aceito em: 5 de março de 2023.

Correspondência: Alan Pantoja-Cardoso, alan_pantoja1996@hotmail.com

Introdução

A Função Executiva (FE) se trata de processos mentais superiores que garantem uma pessoa a engajar em comportamentos do dia-a-dia [1]. A FE inclui habilidades necessárias quando recursos atencionais são requeridos ao longo de uma tarefa, além de ser utilizada para processos cognitivos automáticos e intuitivos [1]. Permite ao indivíduo refletir antes de agir, trabalhar diferentes ideias, resolver desafios inesperados, pensar sob diferentes perspectivas, reconsiderar opiniões divergentes e evitar distrações [2]. O bom funcionamento da FE é essencial para a manutenção da qualidade de vida [3,4]. Dentre os domínios da FE, os mais estudados são o controle inibitório, memória de trabalho e a flexibilidade cognitiva.

O controle inibitório é responsável por inibir processos mentais e comportamentais em detrimento de um objetivo, tais como adequar ações frente a objeções externas; por exemplo, em uma conversa, não falamos tudo que pensamos e sentimos, é preciso escolher o que falar conforme o contexto social [5]. A memória de trabalho, por sua vez, é vista como a manipulação da memória de acordo com a demanda exigida; por exemplo, quando se cozinha seguindo uma receita é necessário seguir etapas na ordem adequada para alcançar o resultado almejado [6]. Por fim, a flexibilidade cognitiva é o processo mental relacionado a adaptação frente a desafios ou acontecimentos, sendo usada para fazer ajustes em ações previamente planejadas ou na criação de algo frente a um contexto; por exemplo, quando possuímos várias opções e precisamos escolher apenas algumas delas para alcançar um resultado [1].

A literatura apresenta várias tarefas para avaliar o controle inibitório. As mais populares são os paradigmas *Go/No-Go* [7], *Flanker* [8] e o *Stroop Color Word Test* (SCWT) [9]. O *Go/No-Go* é uma tarefa com diferentes estímulos, alguns que devem ser respondidos e outros que não. Por exemplo, o avaliado deve reagir ao visualizar uma seta para direita, enquanto não deve reagir ao ver uma seta para esquerda [10]. A tarefa de *Flanker*, por sua vez, é baseada na uso de conjuntos de setas ou símbolos que podem ser congruentes (ex. todas as setas no mesmo sentido “<<<<<”), incongruentes (ex. sentidos diferentes “>><>>”), ou neutras (ex. incluindo setas e outros símbolos “--<--”) [8]. Por fim, o mais usual é o SCWT que é baseado em nomes de cores que são preenchidas pela mesma cor que a palavra indica (congruente) ou uma cor diferente (incongruente) e o avaliado deve indicar a cor de preenchimento não inibindo a leitura do que está escrito [9]. O SCWT possui vasta literatura, porém há divergências quanto a forma de pontuação e reprodutibilidade desse teste [11–14]. Nesse sentido, é preciso avaliar a reprodutibilidade do SCWT de forma computadorizada em indivíduos idosos padronizando sua forma de execução e pontuação.

A memória de trabalho, por sua vez, pode ser avaliada por meio de tarefas verbais ou não verbais. O *N-back test* explora tarefas verbais e não verbais, enquanto o *Teste do Cubos de Corsi* (TCC) é não verbal [15–17]. No *N-back test* o indivíduo deve lembrar de números anteriores ou imagens, podendo ser chamado 1-back (lembrar do número exposto antes do número atual), 2-back (lembrar do número exposto an-

tes dos dois últimos números apresentados), e assim sucessivamente sendo possível avaliar tanto o tempo de resposta quanto a acurácia [15]. O TCC avalia a memória de trabalho visuoespacial, solicitando que a participante selecione quadrados na mesma ordem em que lhe foram apresentados (ordem direta) ou na ordem inversa partindo do último quadrado apresentado até o primeiro. No TCC é possível avaliar o escore composto (sequência x número de acertos) ou somente a sequência de acertos. Entretanto, a literatura ainda diverge sobre qual a melhor pontuação a ser adotada, além de não apresentar bons valores de reprodutibilidade mesmo ao realizar seis ensaios com intervalos de uma semana, principalmente em idosos [18–20].

A flexibilidade cognitiva é entendida como uma resultante do controle inibitório e da memória de trabalho, uma vez que é preciso inibir uma ação premeditada (controle inibitório) e verificar as alternativas para agir de modo diferente comparando com as vivências anteriores (memória de trabalho) [1]. O teste de trilhas (TT) e a Tarefa de classificação de cartas de Wisconsin são duas abordagens para avaliar a flexibilidade cognitiva [1,21,22]. Na tarefa de classificação de cartas de Wisconsin o participante deve combinar as cartas de um baralho que totalizam 128 com quatro cartas-alvo distribuídas sobre a mesa. As cartas podem ser combinadas com base em suas cores “vermelho, azul, amarelo ou verde” ou nas formas geométricas “cruzes, círculos, triângulos ou estrelas”. O teste consiste na combinação de 10 cartas com base nas cores ou formas geométricas [23]. O TT, por sua vez, consiste em uma tarefa dividida em duas partes, A e B, o TT-A avalia a velocidade de processamento ao considerar o tempo que a participante usa para ligar 25 pontos em ordem numérica crescente. O TT- B representa a busca visual e a flexibilidade cognitiva ao avaliar a ligação de números e letras em ordem crescente e intercalada (ex. um número e uma letra) dispostos aleatoriamente. Desse modo, o TT-B inclui o controle inibitório ao verificar a não ligação de letra com letra ou número com número, e a memória de trabalho ao precisar lembrar a sequência numérica e alfabética crescente após cada ligação. Dentre as formas de análise da pontuação no TT está a diferença (B-A) e a razão (B/A) no tempo de execução [19, 24, 25]. Nesse sentido, o estudo de Wang *et al.* [25] apresentou moderada reprodutibilidade do TT-A e excelente no TT-B em indivíduos idosos, porém, não abordam outras medidas como a diferença (B-A) e a razão (B/A), além da literatura não apresentar um consenso na utilização para o público de mulheres idosas e o intervalo entre as aplicações do teste e reteste.

Nesse sentido, é preciso analisar o que é mais relevante pensando na avaliação da FE, avaliar apenas um domínio de forma isolada ou aplicar diferentes testes para analisar diferentes domínios. Consequentemente, é possível que a aplicação de diferentes testes para a FE em sequência, bem como o intervalo de reaplicação e o público alvo possam afetar a reprodutibilidade de testes para a FE. Desse modo, nosso objetivo foi testar a reprodutibilidade do SCWT, TCC e TT em mulheres idosas de forma sequencial usando um intervalo de sete dias entre as mensurações. Acreditamos que ao considerar a amostra envolvida no estudo, o período de sete dias é o mais adequado para minimizar o efeito aprendizagem e garantir uma maior reprodutibilidade

nos testes. Adicionalmente, acreditamos que mesmo ao serem aplicados de forma sequencial os testes apresentarão uma boa reprodutibilidade quando comparado aos valores apresentados pela literatura permitindo uma avaliação consistente dos principais domínios da FE.

Métodos

Participantes

Foram recrutadas um total de 70 mulheres por meio de panfletagem no entorno do campus Prof. José Aloísio de Campos da Universidade Federal de Sergipe na cidade de São Cristovão. Os critérios de inclusão foram: possuir no mínimo 12 pontos *Montreal Cognitive Assessment* (MoCA); ser fisicamente independente; ter idade entre 60 a 79 anos; ser alfabetizada. Por sua vez os critérios de exclusão foram: possuir daltonismo; doenças neurológicas e/ou psiquiátricas (ex.: doença de Parkinson); deficiência auditiva ou visual incompatível com a neuropsicologia dos testes; e não apresentar deficiência motora fina que poderia interferir com a execução das tarefas cognitivas e motoras.

Após a triagem, 40 participantes atenderam os critérios de inclusão, dos quais, 35 participantes realizaram os três testes propostos no estudo de forma sequencial e com um intervalo de sete dias entre o teste e o reteste (Figura 1). Previamente à coleta de dados, as participantes assinaram o termo de consentimento livre e esclarecido (TCLE) mediante a explicação de todos os procedimentos a serem realizados. A pesquisa foi submetida ao comitê de ética da instituição, aprovada sob o parecer 3.225.938 e está de acordo com a Declaração de Helsinki para pesquisas com seres humanos.

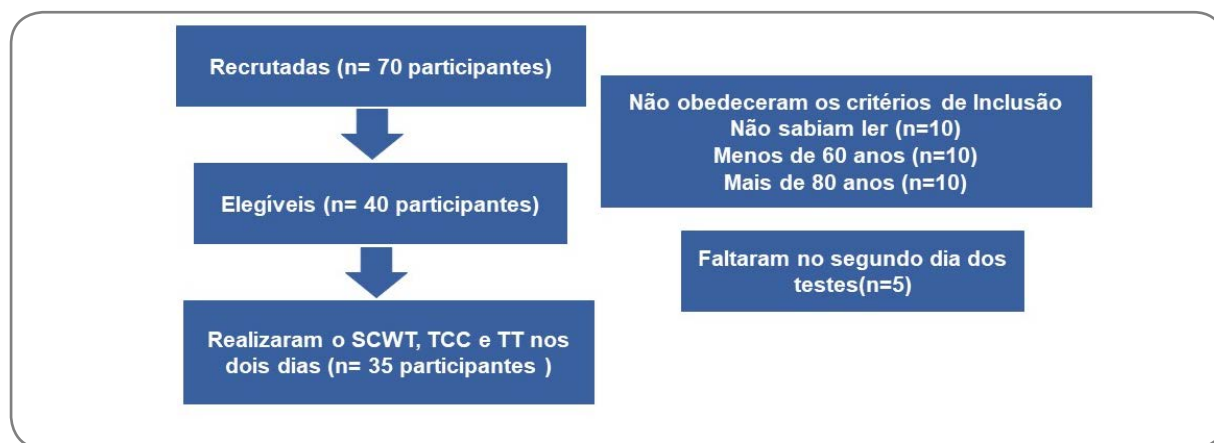


Figura 1- Fluxograma das participantes

Protocolo da função executiva

Inicialmente foram obtidas as medidas de massa corporal e estatura para calcular o índice de massa corporal (IMC) e aplicação do questionário MoCA que envolve a FE, memória de trabalho visuoespacial, memória episódica e atenção para avaliar a cognição global da pessoas idosa [26,27].

Cada participante visitou o laboratório em três sessões distintas: a primeira para caracterização da amostra e duas com um intervalo de sete dias entre si para execução dos testes no horário da manhã. Cada sessão teve uma duração de 30 minutos. Objetivando a manutenção das participantes um lembrete foi feito três dias antes das sessões para confirmar a participação. Previamente à realização das medidas, as participantes foram familiarizadas com os aparatos utilizados para a realização dos testes.

No dia anterior aos testes, as participantes foram orientadas por meio de ligação e mensagem a abster-se do uso de álcool e atividade física vigorosa nas 24 horas, além de não fumar ou ingerir cafeína em até duas horas antes do experimento. Os testes foram conduzidos entre março e novembro de 2022 e foram sempre aplicados pelo mesmo avaliador.

Os testes SCWT e o TCC foram executados em computadores com tela de 15 polegadas. Para a construção dos estímulos e a montagem do experimento foi utilizado o programa *PsychoPy*® versão 2022 1.3 (<https://www.psychopy.org/>) e a sua disponibilização online foi realizada por meio da plataforma *Pavlovia* (<https://pavlovia.org/>). As participantes utilizaram teclados com adesivos das cores amarelo, azul, verde e vermelho nas teclas A, D, J, L para realizar os comandos durante a realização dos testes.

Durante os testes, inicialmente a participante permanecia em repouso durante cinco minutos e em seguida era iniciada a aplicação dos testes. Para tal, a participante permanecia sentada de frente para um monitor a uma distância de 50 cm. Em seguida era iniciada a aplicação dos testes na seguinte ordem: SCWT, TCC e TT. As instruções para cada tarefa foram fornecidas verbalmente e de forma escrita na tela do computador.

Stroop Color-Word Test (SCWT)

O SCWT avalia o controle inibitório [11]. O teste possui respostas congruentes (significado da palavra igual a cor da sua fonte) e incongruentes (significado palavra e cor da fonte divergente). Primeiro a participante executou 10% dos ensaios para a familiarização com o experimento, resultando em 12 ensaios de um total de 120. Em seguida, as participantes executaram 120 ensaios, sendo 60 congruentes e 60 incongruentes. Durante a realização do teste as participantes foram solicitadas a responderem o mais rápido possível. Foram analisados o tempo de resposta (TR) para estímulos congruentes e o TR para estímulos incongruentes que expressa o controle inibitório. Além disso, foi analisada a diferença média no desempenho entre tentativas congruentes e incongruentes que é comumente chamada de efeito Stroop, que é mais uma medida do controle inibitório [14]. O teste foi considerado válido quando a participante obteve uma acurácia de no mínimo 80%.

Teste dos Cubos de Corsi (TCC)

Este teste avalia memória de trabalho visuoespacial [19]. No início do teste, haviam quatro ensaios de familiarização com apenas dois quadrados, no qual ob-

tinham feedback de acerto ou erro. Nosso ensaio foi composto por nove quadrados (2cm x 2 cm) na cor azul, sendo que a cada 500 ms um quadrado mudava de cor, se tornava amarelo e então retornava a cor azul aleatoriamente. Em seguida, a participante era solicitada a indicar quais foram os quadrados que mudaram de cor na mesma ordem em que as mudanças ocorreram (ordem direta). As participantes não receberam nenhum feedback quanto aos acertos e erros na realização do teste. Caso a participante acertasse a sequência, se progredia no teste aumentando o número de extensão dos quadrados. Por outro lado, caso a participante errasse por duas vezes seguidas o ensaio era finalizado. Neste teste o aplicador auxiliou as participantes realizando a utilização do mouse para selecionar a sequência indicada pela participante, visto que as participantes não tinham familiaridade com o mouse. Os valores referentes à sequência que a participante conseguiu alcançar em um determinado ensaio e o escore composto calculado pela multiplicação do número de acertos obtidos em todos os ensaios pela pontuação da sequência foram utilizados para análise.

Teste de Trilhas (TT)

Este teste avalia a flexibilidade cognitiva [22]. O TT consistiu em duas partes: na parte A, as participantes foram solicitadas a ligar de forma contínua, usando uma caneta esferográfica, números de 1 a 25 dispostos aleatoriamente em uma folha de papel. Na parte B, as participantes foram solicitadas a ligar de forma contínua números e letras alternadamente (por exemplo, 1-A, 2-B, etc.). A pontuação em ambas as partes é definida pelo tempo gasto para executar o teste corretamente. Então, a diferença (B-A) é tomada como um índice de flexibilidade cognitiva, sendo que quanto maior a pontuação menor é a flexibilidade cognitiva da participante [28]. Além disso, foi calculada a razão (B/A) que também é uma estimativa da flexibilidade cognitiva. Na aplicação do teste, seguimos a recomendação de Reitan [22] na qual os erros não foram contabilizados e em caso de erro o avaliador indicava para que a participante voltasse ao número ou letra anterior e continuasse o teste [28].

Análise estatística

O tamanho amostral foi calculado por meio do software *G*Power 3.1.9.7* com base em um estudo piloto não publicado, considerando um erro alfa de 0,05, poder de 0,95, e razão entre as hipóteses alternativa e nula equivalente a 0,35 resultando em uma amostra mínima de 27 participantes [29,30]. Esse método de cálculo amostral foi utilizado previamente por Fontes *et al.* [31]. Todos os dados foram analisados no software JAMOVI, versão 2.3.16. A normalidade dos dados foi testada por meio do teste de *Shapiro-Wilk*. A reprodutibilidade do SCWT, TCC e TT foi determinada pelo coeficiente de correlação intraclasse de duas vias (CCI). O CCI foi interpretado conforme o sistema de classificação de Koo *et al.* [32] para reprodutibilidade: < 0,50 = pobre; 0,50- 0,75 = moderada; 0,75-0,90 = boa; e > 0,90 = excelente. Além disso, o coeficiente de variação (CV) e o erro padrão da medida (EPM) foram calculados. O nível de concordância entre as sessões foi analisado por meio do gráfico de Bland-Altman,

considerando o viés sistemático e seus limites de concordância de 95% (LoA = Viés) [33]. Adicionalmente, os dados da soma das diferenças das médias nos dois dias de avaliação foram analisados para uma melhor visualização da concordância entre as medidas. Os gráficos foram construídos por meio do software *GraphPad Prism* versão 8.

Resultados

A saber, a amostra analisada apresentou em média uma idade de $66,4 \pm 5,4$ anos, massa corporal de $67,1 \pm 11,5$ kg, estatura de $1,55 \pm 0,05$ m, IMC de $28,0 \pm 4,2$ kg/m². Além disso, as participantes apresentaram um escore médio de $21,9 \pm 3,83$ pontos no MoCA.

Em relação ao TR Congruente e Incongruente do SCWT, observou-se um CCI excelente, CV baixo e EPM dentro do esperado (Tabela I). Detectamos baixo viés para as duas medidas com base na análise de concordância com a extrapolação de apenas dois indivíduos para além do intervalo de concordância (Figura 2). Em relação ao ES, observamos um CCI moderado, EPM dentro do esperado, mas um elevado CV (Tabela I). Além disso, a concordância entre as medidas apresentou um viés próximo a zero e apenas três indivíduos fora dos limites de concordância (Figura 2).

Tabela I - Valores dos testes e indicadores de reprodutibilidade dos testes da FE entre as sessões

Testes da função executiva		Dia 1	Dia 2	Dia 2-1	Dia 2 e 1		
		Média DP	Média DP	DM (IC 95%)	CCI (IC 95%)	CV%	EPM
SCWT	TR	1173,92	1074,54	99,37	0,92	5,56	87,64
	Congruente (ms)	$\pm 319,10$	$\pm 264,19$	(56,79 – 141,95)	(0,75 \pm 0,96)		
	TR	1322,04	1202,42	119,62	0,91	6,94	122,07
	Incongruente (ms)	$\pm 379,29$	$\pm 351,41$	(60,31 – 178,92)	(0,77 \pm 0,96)		
	Efeito Stroop	147,03	128,57	18,45	0,59	60,55	100,00
		$\pm 141,27$	$\pm 122,27$	(-30,41 – 67,33)	(0,28 \pm 0,77)		
TCC	Sequência	3,00	3,29	-0,28	0,72	12,46	0,50
		$\pm 0,76$	$\pm 0,82$	(-0,53, – 0,04)	(0,51 \pm 0,84)		
	Escore Composto	51,60	61,77	-10,17	0,79	24,42	18,95
		$\pm 34,64$	$\pm 33,23$	(-19,37 – 0,96)	(0,63 \pm 0,88)		
TT	A (s)	54,77	45,87	8,89	0,78	12,39	12,39
		$\pm 23,01$	$\pm 21,26$	(2,87 – 14,91)	(0,57 \pm 0,88)		
	B (s)	122,39	128,21	-5,81	0,84	24,43	43,67
		$\pm 86,60$	$\pm 82,61$	(-27,24 – 15,61)	(0,72 \pm 0,91)		
	Diferença (B-A)	67,62	82,33	-14,70	0,77	42,21	42,21
		$\pm 70,16$	$\pm 71,52$	(-35,21 – 5,80)	(0,61 \pm 0,87)		
	Razão (B/A)	2,22	2,04	-0,43	0,59	21,89	0,65
		$\pm 0,88$	$\pm 0,90$	(-0,75 – -0,11)	(0,28 \pm 0,76)		

DP = desvio padrão; DM = diferença média; IC = intervalo de confiança; CCI = coeficiente de correlação intraclassa; CV = coeficiente de variação; EPM = erro padrão da medida; TR = tempo de resposta; SCWT = Stroop Color-Word Test; TT = teste de trilhas; TCC = teste dos cubos de Corsi

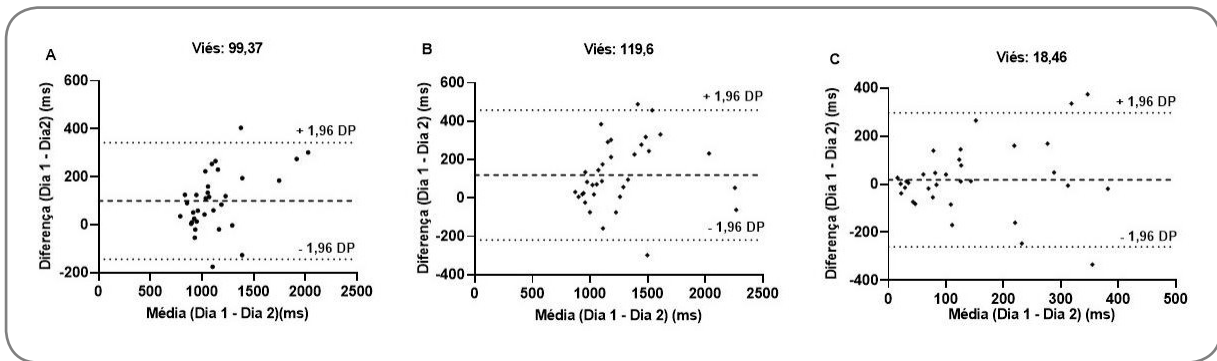


Figura 2 - Gráficos de Bland-Altman das diferenças entre o Dia 1 e o Dia 2 em função da média das medidas pareadas para o TR Congruente(A) e o TR Incongruente(B) e o efeito Stroop(C). A linha pontilhada representa o viés sistêmico e as linhas tracejadas representam os limites de concordância superior e inferior

Em relação ao TCC, os resultados da análise de sequência apresentaram CCI moderado, CV baixo e EPM dentro do esperado (Tabela I). Na concordância entre as medidas houve um viés próximo a zero e apenas um indivíduo extrapolou os limites de concordância (Figura 3). Por sua vez, o escore composto demonstrou um bom CCI, CV baixo e EPM dentro do esperado (Tabela I). Por fim, a concordância entre as medidas teve viés próximo a zero e apenas um indivíduo ficou fora dos limites de concordância (Figura 3).

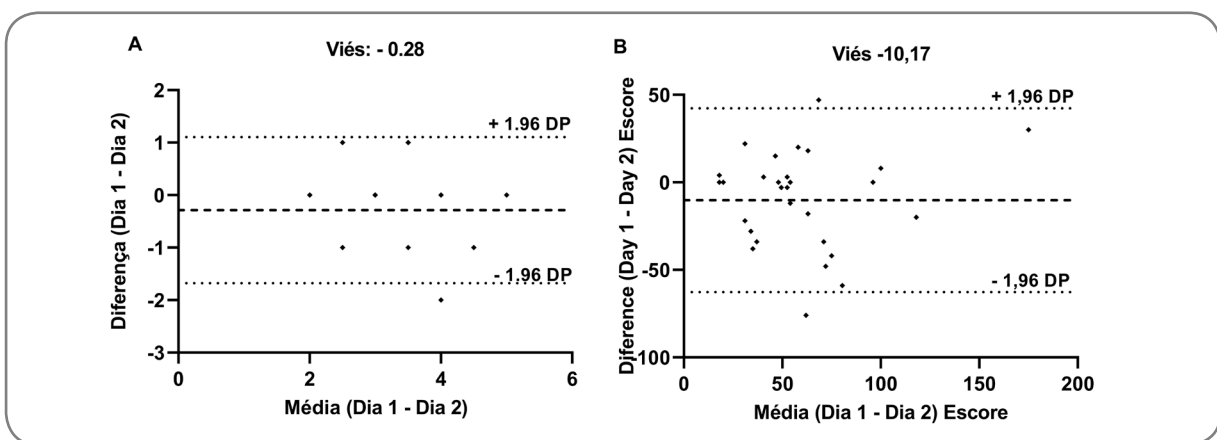


Figura 3 - Gráficos de Bland-Altman das diferenças entre o Dia 1 e o Dia 2 em função da média das medidas pareadas para a sequência do TCC(A) e o escore composto do TCC(B). A linha pontilhada representa o viés sistêmico e as linhas tracejadas representam os limites de concordância superior e inferior

Em relação ao TT-A e o TT-B foi verificado um CCI bom, CV baixo e EPM dentro do esperado (Tabela I). Na concordância entre as medidas verificamos um viés próximo a zero nas duas variáveis com dois indivíduos extrapolando o limite de concordância no TT-B (Figura 4). Usando outras medidas da flexibilidade cognitiva, especificamente a diferença (B-A), foi observado um bom CCI, CV alto e EPM dentro do esperado (Tabela I). A concordância entre as medidas apresentou um viés próximo a zero com dois indivíduos fora do limite de concordância. Já na razão (B/A) foi detectado um CCI moderado, CV baixo e EPM dentro do esperado (Tabela I). A concordância entre as medidas apresentou um viés próximo a zero e apenas um indivíduo fora dos limites de concordância (Figura 4).

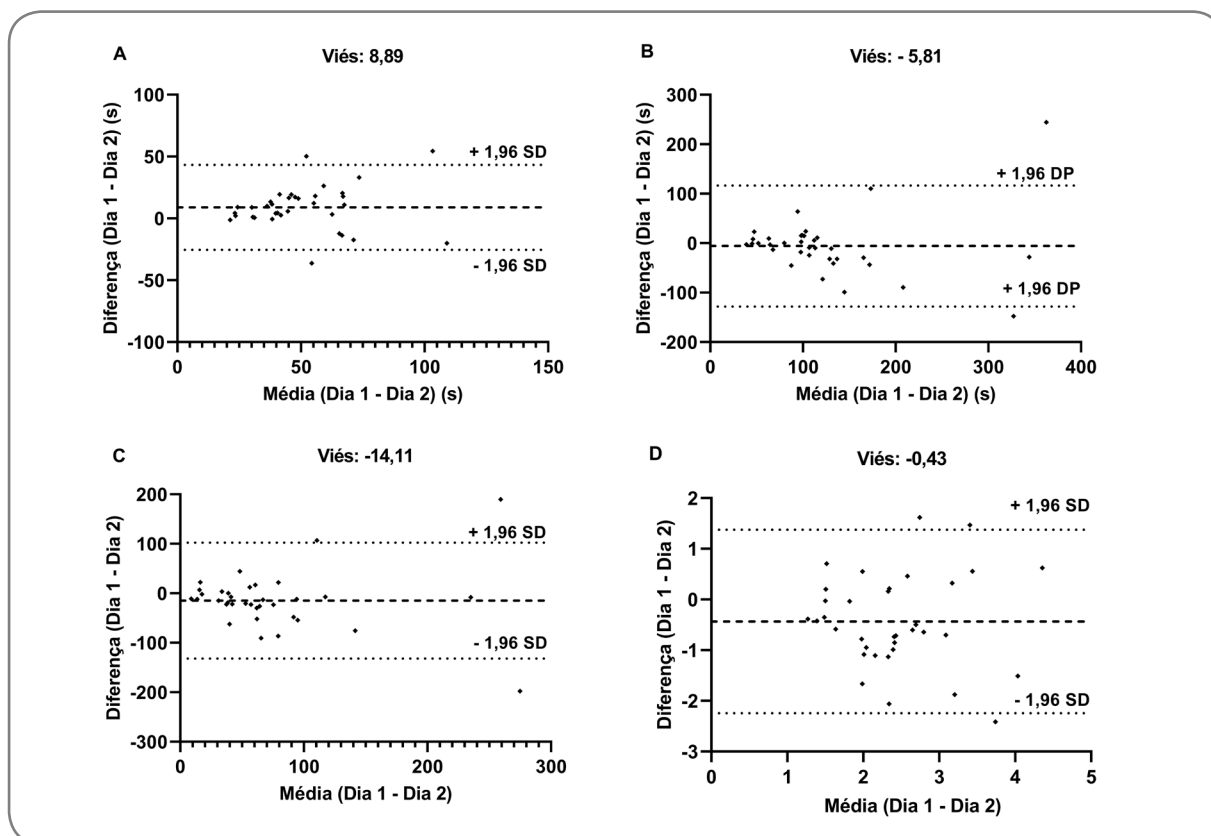


Figura 4 - Gráficos de Bland-Altman das diferenças entre o Dia 1 e o Dia 2 em função da média das medidas pareadas para o TT A (A) e o TT B (B) e diferença (B-A) (C) e a razão (B/A) (D). A linha pontilhada representa o viés sistemático e as linhas tracejadas representam os limites de concordância superior e inferior

Discussão

Os achados do presente estudo corroboram parcialmente a nossa hipótese, uma vez que alguns dos resultados obtidos em cada teste se mostraram reprodutíveis em mulheres idosas. Especificamente, os resultados do TR congruente e incongruente para o SCWT, valores do escore composto para o TCC, além das medidas do TT-A, TT-B e razão (B/A). Ademais, o intervalo de tempo utilizado e a aplicação dos testes em sequência parecem não prejudicar a reprodutibilidade das medidas. Desse modo, nossos achados são úteis para o delineamento de pesquisas investigando a FE de mulheres idosas [34].

Encontramos uma reprodutibilidade excelente no TR congruente e incongruente do SCWT, baixo CV e baixo viés nas análises de concordância. Contudo, o efeito Stroop apresentou moderada reprodutibilidade e alto CV. Esses achados corroboram os apresentados por Wang *et al.* [24], que avaliaram a reprodutibilidade em pessoas idosas no TR congruente e incongruente e demonstraram valor classificado como excelente (CCI = 0,91) com período entre teste e reteste de três a sete dias. De modo interessante, Wang *et al.*[24] aplicaram o SCWT usando lápis e papel enquanto nós realizamos por meio de computadores. Desse modo, é possível que não haja grande impacto na mensuração do controle inibitório com diferentes formas de aplicação. Entretanto, a aplicação por meio de computadores possibilita a maior facilidade des-

de a aplicação até a avaliação e número de testes aplicados [20,35]. Vale ressaltar que esses achados se aplicam a mulheres idosas, uma vez que outros estudos com adultos jovens encontraram valores abaixo dos apresentados no presente estudo [12].

Em relação aos valores do TCC, foram analisados a sequência e o escore composto, demonstrando que ambas as variáveis possuem boa reprodutibilidade. Esses valores diferem do estudo de White *et al.* [20] em que o TCC de ordem direta foi aplicado em 30 homens idosos saudáveis, apresentando reprodutibilidade pobre nas medidas de sequência e de escore composto [20]. Uma possível explicação pode se dar pelo auxílio do aplicador no manuseio do mouse, sendo um importante aspecto ao considerar a aplicação desse teste de forma computadorizada para garantir a qualidade da medida, uma vez que a população idosa tende a apresentar déficits de controle motor fino e baixa familiarização com o uso do mouse [36].

No que remete a flexibilidade cognitiva, os valores referentes ao TT-A, B e diferença (B-A) apresentaram boa classificação no CCI, já a razão (B/A) apresentou classificação moderada do CCI. Importante ainda apontar que o CV para TT-A, TT-B e a razão (B/A) foram classificados como baixo. Esses achados corroboram parcialmente com outros estudos que analisaram a mesma população, tais como os achados de Park e Shott [37] que avaliaram medidas do TT-A e TT-B em idosos encontrando CCI excelente, porém nesse estudo os autores consideraram idosos pessoas acima de 50 anos. Outro estudo aplicando a versão chinesa do TT abordou a reprodutibilidade de teste em pessoas idosas e demonstrou um bom CCI no TT-A e excelente no TT-B usando um intervalo de avaliação semelhante ao do presente estudo, de três a sete dias [25]. Uma possível razão para as diferenças é a diversidade de escolaridade na amostra entre os estudos, uma vez que não exigimos uma escolaridade mínima. Outro ponto importante do nosso estudo é a padronização no intervalo entre aplicação. Ainda vale destacar que mantivemos a realização desse teste com caneta e papel visto que a literatura preconiza a aplicação dessa forma [38,39].

Apesar de os testes usados sozinhos serem reportados na literatura como indicadores gerais da FE, cada um desses testes avalia isoladamente um domínio. Um ponto forte do nosso estudo foi uma abordagem integrada, utilizando o SCWT para avaliar o controle inibitório, o TCC para a memória de trabalho e o TT para verificar a flexibilidade cognitiva, dessa forma favorecendo a interpretação sobre o estado global da FE [11]. Por sua vez, adotamos a aplicação do SCWT e do TCC de modo computadorizado com base em protocolos e softwares de acesso gratuito o que facilita a reprodução do método empregado na prática clínica e na pesquisa científica. Além de inovar ao trazer a reprodutibilidade de testes neurocognitivos em formato computadorizado em pessoas idosas, esse se apresenta relativamente escasso na literatura [20]. Assim, nossos achados fornecem importantes insights para uma avaliação compreensiva e o acompanhamento da FE em mulheres idosas.

Dentre as limitações do presente estudo, podemos apontar a possibilidade do efeito aprendizagem, uma vez que foram realizadas apenas duas medidas para teste e reteste. Entretanto, acreditamos que o intervalo de sete dias entre as medidas

é um intervalo suficiente para minimizar esse efeito. Além disso, na perspectiva de minimizar o efeito aprendizagem, os testes SCWT e TCC foram planejados para que houvesse uma aleatorização das sequências de palavras e blocos apresentados entre os dias de teste e reteste.

Outra limitação é o tamanho amostral que é pequeno o que pode aumentar a chance de erro tipo I ou II, apesar de termos atendido o nosso cálculo amostral. Nesse sentido, não há um consenso na literatura sobre a melhor forma de realizar o cálculo amostral em estudos de reprodutibilidade. Além disso, a maior parte dos estudos utilizou dois grupos e nós utilizamos somente um. Desse modo, é possível que haja diferenças frente a outros grupos. De qualquer forma, considerando a normalidade dos dados acreditamos que os resultados observados no presente estudo contribuem com a literatura a respeito de testes para a FE em idosas, uma vez que fornecemos informações detalhadas sobre as características das tarefas, instruções, estímulos e métodos de pontuação, apresentando-se como importante diferencial para outros estudos da área [40]. Além disso, fornecemos valores de pontuação que podem ser considerados em outros estudos científicos e na prática clínica.

Conclusão

A avaliação do TR congruente e incongruente no SCWT para controle inibitório, da sequência e escore composto do TCC para memória de trabalho visuoespacial e do TT-A, TT-B e a razão (B/A) no TT para flexibilidade cognitiva são métodos reprodutíveis para avaliação da FE de mulheres idosas. Além disso, a realização dos testes de forma sequencial e com um intervalo de uma semana se mostra uma abordagem efetiva para garantir a reprodutibilidade dessas avaliações.

Vínculo acadêmico

Este artigo representa parte da dissertação de mestrado de Alan Pantoja Cardoso, orientada pelo professor Dr. Marzo Edir Da Silva Grigoletto da Universidade Federal de Sergipe, Brasil.

Conflito de Interesse

Não há conflito de interesse

Fonte de Financiamento

Parte deste estudo é financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)

Contribuições dos autores

Concepção e desenho da investigação: Pantoja-Cardoso A, Faro HKC; **Coleta de dados:** Pantoja-Cardoso A, Dos-Santos AC, Santos PJ; **Análise e interpretação de dados:** Pantoja-Cardoso A, Aragão-Santos JS; **Redação do manuscrito:** Pantoja-Cardoso A, Aragão-Santos JS, Monteiro MCP, Santos PJ, Heredia-Elvar JR, Dos-Santos AC; **Revisão crítica do manuscrito quanto conteúdo intelectual importante:** Fortes LS, Silva-Grigoletto ME.

Referências

1. Diamond A. Executive Functions. *Annu Rev Psychol.* 2013;64:135-168. doi: 10.1146/annurev-psy-113011-143750
2. Diamond A. Why improving and assessing EF early in life is critical. In: Griffin JA, McCardle P, Freund LS, eds. *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research.* APA. 2016;11-43. doi: 10.1037/14797-002
3. Freund LS, eds. *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research.* APA. 2016;11-43. doi: 10.1037/14797-002
4. Harada CN, Love MCN, Triebel KL. Normal cognitive aging. *Clin Geriatr Med.* 2013;29(4):737-52. doi: 10.1016/j.cger.2013.07.002
5. Ho H-T, Lin S-I, Guo N-W, Yang YC, Lin MH, Wang CS. Executive function predict the quality of life and negative emotion in older adults with diabetes: A longitudinal study. *Prim Care Diabetes.* 2022;16:537-42. doi: 10.1016/j.pcd.2022.05.002
6. Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn Sci.* 2014;18(4):177-85. doi: 10.1016/j.tics.2013.12.003
7. Jonides J, Lewis RL, Nee DE, Lustig CA, Berman MG, Moore KS. The mind and brain of short-term memory. *Annu Rev Psychol.* 2008;59:193-224. doi: 10.1146/annurev.psych.59.103006.093615
8. Donders FC. On the speed of mental processes. *Acta Psychol (Amst).* 1969;30:412-431. doi:10.1016/0001-6918(69)90065-1
9. Servant M, Logan GD. Dynamics of attentional focusing in the Eriksen flanker task. *Atten Percept Psychophys.* 2019;81(8):2710-2721. doi: 10.3758/s13414-019-01796-3
10. Stroop JR. Studies of interference in serial verbal reactions. *J. Exp. Psychol. Gen.* 1935;18:643-62. doi: 10.1037/h0054651
11. Verbruggen F, Logan GD. Response inhibition in the stop-signal paradigm. *Trends Cogn Sci.* 2008;12:418-24. doi: 10.1016/j.tics.2008.07.005
12. Scarpina F, Tagini S. The stroop color and word test. *Front Psychol.* 2017;8:1-8. doi: 10.3389/fpsyg.2017.00557
13. Strauss GP, Allen DN, Jorgensen ML, Cramer SL. Test-retest reliability of standard and emotional stroop tasks: an investigation of color-word and picture-word versions. *Assessment.* 2005;12(3):330-7. doi: 10.1177/1073191105276375
14. Periañez JA, Lubrini G, García-Gutiérrez A, Ríos-Lago M. Construct validity of the stroop color-word test: influence of speed of visual search, verbal fluency, working memory, cognitive flexibility, and conflict monitoring. *Arch Clin Neuropsychol.* 2021;36(1):99-111. doi: 10.1093/arclin/acia034
15. Ward N, Hussey E, Alzahabi R, Gaspar JG, Kramer AF. Age-related effects on a novel dual-task Stroop paradigm. *PLoS ONE* 2021;16:e0247923. doi: 10.1371/journal.pone.0247923
16. De Nardi T, Sanvicente-Vieira B, Prando M, Stein LM, Fonseca RP, Grassi-Oliveira R. Tarefa N-back Auditiva: Desempenho entre diferentes grupos etários. *Psicol Reflex Crit.* 2013;26(1): 151-159. doi: 10.1590/S0102-79722013000100016
17. Gonçalves VT, Mansur LL. N-Back auditory test performance in normal individuals. *Dement Neuropsychol.* 2009;3(2): 114-117. doi: 10.1590/S1980-57642009DN30200008
18. Nyberg L, Dahlin E, Stigsdotter Neely A, et al. Neural correlates of variable working memory load across adult age and skill: Dissociative patterns within the fronto-parietal network. *Scand J Psychol.* 2009;50(1):41-6. doi: 10.1111/j.1467-9450.2008.00678.x
19. Arce T, McMullen K. The Corsi Block-Tapping Test: Evaluating methodological practices with an eye towards modern digital frameworks. *Comput. Hum. Behav. Reports [Internet].* 2021 [citado 2022 dez 12];4:100099. Disponível em: <https://par.nsf.gov/servlets/purl/10292489>
20. Corsi PM. Human memory and the medial temporal region of the brain [Internet]. *Dissertation Abstracts International* 1973;34(2-B): 891. Disponível em: <https://psycnet.apa.org/record/1976-04900-001>
21. White N, Flannery L, McClintock A, Machado L. Repeated computerized cognitive testing: Performance shifts and test-retest reliability in healthy older adults. *J Clin Exp Neuropsychol* 2019;41:179-191. doi: 10.1080/13803395.2018.1526888
22. Milner B. Some cognitive effects of frontal-lobe lesions in man. *Philos Trans R Soc Lond B Biol Sci.* 1982 Jun 25;298(1089):211-26. doi: 10.1098/rstb.1982.0083
23. Miguel FK. Teste Wisconsin de Classificação de Cartas. *Aval. Psicol.* 2005;4:203-4.

24. Suzuki H, Sakuma N, Kobayashi M, Ogawa S, Inagaki H, Eda Hiro A, et al. Normative data of the Trail Making Test among urban community-dwelling older adults in Japan. *Front Aging Neurosci.* 2022;14:832158. doi: 10.3389/fnagi.2022.832158
25. Wang R-Y, Zhou J-H, Huang Y-C, Yang Y-R. Reliability of the Chinese version of the Trail Making Test and Stroop Color and Word Test among older adults. *Int. J. Gerontol.* 2018;12:336-9. doi: 10.1016/j.ijge.2018.06.003
26. Cesar KG, Yassuda MS, Porto FHG, Brucki SMD, Nitrini R. MoCA Test: normative and diagnostic accuracy data for seniors with heterogeneous educational levels in Brazil. *Arq Neuro-Psiquiatr.* 2019;77(11): 775-781. doi: 10.1590/0004-282X20190130
27. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53(4):695-9. doi: 10.1111/j.1532-5415.2005.53221.x
28. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. *Nat Protoc.* 2006;1(5):2277-81. doi: 10.1038/nprot.2006.390
29. Faul F, Erdfelder E, Buchner A, et al. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods.* 2009;41:1149-60. doi: 10.3758/BRM.41.4.1149
30. Shieh G. Sample size requirements for the design of reliability studies: precision consideration. *Behav Res Methods.* 2014;46(3):808-22. doi: 10.3758/s13428-013-0415-1
31. Fontes AS, Santos MS, Almeida MB, Marín PJ, Silva DRP, Silva-Grigoletto MES. Inter-day reliability of the Upper Body Test for shoulder and pelvic girdle stability in adults. *Braz J Phys Ther.* 2020;24:161-6. doi: 10.1016/j.bjpt.2019.02.009
32. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155-63. doi: 10.1016/j.jcm.2016.02.012
33. Martin Bland J, Altman Douglas G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;327:307-10. doi: 10.1016/S0140-6736(86)90837-8
34. Harvey PD. Domains of cognition and their assessment. *Dialogues Clin Neurosci.* 2019;21:227-37. doi: 10.31887/DCNS.2019.21.3/pharvey
35. Collie A, Maruff P, Darby DG, McStephen M. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *J Int Neuropsychol Soc.* 2003;9(3):419-28. doi: 10.1017/S1355617703930074
36. Hoogendam YY, van der Lijn F, Vernooij MW, Hofman A, Niessen WJ, van der Lugt A, et al. Older age relates to worsening of fine motor skills: a population-based study of middle-aged and elderly persons. *Front Aging Neurosci.* 2014;6:259. doi: 10.3389/fnagi.2014.00259
37. Park S-Y, Schott N. The trail-making-test: Comparison between paper-and-pencil and computerized versions in young and healthy older adults. *Appl Neuropsychol Adult.* 2022;29(5):1208-20. doi: 10.1080/23279095.2020.1864374
38. Miller JB, Barr WB. The technology crisis in neuropsychology. *Arch Clin Neuropsychol.* 2017;32:541-54. doi: 10.1093/arclin/acx050
39. Rumpf U, Menze I, Müller NG, Schmicker M. Investigating the potential role of ecological validity on change-detection memory tasks and distractor processing in younger and older adults. *Front Psychol.* 2019;10:1046. doi: 10.3389/fpsyg.2019.01046
40. Zanini GAV, Miranda MC, Cogo-Moreira H, et al. An Adaptable, open-access test battery to study the fractionation of executive-functions in diverse populations. *Front Psychol.* 2021;12:627219. doi: 10.3389/fpsyg.2021.627219

