



UNIVERSIDADE FEDERAL DE SANTA CATARINA

Centro de Filosofia e Ciências Humanas

Programa de Pós-Graduação em Psicologia

VANESSA BRANDELERO ABI-ABIB

**EVIDÊNCIAS DE VALIDADE DO TESTE STROOP DE CORES E PALAVRAS  
PARA ADULTOS**

FLORIANÓPOLIS, SC

2023

VANESSA BRANDELERO ABI-ABIB

**EVIDÊNCIAS DE VALIDADE DO TESTE STROOP DE CORES E PALAVRAS  
PARA ADULTOS**

Tese apresentada como requisito à obtenção de grau de Doutora em Psicologia, Programa de Pós-Graduação em Psicologia, Mestrado/Doutorado, Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina.

Orientador: Prof. Dr. Carlos Henrique Sancineto da Silva Nunes

Coorientador: Prof. Dr. Roberto Moraes Cruz

FLORIANÓPOLIS, SC

2023

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Abi-Abib, Vanessa Brandelero  
Evidências de validade do Teste Stroop de Cores e  
Palavras para adultos / Vanessa Brandelero Abi-Abib ;  
orientador, Carlos Henrique Sancineto da Silva Nunes,  
coorientador, Roberto Moraes Cruz, 2023.  
139 p.

Tese (doutorado) - Universidade Federal de Santa  
Catarina, Centro de Filosofia e Ciências Humanas, Programa  
de Pós-Graduação em Psicologia, Florianópolis, 2023.

Inclui referências.

1. Psicologia. 2. teste Stroop. 3. psicometria. 4.  
validade. 5. neuropsicologia. I. Nunes, Carlos Henrique  
Sancineto da Silva . II. Cruz, Roberto Moraes . III.  
Universidade Federal de Santa Catarina. Programa de Pós  
Graduação em Psicologia. IV. Título.

Vanessa Brandelero Abi-Abib

Evidências de validade do Teste Stroop de Cores e Palavras para adultos

O presente trabalho em nível de doutorado foi avaliado e aprovado, em 11 de maio de 2023  
pela banca examinadora composta pelos seguintes membros:

Prof.<sup>a</sup> Dr.<sup>a</sup> Monalisa Muniz Nascimento  
Universidade Federal de São Carlos-UFSCar

Prof. Dr. Leandro Kruszielski  
Universidade Federal do Paraná - UFPR

Prof.<sup>a</sup> Dr.<sup>a</sup> Natália Dias  
Universidade Federal de Santa Catarina – UFSC

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado  
adequado para obtenção do título de doutora em Psicologia.

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Ana Lúcia Mandelli De Marsillac  
Coordenadora do Programa de Pós-Graduação em Psicologia da UFSC

---

Prof. Dr. Carlos Henrique Sancineto da Silva Nunes  
Orientador do Programa de Pós-Graduação em Psicologia da UFSC

Florianópolis, 2023

## AGRADECIMENTOS

*“Não há no mundo exagero mais belo que a gratidão”*

*Jean de la Bruyere*

À minha amada família:

À minha querida mãe, Elda, por ser exemplo em tantos caminhos dos quais venho percorrendo. Mulher, mãe, esposa, docente, trabalhadora, que desde sempre me ensinou que nenhum ganho vem sem sacrifício e que a persistência é uma virtude.

Ao meu pai, Olivio, que me ensinou sobre a importância da humildade em todas as esferas da vida, do empenho nos estudos e da dedicação ao trabalho, valores que certamente passarei aos meus.

Ao meu irmão, Gladson, que sempre zeloso e bondoso, honrou e honra o significado semântico de fraternidade.

Ao meu marido, Mikeil, pelo incessante incentivo em minha carreira profissional e acadêmica. Que dividiu comigo a responsabilidade simultânea de ingressar no doutorado e me tornar mãe. Que assumiu, nas minhas ausências, as noites em claro, as responsabilidades e proporcionou todos os arranjos possíveis para que eu pudesse prosseguir.

À minha sogra, Mara, por toda a colaboração oferecida ao longo desses anos. Por ser presente e se mostrar sempre disposta a oferecer ajuda a todo tempo.

Ao meu amado filho, Vicente, que esteve comigo desde as etapas iniciais de seleção do doutorado. Que dentro do meu ventre acompanhou as aulas, as viagens e que, depois do nascimento, aprendeu cedo a lidar com o sentimento de ausência, mas, ilustremente - e com tão pouca idade - mostrou para mim que era possível continuar.

À minha amada filha, Isabel, que acompanhou os momentos finais desta trajetória dentro de mim. Que foi, sem dúvidas, força motriz para a etapa de conclusão e a quem eu

dividirei orgulhosamente esta experiência de ser mãe, mulher, docente e pesquisadora (tal qual aprendi, igualmente, com a minha mãe).

Aos mestres:

Agradeço ao meu professor e orientador Dr. Carlos Henrique Sancineto da Silva Nunes, pelo papel fundamental exercido durante estes anos. Exímio docente, seus apontamentos e orientações foram basilares em todas as etapas da construção deste trabalho. Sempre disponível para ensinar, ofereceu-me suporte e transmitiu seu conhecimento com maestria, gerando, para mim, uma fonte de inspiração e um exemplo a seguir.

Ao meu coorientador, prof. Dr. Roberto Moraes Cruz, por todos os ensinamentos compartilhados, tanto no contexto acadêmico, quanto profissional. Externo aqui minha profunda consideração e respeito por todo o trabalho que o senhor exerce para o avanço da Psicologia nas mais diversas áreas de atuação.

Ao professor Dr. Plínio Marco de Toni, a quem serei eternamente grata. Meu orientador de mestrado, me abriu portas no campo profissional e sempre me incentivou a continuar avançando na carreira acadêmica.

Aos membros da banca de qualificação e de defesa, por aceitarem o convite e contribuírem com tamanha sabedoria no aprimoramento deste trabalho.

E, por fim, agradeço a Deus, por conceder infinitas bençãos a mim e à minha família. Por me presentear com oportunidades únicas e por inserir as pessoas certas no meu caminho. Pela proteção no dia a dia e por estar ao meu lado em todos os momentos da minha vida.

Vanessa Brandelero Abi-Abib. Evidências de validade do Teste Stroop de Cores e Palavras para adultos. Florianópolis, 2023. Tese de Doutorado em Psicologia – Programa de Pós-Graduação em Psicologia. Universidade Federal de Santa Catarina.

Orientador: Prof. Dr. Carlos Henrique Sancineto da Silva Nunes

Coorientador: Prof. Dr. Roberto Moraes Cruz

Data da defesa: 11 de maio de 2023

## RESUMO

O Teste Stroop de Cores e Palavras é um teste neuropsicológico desenvolvido em 1935 por John Ridley Stroop, que compreende a apresentação conflitante, ao sujeito, de dois estímulos simultâneos e concorrentes: a leitura de palavras e a nomeação de cores. Classicamente, o teste é realizado em três etapas. A primeira consiste na leitura do nome das cores, impressas em tinta preta. A segunda requer a nomeação de cores, representadas por manchas de tinta. Na terceira, o nome das cores é impresso em uma tinta não correspondente (ex: a palavra “vermelho” é impressa em tinta verde). Deste modo, nesta situação considerada incongruente, os participantes devem nomear a cor de tinta em vez de ler a palavra, executando uma tarefa menos automatizada - nomeação das cores - enquanto inibem a interferência resultante do processo mais automatizado - leitura da palavra. Tradicionalmente, o Teste Stroop é utilizado na avaliação das funções executivas (FE) e também de alguns de seus componentes, como atenção seletiva, inibição, flexibilidade, memória de trabalho e velocidade de processamento. Entretanto, até o momento, há na literatura uma escassez de referências de evidências de validade das inferências propostas a partir do teste para a população adulta brasileira. Por este motivo, o presente projeto teve por objetivo realizar a busca de evidências de validade do Teste Stroop de Cores e Palavras para adultos. Para isso, inicialmente, foi realizada uma revisão da literatura com o intuito de apresentar o Teste Stroop e identificar as teorias que explicam o efeito obtido pelo teste, bem como os estudos empíricos que o utilizam como meio de avaliação. Posteriormente, foi realizada uma revisão integrativa da literatura com o objetivo de levantar quais construtos psicológicos são avaliados pelo Stroop, além das versões empregadas e o tipo de participantes/amostras avaliados por ele. Na sequência, foram desenvolvidos estudos de evidências de validade convergente, divergente e por construtos relacionados entre o Stroop e os testes FDT - Teste dos Cinco Dígitos, Wisconsin Teste de Classificação de Cartas – WCST, Teste de Avaliação da Atenção Seletiva – EASV, BETA-III – Subteste Códigos e Hayling

Sentence Completion Test (HSCT). Dentre os resultados, foram identificados altos valores de coeficientes de correlação entre indicadores do Stroop e as medidas de velocidade de processamento, caracterizando evidências de validade convergente. Por construtos relacionados, aponta-se a relação moderada entre o Stroop e a medida de inibição do Hayling e, em menor magnitude, com valor limítrofe, a relação entre variáveis do Stroop e escores de flexibilidade e inibição do Wisconsin. Com fracos coeficientes de correlação, têm-se as comparações entre o Stroop e o EASV. Ademais, os dados sociodemográficos dos respondentes foram analisados e indicaram correlações estatisticamente significativas entre variáveis do Stroop e a idade dos participantes, sinalizando melhor desempenho entre os mais jovens em relação ao escore de interferência e o número de erros cometidos. Como conclusão, aponta-se a necessidade de estudos adicionais para verificação de hipóteses residuais levantadas a partir dos resultados obtidos, que dizem respeito à sensibilidade do teste para diferentes faixas etárias e à utilização de outros modelos estatísticos que possam dirimir os problemas das impurezas das tarefas neuropsicológicas.

Palavras-chave: teste Stroop; psicometria; validade; neuropsicologia.

### **ABSTRACT**

The Stroop Color-Word Test is a neuropsychological test developed in 1935 by John Ridley Stroop. It involves the conflicting presentation of two simultaneous and competing stimuli to the subject: reading words and naming colors. Classically, the test is conducted in three stages. The first stage involves reading the names of colors printed in black ink. The second stage requires naming colors represented by ink blots. In the third stage, color names are printed in non-matching ink colors (e.g., the word "red" is printed in green ink). In this incongruent situation, participants must name the ink color instead of reading the word, engaging in a less automated task (color naming) while inhibiting interference from the more automated process (word reading). Traditionally, the Stroop Test is used to assess executive functions (EF) and some of its components, such as selective attention, inhibition, flexibility, working memory, and processing speed. However, up to this point, there's a scarcity of references to evidence of validity for inferences drawn from the test within the Brazilian adult population. For this reason, the present project aimed to search for evidence of validity for the Stroop Color-Word Test in adults. To achieve this, a literature review was initially conducted to introduce the Stroop Test,

identify theories explaining the obtained effects, and explore empirical studies that utilize it as an assessment tool. Subsequently, an integrative literature review was performed to identify the psychological constructs evaluated by the Stroop Test, its employed versions, and the types of participants/samples it has been applied to. Following this, studies were developed to provide evidence of convergent, divergent, and construct-related validity between the Stroop Test and other tests, including the Five-Digit Test (FDT), Wisconsin Card Sorting Test (WCST), Test of Selective Attention Assessment (EASV), BETA-III – Codes Subtest, and Hayling Sentence Completion Test (HSCT). Among the results, high correlation coefficients were identified between Stroop indicators and measures of processing speed, indicating evidence of convergent validity. Regarding construct-related validity, a moderate relationship was found between the Stroop Test and Hayling's inhibition measure. To a lesser extent, with borderline significance, a relationship between Stroop variables and flexibility and inhibition scores from the Wisconsin test was observed. Weak correlation coefficients were observed in comparisons between the Stroop Test and EASV. Furthermore, demographic data of respondents were analyzed, revealing statistically significant correlations between Stroop variables and participants' age, indicating better performance among younger individuals in relation to interference scores and the number of errors committed. In conclusion, there is a need for additional studies to verify residual hypotheses arising from the obtained results. These hypotheses concern the test's sensitivity across different age groups and the use of alternative statistical models to address the challenges posed by impurities in neuropsychological tasks.

Keywords: Stroop test; psychometry; evidence of validity; neuropsychology.

## LISTA DE FIGURAS

Figura 1: Exemplificação do Teste Stroop .....	22
Figura 2: Modelo de unidade/diversidade de Miyake e Friedman (2012): fatores comuns e fatores específicos .....	33
Figura 3: Fluxograma da estratégia de pesquisa.....	38
Figura 4: Condições “color-stroop” e “counting-stroop” .....	44
Figura 5: Evidências baseadas nas relações com variáveis externas.....	54
Figura 6: Distribuição da amostra em relação a idade dos participantes.....	59
Figura 7: Histograma de normalidade para o teste Stroop .....	66
Figura 8: Histograma e gráfico de normalidade para variável escore de interferência logaritimizada .....	67
Figura 9: Histogramas de distribuição do teste EASV .....	68
Figura 10: Histogramas de distribuição dos testes FDT, Hayling, BETA – III – Códigos, e Wisconsin .....	69

## LISTA DE TABELAS

Tabela 1: Construtos agrupados e suas respectivas frequências.....	40
Tabela 2: Participantes com maiores frequências.....	45
Tabela 3: Distribuição da amostra em relação à escolaridade.....	60
Tabela 4: Teste de normalidade para o Teste Stroop de Cores e Palavras.....	66
Tabela 5: Teste de normalidade após logaritimização da variável escore de interferência do teste Stroop.....	67
Tabela 6: Teste de normalidade para demais testes.....	68
Tabela 7: Correlação entre idade e escore de interferência no teste Stroop.....	70
Tabela 8: Correlações entre tempo de execução e a idade dos participantes.....	70
Tabela 9: Correlações entre número de erros e a idade dos participantes.....	71
Tabela 10: Estatísticas descritivas da amostra por faixas etárias.....	72
Tabela 11: Anova para variável Escore de Interferência em relação às faixas etárias.....	72
Tabela 12: Teste de Homogeneidade de variâncias de Levene para o escore de interferência....	73
Tabela 13: Comparações múltiplas – Teste de Post-Hoc de Bonferroni.....	72
Tabela 14: Estatísticas descritivas em relação ao gênero.....	74
Tabela 15: Teste t de amostras independentes: diferença entre gêneros e escore de interferência do Stroop.....	74
Tabela 16: Estatísticas descritivas da amostra por escolaridade do escore de interferência do teste Stroop.....	75
Tabela 17: Anova para variável Escore de Interferência em relação às escolaridades.....	75
Tabela 18: Teste de Homogeneidade de variâncias de Levene para o escore de interferência pela escolaridade.....	75
Tabela 19: Comparações múltiplas – Teste de Post-Hoc de Bonferroni para escolaridade.....	74
Tabela 20: Coeficientes de correlação de Spearman entre os tempos de execução do Teste Stroop e do Teste FDT.....	757
Tabela 21: Coeficientes de correlação entre o Teste Stroop e o Teste FDT: amostra total e grupo com 49 anos ou mais.....	77
Tabela 22: Coeficientes de correlação de Spearman entre o Teste Stroop e o Teste BETA III – Subteste Códigos: amostra total.....	80
Tabela 23: Coeficientes de correlação de Spearman entre as variáveis de tempo do Teste Stroop e a Escala de Atenção Seletiva Visual – EASV (pontuação total).....	81

Tabela 24: Coeficientes de correlação de Spearman entre o escore de interferência do Teste Stroop e a Escala de Atenção Seletiva Visual – EASV (pontuação básica).....	81
Tabela 25: Coeficientes de correlação de Spearman entre o variáveis de acertos e erros do Teste Stroop e a Escala de Atenção Seletiva Visual – EASV (pontuação total).....	81
Tabela 26: Coeficientes de correlação de Spearman entre os tempos de execução do Teste Stroop e do Teste Hayling para amostra total.....	83
Tabela 27: Coeficientes de correlação de Spearman entre os tempos de execução do Teste Stroop e do Teste Hayling: amostra de pessoas com 49 anos de idade ou mais.....	84
Tabela 28: Coeficientes de correlação de Spearman entre o Teste Stroop e o Teste Wisconsin.....	86
Tabela 29: Coeficientes de correlação de Spearman entre as variáveis de acertos e erros do Teste Stroop e do Teste Wisconsin.....	87
Tabela 30: Síntese dos maiores coeficientes de correlação (estatisticamente significativos)...	88

## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO</b> .....	<b>14</b>
<b>2.</b>	<b>OBJETIVOS</b> .....	<b>17</b>
2.1	OBJETIVO GERAL .....	17
2.2	OBJETIVOS ESPECÍFICOS .....	17
2.3	APRESENTAÇÃO ESTRUTURAL DA TESE .....	18
<b>3.</b>	<b>CAPÍTULO I</b> .....	<b>19</b>
3.1	TESTE STROOP DE CORES E PALAVRAS .....	19
3.2	AVALIAÇÃO NEUROPSICOLÓGICA E FUNÇÕES EXECUTIVAS.....	27
<b>4.</b>	<b>CAPÍTULO II: TESTE STROOP DE CORES E PALAVRAS: UMA REVISÃO INTEGRATIVA DE LITERATURA</b> .....	<b>37</b>
4.1	INTRODUÇÃO .....	37
4.2	MÉTODO .....	37
4.3	RESULTADOS .....	39
4.4	DISCUSSÃO .....	47
4.5	CONCLUSÃO .....	50
<b>5.</b>	<b>CAPÍTULO III: ESTUDOS EMPÍRICOS DE EVIDÊNCIAS DE VALIDADE DO TESTE STROOP DE CORES E PALAVRAS</b> .....	<b>52</b>
5.1	MÉTODO .....	56
<b>5.1.1</b>	<b>Caracterização da pesquisa</b> .....	<b>57</b>
<b>5.1.2</b>	<b>Caracterização do campo de pesquisa</b> .....	<b>58</b>
<b>5.1.3</b>	<b>Procedimentos éticos</b> .....	<b>58</b>
<b>5.1.4</b>	<b>Procedimentos de coleta de dados</b> .....	<b>58</b>
<b>5.1.5</b>	<b>Participantes</b> .....	<b>59</b>
<b>5.1.6</b>	<b>Instrumentos</b> .....	<b>60</b>
<b>5.1.7</b>	<b>Análise de dados</b> .....	<b>63</b>
<b>5.1.8</b>	<b>Hipóteses de estudo</b> .....	<b>64</b>
5.2	RESULTADOS .....	65
<b>5.2.1</b>	<b>Normalidade da amostra</b> .....	<b>65</b>
<b>5.2.2</b>	<b>Relações entre as variáveis do Teste Stroop de Cores e Palavras e dados sociodemográficos</b> .....	<b>69</b>
5.2.2.1	Estudos correlacionais entre a variável idade e os escores no teste Stroop.....	70

5.2.2.2	Comparação entre as faixas etárias e os escores no teste Stroop.....	71
5.2.2.3	Estudos comparativos entre a variável gênero e os escores no teste Stroop.....	73
5.2.2.4	Estudos de comparação entre a variável escolaridade e os escores no teste Stroop.....	73
<b>5.2.3</b>	<b>Estudos correlacionais de evidências de validade do Teste Stroop de Cores e Palavras com os demais testes e tarefas cognitivas .....</b>	<b>75</b>
5.2.3.1	Estudos de correlação entre o teste FDT – Teste dos Cinco Dígitos e o Teste Stroop de Cores e Palavras.....	75
5.2.3.2	Estudos de correlação entre o Teste Não Verbal de Inteligência Geral - BETA-III – Subteste Códigos e o Teste Stroop de Cores e Palavras .....	78
5.2.3.3	Estudos de correlação entre a Escala de Atenção Seletiva Visual – EASV e o Teste Stroop de Cores e Palavras .....	79
5.2.3.4	Estudos de correlação entre o Hayling Sentence Completion Test (HSCT) e o Teste Stroop de Cores e Palavras .....	82
5.2.3.5	Estudos de correlação entre o Wisconsin Teste de Classificação de Cartas - WCST e o Teste Stroop de Cores e Palavras.....	84
5.3	DISCUSSÃO .....	87
<b>5.3.1</b>	<b>Análises e discussões dos estudos entre o Teste Stroop de Cores e Palavras e as variáveis sociodemográficas.....</b>	<b>88</b>
<b>5.3.2</b>	<b>Análises e discussões dos estudos entre o Teste Stroop de Cores e Palavras e os demais testes e tarefas cognitivas .....</b>	<b>92</b>
5.3.2.1	Estudos correlacionais entre o teste FDT – Teste dos Cinco Dígitos e o Teste Stroop de Cores e Palavras.....	92
5.3.2.2	Estudos correlacionais entre o Teste não verbal de inteligência geral - BETA-III – Subteste Códigos e o Teste Stroop de Cores e Palavras .....	95
5.3.2.3	Estudos correlacionais entre a Escala de Atenção Seletiva Visual – EASV e o Teste Stroop de Cores e Palavras .....	97
5.3.2.4	Estudos correlacionais entre o teste Hayling e o Teste Stroop de Cores e Palavras ....	99
5.3.2.5	Estudos correlacionais entre o Teste Wisconsin de Classificação de Cartas Modificado (WCSTM-48) e o Teste Stroop de Cores e Palavras .....	100
5.3.2.6	Síntese dos coeficientes de correlação entre as tarefas neuropsicológicas e o teste Stroop.....	102
5.3.2.7	Testes neuropsicológicos e o problema da impureza das tarefas.....	105
5.3.2.8	Hipóteses a respeito das análises correlacionais entre o grupo de pessoas com idade igual ou superior a 49 anos .....	106

5.4	CONCLUSÃO.....	108
	REFERÊNCIAS.....	111
	APÊNDICES.....	127
	APÊNDICE 1.....	128
	APÊNDICE 2.....	129
	APÊNDICE 3.....	130
	APÊNDICE 4.....	131
	ANEXOS.....	134
	ANEXO 1.....	135

## 1. INTRODUÇÃO

Desenvolvido em 1935 por John Ridley Stroop o Teste Stroop de Cores e Palavras é um teste neuropsicológico amplamente utilizado no cenário nacional e internacional que compreende a apresentação simultânea e conflitante, ao sujeito, de dois estímulos concorrentes: a leitura de palavras e a nomeação de cores. Ao respondente, solicita-se o processo de inibição da resposta automática (leitura da palavra) em favor da não automática (nomeação da cor). Deste modo, a execução sincrônica destas demandas conflitantes possui o intuito de gerar interferências e estímulos distratores, avaliando a habilidade de inibição de respostas involuntárias (Spreeen, Sherman & Strauss, 2006).

A versão do teste criada por Stroop compreende três etapas. As duas primeiras correspondem a condições congruentes e a terceira, a uma condição incongruente. A primeira tarefa consiste na leitura do nome das cores, impressas em tinta preta. A segunda, na nomeação das cores, representadas por manchas de tinta. Na terceira, o nome das cores é impresso em uma tinta não correspondente (ex: a palavra "vermelho" é impressa em tinta verde). Nesta situação considerada incongruente, os participantes devem nomear a cor de tinta em vez de ler a palavra, executando uma tarefa menos automatizada (nomeação das cores) enquanto inibem a interferência resultante do processo mais automatizado (leitura da palavra) (MacLeod & Dunbar, 1988). Esta condição conflitante foi denominada “Efeito Stroop”.

Tradicionalmente, o Teste Stroop é utilizado na avaliação das funções executivas (FE) e também de seus componentes, como atenção seletiva, inibição (Sweet, 1999; Gindri, Zibetti & Fonseca, 2008), flexibilidade (Spreeen & Strauss, 1998), memória de trabalho (Nowakowska-Domagaa, Jabkowska-Górecka, Mokros, Koprowicz & Pietras, 2017) e velocidade de processamento (Kim, Lee, Jo & Lee, 2017). Já considerado anteriormente como um instrumento de avaliação “padrão ouro” em neuropsicologia (MacLeod, 1991), é referenciado pela sua alta confiabilidade em identificar diferenças individuais nos testados. Entretanto, até o momento, há uma escassez, na literatura, de referências de evidências de validade publicadas para a população adulta brasileira, fato que justifica a necessidade de condução de estudos em direção a esta finalidade.

Tratando-se da área da neuropsicologia, mesmo com um cenário de desenvolvimento progressivo, ainda se considera evidente a carência de instrumentos padronizados - para a população brasileira - que avaliem as funções executivas e que sejam sensíveis a alterações

neuropsicológicas leves, ou ainda, que indiquem déficits cognitivos de modo mais específico (Pawlowski, Trentini & Bandeira 2007). Como resultado deste crescente interesse pela área de avaliação neuropsicológica no Brasil, julga-se necessário incorporar instrumentos eficientes de avaliação. Logo, torna-se primordial que pesquisas sobre validação, adaptação, padronização e normatização de testes, além de pesquisas sobre construção de instrumentos, sejam realizadas, a fim de amparar a comunidade de profissionais que utilizam os testes como ferramentas de seu trabalho (Pawlowski, Trentini & Bandeira 2007; Casarin 2010).

Neste sentido, é emergente a utilização da psicometria para subsidiar ferramentas, técnicas ou meios para a construção de novas evidências na utilização dos testes no campo da neuropsicologia. Ao reconhecerem a necessidade de aprimoramento na qualidade dos testes utilizados, Reppold et al. (2015) apontam certo avanço na construção, adaptação, análise de dados e na investigação da validade dos escores derivados dos testes neuropsicológicos, contudo, ponderam que o uso da psicometria é ainda exíguo, considerando o potencial que dela pode ser extraído para o aperfeiçoamento dos instrumentos de avaliação.

Ramos e Hamdan (2016) indicam que pesquisas no campo da neuropsicologia relacionadas ao desenvolvimento de novos testes, ou mesmo de adaptações de instrumentos tradicionais, não figuraram como um tema de interesse pelos pesquisadores, representando menos de 15% das publicações em neuropsicologia no Brasil. Isso é contraditório frente ao crescente interesse pelo uso e desenvolvimento de testes neuropsicológicos no país. Ainda, os autores apontam a importância de que a avaliação neuropsicológica seja realizada por intermédio de testes neuropsicológicos que possuam boas qualidades psicométricas. Este critério de qualidade, por sua vez, permite fazer inferências sobre os escores obtidos pelo paciente/testando de modo preciso e com rigor metodológico, características consideradas essenciais na prática clínica e também na pesquisa. Baseados nessas premissas, entende-se como procedente e relevante esta pesquisa, que tem, como objetivo central, realizar estudos de evidências de validade do teste Stroop.

Para alcançar o objetivo supramencionado, faz-se necessário, a priori, identificar os construtos teóricos frequentemente indicados na avaliação do teste. Em decorrência da atual diversidade de indicações para sua utilização prática, tornou-se necessário realizar uma revisão integrativa de literatura com o objetivo de levantar quais construtos são citados como alvo em sua avaliação, bem como as versões utilizadas e os contextos nos quais o teste foi empregado.

A partir destas verificações, os demais propósitos se destinam aos estudos de evidências de validade pertinentes ao projeto.

O conceito de validade é considerado um tópico central na teoria psicométrica e trata-se de um ponto medular neste trabalho. É definido como o conjunto de evidências capazes de justificar as inferências feitas com base nas pontuações do teste para seus usos propostos (American Educational Research Association AERA, American Psychological Association APA, National Council on Measurement in Education NCME, 1999). Neste projeto, serão realizados procedimentos com a finalidade de buscar evidências de validade pela relação com variáveis externas, também denominadas evidências baseadas nas relações com outras variáveis. Estudos com esta pretensão buscam identificar relações entre os escores dos testes com variáveis externas a ele (American Educational Research Association AERA, American Psychological Association APA, National Council on Measurement in Education NCME, 2014). Nesta pesquisa, a metodologia envolverá a averiguação de evidências de validade convergente, divergente e de construtos relacionados.

## 2. OBJETIVOS

### 2.1 OBJETIVO GERAL

Realizar a busca de evidências de validade do Teste Stroop de Cores e Palavras, em versão impressa, para adultos.

### 2.2 OBJETIVOS ESPECÍFICOS

- a. Apresentar, por meio de uma revisão de literatura, o Teste Stroop de Cores e Palavras.
- b. Apresentar, por meio de uma revisão de literatura, estudos empíricos que utilizam o Teste Stroop como meio de avaliação.
- c. Levantar, por meio de revisão integrativa de literatura, quais construtos psicológicos são avaliados pelo Teste Stroop de Cores e Palavras;
- d. Analisar as relações entre as variáveis do Teste Stroop de Cores e Palavras com dados sociodemográficos;
- e. Desenvolver estudos de evidências de validade convergente do instrumento;
- f. Desenvolver estudos de evidências de validade discriminante do instrumento;
- g. Desenvolver estudos de evidências de validade por construtos relacionados do instrumento;

## 2.3 APRESENTAÇÃO ESTRUTURAL DA TESE

Esta tese está organizada em três capítulos. O Capítulo I compreende, inicialmente, a apresentação do Teste Stroop de Cores e Palavras, de seu funcionamento, das principais teorias utilizadas para sua interpretação, além de estudos empíricos que utilizam o instrumento como meio de avaliação, bem como, na apresentação da versão Victoria, que será utilizada para os estudos empíricos desta pesquisa. Na sequência, é apresentado o referencial teórico adotado neste trabalho sobre o tema funções executivas, em uma seção intitulada “funções executivas: modelo psicométrico de Miyake e Friedman”. Esta tem particular importância, pois a utilização do teste Stroop, de acordo com a literatura, se respalda na presunção de que ele avalia funções executivas ou alguns de seus domínios. Deste modo, é fundamental discorrer sobre o tema para que a partir do campo teórico seja possível compreender operacionalmente os construtos indicados como alvo na utilização do teste.

O Capítulo II apresenta o artigo “Teste Stroop de Cores e Palavras: uma revisão integrativa de literatura”. Esta seção retrata o resultado de um dos objetivos específicos desta tese, que foi levantar, por meio de revisão integrativa de literatura, quais construtos psicológicos são presumidamente avaliados pelo Teste Stroop de Cores e Palavras. Além dos construtos, o capítulo aborda os contextos, tipo de participantes e versões citadas pelos artigos.

O Capítulo III, é intitulado “Estudos empíricos de evidências de validade do teste Stroop de Cores e Palavras”. Nele, são apresentados os estudos empíricos que foram realizados com o objetivo de verificar as evidências de validade do teste a partir da busca de evidências baseadas nas relações com variáveis externas, como validade convergente, divergente e construtos relacionados. Tal capítulo compreende a metodologia destes estudos, indicando a caracterização do delineamento de pesquisa, dos campos de coleta de dados, dos participantes, instrumentos, procedimentos de coleta de dados, procedimentos éticos, análise de dados, hipóteses de estudo, resultados obtidos e discussão destes de acordo com a literatura científica.

### 3. CAPÍTULO I

#### 3.1 TESTE STROOP DE CORES E PALAVRAS

O Teste Stroop de Cores e Palavras é um teste neuropsicológico desenvolvido em 1935 por John Ridley Stroop. Sua ampla utilização se destina à avaliação da capacidade de inibição da interferência cognitiva (Stroop, 1935), que ocorre quando o processamento de um dado estímulo afeta o processamento de outro estímulo, se apresentados de modo simultâneo.

A versão do teste criada por Stroop compreende três etapas. As duas primeiras correspondem a condições congruentes e a terceira, a uma condição incongruente. A primeira tarefa consiste na leitura do nome das cores, impressas em tinta preta. A segunda, na nomeação das cores, representadas por manchas de tinta. Na terceira, os nomes das cores são impressos em uma tinta não correspondente (ex: a palavra “vermelho” é impressa em tinta verde). Deste modo, nesta situação considerada incongruente, os participantes devem nomear a cor de tinta em vez de ler a palavra, executando uma tarefa menos automatizada (nomeação das cores), enquanto inibem a interferência resultante do processo mais automatizado (leitura da palavra) (MacLeod & Dunbar, 1988).

O modo conflitante da apresentação dos cartões ao sujeito - que constituem as etapas do instrumento - tem como intuito gerar interferências e estímulos distratores, avaliando a habilidade de inibir uma resposta automática em detrimento de outra menos utilizada (Spren, Sherman & Strauss, 2006). O efeito que o teste busca avaliar é considerado um dos fenômenos cognitivos mais robustos disponíveis para avaliação neuropsicológica (McLeod, 1991) e foi denominado “Efeito Stroop”.

Extensivamente utilizado tanto para propósitos clínicos quanto experimentais, tem sido empregado para avaliar principalmente o controle inibitório e processos de atenção (Haldane & Frangou, 2005; Mesquita, Coutinho & Matos, 2009; Malloy-Diniz, Paula, Loschiavo-Alvares, Fuentes & Leite, 2010). Entretanto, atualmente, versões do Teste Stroop são utilizadas para avaliar vários aspectos das funções executivas, incluindo, entre outros, construtos como atenção, interferência, inibição (Sweet, 1999; Gindri, Zibetti & Fonseca, 2008), flexibilidade

(Spreeen & Strauss, 1998), velocidade de processamento (Kim, Lee, Jo, & Lee, 2017) e memória de trabalho (Nowakowska-Domagaa, Jabkowska-Góreck, Mokros, Koprowicz & Pietras, 2017).

Com o intuito de explicar o paradigma do teste, a Teoria da Velocidade Relativa de Processamento considera que o padrão de interferência será definido pela relação de tempo que envolve o processamento dos estímulos, já que as respostas induzidas pelo teste (correta e errada) são processadas em paralelo. Segundo esta teoria, a leitura das palavras é mais rápida do que a nomeação das cores, e no momento conflitante em que os dois estímulos incongruentes (cor e palavra) são apresentados ao indivíduo, é necessário que este supere a automaticidade da leitura. É por este processo que o teste Stroop se propõe a avaliar a inibição, pois o indivíduo deverá ignorar a primeira resposta, considerada automática e errada, para então verbalizar a resposta correta (Kulaif, 2005).

Ainda no campo das teorias que se propõe a explicar o efeito Stroop, destaca-se a teoria da automaticidade, que postula a ideia básica de que o processamento de uma determinada dimensão requer mais atenção do que o processamento de outra dimensão. Assim, a nomeação de cores demanda um maior peso atencional do que a leitura de palavras. Esta teoria reforça que a leitura de palavras é um processo automático desenvolvido por um longo hábito de leitura (MacLeoad, 1991; Shiffrin & Schneider, 1977; Logan, 1978). Processos automáticos, segundo Macleod (1991), derivam de aprendizado e prática, e quanto mais automático um processo se tornar por meio de sua práxis, mais ele será capaz de causar interferências em processamentos menos automáticos.

Sob essa visão, a assimetria que é a característica fundamental da tarefa Stroop ocorre. As palavras são lidas muito automaticamente e as cores, por sua vez, requerem mais atenção para serem nomeadas. Por esta perspectiva, processamentos automáticos podem interferir em um processamento menos automático, entretanto, o contrário não tende a ocorrer. Esta hipótese é mencionada como uma das explicações do efeito Stroop e também das suas variações (MacLeoad, 1991; Shiffrin & Schneider, 1977; Logan, 1978).

Para a possibilidade de inibir a resposta automática em favor da não automática, é evidente a necessidade de ignorar a palavra para poder nomear a cor de tinta em que ela está impressa. Dificuldades no processo de inibição se manifestam como erros que ocorrem quando

a palavra é lida em detrimento da correta nomeação da cor de tinta em que está impressa, o que provavelmente representa a maior característica indicadora de alguma patologia na execução da tarefa Stroop (Sweet, 1999).

O teste Stroop possui diferentes versões, segundo Spreen, Sherman e Strauss (2006). De modo geral, as versões diferem em alguns aspectos, como o número e as cores usadas, o formato do estímulo utilizado para representar as cores na etapa “Cor”, a apresentação dos itens em sequências de linhas ou colunas, a metodologia para formação do escore do teste (Golden & Freshwater, 2002), ou ainda, o idioma utilizado.

A versão original do Stroop empregava cinco cores distintas. Golden e Freshwater (2002), através de um levantamento, encontraram versões que utilizavam três, quatro e, igualmente ao original, cinco diferentes cores. Com o objetivo comparar o desempenho dos testados em três versões diferentes do teste - que diferiam na quantidade de cores apresentadas - os autores não encontraram evidências de diferenças significativas entre as versões, o que pôde ser compreendido como um indicativo de que a quantidade de cores não influenciou o efeito de interferência na tarefa.

A versão Victoria - escolhida nesta pesquisa - tem apresentado algumas vantagens. Cita-se que a principal delas se refere a ser de domínio público, permitindo adaptações conforme cada pesquisador (Strauss, Sherman & Spreen, 2006) no que tange, por exemplo, o idioma para qual será traduzido, ou mesmo à forma de apresentação de estímulos.

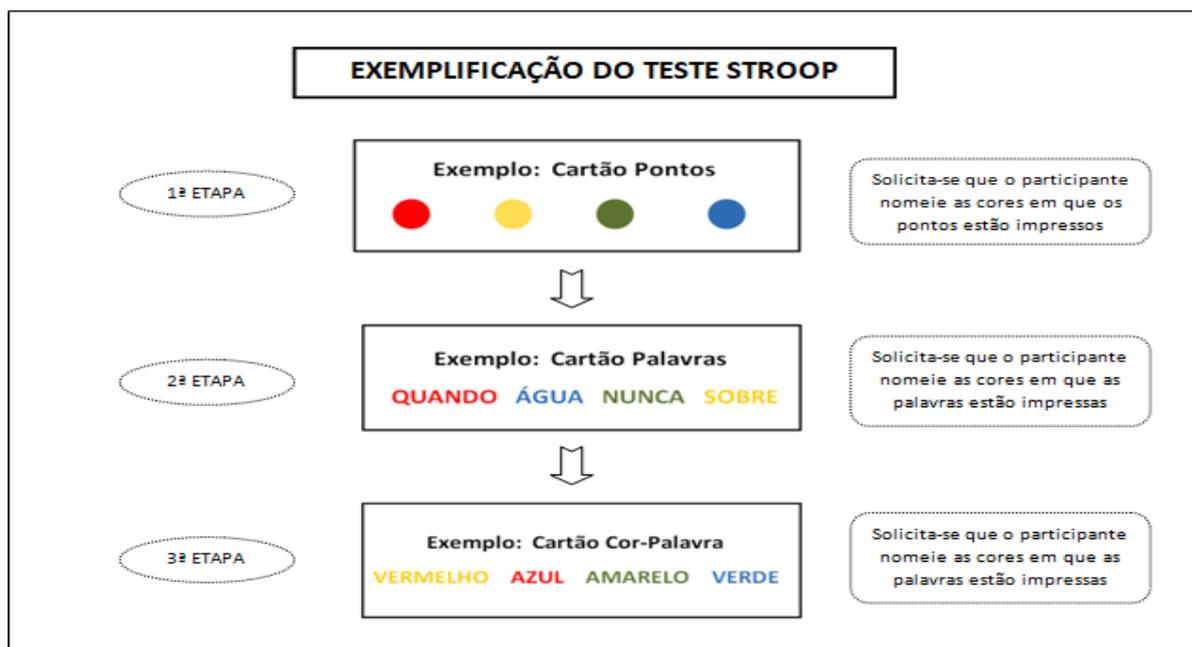
Com objetivo de verificar possíveis interferências na tradução dos itens do Teste Stroop, versão Golden e Freshwater (2002), Brandelero e De Toni (2015) realizaram estudos para verificar a influência da extensão da palavra *red* quando traduzida para a língua portuguesa, já que na língua original as palavras *red*, *green*, e *blue* (cores definidas pela versão Golden e Freshwater) possuem uma quantidade similar de letras, o que passa a ser diferente quando traduzidas (vermelho, verde e azul). Assim, foram utilizadas duas versões do Stroop, onde a única modificação foi a substituição da palavra VERMELHO para a palavra ROSA em uma das versões. Por meio da análise de dados, não foram verificadas diferenças estatisticamente significativas entre as duas versões em relação ao número de acertos obtidos pelos participantes em 45 segundos, podendo, assim, optar tanto por uma quanto pela outra.

A forma de avaliação do teste, conforme a versão utilizada pode ser refletida por duas variáveis, sendo uma delas o tempo gasto pelo indivíduo em cada cartão e a outra o número de erros cometidos não autocorrigidos. Os erros que são corrigidos pelo participante espontaneamente não são computados como erros, mas sim como acertos (Kulaif, 2005).

A versão de Victoria, denominada *Victoria Stroop Test*, é descrita em detalhes por Spreen e Strauss (1998). O teste consiste em três cartões/tarefas, que seguem a seguinte ordem: Cartão Pontos, Cartão Palavra e Cartão Cor-Palavra, cada um contendo seis linhas de quatro itens. Em cada uma das três tarefas, os participantes são solicitados a examinar os itens nas linhas da esquerda para a direita e realizar o instrumento de forma mais rápida e precisa possível. Na primeira etapa, Cartão Pontos, os itens são pontos coloridos, impressos nas cores verde, vermelho, azul e amarelo (ver em 1ª etapa, Figura 1), em que se solicita que o participante nomeie as cores dos pontos. Na segunda etapa, Cartão Palavras, palavras neutras, comuns (ex: água, quando, difícil, sobre) estão impressas nas mesmas cores já apresentadas no cartão anterior (verde, vermelho, azul e amarelo). Aqui, os participantes devem nomear rapidamente a cor em que as palavras foram impressas (ver em 2ª etapa, Figura 1). Na terceira e última tarefa, Cartão Cor-Palavra - onde é observado o efeito Stroop - os itens são nomes de cores (ou seja, as palavras vermelho, azul, amarelo e verde) impressas em cores que não correspondem às próprias palavras (ou seja, a palavra VERMELHO estará impressa na cor VERDE). Nesta etapa, os participantes deverão nomear a cor de tinta em que a palavra está impressa, situação que não será congruente (ex: a palavra VERMELHO não estará impressa na cor vermelha, ver em 3ª etapa, Figura 1). O tempo de execução para cada cartão será registrado, além do número de erros para cada participante e a diferença de tempo e erro entre os cartões serão calculados.

### **Figura 1**

*Exemplificação do Teste Stroop.*



Em relação ao potencial avaliativo, a literatura científica cita diversas funções que podem ser mensuradas por meio do Teste Stroop. A fim de levantar tais funções, procedeu-se uma revisão integrativa de literatura, apresentada no Capítulo 2 desta tese, com o objetivo de elencá-las, para então, posteriormente, definir os construtos alvo que constituirão os estudos de evidências de validade do instrumento.

Já considerado como um instrumento de avaliação “padrão ouro” em neuropsicologia, o Teste Stroop é referenciado pela consistência encontrada no padrão de assimetria da interferência, bem como pelo crescente interesse teórico (MacLeod, 1991). A assimetria de interferência se refere ao resultado em que palavras incongruentes retardam a nomeação de cores, mas cores incongruentes não retardam a leitura das palavras.

No campo neuropsicológico, o teste vem sendo utilizado em diferentes amostras e na avaliação de diversos aspectos. Troyer, Leach e Strauss (2007) propuseram um estudo com o objetivo de verificar a relação entre o fator idade e o aumento na dificuldade na inibição de respostas, mensurada por meio do Stroop versão Victoria. Para isto, avaliaram 272 adultos saudáveis, com idades entre 18 e 94 anos de idade. Como resultados, obtiveram dados que indicam correlações positivas entre idade e o tempo requerido para completar as três etapas do instrumento ( $r=0.44$  na etapa Pontos,  $r=0.58$  na etapa Cores Neutras e  $r=0.62$  na etapa Cor-Palavra, com  $p's < 0.001$ ), indicando que pessoas com maior idade demandaram maior quantidade de tempo para completar as três tarefas. Os escores de erros na tarefa Stroop também

foram positivamente correlacionados, entretanto com baixa magnitude de efeito, indicando menor precisão na execução do teste conforme o avanço da idade ( $r=0.17$ ,  $p=0.006$ ). De acordo com os autores, estes achados sugerem que o aumento da interferência nas tarefas Stroop com a idade não pode ser simplesmente explicado pela desaceleração cognitiva comum nessa faixa etária. Em vez disso, provavelmente reflete outras mudanças cognitivas, como a diminuição da capacidade de inibição de respostas.

Diferentemente da ideia anterior, os dados de Zysset, Schroeter, Neumann e Craon (2007) apresentaram resultados que indicam que adultos mais velhos são mais lentos em comparação com adultos mais jovens em relação ao tempo de processamento da informação no teste Stroop ( $F(1, 45) = 29.4$ ;  $p < 0.001$ ), tanto na etapa neutra quanto na etapa de interferência. No entanto, esta diferença de tempo não aumenta na etapa incongruente, apontando que a condição incongruente não é mais afetada pela idade do que a condição neutra. Os autores afirmam que os dados dão suporte a presunção de que a desaceleração geral produz efeitos em diversas capacidades cognitivas relacionadas à velocidade de processamento, e não necessariamente no processo de inibição.

Em relação ao gênero, Duncan (2006) não encontrou diferenças significativas entre estudantes do sexo masculino e feminino, com idades entre 12 a 14 anos, nos tempos de execução dos cartões I, II e III do Stroop (respectivamente palavra, cor, cor-palavra). Assim, no cartão *Palavra*, os valores médios para o sexo feminino foram  $14,38 \pm 3,87$  (IC de 95%) e para o sexo masculino,  $14,36 \pm 2,88$ . Para o cartão *Cor*, os valores médios para o sexo feminino foram  $17,14 \pm 4,62$  e para o masculino,  $17,83 \pm 3,88$  e, por fim, no cartão *Cor-palavra* os valores médios encontrados foram  $24,39 \pm 6,39$  para as meninas e  $25,76 \pm 8,3$  para os meninos.

Grande (2013) utilizou o teste Stroop como medida de controle inibitório em dois grupos de pacientes, comparando o desempenho entre 18 idosos com diagnóstico de Comprometimento Cognitivo Leve e 12 idosos diagnosticados com Demência tipo Alzheimer. Os resultados obtidos indicaram que os participantes com Demência tipo Alzheimer apresentaram maior dificuldade na inibição de respostas prepotentes quando solicitados a responder com rapidez, e também, apresentaram maior número de erros durante a etapa incongruente em comparação ao outro grupo. O autor justifica que os achados vão ao encontro com o previsto na literatura, que indica que tais pacientes demonstram maiores dificuldades em tarefas que demandam comportamentos inibitórios, tendo em vista que tais habilidades estão

associadas às áreas pré-frontais orbitais e do cíngulo anterior, regiões que se mostram afetadas já nas fases iniciais da doença de Alzheimer.

Westerhausen, Kompus e Hugdahl (2011) postulam a associação entre esquizofrenia e comprometimento no funcionamento executivo, mais especificamente, no subcomponente de inibição cognitiva. Por meio de uma meta-análise com 36 estudos, englobando 1081 pacientes com esquizofrenia e 1026 pacientes saudáveis, demonstraram que os pacientes esquizofrênicos exibem um efeito de interferência aumentado na tarefa Stroop, tanto em relação ao tempo de resposta ( $M(g) = 0.43$ ; 95% intervalo de confiança, IC95%: 0.35–0.52), quanto na precisão, medida pelo número de erros ( $M(g) = 0.62$ ; IC95%: 0.47–0.77)

No trabalho de De Moura Kolling, da Silva, Carvalho, da Cunha e Kristensen (2007) os autores investigaram se o consumo de álcool e drogas produzem alterações neuropsicológicas em pacientes usuários. Para isso, os pesquisadores avaliaram as funções executivas utilizando instrumentos, dentre eles o Teste Stroop, com o objetivo de avaliar possíveis disfunções no controle inibitório. Como resultado, encontraram maior prejuízo nessas funções entre os participantes usuários de álcool (média=183,0 dp=44,22) em detrimento dos usuários de crack (média 145,16, dp= 32,94).

Ainda em relação ao desempenho no Teste Stroop em participantes usuários de drogas, Fernández-Serrano, Perales, Moreno-López, Pérez-García e Vardejo-Garcia (2011) realizaram uma pesquisa com o objetivo de examinar as diferenças no desempenho em testes neuropsicológicos que avaliam inibição, entre eles o Teste Stroop de Cores e Palavras, em indivíduos dependentes de cocaína e indivíduos saudáveis. Os indivíduos usuários de cocaína tiveram significativamente um maior número de erros na execução do teste, em comparação com os indivíduos saudáveis ( $U=1,020$ ;  $p=0,013$ ;  $d=0,36$ ).

Motivados pela hipótese de que existe uma relação entre dependência de internet e controle de impulso, e que sujeitos com *internet addiction disorder* (transtorno de dependência de internet) mostrariam uma capacidade de controle executivo prejudicada, Dong, Zhou e Zhao (2011) utilizaram o teste Stroop como medida neuropsicológica, juntamente com uma medida eletroencefalográfica, em um estudo comparativo entre estudantes com *internet addiction disorder* - IAD e estudantes controle. Como resultados, apontaram que o tempo de resposta na etapa incongruente foi significativamente mais longo no grupo de participantes com transtorno de dependência de internet ( $F(32) = 5.02$ ,  $p < 0.01$ ). Nas demais condições, congruentes, não

ocorreram diferenças entre os grupos. A respeito do número de erros, o grupo com IAD mostrou resultados significativamente mais altos do que o grupo controle ( $F(32) = 3.80, p < 0.01$ ). Ainda, o grupo IAD mostrou uma amplitude média reduzida na medida eletroencefalográfica do que o grupo normal em condições incongruentes [ $F(32) = 7,55, p = 0,01$ ], o que sugere que a menor amplitude média nas pessoas com IAD em condições incongruentes refletem déficits na capacidade de controle executivo.

No mesmo sentido, Albein-Urios, Martinez-González, Lozano, Clark e Verdejo-García (2012) compararam o desempenho cognitivo de indivíduos dependentes de cocaína, jogadores patológicos e um grupo controle, por meio de uma bateria de testes, em que o Stroop foi utilizado como medida de inibição. A escolha destes dois grupos baseou-se na hipótese de que a dependência de cocaína e o jogo patológico compartilham vulnerabilidades neurobiológicas relacionadas à adição. Ambos os grupos – jogadores patológicos (JP) e dependentes de cocaína (DC) – apresentaram menor inibição no Stroop em comparação aos controles ( $F(2,68) = 4.04, p = 0.022$ ). Adicionalmente, mostraram que o desempenho de inibição de Stroop foi significativamente correlacionado com o pico de uso de cocaína ( $r = 0,48, p = 0,014$ ) e com o uso regular de álcool ( $r = 0,36, p = 0,01$ ) em ambos os grupos JP e DC, significando que o maior uso correlacionou-se com maior tempo necessário para inibição, ou seja, pior desempenho do teste. Também foram realizadas análises não paramétricas sobre o número de erros de inibição. Os resultados mostraram diferenças significativas entre os grupos, em que os grupos de jogadores patológicos e dependentes de cocaína apresentaram maior número de erros do que o grupo controle.

O teste Stroop é, também, frequentemente utilizado em pesquisas que envolvem transtornos alimentares (TA), tanto em sua versão original quanto em versões modificadas (Montagner & Esteves, 2013). Duchesne, Mattos, Fontenelle, Veiga, Rizo e Appolinario (2004) realizaram uma revisão de literatura com o objetivo de levantar publicações relacionadas à neuropsicologia dos transtornos alimentares e encontraram menções à utilização do teste em diferentes tipos de TA. Analisando as funções cognitivas relacionadas à anorexia nervosa, identificaram resultados contraditórios. Citam que no estudo de Kingston et al. (1996) e Jones et al. (1991) pacientes com anorexia nervosa (AN) apresentam pior desempenho, quando comparados com pacientes sem AN, na tarefa incongruente do Stroop. Já em Mathias e Kent (1998) e Witt e et al. (1985), não foram identificadas diferenças nos resultados do instrumento entre grupos.

No campo internacional, Colin MacLeod em 1991 citou que, até a data, somavam-se mais de 700 artigos publicados relatando o efeito Stroop. Em uma revisão sistemática de literatura, Ramos e Hamdan (2016) produziram um levantamento nas bases de dados SciELO, PePSIC, LILACS e BDTD sobre a produção científica envolvendo a avaliação neuropsicológica no Brasil. Um dos resultados apontou o teste Stroop como o 10º instrumento mais utilizado neste tipo de avaliações, sendo citado em 34 artigos científicos.

Neste mesmo sentido, Faria, Alves e Charchat-Fichman (2015) realizaram um estudo de revisão sistemática com o objetivo de verificar os instrumentos mais utilizados para a avaliação de funções executivas (FE) em idosos. Para isto, utilizaram os resultados dos últimos cinco anos apresentados na base de dados PubMed, com os termos "*neuropsychological tests*", "*executive functions*" e "*mild cognitive impairment*", sem restrição de idioma, que envolviam o uso de testes neuropsicológicos em adultos saudáveis, ou que apresentassem comprometimentos cognitivos leves e demências. Como resultados, apontaram a utilização do Teste Stroop como o 6º instrumento mais utilizado para avaliação de FE em pessoas acima de 50 anos de idade.

A frequente utilização do Stroop também foi pontuada por Nyongesa et al. (2019), no estudo que teve como uma de suas finalidades identificar quais medidas são comumente empregadas para avaliar a função executiva em adolescentes. A busca abrangeu artigos de pesquisa publicados entre 1º de janeiro de 2002 e 31 de julho de 2017 em três bases de dados bibliográficas (PsycINFO, Ovid Medline e Web of Science) usando os termos-chave "adolescents," "executive functions," and "measures". Das 338 medidas individuais identificadas, foram agrupadas 72 tarefas em 12 paradigmas de Função Executiva predominantes, e dentre elas, o Stroop foi mencionado 216 vezes, estando entre as 10 medidas mais frequentemente utilizadas na avaliação desta faixa etária.

### 3.2 AVALIAÇÃO NEUROPSICOLÓGICA E FUNÇÕES EXECUTIVAS

A neuropsicologia é um campo científico alocado dentro das neurociências, que se dedica a compreender a relação entre o cérebro e o comportamento, no esforço de tentar explicar a maneira pela qual a atividade cerebral é expressa no comportamento observável. Para que seja

possível entender o comportamento humano, nesta perspectiva, parte-se da necessidade de entender o cérebro humano (Beaumont, 2008).

Esse complexo estudo das relações entre o cérebro e o comportamento somente se torna alcançável por meio da inter-relação entre conhecimentos interdisciplinares, como fundamentos de neuroanatomia, neurofisiologia, neuroquímica, neurofarmacologia com demais áreas da psicologia, exemplo psicometria, psicologia clínica, psicologia experimental, psicopatologia e psicologia cognitiva (Hamdan, Pereira & Riechi, 2011). Pinheiro (2005) aponta que atualmente a neuropsicologia tem como enfoque central o estudo da relação entre comportamento, cognição e sistema nervoso e reforça que tal relacionamento é possibilitado pela sua alocação interdisciplinar entre as neurociências e as ciências do comportamento.

Lezak (1995), na década de 90, afirmou que a evolução da área refletia uma sensibilidade crescente entre os clínicos para os problemas que circundam a identificação, a avaliação, os cuidados e o tratamento de pacientes com danos cerebrais. Atualmente, os avanços circundam a introdução e o aperfeiçoamento das técnicas de observação do cérebro, por meio de exames de neuroimagem como a tomografia computadorizada, tomografia por emissão de pósitrons (PET) e ressonância magnética funcional (fMRI); o desenvolvimento de metodologias para a reabilitação e intervenção em pacientes com comprometimentos em funções psíquicas superiores decorrentes de lesões; e o refinamento dos instrumentos de avaliação neuropsicológicas (Pinheiro, 2005).

Ramos e Handam (2016) definem a avaliação neuropsicológica (ANP) como um procedimento de investigação, que:

“utiliza de entrevistas, observações, provas de rastreio e testes psicométricos para identificar rendimento cognitivo funcional e investigar a integridade ou comprometimento de uma determinada função cognitiva. Podem ser destacados, dentre seus objetivos, identificar e descrever prejuízos ou alterações no funcionamento psicológico, clarificar o diagnóstico em casos de alterações não detectadas por neuroimagem, avaliar a evolução de condições neurodegenerativas, correlacionar o resultado dos testes com aspectos neurobiológicos e/ou dados obtidos por neuroimagem, investigar alterações cognitivas e comportamentais que possam relacionar-se a comprometimentos psiquiátricos e/ou neurológicos” (Ramos & Handam, 2016, p. 3).

A avaliação neuropsicológica é um processo de investigação sobre as manifestações cognitivas, comportamentais e emocionais do sistema nervoso em condições normais e patológicas. De acordo com Malloy-Diniz et al. (2016) a realização de uma ANP envolve uma sucessão de etapas que incluem o uso de diferentes técnicas que possibilitam avaliar as hipóteses levantadas. Para a formação de um raciocínio clínico, é fundamental a testagem de hipóteses, que compreende, dentre outros meios, a seleção de testes e tarefas neuropsicológicas. Os autores reforçam que esta seleção deverá estar pautada nos aspectos técnicos de qualidade das medidas, considerando o possível impacto que as decisões diagnósticas indicadas pelos profissionais produzirão no cotidiano dos pacientes.

Na atualidade, a avaliação neuropsicológica tem se dedicado ao estudo de métodos para a avaliação das funções executivas (FEs). O acúmulo de evidências que relacionam a influência e a relevância destas habilidades para o sucesso em inúmeras tarefas desempenhadas ao longo da vida dos indivíduos direcionou substancial atenção para os instrumentos que possibilitam a avaliação de tais habilidades (Dias & Malloy-Diniz, 2020).

As funções executivas são habilidades consideradas essenciais para o ser humano, que permitem a sua interação com o mundo e com seus pares. Essas funções repercutem tanto na saúde mental, quanto física, além de ter influência no sucesso na vida cotidiana como um todo. Assim, pode-se afirmar que as FEs impactam contextos escolares, laborais, de relacionamentos, e, também, o desenvolvimento cognitivo, social e psicológico (Diamond, 2014).

As FEs são qualificadas como um conjunto de habilidades que contribuem, integradamente, por diversos fatores, como controle e direcionamento das ações do indivíduo; escolha das metas a serem alcançadas; planejamento das ações; seleção das respostas mais adequadas; abandono das estratégias ineficazes; acompanhamento da atividade; e verificação dos resultados (Malloy-Diniz, Sedo, Fuentes & Leite, 2008; Lezak 1995).

Chamadas, também, de controle executivo ou controle cognitivo (Diamond, 2014) as FE incluem vários componentes de processamento. Entretanto, mesmo que extensivamente citado, este construto vem sendo definido de diversas formas (Uehara, Charchat-Fichman & Landeira-Fernandez, 2013; Baggetta & Alexander, 2016). Esta vastidão conceitual pode ser observada por meio do trabalho de Goldstein, Naglieri, Princiotta e Otero (2014), que identificaram por meio de um levantamento 32 definições conceituais de Funções Executivas de diferentes autores.

Baggetta e Alexander (2016) conduziram um estudo de revisão sistemática da literatura com diversos objetivos, incluindo a identificação, entre os estudos empíricos contemporâneos, das definições do termo "funções executivas", seus componentes ou processos e os modelos teóricos relacionados. No que diz respeito aos modelos teóricos empregados nesses estudos, os pesquisadores identificaram um total de 48 modelos distintos. Entre esses modelos, o mais citado foi o de Miyake et al. (2000), sendo referenciado em 34 ocasiões.

Miyake et al. (2000) define as FEs como um conjunto de mecanismos de controle, os quais regulam a dinâmica da cognição e da ação humana. Para o autor, três funções ocupam papéis centrais: (a) *updating and monitoring of working memory representations* (acompanhamento e atualização das representações da memória de trabalho); (b) *shifting of mental sets* (traduzido aqui como mudança entre conjuntos mentais); e (c) *inhibition of prepotent responses* (inibição de respostas prepotentes);

Nesta tese, o modelo teórico adotado para a compreensão das funções executivas será o modelo psicométrico de Miyake et al. (2000), e as atualizações decorrentes dos grupos de estudos dos autores e, portanto, as três funções principais desta teoria sobre as FEs serão relatadas a seguir.

Miyake e colegas (2000) denominam "*Updating*", como se referindo a constante atualização e ao monitoramento das representações da memória de trabalho. Traduzida aqui como "acompanhamento e atualização das representações da memória de trabalho", os autores postulam que esta função vai além da simples manutenção das informações relevantes à realização de uma tarefa. Enfatizam que a essência está na manipulação ativa de informações armazenadas na memória de trabalho, e não somente em sua retenção.

Os autores apontam que esta função exige o monitoramento e a codificação de informações recebidas de acordo com a relevância da tarefa em questão e, em seguida, revisa os itens mantidos na memória de trabalho, substituindo aquelas informações que ora não são mais pertinentes, por aquelas consideradas mais relevantes no momento. Para melhor compreender o conceito de *updating*, faz sentido esclarecer o que é a função memória de trabalho.

Diamond (2013) conceitua que a memória de trabalho (MT) é considerada uma função responsável pela manutenção das informações em mente e a capacidade de trabalhar

mentalmente com elas. Existe a diferenciação entre dois tipos de MT pelo seu conteúdo, que são memória de trabalho verbal e memória de trabalho não verbal – visuoespacial. A autora exemplifica que fazer contas matemáticas mentalmente, reorganizar uma lista de tarefas, incorporar novas informações, considerar outras alternativas e relacioná-las e verificar a associação entre ideias ou itens dependem desta função. Ainda, pontua que o raciocínio, por exemplo, não seria possível sem a função de MT.

Traduzida nesta tese como “mudança/alternância entre conjuntos mentais”, originalmente foi denominada por Miyake e colegas (2000) como *shifting between tasks or mental sets*, e traduzida aqui como alternância entre tarefas ou conjuntos mentais. Em sua forma abreviada, foi chamada de “*Shifting*” e diz respeito à habilidade de alternar entre diversas tarefas, operações ou conjuntos mentais. Também referenciada como “*attention switching*” ou “*task switching*” (mudança/troca atencional ou mudança/troca entre tarefas), foi identificada como importante na compreensão de falhas no controle cognitivo de pacientes com danos cerebrais estudados.

Em termos de definição operacional, os autores pontuam que esta função é responsável pelo desengajamento de um conjunto de tarefas irrelevantes e pelo subsequente engajamento ativo em um conjunto de tarefas relevantes. Acrescentam, que tal operação envolve, ainda, a habilidade de realizar uma nova operação diante de uma interferência proativa, superando-a.

Miyake e colaboradores (2000), denominaram, abreviadamente, como inibição o conceito de “*inhibition of prepotent responses*” - que pode ser traduzido como inibição de respostas prepotentes – e definiram que se refere à habilidade de inibir, deliberadamente, respostas prepotentes, dominantes ou automáticas sempre que necessário. Apontaram, entretanto, que tendo em vista a variedade de funções associadas ao termo inibição, em seu trabalho, esta função seria restrita à supressão deliberada e controlada de respostas prepotentes.

Anos depois, Friedman e Miyake (2004), em novo trabalho, expuseram que o conceito de inibição como um construto unitário ou como uma junção de funções relacionadas não era consensual na literatura, e que, embora possam ser potencialmente distintas, em termos conceituais, não havia clareza sobre até que ponto tais funções refletem, ou não, as mesmas habilidades cognitivas. Dotados desta justificativa, se propuseram a examinar, por meio da técnica estatística de análise fatorial confirmatória, as relações entre três funções relacionadas à inibição, potencialmente separadas, denominadas de (a) Inibição de Resposta Prepotente, (b)

Resistência a Interferência de Distratores e (c) Resistência à Interferência Proativa. Tais definições, conforme os autores, seguem a taxonomia conceitual proposta por Joel Thomas Nigg (2000).

Inibição de Resposta Prepotente é tratada como uma competência em suprimir intencionalmente respostas dominantes, prepotentes e/ou automáticas (Friedman & Miyake, 2004; Miyake & cols, 2000). Os autores, na avaliação desta característica, optaram por utilizar os instrumentos *Antisaccade task*, *Stop-signal task* e o Teste Stroop, como medidas para mensuração. Ainda, indicam que a Inibição da Resposta Prepotente é o tipo de inibição mais diretamente associado à supressão ativa e ao funcionamento executivo (Friedman & Miyake, 2004).

Resistência à Interferência de Distratores corresponde à capacidade de resistir ou solucionar interferências de informações no ambiente externo que são irrelevantes para a realização da tarefa em questão. Neste estudo, foram selecionados testes em que os participantes selecionavam alvos que eram apresentados junto a distratores irrelevantes. Para esta modalidade de tarefa, a função estaria mais relacionada com atenção focada ou seletiva dos alvos, e as atividades escolhidas foram: *Eriksen flanker task*, *Word naming* e *Shape matching* (Friedman & Miyake, 2004).

Resistência à Interferência Proativa é a capacidade de resistir a invasões de lembranças mnemônicas, que anteriormente eram consideradas relevantes para a execução da tarefa, mas que se tornaram irrelevantes desde então. Friedman e Miyake (2004) afirmam que embora semelhante ao conceito de Resistência à Interferência de Distratores, no que tange o envolvimento do controle de interferência, dois recursos os diferenciam conceitualmente. Enquanto na Resistência à Interferência de Distratores as informações de distração são apresentadas simultaneamente com as informações de alvo e não possuem relevância para a realização da tarefa, na Resistência à Interferência Proativa, as informações interferentes são apresentadas antes das informações alvo e eram consideradas, anteriormente, relevantes para a tarefa. Os testes escolhidos para avaliar a Resistência à Interferência Proativa foram *Brown–Peterson variant*, *AB–AC–AD* e *Cued recall* (Friedman & Miyake, 2004).

Considerando isto, Friedman e Miyake (2004) examinaram as relações entre as três funções relacionadas à inibição, citadas anteriormente, que seriam potencialmente separáveis. Os resultados encontrados sugeriram relação entre Inibição de Respostas Prepotentes e

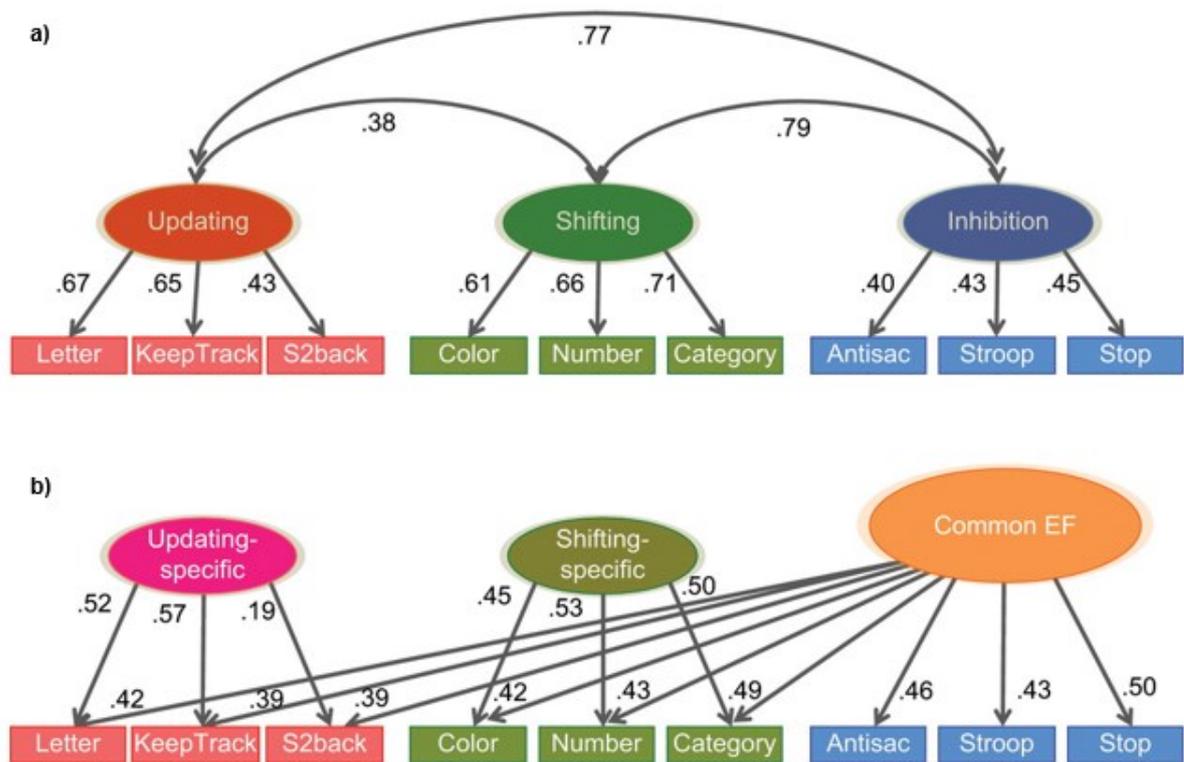
Resistência à Interferência de Distratores, contudo, não com a habilidade de Resistência à Interferência Proativa, indicando a existência de uma variância comum compartilhada entre as duas primeiras habilidades, entretanto, tal fator comum não estaria envolvido na terceira função, Resistência à Interferência Proativa. Por fim, sumarizam que as teorias que postulam a inibição como um mecanismo unificado tendem a ser excessivamente ambiciosas nesta presunção. Tal argumento, enfatiza que inibição, quando definida amplamente, não sustentaria as explicações necessárias tendo em vista a sua complexidade. Indicam que é possível que um tipo de inibição forneça uma estrutura comum, entretanto, esta natureza deve ser melhor especificada por teorias mais aprofundadas.

Em consonância, Magalhães (2013) acrescenta que, no campo clínico, não se pode afirmar que todas as psicopatologias compartilhem do mesmo déficit inibitório, tendo em vista a extensão do termo inibição, bem como de seus paradigmas experimentais. Assinala, ainda, que o questionamento no tocante à validade unitária do construto é promissor, e que alguns pesquisadores já compactuam o entendimento de que o controle inibitório é formado por habilidades implícitas. Indica, também, a importância em especificar e definir os subtipos distintos dos processos de inibição, que, de forma inespecífica, consideram atos de hesitação, conflito, interferência como análogos.

Em trabalho posterior, Miyake e Friedman (2012) revisitam os próprios estudos e reforçaram alguns de seus postulados, dentre eles, a ideia de unidade e diversidade em relação ao construto funções executivas. Os autores concluíram que diferentes FEs se correlacionam entre si, explorando, assim, alguma habilidade comum subjacente – ideia de unidade – mas que também demonstraram algumas diversidades.

## **Figura 2**

*Modelo de unidade/diversidade de Miyake e Friedman (2012)*



Fonte: Miyake e Friedman (2012)

A figura 2a apresenta o que os autores chamam de unidade e diversidade nas FEs. Por meio de análises estatísticas utilizando o método de análise fatorial confirmatória, os dados indicam que as três funções, ou variáveis latentes - *updating*, *shifting*, and *inhibition* - possuem correlações entre si, entretanto, são consideradas distintas, separáveis, pois os valores das correlações estão afastados de 1.0, medida que indicaria total associação. Ainda, demonstram a existência de uma variável latente comum em todos os nove testes de funções executivas (figura 2b). Nesta proposta, sugerem que os componentes *updating* e *shifting* resultam de um fator comum e de um fator específico, o que não ocorre com o fator *inhibition*, que parece ser totalmente explicado por este fator comum.

Dentre os alvos de seus estudos, os autores concluem que as diferenças individuais nas FEs, mensuradas por meio de tarefas simples de laboratório, mostram unidade e diversidade (diferentes FEs são correlacionadas, mas separáveis), refletem contribuições genéticas substanciais, estão relacionadas a várias queixas clínicas e fenômenos socialmente importantes, e mostram alguma estabilidade de desenvolvimento, reforçando a importância da continuidade

de estudos empíricos que versem sobre a temática e que possam esclarecer as dissidências na área.

Para Moreira, Costa, Castro Lima e Vicente (2017) ser capaz de identificar os possíveis déficits nas funções executivas é um procedimento de elevada importância no diagnóstico e no monitoramento de inúmeras situações clínicas, incluindo os distúrbios neurodegenerativos. Deste modo, os variados recursos utilizados neste tipo de avaliação são tópicos de grandes debates, sendo um deles, o uso de ferramentas neuropsicológicas que sejam potencialmente hábeis a fornecer medidas confiáveis na avaliação de FEs.

A respeito deste argumento, têm-se levantado a temática da impureza das tarefas frequentemente utilizadas para tais avaliações. Friedman e Miyake (2004) apontam no estudo destinado à análise das medidas de inibição, já mencionado nesta tese anteriormente, que nenhuma tarefa poderia ser considerada uma medida pura para tal fator (inibição). Descrevem, por exemplo, que o processo de inibição é sempre a inibição de algo, seja uma resposta, pensamento, distração, entre outros, e por tal, o resultado - em termos de performance - envolverá outros processos. Por este motivo, não se poderia justificar que baixas pontuações em uma tarefa seriam derivadas de uma inabilidade na função, ou ainda, que baixas correlações entre diferentes testes podem não necessariamente ser decorrentes de diferentes processos inibitórios, mas sim, justificadas por outras variáveis que podem mascarar as semelhanças atribuíveis (Miyake, Friedman, et al., 2000; Shilling et al., 2002).

Assim, a presença de outros componentes cognitivos além daqueles que estão sendo especificamente investigados em uma determinada tarefa trata-se de um tópico que vem sendo discutido no campo da avaliação neuropsicológica. Van der Sluis e seus colegas (2006) discutem esse conceito em relação ao estudo das funções executivas e destacam que essa impureza das tarefas pode levar a resultados confusos ou ambíguos, uma vez que não é possível determinar se as diferenças observadas nos resultados se devem à função executiva específica sendo investigada ou a outros processos cognitivos envolvidos na tarefa. Em resumo, o conceito de "impureza das tarefas", refere-se à presença de componentes cognitivos adicionais em demandas que podem complicar a interpretação dos resultados específicos das funções ora investigadas.

Para lidar com a impureza das tarefas, os pesquisadores precisam projetar seus estudos cuidadosamente, utilizando testes que sejam específicos para a função que está sendo

investigada e controlando o máximo possível de fatores confundidores. Isso pode envolver o uso de tarefas mais puras, ou seja, medidas que envolvam predominantemente a função executiva de interesse, e a análise cuidadosa dos resultados para separar os efeitos da função específica dos efeitos de outros processos cognitivos. É importante, portanto, minimizar essa impureza ao projetar e analisar estudos de funções executivas (Van der Sluis, et al. 2006).

Conhecer a impureza das tarefas permite selecionar testes mais específicos para a função executiva desejada, planejar estudos e intervenções com maior precisão, realizar avaliações clínicas mais precisas, e avançar na pesquisa ao identificar os testes mais confiáveis. Em suma, o entendimento da impureza das tarefas é crucial para o uso de testes neuropsicológicos de forma eficaz e para interpretar os resultados de maneira precisa. Isso contribui para a confiabilidade, validade e utilidade dos testes na avaliação das funções executivas e de outras habilidades cognitivas.

## 4. CAPÍTULO II

### TESTE STROOP DE CORES E PALAVRAS: UMA REVISÃO INTEGRATIVA DE LITERATURA

#### 4.1 INTRODUÇÃO

O teste Stroop de Cores e Palavras (chamados neste artigo de TSCP, Stroop ou Teste Stroop), desenvolvido por John Ridley Stroop, em 1935, trata-se de uma tarefa mundialmente conhecida no campo das neurociências. Amplamente utilizado em contextos experimentais e clínicos (Stroop, 1935), ganhou status de “padrão ouro” por ser um dos fenômenos cognitivos mais robustos disponíveis na avaliação neuropsicológica (MacLoad, 2002).

A literatura, entretanto, indica que o potencial avaliativo do instrumento não se restringe a um único construto teórico, mas sim, a uma ampla quantidade de funções. Classicamente, o TSCP vem sendo mencionado como uma medida de atenção seletiva (MacLeod, 1991), todavia, é, também, referenciado na avaliação de diversos outros construtos, como por exemplo, controle inibitório (Malloy-Diniz, Paula, Loschiavo-Alvares, Fuentes & Leite, 2010); interferência; inibição (Sweet, 1999); funções executivas (Della Corte et. al. 2018); velocidade de processamento (Kim, Lee, Jo, & Lee, 2017); flexibilidade (Klojčnik, Kavcic, & Bakracevic Vukman, 2017); e memória (Cammissuli & Timpano Sportiello, 2017).

Em decorrência desta diversidade de indicações para a utilização prática do teste, procedeu-se uma revisão integrativa de literatura. Esta teve o objetivo de realizar um levantamento sobre quais construtos teóricos são indicados como alvo na avaliação deste teste, bem como, as versões utilizadas e os contextos nos quais o teste foi empregado.

#### 4.2 MÉTODO

A metodologia empregada na busca por artigos se deu por meio da utilização de associação dos termos “*stroop test and neuropsychological tests*” AND “teste stroop *and* testes neuropsicológicos”, definidos após consulta aos Descritores em Ciências da Saúde - DeCS. As

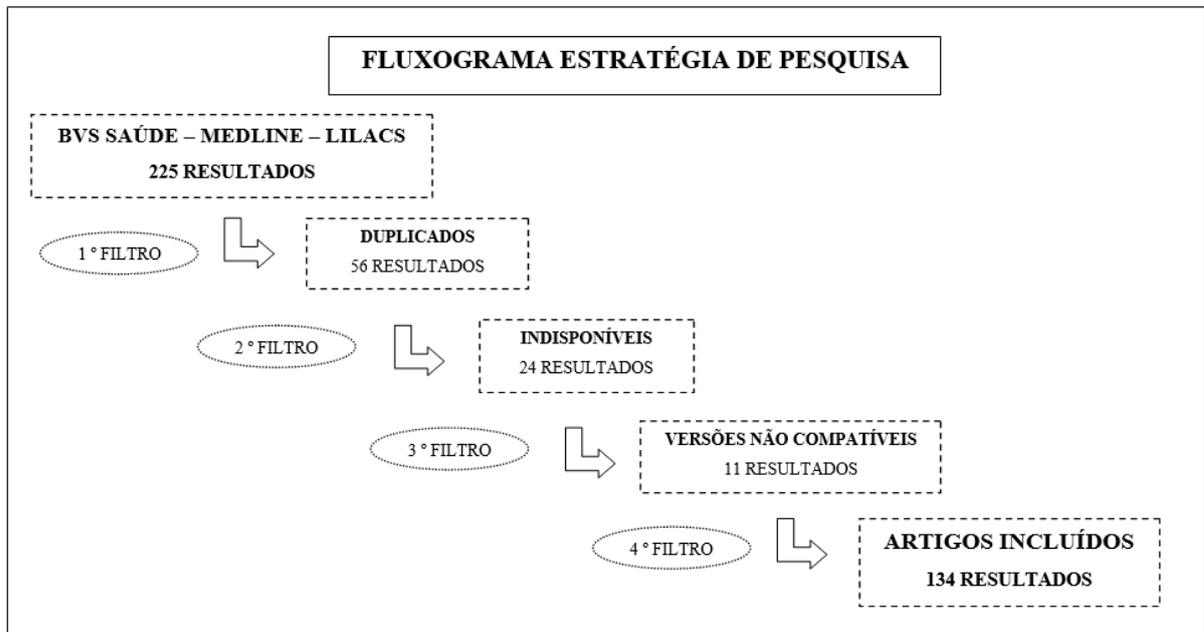
bases de dados utilizadas foram a Biblioteca Virtual em Saúde BRASIL (BVS Saúde), *Medical Literature Analysis and Retrieval System Online* (MEDLINE) e Literatura Latino-americana e do Caribe em Ciências da Saúde (LILACS).

O período cronológico definido nesta pesquisa compreendeu as publicações datadas nos anos 2017, 2018 e 2019. A escolha deste período justifica-se devido ao extenso volume de trabalhos envolvendo o Teste Stroop. Foram definidos como critérios de inclusão artigos disponíveis na íntegra, nos idiomas português, inglês e espanhol, além de trabalhos cujas versões compreendessem a apresentação simultânea de cores e palavras. Foram incluídas versões tanto em formato manual quanto computadorizado. Foram excluídos da amostra, portanto, versões que não representassem como procedimento a aplicação envolvendo a apresentação incongruente entre cores e palavras.

Inicialmente encontrados 225 trabalhos. Destes, 56 foram excluídos por serem duplicados na primeira filtragem. Dos restantes, 24 não estavam disponíveis e 11 não se enquadraram nos critérios de inclusão referentes à compatibilidade das versões. O resultado final contemplou 134 trabalhos incluídos, que foram lidos na íntegra e analisados (Figura 3). Em relação ao idioma, 132 artigos foram escritos na língua inglesa, 01 artigo estava disponível em português e 01 em espanhol.

### **Figura 3**

*Fluxograma da estratégia de pesquisa*



Fonte: autoria própria

### 4.3 RESULTADOS

Com o propósito de expor os dados obtidos a seção resultados foi dividida em três subseções: a) estudo 1: Teste Stroop e construtos indicados; b) estudo 2: versões do Teste Stroop; e c) estudo 3: utilização prática do Teste Stroop em amostras ou casos clínicos.

#### Estudo 1: Teste Stroop e construtos indicados

Nos 134 artigos revisitados integralmente foram encontradas, inicialmente, 31 definições de construtos ou habilidades indicadas como avaliadas pelo teste Stroop. Dentre elas, é possível identificar tanto construtos teóricos específicos, como, por exemplo, controle inibitório, quanto definições operacionais de capacidades, como “habilidade de inibir respostas interferentes” (Karlsson, et al. 2017).

A maior frequência apontou para a utilização do instrumento na avaliação de Funções Executivas, mencionado em 42 trabalhos. Na sequência, a segunda maior quantidade de citações indicou para a capacidade avaliativa da inibição (34 trabalhos). Com a terceira maior ocorrência, apareceu o construto de atenção seletiva (25 trabalhos). O controle inibitório (20 trabalhos) figurou em quarto lugar. Velocidade de processamento (19 trabalhos) ficou a quinta

colocação, e flexibilidade (12 trabalhos) apareceu em sexto lugar em relação à frequência de citações.

Foi possível observar que algumas habilidades mencionadas indicam pertencer a definições operacionais de construtos. Por exemplo: “supressão de resposta” (Ferreira, et al. 2018); “capacidade de resistência a um estímulo disruptivo” (Arslan, Tasdemir, Durmaz, & Tosun, 2018); “habilidade de manutenção de um objetivo em mente e realizar a supressão de uma resposta habitual em favor de uma menos familiar (Masiliūnas, et al. 2017)”; “habilidade de inibir respostas interferentes” (Guevara et al. 2018); e “habilidade de supressão e interrupção de alerta inapropriado e tendência de reação” (Arslan, Tasdemir, Durmaz, & Tosun, 2018) são exemplificações do conceito de inibição. O mesmo ocorreu com a definição indicada como “capacidade de resistir à interferência de estímulos externos (Alizadeh, et al. 2017)” - que está relacionada ao construto de atenção seletiva - e também com a definição “capacidade de intensificação e sustentabilidade da atenção” (Arslan, Tasdemir, Durmaz, & Tosun, 2018) abarcada pelo construto de atenção sustentada.

Ainda, com base na fundamentação teórica levantada a partir da análise dos artigos selecionados pôde-se agrupar construtos equivalentes, como por exemplo, inibição, controle inibitório e controle de interferência. Por meio da análise dos três construtos foi possível identificar que ambos correspondem à mesma habilidade funcional, e por tanto, foram considerados nesta pesquisa como referentes a um único domínio: inibição.

Após análise inicial, optou-se pela organização dos construtos com base em sua similaridade conceitual de acordo com a literatura teórica. Deste modo, procedeu-se com o agrupamento de construtos e funções conforme a tabela 1.

**Tabela 1**

*Construtos agrupados e suas respectivas frequências*

CONSTRUTO/FUNÇÃO	FREQUÊNCIA
Inibição/Controle Inibitório/Controle da Interferência	67
Funções Executivas	42
Atenção Seletiva	28
Velocidade de Processamento	19
Atenção	17

Flexibilidade cognitiva	15
Controle Cognitivo	8
Funções Cognitivas	6
Memória de trabalho	3
Atenção Sustentada	3
Interferência Cognitiva	2
Tempo de reação	2
Memória	3
Eficiência de processamento	1
Estado de alerta	1
<i>Scanning</i>	1
Concentração	1

De acordo com a organização dos dados (Tabela 1) que considerou o agrupamento das funções com base nos conceitos levantados pela literatura, o Teste Stroop foi majoritariamente citado em 67 artigos como um instrumento para avaliação da capacidade de inibição. Em segundo lugar, destacou-se a utilização do instrumento para a avaliação de Funções Executivas, citado nos artigos como um termo genérico que englobava vários processos cognitivos, como memória de trabalho, flexibilidade, raciocínio, capacidade de resolução de problemas, planejamento e inibição.

Em terceiro lugar, o construto atenção seletiva apareceu com 28 citações. Na sequência, a velocidade de processamento foi citada 19 vezes. Em quinto lugar, o conceito de atenção foi indicado em 17 trabalhos. Nestes, a atenção não foi conceitualizada ou tipificada, e por esse motivo, não foi agrupada com o construto atenção seletiva. É justificável manter essas categorias separadas, pois podem, ou não, pertencer à mesma classificação/função atencional.

Em sexto lugar, com 15 citações, apontou-se o conceito de flexibilidade cognitiva. Em menores quantidades, citou-se controle cognitivo (08 trabalhos), funções cognitivas (06), memória de trabalho (03), atenção sustentada (03), memória (03), interferência cognitiva (02), tempo de reação (02), eficiência de processamento (01), estado de alerta (01), atenção dirigida (01), scanning, traduzido como varredura, (01) e concentração (01).

Estudo 2: Versões do Teste Stroop

Previamente determinados como critérios de inclusão e exclusão, os artigos selecionados atenderam a característica de terem formato semelhante ao original em que a representação do efeito Stroop fosse realizada por meio de um conjunto de itens nos sistemas de cor e palavra. Tal sistema deveria estar retratado ao longo do texto, ou, também, citado como fonte referencial. Seguindo estes parâmetros, foi possível identificar vinte e uma (21) versões do Stroop entre os 134 artigos selecionados.

Dentre elas, 16 foram referenciadas com nomenclaturas específicas, sendo elas: *Stroop Color Word Test* por Golden (Yu, Vock, & Barclay, 2018), *Victoria Stroop Test* (Tuokko et al. 2017), *Prague Stroop Test* (Mana et al. 2019), *Stroop Test TBAG version* (Arhan et al. 2017), *The Chinese version* (Cao et al. 2017), *Korean Color-Word Stroop Test* (Kim et al. 2017), *Iranian version* (Rezaei, Hosseini Ramaghani, & Fazio 2017), *The Persian computerized version* (Shoghi Javan, & Ghonsooly, 2018), *The Color Word Interference Test (CWIT)* de *Delis-Kaplan Executive Function System D-KEFS* (Adólfssdóttir, 2017), *Modified Stroop Color-Word Test* (Talarico et al. 2017), *Stroop Neuropsychological Screening Test* de Trenerry, Crosson, DeBoe, e Leber (Borella et al. 2017), *Stroop Color and Word Test Children's Version* adaptado por Golden, Freshwater e Zarabeth, *Short Version* (Eggermont, et al. 2017), *Shortened version* para população italiana por Venneri et al. (Clerici, et al. 2017), *Teste Stroop de Cores e Palavras* adaptado por Zimmermann (Cotrena et al. 2017) e versão traduzida e adaptada por Cunha, Nicastrì, Andrade e Bolla (Milioni et al. 2017).

Ao longo da pesquisa, foram mencionadas versões em formato computadorizado. Destas, uma sinalizou ter inserido modificações no formato do teste, incluindo, além da clássica tarefa cor e palavra, uma etapa que envolvia contagem numérica, denominada condição “*counting-stroop*”. As demais indicaram execução similar ao modelo manual (que consiste na apresentação impressa do teste).

Não nomeadas distintamente, foram identificadas duas versões que incorporaram adaptações no modo funcionamento, entretanto, mantiveram o sistema similar ao original. Pela similaridade com a versão de John Ridley Stroop, foram mantidas no rol de resultados desta pesquisa.

Um fato importante de se destacar é de que aproximadamente 62% dos artigos revisados não indicaram a versão do teste utilizada, mencionando apenas o emprego do Teste Stroop

como medida neuropsicológica. A versão com maior número citações foi a versão Golden, com doze indicações. Na sequência, as versões Victoria, TBAG e versões computadorizadas foram indicadas quatro vezes; as versões *Prague Stroop Test*, *Chinese version*, *Korean Color-Word Stroop Test* e *The Color Word Interference Test (CWIT)* de *Delis- Kaplan Executive Function System (D-KEFS)* foram indicadas três vezes; e a versão reduzida para a população italiana por Venneri et al. foi citada duas vezes.

Entre as versões são perceptíveis variações em diversos aspectos do teste. Em relação às traduções, os artigos indicaram a adaptação para diversos idiomas, como inglês, checo, turco, mandarim, coreano, italiano, português, espanhol e persa. Referente à quantidade de cartões ou tarefas algumas versões indicam duas (Davison, Deeprose & Terbeck, 2018), três etapas (Mana, Vaneckova, Klempíř, Lišková, Brožová, Poláková, Maréchal et al., 2019), quatro (Cao, Chen, Xie, H, Zou, Hu & Zhou, 2017) e cinco etapas (Arslan, Tasdemir, Durmaz & Tosun, 2018).

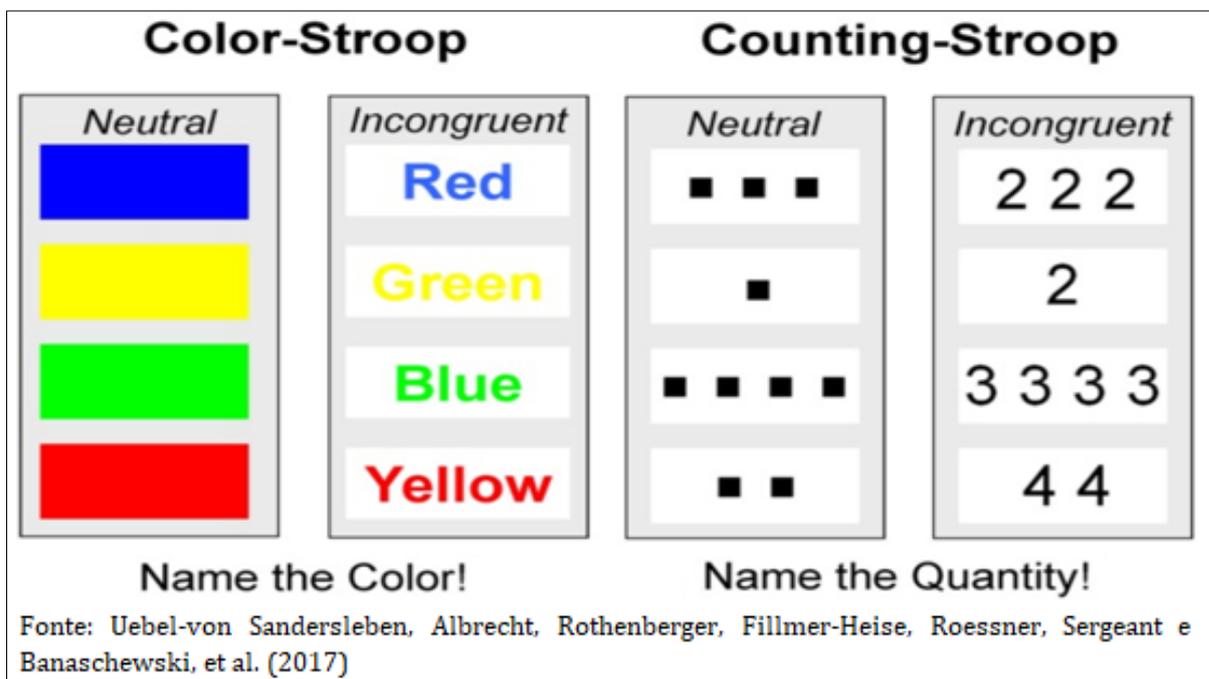
Em relação às cores empregadas, ocorrem variações tanto na quantidade quanto na escolha da palheta. Por exemplo, o *Prague Stroop Test* utiliza quatro cores: verde, azul, amarelo e vermelho. O *TBAG Version* utiliza cinco cores: vermelho, azul, verde, marrom e roxo. Em uma versão não nomeada, os autores citam três cores: vermelho, azul e verde, igualmente identificadas também na versão de Golden. A versão de Victoria foi relatada com a apresentação de quatro cores: azul, verde, vermelho e amarelo, semelhantemente à versão chinesa e também à versão *Stroop color and word test children's version*.

Referente ao formato do teste, algumas versões adicionaram tarefas além das tradicionalmente conhecidas. No subteste *Colour Word Interference* de *Kaplan Executive Function System (D-KEFS)* foi incluída uma quarta condição, chamada de alternância (Adólfssdóttir, Wollschlaeger, Wehling & Lundervold, 2017). Nesta etapa adicional, os sujeitos foram orientados a alternar entre a tarefa menos automatizada (nomeação da cor de tinta incongruente) e a automática (leitura da palavra). No trabalho de Marín-Navarrete, Toledo-Fernández, Villalobos-Gallegos, Pérez-López, e Medina-Mora (2018) foram empregadas duas versões do teste, Stroop A e Stroop B. Na versão A os participantes leram uma lista de palavras impressas em cores incongruentes e, ao ler uma palavra sublinhada, deveriam alternar a tarefa, passando a nomear a cor em que palavra é impressa. Já na versão B, as interferências ocorrem em intervalos fixos, ou seja, os participantes leram uma lista contínua de seis itens de palavras coloridas incongruentes, e alternaram para outra lista de seis itens congruentes.

No trabalho de Uebel-von Sandersleben, Albrecht, Rothenberger, Fillmer-Heise, Roessner, Sergeant e Banaschewski, et al. (2017) o teste foi aplicado em duas condições, ambas computadorizadas. A condição “*color-stroop*” consistia na utilização das cores como alvo, de modo similar à tarefa clássica de incongruência do teste Stroop. A outra condição, nomeada “*counting-stroop*”, utilizava números e quantidades para representar uma tarefa incongruente, como mostra a Figura 4.

**Figura 4**

Condições “*color-stroop*” e “*counting-stroop*”



Ocorreram variações também no formato de correção entre as versões, entretanto, a respeito deste quesito, os trabalhos tendem a dividir-se entre execução com tempo livre, fixação de tempo em segundos e contagem de erros. Em testes com tempo livre, leva-se em consideração o tempo em que o participante leva para completar cada etapa. Em modelos com tempo fixo, conta-se o número de acertos ou itens corretamente identificados. Em relação à contagem de erros, algumas versões indicaram o registro de erros não autocorrigidos, e outras, não os consideram como um dado adicional.

Estudo 3: utilização prática do Teste Stroop em amostras ou casos clínicos

Em relação aos grupos de referência, amostra ou participantes este estudo encontrou, entre os 134 artigos visualizados, 67 amostras distintas de pessoas alvo da avaliação do teste Stroop. Os estudos ocorreram majoritariamente com pacientes adultos saudáveis correspondendo a aproximadamente 25% dos dados, entretanto, houve o registro, também, de pesquisas envolvendo as demais faixas etárias. A Tabela 2 apresenta os tipos de participantes com as maiores frequências entre artigos relacionados.

**Tabela 2**

*Participantes com maiores frequências*

PARTICIPANTES/AMOSTRA	FREQUÊNCIA
Adultos saudáveis	34
Esclerose Múltipla	6
Pacientes com Doença de Parkinson	5
Idosos	5
Pacientes com Esquizofrenia	5
Pacientes com Comprometimento Cognitivo Leve	5
Pacientes com Doença de Huntington	4
Pacientes com Transtorno Bipolar	4
Pacientes com transtornos ou dependência do uso de álcool	3
Pacientes com Transtorno Obsessivo Compulsivo	2
Pacientes com Transtorno Depressivo Maior	2
Pacientes com Transtorno Depressivo	2
Pacientes com Transtorno Bipolar Eutímico	2
Crianças com Transtorno do Déficit de Atenção e Hiperatividade (TDAH)	2
Pacientes com Doença de Alzheimer	2
Pacientes com Comprometimento Cognitivo Subjetivo	2
Pacientes com Demência	2
Crianças população geral	2
Usuários de cocaína e estimulantes	2

A maior frequência de estudos contemplou a aplicação do teste Stroop em uma amostra da população adulta saudável (34 artigos). Nesta categoria, foram incluídos os participantes caracterizados como “população geral”, ou ainda, integrantes de grupos controles. Sobre esta

segunda amostra, os delineamentos consistiam na comparação entre um grupo controle versus variados grupos contrastantes.

Na sequência, 06 artigos se destinaram a avaliar capacidades neurocognitivas por meio do Stroop em pacientes com Esclerose Múltipla; 05 em pacientes com Doença de Parkinson; 05 em idosos saudáveis; 05 em pacientes com diagnóstico de esquizofrenia; 05 em pacientes com Comprometimento Cognitivo Leve; 04 em pacientes com Doença de Huntington; 04 em pacientes com Transtorno Bipolar; e 03 em pacientes com dependência ou transtornos por uso de álcool. Citados por duas vezes aparecem pacientes com diagnósticos de transtornos psiquiátricos, como: Transtorno Depressivo e Depressivo Maior, Transtorno Bipolar Eufímico, Transtorno de Déficit de Atenção e Hiperatividade e Transtorno Obsessivo Compulsivo. Também citados duplamente, as pesquisas apontam a utilização do teste em pacientes com comprometimentos neurocognitivos, como: Doença de Alzheimer, Comprometimento Cognitivo Subjetivo e Demência (sem especificação). A lista de amostras duplamente utilizadas é finalizada com usuários de cocaína e estimulantes e crianças (sem critério de especificação).

Resultados que apareceram somente uma vez no conjunto de textos foram agrupados pela similaridade ou relação entre os participantes. No primeiro grupo, denominado, “amostra de pacientes com quadros diagnósticos psiquiátricos”, figuraram pacientes jovens adultos com TDAH; pacientes com Transtorno do Pânico; pacientes com Transtorno de Personalidade Borderline; idosos com Transtorno Bipolar; crianças com Transtorno de Conduta; pacientes com Transtorno de Estresse Pós-Traumático; pacientes com Anorexia Nervosa; pacientes jogadores compulsivos; pacientes com Transtorno de Jogo Compulsivo pela Internet; pacientes com Psicose precoce; alcoolistas crônicos; pacientes com transtorno por uso de inalantes; e Transtorno Tic Crônico.

O segundo grupo foi composto por pacientes com quadros neurológicos centrais e alterações cognitivas, sendo eles: idosos com queixas de memória; pacientes com Acidente Vascular Cerebral (AVC); pacientes com lesões cerebelares; pacientes com demência frontotemporal; lesão cerebral envolvendo disfunção executiva; Ataxia de Friedreich; Ataxia espinocerebelar; pacientes com Traumatismo Craniano; e pacientes adultos com Epilepsia.

O terceiro, retratou pacientes com outras doenças, como, hipertensão, câncer de esôfago diagnosticados com depressão; dor lombar crônica; psoríase grave; rinossinusite crônica;

distonia cervical; distonia isolada; fibromialgia; cardiomiopatia dilatada primitiva; indivíduos infectados pelo HIV com comprometimento cognitivo leve; adolescentes e adultos com Hiperplasia Adrenal Congênita; pacientes com insuficiência vestibular crônica; doença renal precoce; acromegalia; adolescentes obesas do sexo feminino; e pacientes com Lesão Crônica da Medula Espinhal.

Por fim, seis trabalhos retrataram participantes que não se encaixaram nos agrupamentos anteriores, sendo eles: crianças prematuras; aprendizes iniciantes e avançados de inglês como língua estrangeira; fumantes e não fumantes; pacientes que sobreviveram a envenenamento agudo por metanol e indivíduos propensos à ilusão.

#### 4.4 DISCUSSÃO

A utilização do Teste Stroop e a diversidade de conceitos teóricos relacionados a ele pôde ser corroborado pelos dados obtidos nesta revisão bibliográfica de literatura. Tais resultados levantam possibilidades de questionamentos diversos.

A primeira delas trata-se do questionamento a respeito do Stroop avaliar funções consideradas principais e também funções secundárias; ou ainda, se ele pode se propor a avaliar mais de um construto. Huang, Lu e Xu et al. (2017) ao utilizar o Stroop como medida neuropsicológica indicaram o desmembramento das funções cognitivas envolvidas na execução de cada etapa do teste. Consideraram que as etapas SCWTA (leitura rápida de cores impressas), SCWTB (nomeação rápida de cores) são relacionadas à atenção seletiva e velocidade de processamento e a etapa SCWTC (nomeação de cor de tinta de uma palavra incongruente) foi considerada para avaliação da inibição. O cálculo da diferença de tempo entre a condição C e B foi chamada de interferência SCWT, e identificada como medida de função executiva (FE). Deste modo, pode-se entender que para a avaliação do grande conceito FE é necessário avaliar de modo separado vários de seus componentes, e assim, este único teste avaliaria mais de um construto.

Outro questionamento diz respeito à quantidade de termos mencionados e se é possível considerá-los equivalentes. Os resultados obtidos sinalizaram uma equivalência conceitual entre alguns termos distintos. Procedeu-se a análise conceitual dos construtos citados e pôde-se constatar que mesmo com nomenclaturas diferentes, alguns correspondiam à mesma habilidade funcional, ou seja, quando conceitualizados ou explicados por meio de exemplos práticos,

indicavam avaliar a mesma função. Esta situação ocorreu, por exemplo, com os construtos inibição, controle inibitório e controle de interferência. Pela congruência entre os três conceitos - inibição, controle inibitório e controle de interferência - houve a opção por considerá-los referentes à mesma habilidade funcional.

Em virtude da natureza dos periódicos encontrados, que representam a utilização prática do teste Stroop - e de demais tarefas cognitivas - é importante ressaltar que eles não se propuseram a realizar discussões teóricas a respeito dos construtos utilizados, e nem mesmo objetivam criar novos modelos teóricos sobre as funções cognitivas avaliadas. Ainda, em uma significativa quantidade de artigos, os autores nem mesmo descrevem ou conceitualizam a função medida, procedendo somente com a informação do construto avaliado, sem explicações adicionais. De modo geral, os autores fazem menção a outras referências quando conceituam suas funções alvo.

Ao citarem, por exemplo, o termo Funções Executivas indicaram bibliografias diversas. Entretanto, incluem textos como de Miyake e Friedman (2012), Diamond (2013) e Miyake et al. (2000). Similarmente, a respeito do construto atenção seletiva, pôde-se identificar que nenhum dos textos que referenciaram a avaliação desta função por meio do teste Stroop se propôs a conceitualizá-la. Os autores, entretanto, citaram outros textos que fazem referência à utilização do teste para este propósito. O mesmo ocorreu com os construtos velocidade de processamento e flexibilidade cognitiva. Ambos foram citados como funções cognitivas alvo na utilização do Stroop, no entanto, nem suas definições conceituais quanto operacionais foram descritas nestes trabalhos.

Foi possível identificar que dentre os construtos levantados alguns correspondem aos construtos definidos no modelo hipotético de Funções Executivas (FEs) proposto por Diamond (2013), representado por três unidades básicas: inibição, memória operacional e flexibilidade. Entretanto, a ausência de justificativa teórica ou menção ao modelo teórico adotado, e a omissão da conceitualização dos construtos mencionados não permite ao leitor identificar se o uso do teste se destinaria à mensuração destes subcomponentes, isoladamente, ou se se referem a avaliação do conceito geral de FEs, o que retoma a discussão sobre a necessidade das publicações científicas justificarem suas metodologias de pesquisa com o devido embasamento teórico.

Por ter sido considerado um instrumento padrão ouro na avaliação neuropsicológica, sua utilização na prática clínica se dá em diversos contextos e em diversos grupos de pessoas. Pesquisas apontam que o teste Stroop é uma medida neuropsicológica sensível ao aumento da idade, o que significa que os resultados da medida variam de acordo com mudanças na faixa etária, dado este, que justifica sua frequente aplicação na população idosa. Davison, Deeprouse e Terbeck (2018) demonstraram empiricamente aumento da interferência na tarefa Stroop ao comparar um grupo de adultos idosos (com participantes com 65 anos ou mais) com um grupo de adultos jovens (participantes entre 18 e 25 anos). Os dados apontam que a média de respostas corretas na etapa “cor e palavra” foi de 111.27 (dp 1.32) para o grupo jovem e 92.28 (dp 28.38) para o grupo idoso, indicando uma diferença estatisticamente significativa ( $U = 57$ ,  $z = -4.1$ ,  $p < 0.001$ ), o que representa que adultos jovens obtiveram significativamente maior número de acertos. Demais aspectos como número de erros e tempo de execução não foram avaliados.

Klojčnik, Kavcic, e Bakracevic Vukman (2017) apontaram a relação entre depressão e déficits cognitivos e utilizaram o teste Stroop como integrante de uma bateria para avaliar a hipótese de que estas duas variáveis são negativamente correlacionadas. Como resultados, identificaram que altos escores na escala de depressão BDI (*Beck Depression Inventory* traduzido para o português como Inventário de Depressão de Beck) são correlacionados com baixa performance no teste Stroop ( $r = -0,69$ ). Ademais, os resultados da análise de regressão mostraram que o teste Stroop e o teste Rey-Osterrieth de recordação (traduzido para o português como Teste da Figura Complexa de Rey-Osterrieth) previram significativamente o desempenho na BDI, explicando uma variação de 70% ( $F(2, 69) = 82,14$ ,  $p < 0,0005$ ,  $R^2_{Adj} = 0,70$ ). O preditor mais forte foi o teste de Rey-Osterrieth ( $\beta = -0,67$ ,  $p < 0,0005$ ), seguido pelo desempenho no teste de Stroop ( $\beta = -0,23$ ,  $p = 0,015$ ). Ambos os preditores foram negativamente correlacionados com a variável critério, o que indica que escores mais baixos nos testes de Rey-Osterrieth e Stroop são preditivos de escores mais altos de depressão.

Nowakowska-Domagala, Jabłowska-Górecka, Mokros, Koprowicz, e Pietras, (2017) também indicaram sensibilidade do teste Stroop ao comparar pacientes controle com um grupo de pacientes dependentes de álcool. Encontraram resultados que apontam diferenças significativas entre os grupos em relação à velocidade de desempenho do Stroop na etapa incongruente. Os sujeitos do grupo controle realizaram a execução do subteste em um tempo significativamente menor do que os pacientes dependentes ( $t = 6.87$ ,  $p < 0,001$ ).

Resultados que indicam acurácia na identificação dos mais variados critérios impulsionam a utilização do Teste Stroop em estudos empíricos em diversas áreas do conhecimento. O levantamento realizado nesta pesquisa aponta que a avaliação de funções cognitivas, por meio deste instrumento, é alvo não somente da área da psicologia e neuropsicologia, mas também, da psiquiatria, da medicina, das neurociências, da educação física, da enfermagem, entre outros, fato que pode ser identificado pelo extenso número de contextos e quadros clínicos dos participantes avaliados.

Em relação ao formato de apresentação e execução do teste, como por exemplo, quantidade de cores utilizadas, formato de apresentação das manchas de tinta, etapas e modelos de pontuação, as versões identificadas apareceram tanto de modo idêntico ao original, quanto de maneiras distintas. Tais divergências levantam questões referentes ao construto que se pretende medir. Ao relacionar uma grande variação de versões, suspeita-se de quais modelos possuem maior eficiência no alcance do, ou dos construtos. Entretanto, para tal inquietude, faz-se necessário levantar a efetividade dos diferentes métodos tomando como critério um único construto, e metodologicamente, identificar qual modelo ou versão aponta maior precisão em sua avaliação.

#### 4.5 CONCLUSÃO

O objetivo desta revisão integrativa foi de levantar dados sobre os construtos, contextos em que o teste foi empregado, além das versões utilizadas. Os resultados confirmaram sua ampla utilização em todos os aspectos investigados. Acerca dos construtos, os dados indicaram diversidade em relação as funções ou habilidades avaliadas pela tarefa. Após análise inicial, as funções consideradas sinônimas foram agrupadas, no entanto, ainda se manteve tal pluralidade. Todavia, houve reiteração de construtos como inibição, funções executivas, atenção, velocidade de processamento e flexibilidade cognitiva.

Destacou-se, também, que inúmeros estudos não se propuseram a definir os construtos mencionados, como por exemplo, quando citado como um instrumento para avaliação de funções executivas (FEs) - construto que historicamente é considerado de difícil definição por sua complexidade operacional. Nestes trabalhos, não foi possível identificar se os autores

estariam se referindo as FEs como um todo, de acordo com algum modelo teórico, ou se estariam avaliando seus subcomponentes de modo isolado.

Sobre as versões, dentre os artigos revisados, 62% dos não indicaram a versão do teste utilizada, mencionando apenas o seu emprego como uma medida neuropsicológica. Foram levantadas, também, variações em diversos aspectos do teste, como em relação a adaptações e traduções, ao emprego de palhetas e quantidades de cores, formato de execução, tarefas e método de correção. Em relação ao último objetivo deste estudo, a respeito dos participantes, grupos de referência e amostras o estudo encontrou tipos distintos, no entanto, a maior frequência foi de pacientes adultos saudáveis.

Em suma, os dados reforçam a multiplicidade de construtos, versões e contextos em que o teste é empregado para avaliação de funções neuropsicológicas. Por um lado, mostra-se factual que o Teste Stroop é uma tarefa neuropsicológica constantemente utilizada para fins de pesquisa e, também, de avaliação, entretanto, a diversidade traz consigo questionamentos sobre validade, indagando a eficiência da medida. Assim, este artigo indica a necessidade de estudos que mostrem a relação entre os diferentes construtos e suas versões, sugerindo pesquisas com diferentes grupos de referência. Ademais, propõe-se verificar correlações entre o teste e outras medidas neuropsicológicas que avaliam os mesmos construtos citados, objetivando identificar a similaridade, ou discrepância, entre as diferentes tarefas.

## 5. CAPÍTULO III

### **ESTUDOS EMPÍRICOS DE EVIDÊNCIAS DE VALIDADE DO TESTE STROOP DE CORES E PALAVRAS**

Os objetivos deste capítulo compreenderam a realização de estudos empíricos de evidências de validade do teste Stroop de Cores e Palavras. A construção metodológica desses ensaios baseou-se nos dados descritos nos capítulos I e II desta tese. O Capítulo I apresentou aspectos teóricos do teste Stroop e também a versão escolhida, referida na literatura como versão Victoria. Essa seção também exibiu uma revisão de literatura sobre resultados de pesquisas em diferentes contextos e participantes nos quais o teste é frequentemente utilizado, e demarcou seu campo de utilização, a avaliação neuropsicológica de funções executivas.

O Capítulo II apresentou o artigo “Teste Stroop de Cores e Palavras: uma revisão integrativa de literatura”, que indicou os construtos considerados mensuráveis na utilização do teste. A partir dos dados obtidos nessa pesquisa bibliográfica, foram selecionados os construtos que seriam empregados nos estudos de validade convergente, divergente e de construtos relacionados. Com base na definição desses construtos, foram escolhidas as medidas neuropsicológicas que possibilitaram a realização dos estudos empíricos de validade. Além disso, o capítulo apresentou as diferentes versões do teste Stroop e seus respectivos indicadores para a formação de escores, o que reforçou a adoção dos critérios previamente estabelecidos para a análise dessa pesquisa.

Este capítulo, ora apresentado, intitulado “Capítulo III - Estudos empíricos de evidências de validade do teste Stroop de Cores e Palavras” corresponde, portanto, à realização de estudos empíricos com o objetivo de verificar evidências de validade do teste Stroop. Essa divisão compreendeu os objetivos específicos: desenvolver estudos de evidências de validade convergente do instrumento; desenvolver estudos de evidências de validade divergente; e desenvolver estudos de evidências de validade por construtos relacionados.

Considerada como tópico central na teoria da medida, a validade é atualmente referida como “grau em que evidência e teoria suportam as interpretações das pontuações dos testes para

os usos propostos” (*American Educational Research Association AERA, American Psychological Association APA, National Council on Measurement in Education NCME, 1999*). De acordo com o modelo atual, consagrado no *Standards for educational and psychological testing* (AERA, APA, NCME, 1999), toda validação é tratada como um processo de teste hipóteses, alterando o modo como se pensava a validade. Antes, questionava-se qual o “tipo de validade” um determinado estudo apoiava, agora, verifica-se se as hipóteses fazem sentido a partir do referencial em que o teste foi projetado e se os resultados permitem fazer inferências sobre as pessoas (Streiner, Norman & Cariney, 2015).

O modelo proposto pelo *Standards for educational and psychological testing* objetiva o levantamento de evidências. Essas figuram como pontos basilares para as interpretações dos escores dos testes. Ressalta-se, ainda, que neste processo, o que se avalia são as evidências que sustentam as interpretações dos escores obtidos pelo instrumento, e não se o teste psicológico pode ser considerado válido (AERA, APA, NCME, 1999). Se um instrumento não possui estudos indicando boas evidências de validade, não há segurança de que as interpretações realizadas com base nas respostas obtidas sejam plausíveis (Lins & Borsa, 2017).

Primi, Muniz e Nunes (2009) ilustram o conceito de validade através da exemplificação das evidências que justificam a associação entre os comportamentos avaliados pelos testes e os traços latentes. Para eles, os instrumentos de avaliação demandam aos testandos a realização de tarefas específicas. A partir do registro e da observação dos comportamentos, os profissionais realizam inferências sobre a presença de determinadas características psicológicas. Com base na hipotetização de que os comportamentos dos testandos são determinados pelas suas características psicológicas, denominadas de traços latentes, supõe-se que as respostas dos testes permitem conhecer, portanto, tais características. A confirmação ou não desta associação depende de estudos fortemente embasados teoricamente que busquem verificar empiricamente se as relações entre variáveis ou resultados de indicadores externos são obtidos conforme esperado. Deste modo, quando se confirma a legitimidade das interpretações, por meio de um conjunto de evidências, diz-se que o teste possui validade para aquelas interpretações, não podendo ser considerado válido para todos os tipos de contextos, propósitos ou públicos.

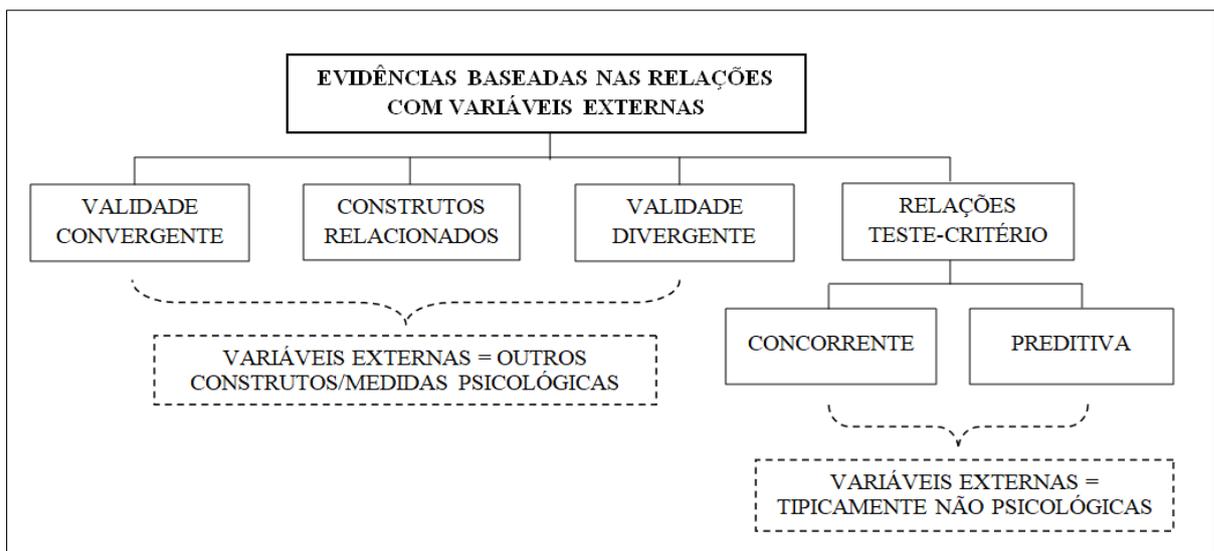
O *Standards for educational and psychological testing* (AERA, APA, NCME, 2014) cita cinco fontes a serem utilizadas como evidências de validade, que são: a) evidências baseadas no conteúdo; b) evidências baseadas no processo de resposta; c) evidências baseadas

na estrutura interna; d) evidências baseadas nas relações com outras variáveis; e e) evidências baseadas nas consequências da testagem.

As evidências baseadas nas relações com variáveis externas (também denominada evidências baseadas nas relações com outras variáveis) (AERA, APA, NCME, 2014), parâmetro utilizado para as análises desta tese, buscam identificar se os resultados empíricos convergem com o que teoricamente é esperado. Classicamente, os procedimentos envolvem a verificação de evidências de validade convergente, divergente e as relações teste-critério (Reeves & Marbach, 2016).

### Figura 5

*Evidências baseadas nas relações com variáveis externas*



Por validade convergente, entendem-se as relações significativas entre os escores do teste alvo e outras medidas que já detenham evidências consistentes de validade. A convergência pode se destinar a avaliar a correlação entre testes que mensuram os mesmos construtos (AERA, APA, NCME, 2014).

Quando a validade convergente utiliza como método a correlação entre dois testes que se destinam ao alcance do mesmo construto, espera-se que estes dois instrumentos apresentem empiricamente altas correlações entre si. Neste caso, os coeficientes de correlação esperados devem ser de elevada magnitude, acima de 0,509 (Nunes & Primi, 2010). Nesta pesquisa, por

exemplo, serão utilizados simultaneamente testes ou tarefas que avaliam funções ditas como equivalentes em termos de construto teórico ao teste Stroop.

Entretanto, quando o teste que será utilizado como medida comparativa ao teste alvo não se destina a avaliação do mesmo construto, denomina-se evidências de validade por construtos relacionados. Nesta proposta, os coeficientes de correlação esperados são de moderada magnitude, variando entre 0,20 a 0,50 (Nunes & Primi, 2010).

Já a validade divergente ou discriminante, ocorre quando as relações entre os escores do teste e as medidas - consideradas diferentes ou não relacionadas ao construto alvo - mostram-se discriminativas (AERA, APA, NCME, 2014). Opostamente aos estudos de validade convergente, ocorre a busca de evidências de que os testes avaliam construtos distintos, antecipando a hipótese de ausência de relação entre as medidas (Primi, Muniz & Nunes, 2009). Em se tratando do dado estatístico empírico, esperam-se coeficientes de correlação baixos, inferiores a 0,20 (Nunes & Primi, 2010).

Nunes e Primi (2010) indicam que, na prática, os estudos que buscam evidências de validade por meio de comparações entre testes psicológicos, exploram simultaneamente evidências convergentes e discriminantes. Os autores apontam, ainda, que ao encontrar resultados compatíveis com os esperados teoricamente, sustenta-se a hipótese de que, além de convergir com medidas semelhantes, o teste possui nível de especificidade adequado.

Outra forma de obter dados de validade pode ser por meio de relações teste-critério. Nesta perspectiva, uma, ou um conjunto, de variáveis (tipicamente não psicológicas) são identificadas como critérios externos e, em acordo com o referencial teórico, empírico e lógico, se associam ao construto avaliado. Consistem em eventos importantes, ou aspectos relevantes na vida das pessoas, estreitamente relacionados com os construtos avaliados pelo teste. De acordo com Primi, Muniz e Nunes (2009, p.255) “a observação de associações significativas entre teste-critério pode adjudicar validade ao instrumento, especialmente sobre a relevância e utilidade em prever certas situações importantes em ambientes aplicados específicos”.

Uma forma de alcançar este tipo de evidência de validade é a diferenciação de grupos de pessoas (AERA, APA, NCME, 2014). Pasquali (2017) cita que esta pode ser obtida por diversas formas, e que, comumente, utilizam-se indicadores como desempenho acadêmico, desempenho em treinamento especializado, desempenho profissional, diagnóstico psiquiátrico

e diagnóstico subjetivo como variáveis de comparação. Neste caso, por exemplo, estudos poderiam ser conduzidos a fim de verificar a capacidade de um dado teste identificar pessoas que apresentam um determinado quadro psicopatológico (Nunes & Primi, 2010).

A relação teste-critério historicamente se distingue entre concorrente e preditiva. Por validade concorrente, entende-se as relações estabelecidas entre o teste e a variável externa quando as duas ocorrem ao mesmo tempo (Reeves & Marbach, 2016). Diferentemente da situação anterior - quando a aplicação do teste e a verificação do critério não ocorrem no mesmo momento - denomina-se validade preditiva. Um estudo preditivo indica, portanto, a força da relação entre os resultados do teste e os critérios, estes últimos, obtidos a posteriori. Um exemplo pode ser a administração de um teste de habilidade matemática em estudantes do Ensino Fundamental e a posterior verificação das notas obtidas ao final do ano letivo, variável definida como critério (AERA, APA, NCME, 2014).

## 5.1 MÉTODO

Este capítulo, intitulado “Estudos empíricos de evidências de validade do teste Stroop de Cores e Palavras” apresenta os estudos empíricos realizados com o objetivo de verificar as evidências de validade do teste a partir da busca de evidências baseadas nas relações com variáveis externas, como validade convergente, divergente e construtos relacionados. Para este último, fora planejado um delineamento de pesquisa que contemplasse amostras de pessoas com características clínicas e não clínicas diversas, como por exemplo, usuários dependentes de drogas/álcool, pessoas com diagnóstico de Transtorno do Déficit de Atenção e Hiperatividade – TDAH, jogadores patológicos, idosos e, também, a população em geral.

Contudo, o cronograma de coleta de dados foi atravessado pelo panorama nacional e internacional decorrente da pandemia de SARS-CoV-2 - nome oficial do vírus que causa a Covid-19. Em decorrência das medidas de isolamento e distanciamento social impostas, a coleta de dados foi adiada, já que toda a aplicação da pesquisa compreenderia a administração de testes em modo presencial com os participantes. Tais instrumentos de medida escolhidos para execução deste trabalho possuem aplicação estritamente presencial, não sendo possível realizar quaisquer adaptações que pudessem viabilizar esta pesquisa em formato remoto. Durante o

período em que as medidas de distanciamento social se encontravam mais rígidas, esta etapa, portanto, ficou suspensa.

No momento em que as medidas de distanciamento social se tornaram mais brandas, a coleta de dados deu-se início. Entretanto, foi necessária a alteração do grupo amostral previamente selecionado para participação - que contemplariam as etapas de busca de evidências de validade baseados em critérios externos, mais especificamente, os grupos contrastantes ou clínicos - e assim, optou-se pela realização da aplicação somente com a população geral. Tal escolha tornou-se necessária tendo em vista que, por exemplo, o ingresso em instituições como casas de repouso para idosos, hospitais, clínicas de reabilitação para dependentes químicos e para dependentes de álcool, e os demais locais citados nos campos de pesquisa, no cenário atual pandêmico, tornaram-se desaconselháveis por questões de biossegurança. Por fim, após o abrandamento dos índices de transmissão comunitária da Covid-19 e as necessárias reformulações da pesquisa, a coleta de dados foi iniciada com tais modificações.

### **5.1.1 Caracterização da pesquisa**

Os objetivos deste capítulo relacionam-se à busca de evidências de validade pela relação com variáveis externas, mais especificamente, validade divergente, validade convergente e por construtos relacionados. Para isto, proceder-se-á à coleta de dados e esta será baseada na mensuração dos construtos por meio de instrumentos psicológicos padronizados - procedimento de natureza quantitativa - na qual o produto destes dados serão representados por números e analisados com base em métodos estatísticos e psicométricos. As análises quantitativas serão interpretadas com base nas hipóteses iniciais e/ou estudos anteriores, buscando uma perspectiva de como os resultados obtidos dialogam com o conhecimento existente (Sampieri, Callado & Lucio, 2013).

Prioritariamente, as análises compreenderam estudos correlacionais, que circundam sobre a busca pela compreensão de relações e associações entre variáveis. Tais estudos se propõe a analisar a força das relações, bem como sua direção e magnitude, não tendo como alvo a identificação de causalidade e efeito (Sousa, Driessnack & Mendes, 2007).

### **5.1.2 Caracterização do campo de pesquisa**

A coleta de dados foi realizada nas cidades de Curitiba-PR e Colombo-PR, município pertencente à região metropolitana da capital. Todos os locais foram escolhidos por conveniência após a reestruturação metodológica decorrente da pandemia de SARS-CoV-2. A aplicação ocorreu em duas clínicas de psicologia e medicina que se destinam a avaliação médica e psicológica de candidatos a obtenção/renovação da Carteira Nacional de Habilitação, ambas localizadas na cidade de Colombo-PR (região metropolitana de Curitiba) e em uma instituição privada de ensino superior localizada na cidade de Curitiba-PR (APÊNDICE 1, 2 e 3, respectivamente). As pessoas que compuseram a amostra foram convidadas a participar da pesquisa após prévia autorização dos campos referidos (APÊNDICE 4).

### **5.1.3 Procedimentos éticos**

Esta pesquisa foi submetida ao Comitê de Ética e Pesquisas com Seres Humanos da Universidade Federal de Santa Catarina. Os participantes que sinalizaram interesse em participar foram orientados sobre os cuidados éticos que envolvem a sua participação, como, por exemplo, quanto ao anonimato de seus dados de identificação; aos possíveis riscos e benefícios; ao caráter voluntário e revogável de sua participação durante qualquer momento da etapa; à toda e qualquer informação sobre a condução e sobre os responsáveis da pesquisa; e demais informações que estarão contidas no Termo de Consentimento Livre e Esclarecido, que foi assinado mediante apresentação detalhada e anterior a realização da pesquisa. Somente após a assinatura deste, deu-se início a coleta de dados.

### **5.1.4 Procedimentos de coleta de dados**

Inicialmente, foi encaminhado aos locais de coleta de dados uma carta de solicitação para a realização da pesquisa em suas respectivas dependências. Após o recebimento das autorizações, o projeto foi submetido ao Comitê de Ética, recebendo parecer favorável, n.4.137.183, em anexo (ANEXO 1). Sequencialmente à autorização do Comitê de Ética, a amostra foi selecionada por conveniência. Os participantes que contemplaram a amostra

estavam nas dependências físicas dos locais de aplicação, e foram convidados para a realização da pesquisa.

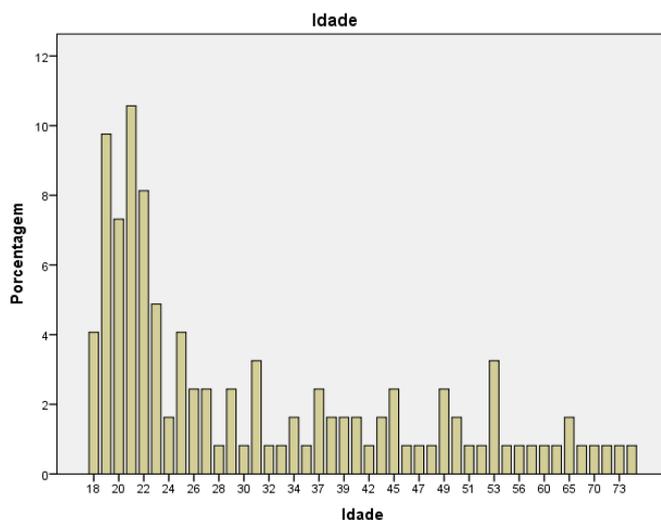
Ao participante foi solicitada a leitura e assinatura do Termo de Consentimento Livre e Esclarecido e somente após o consentimento deste a aplicação dos testes teve início. Foram esclarecidos aos participantes o objetivo da pesquisa e reforçado o seu caráter ético e sigiloso. A aplicação ocorreu de modo individual, com alternância na ordem de execução dos instrumentos, com o objetivo de não enviesar os dados caso uma ordem fixa fosse mantida. A aplicação durou, em média, 50 minutos, devido à extensão do número de instrumentos utilizados para cada participante, em sessão única.

### 5.1.5 Participantes

A amostra foi composta por 128 participantes com idades entre 18 e 85 anos. A média de idade dos participantes foi de 32,5 anos (DP=15,21). Em relação ao gênero, 81 pessoas se identificaram com o gênero feminino e 47 com o gênero masculino. Nenhum participante se identificou com outra nomenclatura de gênero. A respeito da escolaridade, a amostra foi composta por participantes com ensino fundamental, médio, superior e pós-graduação, tanto completos como incompletos.

#### Figura 6

*Distribuição da amostra em relação a idade dos participantes*



**Tabela 3***Distribuição da amostra em relação à escolaridade*

	<b>Escolaridade</b>	<b>Frequência</b>	<b>Porcentagem</b>	<b>% válida</b>	<b>% acumulativa</b>
Válido	Ensino Fundamental	5	3,9	4,2	4,2
	Ensino Médio	31	24,2	26,1	30,3
	Superior Incompleto	44	34,4	37,0	67,2
	Superior Completo	39	30,5	32,8	100,0
	Total	119	93,0	100,0	
Ausente	Sistema	9	7,0		
Total		128	100,0		

Para serem incluídos os participantes atenderam aos seguintes critérios de inclusão, sendo eles: a) ter acuidade visual satisfatória (com ou sem o uso de lentes corretivas) que permitam a leitura de palavras, a identificação de figuras, desenhos e demais estímulos que compõem os testes padronizados. Assim, excluiu-se, portanto, pessoas com deficiência visual total ou parcial que façam uso de instrução em Braille (sistema de escrita por pontos em relevo) ou por *softwares* de leitura de textos. b) ser alfabetizado e fluente na leitura de palavras simples; e c) ser capaz de identificar cores por meio de captação visual. Excluiu-se com base neste critério pessoas detentoras de daltonismo ou que sejam incapazes de diferenciar as cores vermelha, verde, azul e amarela. Tais informações foram obtidas por meio de uma entrevista inicial, em que os questionamentos se direcionaram a identificar se os critérios seriam satisfatoriamente preenchidos para a realização de toda a bateria de testagem.

### 5.1.6 Instrumentos

O Teste Stroop de Cores e Palavras, versão Victoria, consiste em três cartões/tarefas, organizados na seguinte ordem: 1º Cartão Pontos, 2º Cartão Palavra e 3º Cartão Cor-Palavra. Cada cartão contém seis linhas com quatro itens por linha. Em todas as tarefas, é solicitado que o participante examine todos os itens nas linhas, realizando-as da esquerda para a direita de modo rápido e preciso. Na primeira tarefa, “Cartão Pontos”, os itens são pontos coloridos, impressos nas cores verde, vermelho, azul e amarelo, e é solicitado que o participante diga a cor em que os pontos estão impressos. Na segunda tarefa, “Cartão Palavras”, os itens são palavras comuns (ex: água, quando, difícil, sobre) impressas nas mesmas cores já apresentadas

no cartão anterior (verde, vermelho, azul e amarelo). Novamente, os participantes devem nomear a cor das palavras. Na terceira e última tarefa, “Cartão Cor-Palavra” (onde é observado o efeito Stroop), os itens são nomes de cores (ou seja, as palavras vermelho, azul, amarelo e verde) impressas em cores que não correspondem às próprias palavras (por exemplo, a palavra VERMELHO estará impressa na cor VERDE). Nesta etapa, os participantes precisam nomear a cor de tinta em que a palavra está impressa, situação que não será congruente (ex: a palavra VERMELHO não estará impressa na cor vermelha). Nesta tese, foram calculados os tempos de execução em segundos, acertos e erros para cada uma das três etapas. Adicionalmente, foi realizada a média aritmética entre os tempos do cartão cor e do cartão palavras, valor este que foi subtraído do tempo total de execução do cartão cor-palavra. Tal subtração gerou o resultado denominado escore de interferência (tempo de execução da última etapa menos a média aritmética das duas primeiras etapas). Além do escore de interferência também foi registrado um escore derivado que consistiu na subtração do tempo gasto na etapa três menos o tempo gasto na etapa um.

Os instrumentos que foram utilizados para o desenvolvimento dos estudos de validade serão: Teste dos Cinco Dígitos - FDT; Wisconsin Teste de Classificação de Cartas - WCST; Teste de Avaliação da Atenção Seletiva - EASV; BETA-III – Subteste Códigos; e *Hayling Sentence Completion Test* (HSCT). Os quatro primeiros - FDT, WCST, EASV e BETA - são instrumentos validados e padronizados para a população brasileira. O quinto, é amplamente referenciado na literatura estrangeira para a avaliação de seus respectivos construtos.

O Teste dos Cinco Dígitos (FDT) trata-se de um instrumento para a avaliação do efeito de interferência atencional. O teste é realizado em quatro etapas. Na primeira etapa denominada Leitura, o examinando deve ler o algarismo/número contido dentro de um quadrado em uma condição congruente. Na segunda, Contagem, é necessário realizar a contagem de asteriscos dentro dos quadrados. A terceira etapa, Escolha, exige que o participante conte a quantidade de números no quadrado alvo e inibir a leitura dos mesmos - envolvendo a capacidade de controle de componentes executivos, como o controle inibitório. A quarta e última etapa, Alternância, corresponde a variação alternada entre a contagem dos números que estão dentro dos quadrados e a leitura/nomeação dos algarismos. Para o escore, considera-se o tempo necessário para a execução da tarefa e os erros cometidos pelo indivíduo. Para o cálculo dos escores de interferência, subtrai-se o tempo de Leitura do tempo de Escolha e do tempo de Alternância, criando, então, os índices de Inibição e Flexibilidade (Malloy-Diniz; Costa; Querino; de Paula,

2014). Neste trabalho, os seis índices possíveis do FDT supramencionados foram utilizados nas análises correlacionais. O tempo médio de realização foi de cinco minutos.

O Wisconsin de Classificação de Cartas - WCST - é um teste largamente utilizado na avaliação de funções executivas, baseado na flexibilidade de pensamento do sujeito. Almeja além de monitorar, regular e inibir comportamentos automatizados, flexibilizar o curso do pensamento a partir do planejamento de estratégias para a solução de problemas (Cunha, Trentini, Argimon, Oliveira, Werlang & Prieb, 2005). A versão utilizada neste trabalho trata-se de uma versão reduzida de 48 cartas, proposta por Zimmermann, Hermes-Pereira e Fonseca (2017), composta por quatro cartas-estímulo e 48 cartas-resposta, que apresentam figuras de variadas formas geométricas (cruzes, triângulos, círculos e estrelas), com quatro cores distintas (vermelho, azul, amarelo ou verde) e de diferentes quantidades (uma, duas, três ou quatro figuras em cada). A correção compreende sete escores, que são: categorias completadas, (considerada uma medida geral do teste), acertos, ensaios administrados, erros totais, erros do tipo perseverativo, erros não perseverativos e rupturas. Todos os índices do WCST foram utilizados para as análises correlacionais.

A Escala de Atenção Seletiva Visual (EASV) avalia a capacidade do indivíduo de selecionar estímulos visuais identificando a semelhança entre o modelo apresentado e os demais estímulos disponíveis para seleção. A aplicação pode ocorrer tanto na forma individual quanto coletiva e possui tempo de resolução dos itens do teste de 9 minutos. Compreende um conjunto de 30 itens representados por formas geométricas. Solicita-se que o participante observe a primeira figura modelo, da primeira linha, e localize a opção mais próxima a ela entre quatro possíveis respostas. Para a pontuação do teste, os acertos equivalem a um ponto e os erros, a zero pontos (Sisto & Castro, 2010). Os resultados fornecem quatro escores de acordo com o nível de dificuldade dos itens, sendo eles, “capacidade seletiva visual básica”, “capacidade seletiva visual mediana”, “capacidade seletiva visual superior” e, o somatório denominado “pontuação total” se caracteriza pela soma de pontos obtidos (total de acertos), independentemente do nível de exigência envolvido nos diferentes itens do teste. Os quatro escores possíveis foram utilizados nas análises correlacionais deste trabalho.

BETA-III – Subteste Códigos. Este instrumento configura-se como uma medida de avaliação da velocidade de processamento. Nele, o modelo que deve ser observado consiste em símbolos que correspondem a determinados números. O examinador deve instruir os

respondentes a completarem os espaços em branco com os números respectivos, indicados no modelo, em um tempo de 2 minutos. Cada resposta correta corresponde a um ponto, e o escore total, único índice de pontuação, pode variar de 3 a 137 pontos. O teste pode ser administrado no formato individual e, também, coletivo (Rabelo, Pacanaro, Leme, Ambiel & Alves, 2011).

Hayling Sentence Completion Test (HSCT) é um teste para a avaliação do controle inibitório, considerado como uma medida de iniciação e supressão de resposta. Consiste em dois conjuntos de 15 frases, em que a última palavra da frase se encontra ausente. Na primeira sessão (condição automática), o examinador lê cada sentença em voz alta e o participante precisa simplesmente completar as sentenças. Na segunda sessão (condição de inibição), solicita-se que o participante complete as frases com uma palavra que não usual, ou que não se encaixa, o mais rápido possível, fornecendo medidas de capacidade de supressão de resposta e tempo de reflexão, em que o respondente precisa inibir uma resposta fortemente ativada, ou automática, e gerar uma nova resposta. Os tempos de latência de resposta (de ambas sessões) são registrados em segundos e transformados em escores. A latência de tempo na condição automática avalia a rapidez em iniciar uma resposta usual; já na condição de inibição fornece informações sobre o tempo necessário para inibir a resposta dominante e encontrar uma resposta incomum (Bayard, Gély-Nargeot, Raffard, Guerdoux-Ninot, Kmara, Gros-Balthazard, 2017). Os índices utilizados para as análises foram os tempos, em segundos, das etapas A e B, e, também, a subtração do tempo executado em B menos o tempo executado em A, denominado escore de interferência. A execução durou, em média, entre cinco e dez minutos.

### **5.1.7 Análise de dados**

Os dados coletados com a aplicação do Teste Stroop de Cores e Palavras e demais testes foram submetidos à análise estatística descritiva e inferencial, considerando erro tipo *alpha* (ou erro tipo I) de 0,05. As análises realizadas objetivaram a verificação das hipóteses levantadas sobre as correlações entre os escores do teste Stroop e as demais medidas do presente estudo. Os dados foram analisados por meio do pacote estatístico SPSS versão 22.0.

A escolha de procedimentos paramétricos ou não paramétricos foi pautada na verificação da distribuição das principais variáveis analisadas, tanto do teste Stroop, quanto das demais tarefas utilizadas. A normalidade das distribuições foi verificada pela observação do

histograma e dos dados descritivos, que consideraram os índices de curtose e assimetria ideais próximos à zero. Por curtose entende-se a elevação ou achatamento da distribuição quando comparada à distribuição normal, e o indicador de assimetria é uma medida da falta de simetria na distribuição de frequência dos dados observados, também comparada à curva normal (Hair et al., 2009). Distribuições semelhantes à normal apresentam assimetria com valores próximos a zero.

A análise estatística inferencial compreendeu a utilização dos testes de correlação de Pearson - quando os dados apresentaram distribuição paramétrica - e de Spearman - condições não paramétricas - para realizar as análises correlacionais entre os escores do Stroop e das demais medidas/testes utilizados. A análise da Variância (ANOVA) - medida estatística utilizada para comparar as médias de três ou mais grupos diferentes e verificar se as diferenças observadas são estatisticamente significativas - foi empregada na verificação das diferenças entre escolaridade e faixas etárias da amostra. O teste *t de Student* – usado para determinar se há uma diferença estatisticamente significativa entre as médias de duas amostras independentes - foi empregue para comparação entre os gêneros dos participantes.

### **5.1.8 Hipóteses de estudo**

- a. Correlações significativas e de elevadas magnitudes entre os índices do Teste Stroop e os índices do Teste dos Cinco Dígitos – FDT. Tal hipótese se amparou, primeiramente, pelo motivo do FDT ter sido originalmente desenvolvido com base no efeito de interferência existente no Stroop. Ademais, o FDT possui um formato de execução e de apresentação de estímulos semelhante ao teste Stroop, bem como, possui também similaridade nos índices que contemplam os escores.
- b. Correlações significativas e de moderadas magnitudes entre o Teste Stroop e o Wisconsin Teste de Classificação de Cartas – WCS, considerando que ambos os testes, de acordo com o levantamento bibliográfico indicado no Capítulo II, são indicados como instrumentos para avaliação das Funções Executivas.
- c. Correlações significativas de moderadas ou elevadas correlações magnitudes entre o Teste Stroop e o Teste de Avaliação da Atenção Seletiva – EASV, considerando que, na literatura, ambos os testes se propõem a avaliar atenção seletiva.
- d. Correlações significativas de moderadas magnitudes entre o Teste Stroop (em sua totalidade ou em cartões específicos) e o BETA-III – Subteste Códigos. Esperava-se

obter moderados coeficientes de correlação considerando que, na literatura, ambos os testes se propõem a avaliar velocidade de processamento. Nesta análise, dar-se-á atenção para a verificação dos coeficientes de correlação entre os três cartões do Stroop, considerando a hipótese de que maiores correlações serão encontradas nas duas etapas iniciais do teste, “Cartão Cor” e “Cartão Palavra”, e na análise do tempo de execução de cada cartão.

- e. Correlações significativas de moderadas magnitudes entre o Teste Stroop e o teste *Hayling Sentence Completion Test* (HSCT), já que ambos os testes se propõem a avaliar inibição/controle inibitório. Adicionalmente, a expectativa de tais correlações foi baseada no formato de execução dos testes e nas tarefas exigidas. Em ambos, uma das demandas é considerada como um processamento automático e a outra um processamento não automático. A correção dos testes, similarmente, também compreende a subtração dos tempos de execução destas etapas, fornecendo um escore de interferência.

## 5.2 RESULTADOS

### 5.2.1 Normalidade da amostra

Primeiramente foram realizadas análises com o intuito de verificar a distribuição dos dados em relação à curva normal. Além dos valores de média e desvio padrão, foram analisados os valores de assimetria e curtose de todas as variáveis analisadas. Após testada a normalidade de dois escores do teste Stroop, os resultados, de acordo com a análise pelo Teste de Shapiro-Wilk, demonstraram valores de  $p$  menores que 0,05 ( $p < 0,001$ ), o que indicam distribuição não paramétrica (Tabela 4 e Figura 7). Tal resultado é convergente com os valores de curtose e assimetria verificados.

**Tabela 4**

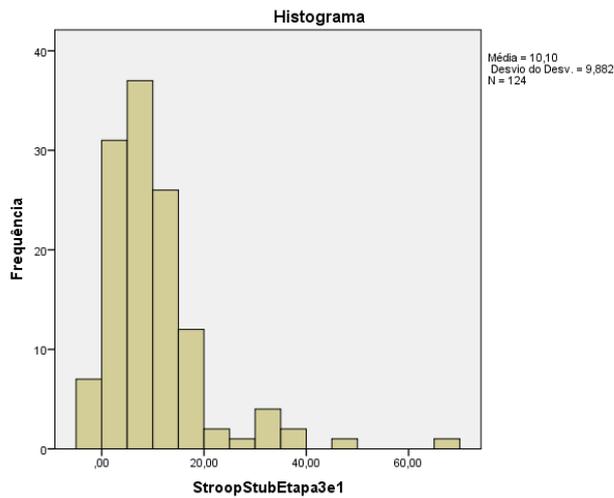
*Teste de normalidade para o Teste Stroop de Cores e Palavras*

Teste	Shapiro-Wilk	Média	DP	Assimetria	Curtose
	W	df	$p$		

Stroop etapa 3 – etapa 1	0,78	124	0,001	10,10	9,88	2,52	9,39
Stroop – escore interferência	0,79	124	0,001	8,87	8,31	2,33	7,74

**Figura 7**

*Histograma de normalidade para o teste Stroop*



A partir destes dados foi realizada a logaritimização da variável escore de interferência, com objetivo de verificar se tal transformação influenciaria a distribuição dos dados, permitindo a utilização de testes estatísticos paramétricos. Novamente testada a normalidade, os resultados se mostraram satisfatórios, atendendo aos pré-requisitos de uma distribuição gaussiana, como pode ser observado na Tabela 5 e na Figura 8.

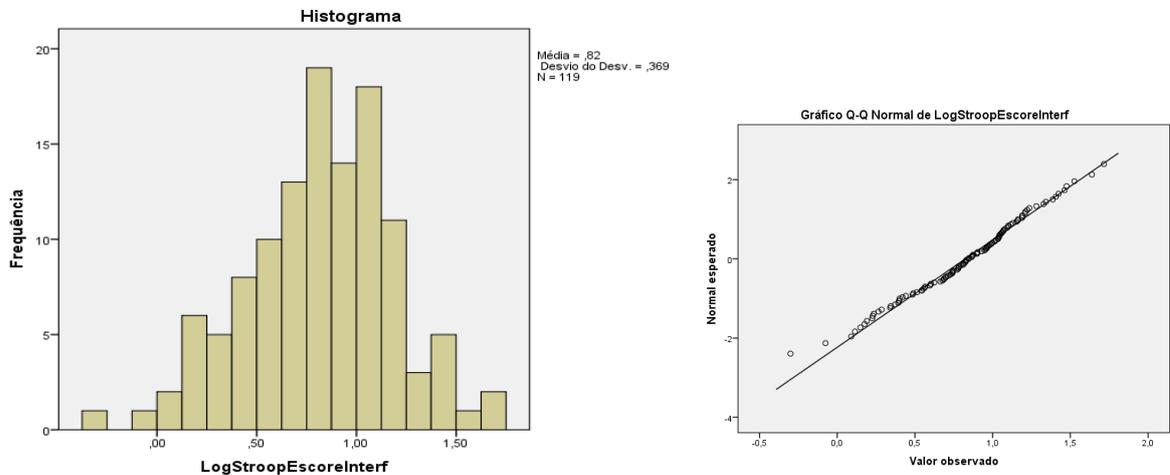
**Tabela 5**

*Teste de normalidade após logaritimização da variável escore de interferência do teste Stroop*

Teste	Shapiro-Wilk			Média	DP	Assimetria	Curtose
	W	df	p				
<b>Stroop – escore de interferência</b>	0,99	119	0,533	0,82	0,36	-0,31	0,16
<b>Log</b>							

**Figura 8**

*Histograma e gráfico de normalidade para variável escore de interferência logaritimizada*



Também foi verificada a distribuição dos dados em relação à curva normal para os demais testes utilizados. Como pode ser observado na Tabela 6 e na Figura 10, os escores dos testes FDT, Hayling, BETA-III e Wisconsin demonstraram distribuições não paramétricas. Em decorrência destes resultados, optou-se por utilizar testes estatísticos não paramétricos quando realizadas correlações entre eles e o teste Stroop. Diferentemente dos demais, o teste EASV apresentou distribuição normal, e por isso, o histograma que apresenta graficamente sua distribuição foi apresentado em imagem separada (Figura 9).

**Tabela 6**

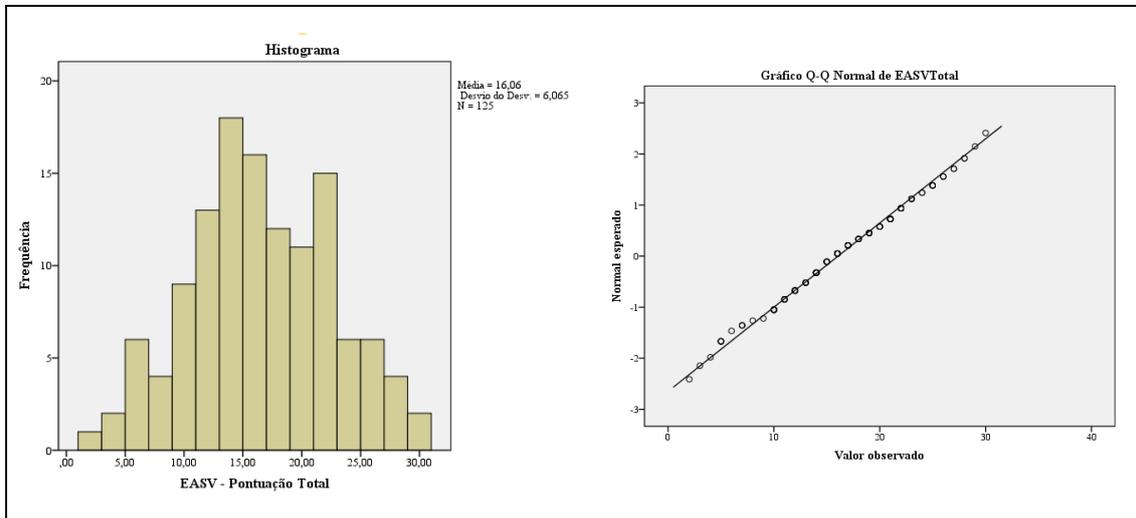
*Teste de normalidade para demais testes*

Teste	Shapiro-Wilk			Média	Assimetria	Curtose
	W	df	p			
FDT - Inibição	0,95	128	0,001	14.16	0,21	0,42
FDT - Flexibilidade	0,88	128	0,001	23.64	0,21	0,42
Hayling	0,77	126	0,001	26.38	0,21	0,42
BETA - III	0,95	119	0,001	60.58	0,22	0,44
Wisconsin – categorias completadas	0,86	117	0,001	5.66	0,22	0,44
EASV*	0,98	125	0,443	16.06	0,21	0,43

\* Valores indicam distribuição normal

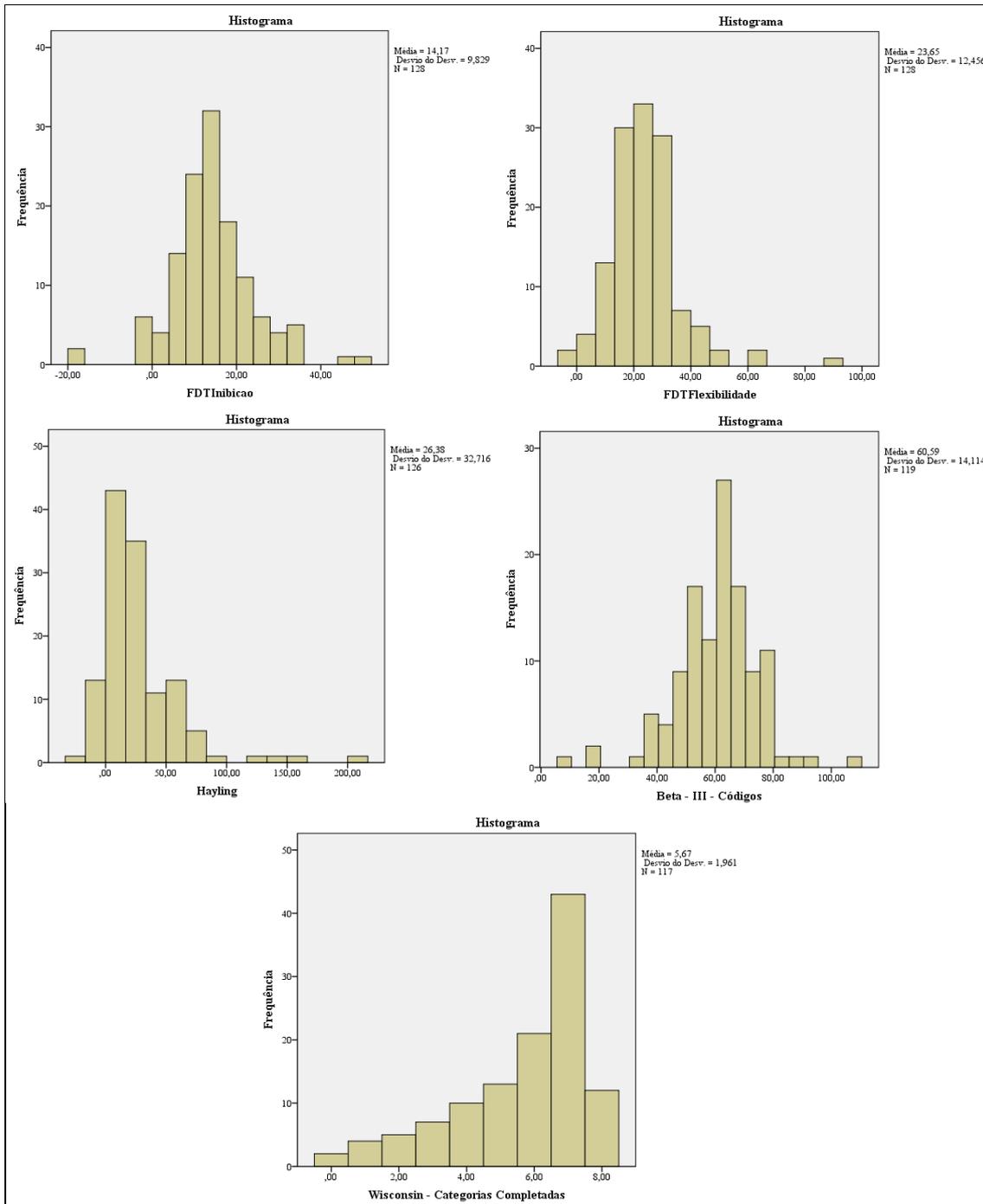
**Figura 9**

*Histogramas de distribuição do teste EASV*



**Figura 10**

*Histogramas de distribuição dos testes FDT, Hayling, BETA – III – Códigos, e Wisconsin*



### 5.2.2 Relações entre as variáveis do Teste Stroop de Cores e Palavras e dados sociodemográficos

### 5.2.2.1 Estudos correlacionais entre a variável idade e os escores no teste Stroop

Os dados indicam haver uma correlação positiva ( $\rho=0,44$   $p<0,001$ ) entre a idade dos participantes e o escore de interferência no teste Stroop (Tabela 7), o que significa que quanto maior é a idade do respondente, maior é o escore de interferência, calculado pela subtração do tempo de execução na etapa cor-palavra da média aritmética entre os tempos das etapas nomeação de cores e leitura de palavras comuns. A medida do teste Stroop denominada “escore de interferência” indica que quanto maior é o resultado da interferência, mais complexa é a realização da tarefa para o respondente. O mesmo pode ser percebido nas comparações entre o tempo de execução das três etapas do Stroop e a idade. De acordo com a Tabela 8, os coeficientes de correlação apresentam relacionamentos positivos entre as variáveis, indicando que quanto maior é a idade dos participantes, maior é o tempo de execução para cada um dos cartões ( $\rho=0,26$ ,  $p=0,004$  entre tempo da etapa 1 e idade;  $\rho=0,46$ ,  $p<0,001$  entre tempo da etapa 2 e idade; e  $\rho=0,49$ ,  $p<0,001$  entre tempo da etapa 3 e idade).

**Tabela 7**

*Correlação entre idade e escore de interferência no teste Stroop*

		<b>Escore de Interferência</b>
<b>Idade</b>	Coefficiente de correlação de Spearman	<b>0,44**</b>
	<i>p</i>	0,001
	n	114

\*\*A correlação é significativa no nível 0,01

**Tabela 8**

*Correlações entre tempo de execução e a idade dos participantes*

		<b>Idade</b>
<b>Tempo etapa 1</b>	Coefficiente de correlação de Spearman	<b>0,26**</b>
	<i>p</i>	0,004
	n	119
<b>Tempo etapa 2</b>	Coefficiente de correlação de Spearman	<b>0,46**</b>
	<i>p</i>	0,000
	n	119
<b>Tempo etapa 3</b>	Coefficiente de correlação de Spearman	<b>0,49**</b>
	<i>p</i>	0,000
	n	119

\*\*A correlação é significativa no nível 0,01

Em relação às análises que comparam o número de erros e a idade dos participantes, descritas na Tabela 9, os dados demonstram que nas etapas 1 e 2 as correlações foram não significativas. Entretanto, na etapa 3 e nos erros totais, houve uma correlação positiva e estatisticamente significativa entre o número de erros e a idade dos respondentes, indicando que pessoas mais velhas cometeram mais erros na etapa incongruente, e também, no somatório de erros totais do teste ( $\rho=0,18$ ,  $p=0,004$  e  $\rho=0,21$ ,  $p=0,02$ , respectivamente).

**Tabela 9**

*Correlações entre número de erros e a idade dos participantes*

		<b>Idade</b>
<b>Erros etapa 1</b>	Coeficiente de correlação de Spearman	0,12
	<i>p</i>	0,185
	n	119
<b>Erros etapa 2</b>	Coeficiente de correlação de Spearman	0,10
	<i>p</i>	0,23
	n	119
<b>Erros etapa 3</b>	Coeficiente de correlação de Spearman	<b>0,18*</b>
	<i>p</i>	0,04
	n	119
<b>Erros totais</b>	Coeficiente de correlação de Spearman	<b>0,21*</b>
	<i>p</i>	0,02
	n	119

\*A correlação é significativa no nível 0,05

#### 5.2.2.2 Comparação entre as faixas etárias e os escores no teste Stroop

Na sequência, para investigar quais faixas etárias apresentam diferenças nos resultados do teste Stroop a amostra foi dividida em três grupos em relação à idade dos participantes. O primeiro grupo contemplou pessoas com idade entre 18 e 28 anos ( $n=64$ ), o segundo, com idade entre 29 e 48 anos ( $n=28$ ), e o terceiro, conteve participantes com idades iguais e superiores a 49 anos ( $n=22$ ).

Após esta divisão, por meio da utilização do teste estatístico ANOVA, foi possível identificar diferenças estatisticamente significativas entre os grupos (Tabela 11). O Teste de Post-Hoc de Bonferroni mostrou que a diferença entre os grupos se dá somente na comparação

entre o 49+ e os demais. A comparação entre 18-28 e 29-48 não mostrou diferenças entre os grupos (Tabela 13).

**Tabela 10**

*Estatísticas descritivas da amostra por faixas etárias*

Faixa etária	N	Média	DP	Mínimo	Máximo
18-28	64	0,72	0,35	-0,30	1,72
29-48	28	0,83	0,34	0,09	1,48
49 +	22	1,15	0,22	0,82	1,63
Total	114	0,83	0,36	-0,30	1,72

**Tabela 11**

*Anova para variável Escore de Interferência em relação às faixas etárias*

	Soma dos quadrados	df	Quadrado Médio	F	p
Entre grupos	3,01	2	1,50	13,60	0,001
Nos grupos	12,29	111	,11		
Total	15,30	113			

**Tabela 12**

*Teste de Homogeneidade de variâncias de Levene para o escore de interferência*

Estatística de Levene	df1	df2	p
1,61	2	111	0,20

**Tabela 13**

*Comparações múltiplas – Teste de Post-Hoc de Bonferroni*

(I) Faixa Etária	(J) Faixa Etária	Diferença da média (I-J)	Erro Padrão	p
18-28	29-48	-0,11	0,07	0,39
	49 +	-0,42*	0,08	0,001
29-48	18-28	0,11	0,07	0,39
	49 +	-0,31*	0,09	0,004
49 +	18-28	0,42*	0,08	0,001
	29-48	0,31*	0,09	0,004

\* A diferença média é significativa no nível 0,05

### 5.2.2.3 Estudos comparativos entre a variável gênero e os escores no teste Stroop

Para realizar as comparações entre a variável gênero e os escores no Stroop foi realizado um teste  $t$  (Tabela 15). Com esta análise, verificou-se os resultados para o teste  $t$  na condição “variâncias iguais assumidas” em que  $t= 0,29$  e  $p=0,76$ , indicando que não houve diferenças nas médias entre o grupo feminino e masculino. Como pressuposto para esta análise estatística, foi realizado o teste de Levene para averiguar a homogeneidade das variâncias. O teste de Levene para igualdade das variâncias mostrou que os grupos foram homogêneos ( $p=0,11$ ).

**Tabela 14**

*Estatísticas descritivas em relação ao gênero*

	Sexo	N	Média	DP
Escore de Interferência	Feminino	75	0,83	0,33
	Masculino	44	0,81	0,43

**Tabela 15**

*Teste  $t$  de amostras independentes: diferença entre gêneros e escore de interferência do Stroop*

			Teste de Levene		Teste $t$ para igualdade das médias			95% intervalo de confiança da diferença	
			Z	p	t	df	p	Inferior	Superior
Escore de Interferência	de	Variâncias iguais assumidas	2,55	0,11	0,29	117	0,76	-0,11	0,16
		Variâncias iguais não assumidas			0,27	72,91	0,78	-0,12	0,17

### 5.2.2.4 Estudos de comparação entre a variável escolaridade e os escores no teste Stroop

Usando a variável Escore Interferência logaritmizada foi identificado que não houve diferenças em relação à escolaridade nos escores obtidos pelos participantes da amostra (Tabela 17). A Análise de Variância (ANOVA) indicou que os grupos não possuem médias diferentes e o Teste de Levene apresentou resultado indicando que não há homogeneidade nas variâncias (Tabela 18). A Tabela 19 apresenta os resultados do teste de *post-hoc* de Bonferroni para escolaridade, em que é possível identificar as ausências de diferenças entre as diferentes escolaridades dos participantes, já apontadas pela ANOVA.

**Tabela 16**

*Estatísticas descritivas da amostra por escolaridade do escore de interferência do teste Stroop.*

<b>Escolaridade</b>	<b>N</b>	<b>Média</b>	<b>DP</b>	<b>Mínimo</b>	<b>Máximo</b>
Fundamental	5	1,03	0,75	-0,30	1,42
Médio	30	0,95	0,39	0,11	1,72
Superior Incompleto	39	0,73	0,30	-0,08	1,46
Superior Completo	36	0,80	0,32	0,15	1,48
Total	110	0,83	0,36	-0,30	1,72

**Tabela 17**

*Anova para variável Escore de Interferência em relação às escolaridades*

	<b>Soma dos quadrados</b>	<b>df</b>	<b>Quadrado Médio</b>	<b>F</b>	<b>p</b>
Entre grupos	1,04	3	0,34	2,66	0,051
Nos grupos	13,77	106	0,13		
Total	14,81	109			

**Tabela 18**

*Teste de Homogeneidade de variâncias de Levene para o escore de interferência pela escolaridade*

<b>Estatística de Levene</b>	<b>df1</b>	<b>df2</b>	<b>p</b>
2,99	3	106	0,03

**Tabela 19**

*Comparações múltiplas – Teste de Post-Hoc de Bonferroni para escolaridade*

<b>(I) Escolaridade</b>	<b>(J) Escolaridade</b>	<b>Diferença da média (I-J)</b>	<b>Erro Padrão</b>	<b>p</b>
Fundamental	Médio	0,08	0,17	1,00
	Superior Incompleto	0,30	0,17	0,46
	Superior Completo	0,23	0,17	1,00
Médio	Fundamental Incompleto	-0,08	0,17	1,00
	Superior Incompleto	0,21	0,08	0,08
	Superior Completo	0,14	0,08	0,63
Superior Incompleto	Fundamental Incompleto	-0,30	0,17	0,46
	Médio	-0,21	0,08	0,08
	Superior Completo	-0,07	0,08	1,00
Superior Completo	Fundamental Incompleto	-0,23	0,17	1,00

Médio	-0,14	0,08	0,63
Superior Incompleto	0,07	0,08	1,00

### 5.2.3 Estudos correlacionais de evidências de validade do Teste Stroop de Cores e Palavras com os demais testes e tarefas cognitivas

Considerando os valores de assimetria e curtose próximos aos limites de normalidade, optou-se pela logaritimização da variável escore de interferência do teste Stroop para buscar obter uma distribuição que se aproxime da normal. Após a aplicação da logaritimização, os dados assumiram uma distribuição normal, permitindo o uso de testes paramétricos para algumas das análises de dados. Entretanto, como a única variável logaritimizada foi o escore de interferência, os procedimentos estatísticos paramétricos foram utilizados somente quando esta era a única variável do Stroop de interesse. Nos demais casos, optou-se por utilizar a variável de interferência sem a transformação logarítmica, já que os demais escores do teste também foram analisados em conjunto, e assim, em uma mesma análise, utilizou-se de um único teste estatístico, o não paramétrico.

#### 5.2.3.1 Estudos de correlação entre o teste FDT – Teste dos Cinco Dígitos e o Teste Stroop de Cores e Palavras

O delineamento de pesquisa desta tese foi projetado com o objetivo de levantar possíveis evidências de validade do Teste Stroop de Cores e Palavras. Para isto, foram elencadas outras medidas de avaliação, e dentre elas, o Teste dos Cinco Dígitos – FDT foi uma das ferramentas escolhidas. A ficha síntese do manual técnico do FDT apresenta que o objetivo do instrumento consiste em medir a velocidade de processamento, a atenção e as funções executivas por meio de seus subcomponentes controle inibitório e flexibilidade cognitiva (Sedó; Paula & Malloy-Diniz, 2015). Considerando esta indicação, as hipóteses iniciais traçadas presumiam elevadas correlações entre o teste Stroop e o FDT.

As análises para verificar a correlação entre as medidas foram realizadas considerando a amostra total (n=124), e posteriormente, selecionando os indivíduos com idade igual e superior a 49 anos (n=22), como pode ser identificado na Tabela 20. As variáveis

correlacionadas foram os tempos de execução entre todas as etapas do FDT (leitura, contagem, escolha e alternância) e, igualmente, entre todas as etapas do Stroop (cartão pontos, cartão palavras e cartão cor-palavra). Adicionalmente, foram realizadas análises entre os escores derivados do FDT (inibição e flexibilidade) e o escore de interferência do Stroop.

Considerando a amostra total, os resultados apresentaram correlações estatisticamente significativas, de magnitude moderada, entre os tempos de execução de todas as etapas do FDT com todas as etapas do teste Stroop (Tabela 20). Entretanto, de acordo com a Tabela 21, os escores derivados de inibição e flexibilidade do FDT, quando correlacionados com o escore de interferência do Stroop, apresentaram correlações fracas ( $\rho=0,24$  e  $\rho=0,23$ , respectivamente,  $p=0,001$ ).

Após a seleção de casos, utilizando como critério a separação do grupo com idade igual/superior a 49 anos, os dados apresentam aumento dos coeficientes de correlação. Em relação às variáveis de tempo, é possível identificar que algumas das correlações, antes moderadas, agora assumem magnitudes elevadas. Ao selecionar as relações entre variáveis consideradas idênticas em termos do tipo de tarefa exigida, percebe-se o aumento dos valores dos coeficientes de correlação. As duas primeiras etapas do FDT são ditas processos automáticos, de leitura e contagem, e em termos da natureza da tarefa, podem ser considerados similares às duas primeiras etapas do Stroop, cartão pontos e palavras, em que a demanda, também, concerne a um processamento automático (nomeação de cores). Já as etapas finais do FDT e a última etapa do Stroop se referem a processamentos controlados. Nestes, observa-se, do mesmo modo, o aumento nos coeficientes de correlação (ver Tabela 20).

O mesmo pode ser observado em relação aos escores derivados dos testes. Na amostra total, o coeficiente de correlação entre o escore de inibição do FDT com o escore de interferência do Stroop eram de  $\rho=0,24$  ( $p=0,006$ ), e após a seleção dos casos, passou a ser de  $\rho=0,56$  ( $p=0,007$ ). O escore de flexibilidade do FDT, quando correlacionado com o escore de interferência do Stroop passou de  $\rho=0,23$  ( $p=0,008$ ) – na amostra total – para  $\rho=0,42$  ( $p=0,054$ ) para idade superior a 49 anos (Tabela 21).

## **Tabela 20**

*Coefficientes de correlação de Spearman entre os tempos de execução do Teste Stroop e do Teste FDT*

Dados de correlação		Stroop – cartão pontos		Stroop – cartão palavras		Stroop – cartão cor-palavra	
		Amostra total	49 +	Amostra total	49 +	Amostra total	49 +
<b>FDT leitura</b>	coef. correlação	0,44**	0,58**	0,53**	0,62**	0,42**	0,51*
	<i>p</i>	0,001	0,005	0,001	0,002	0,001	0,015
	<i>n</i>	124	22	124	22	124	22
<b>FDT contagem</b>	coef. correlação	0,40**	0,51*	0,55**	0,56**	0,44**	0,44*
	<i>p</i>	0,001	0,01	0,001	0,006	0,001	0,036
	<i>n</i>	124	22	124	22	124	22
<b>FDT escolha</b>	coef. correlação	0,38**	0,53*	0,58**	0,82**	0,53**	0,73**
	<i>p</i>	0,001	0,01	0,001	0,001	0,001	0,001
	<i>n</i>	124	22	124	22	124	22
<b>FDT alternância</b>	coef. correlação	0,42**	0,47*	0,52**	0,87**	0,52**	0,69**
	<i>p</i>	0,001	0,025	0,001	0,001	0,001	0,002
	<i>n</i>	124	22	124	22	124	22

\*\* A correlação é significativa no nível 0,01

\* A correlação é significativa no nível 0,05

**Tabela 21**

*Coefficientes de correlação entre o Teste Stroop e o Teste FDT: amostra total e grupo com 49 anos ou mais*

Dados de correlação		Stroop – escore de interferência	
		Amostra total	49 +
<b>FDT - escore inibição</b>	$\rho$ de Spearman	0,24**	0,56*
	<i>p</i>	0,006	0,007
	<i>n</i>	124	22
<b>FDT - escore flexibilidade</b>	$\rho$ de Spearman	0,23**	0,42*
	<i>p</i>	0,008	0,052
	<i>n</i>	124	22

\*\* A correlação é significativa no nível 0,01

\* A correlação é significativa no nível 0,05

### 5.2.3.2 Estudos de correlação entre o Teste Não Verbal de Inteligência Geral - BETA-III – Subteste Códigos e o Teste Stroop de Cores e Palavras

O Subteste Códigos é uma medida de avaliação que compõe o Teste Não Verbal de Inteligência Geral - BETA-III, instrumento aprovado pelo Conselho Federal de Psicologia para utilização para fins profissionais. Sua finalidade destina-se à mensuração do construto velocidade de processamento. Tendo em consideração que a revisão integrativa de literatura, apresentada no Capítulo II deste trabalho, verificou que este construto foi relacionado como um dos potenciais construtos alvo do Stroop, optou-se por utilizá-lo, também, para a busca de evidências de validade. Como hipóteses, esperava-se encontrar correlações moderadas entre os tempos de execução dos três cartões do Stroop e o escore geral do Subteste Códigos do BETA-III. Mais precisamente, considerando a natureza e as demandas do teste Stroop, era previsto que maiores coeficientes de correlação derivassem das comparações entre as duas primeiras etapas do Stroop (cartão cor e cartão palavra) – etapas estas que requerem a simples nomeação de cores e não contemplam a apresentação simultânea de tarefas concorrentes – e o escore geral do BETA-III.

Considerando a distribuição não paramétrica dos resultados dos testes, foi utilizado o coeficiente de correlação de Spearman. Primeiramente, as análises contemplaram a amostra total ( $n=115$ ) de participantes, e posteriormente, foram selecionados os casos em que os participantes tivessem idade igual ou superior a 49 anos ( $n=20$ ).

Os valores foram de  $r_{ho} = -0,53$ ,  $-0,65$  e  $-0,54$ , respectivamente (valores de  $p < 0,001$ ), para as comparações entre cartão pontos, cartão palavra, cartão cor-palavra e o escore geral do Subteste Códigos (Tabela 22). A média de tempo das duas primeiras etapas do Stroop também foi correlacionada com o escore geral do BETA, com coeficiente igual a  $r_{ho} = -0,64$  ( $p < 0,001$ ). Nota-se que o sentido das correlações foi negativo, indicando que altas pontuações no BETA-III foram relacionadas a menores tempos de execução do Stroop, relação esperada, já que o melhor desempenho do Stroop é equivalente a baixos tempos de execução. No BETA, o melhor desempenho, para o participante, corresponde a altas pontuações de acertos.

Os resultados obtidos indicaram, portanto, correlações estatisticamente significativas - de grau moderado - entre os tempos de execução das três etapas do teste Stroop, cartão pontos, cartão palavra e cartão cor-palavra, e o escore geral do Subteste Códigos do BETA-III, quando

considerada a amostra total. Quando selecionados os adultos com idade igual ou superior a 49 anos, os dados mantiveram-se estatisticamente significativos, aumentando os valores dos coeficientes de correlação. Nesta nova análise, foram obtidos coeficientes entre  $\rho = -0,62$  e  $-0,75$  entre as medidas de tempo de execução das etapas do Stroop e o escore geral do Subteste Códigos do BETA ( $p < 0,001$ ).

**Tabela 22**

*Coefficientes de correlação de Spearman entre o Teste Stroop e o Teste BETA III – Subteste Códigos: amostra total*

Dados de correlação		BETA III - Códigos	
		Amostra total	49 anos +
<b>Stroop – cartão pontos (etapa 1)</b>	coeficiente de correlação	-0,53**	-0,62**
	<i>p</i>	0,001	0,003
	<i>n</i>	115	20
<b>Stroop – cartão palavras (etapa 2)</b>	coeficiente de correlação	-0,65**	-0,72**
	<i>p</i>	0,001	0,001
	<i>n</i>	115	20
<b>Stroop – cartão cor-palavra (etapa 3)</b>	coeficiente de correlação	-0,54**	-0,53**
	<i>p</i>	0,001	0,001
	<i>n</i>	115	20
<b>Stroop – média etapas 1 e 2</b>	coeficiente de correlação	-0,64**	-0,75**
	<i>p</i>	0,001	0,001
	<i>n</i>	115	20
<b>Stroop escore de interferência</b>	coeficiente de correlação	-0,30**	-0,51*
	<i>p</i>	0,001	0,001
	<i>n</i>	115	20

\*\* A correlação é significativa no nível 0,01

\* A correlação é significativa no nível 0.05

### 5.2.3.3 Estudos de correlação entre a Escala de Atenção Seletiva Visual – EASV e o Teste Stroop de Cores e Palavras

A Escala de Atenção Seletiva Visual, elaborada por Sisto e Castro (2011), avalia a capacidade do indivíduo de selecionar estímulos visuais identificando regularidades entre o modelo apresentado e os estímulos disponíveis para seleção. A EASV pretende avaliar a atenção seletiva partindo do modelo da seleção com base na localização espacial do objeto. Tal

instrumento foi selecionado para esta pesquisa por figurar como um teste psicológico aprovado e reconhecido pelo Conselho Federal de Psicologia como uma medida destinada à mensuração do construto de atenção seletiva visual.

De acordo com os dados observados na Tabela 23 foi possível identificar que em relação à amostra total, os coeficientes de correlação se mostraram fracos, ou mesmo, inexistentes entre o tempo de execução dos diferentes cartões do teste Stroop e a pontuação total do EASV. Observa-se que os valores com maiores índices de correlação, para estas comparações, foram iguais a  $\rho = -0,19$  ( $p = 0,029$  e  $0,036$ ) entre as etapas cartão palavras e cartão cor-palavra com a pontuação total do EASV. Em relação às demais medidas do EASV, foi identificado um coeficiente de correlação estatisticamente significativo somente na comparação entre o escore de interferência do Stroop e a capacidade seletiva visual básica do EASV ( $r = -0,21$ ,  $p = 0,019$ ), novamente de baixa magnitude (Tabela 24). O escore “pontuação básica” se refere aos itens considerados como de baixa complexidade da escala. Nesta análise, o coeficiente de correlação utilizado foi o de Pearson, considerando a normalidade dos dados da EASV e a utilização da variável logaritimizada do Stroop escore de interferência. Para as demais correlações, manteve-se o coeficiente de correlação de Spearman, por não terem sido logaritimizadas todas as variáveis do Stroop.

Seguindo a proposta das demais análises já relatadas, a amostra foi dividida e foram selecionados os participantes com idade igual e superior a 49 anos. Nesta nova operação, os coeficientes de correlação se mostraram mais elevados, com resultados de  $\rho = -0,66$ ,  $-0,58$  e  $-0,55$  entre EASV pontuação total e cartão palavras, cartão cor-palavra e escore de interferência, respectivamente ( $p < 0,001$ ,  $p = 0,006$  e  $p = 0,012$ ).

### Tabela 23

*Coefficientes de correlação de Spearman entre as variáveis de tempo do Teste Stroop e a Escala de Atenção Seletiva Visual – EASV (pontuação total)*

		EASV – pontuação total	
		Amostra total	49 anos ou +
<b>Stroop – cartão pontos</b> (etapa 1)	coeficiente de correlação	-0,15	-0,38
	<i>p</i>	0,100	0,092
	<i>n</i>	121	20
<b>Stroop – cartão palavras</b> (etapa 2)	coeficiente de correlação	-0,19*	-0,66**
	<i>p</i>	0,029	0,001

	n	121	20
<b>Stroop – cartão cor-</b>	coeficiente de correlação	-0,19*	-0,58**
<b>palavra (etapa 3)</b>	<i>p</i>	0,036	0,006
	n	121	20
<b>Stroop escore de</b>	coeficiente de correlação	-0,15	-0,55*
<b>interferência</b>	<i>p</i>	0,090	0,012
	n	121	20

\*\* A correlação é significativa no nível 0,01

\* A correlação é significativa no nível 0,05

**Tabela 24**

*Coefficientes de correlação de Spearman entre o escore de interferência do Teste Stroop e a Escala de Atenção Seletiva Visual – EASV (pontuação básica)*

				<b>EASV – pontuação básica</b>	
				Amostra total	49 anos ou +
<b>Stroop escore de</b>	coeficiente de correlação			-0,21*	-0,57*
<b>interferência</b>	<i>p</i>			0,019	0,008
	n			121	20

\* A correlação é significativa no nível 0.05

Similarmente, pode-se identificar correlações não significativas entre erros e acertos no Stroop e a pontuação total do EASV na amostra total. Entretanto, ao selecionar os participantes acima de 49 anos, os coeficientes se mostraram significativos e de magnitude moderada, indicando valores como -0,51 entre EASV e acertos no cartão palavras e -0,49 entre EASV e acertos no cartão cor-palavra ( $p= 0,021$  e  $0,028$ , respectivamente).

**Tabela 25**

*Coefficientes de correlação de Spearman entre o variáveis de acertos e erros do Teste Stroop e a Escala de Atenção Seletiva Visual – EASV (pontuação total)*

				<b>EASV – pontuação total</b>	
				Amostra total	49 anos ou +
<b>Stroop Erros – cartão</b>	coeficiente de correlação			-0,06	0,144
<b>pontos (etapa 1)</b>	<i>p</i>			0,477	0,544
	n			121	20
<b>Stroop Erros – cartão</b>	coeficiente de correlação			-0,091	-0,51*
<b>palavras (etapa 2)</b>	<i>p</i>			0,320	0,021

	n	121	20
<b>Stroop Erros – cartão cor-palavra (etapa 3)</b>	coeficiente de correlação	-0,15	-0,49*
	<i>p</i>	0,091	0,028
	n	121	20
<b>Stroop Acertos – cartão pontos (etapa 1)</b>	coeficiente de correlação	0,065	-0,14
	<i>p</i>	0,477	0,544
	n	121	20
<b>Stroop Acertos – cartão palavras (etapa 2)</b>	coeficiente de correlação	0,11	-0,51*
	<i>p</i>	0,205	0,021
	n	121	20
<b>Stroop Acertos – cartão cor-palavra (etapa 3)</b>	coeficiente de correlação	0,14	0,49*
	<i>p</i>	0,122	0,028
	n	121	20

\* A correlação é significativa no nível 0,05

#### 5.2.3.4 Estudos de correlação entre o Hayling Sentence Completion Test (HSCT) e o Teste Stroop de Cores e Palavras

O *Hayling Sentence Completion Test* (HSCT) consiste em uma tarefa cognitiva que demanda que o respondente complete frases incluindo somente uma palavra ao final da sentença. É dividido em duas etapas: parte A e parte B. Na parte A solicita-se que o participante emita uma palavra que faça sentido ao contexto da frase. Nesta, os construtos relacionados indicados, de acordo com Siqueira, Scherer, Reppold e Fonseca (2010), são atenção concentrada, iniciação verbal, velocidade de processamento e estratégias de busca de palavras relacionadas a pré-ativação da rede semântica. Já na parte B, é solicitado que o sujeito emita uma palavra que não faz sentido com o contexto da sintaxe e por tal, presume-se que sejam mobilizados componentes mais complexos das funções executivas, como inibição verbal e planejamento, pois o sujeito deve inibir o conteúdo semântico da frase e desenvolver alternativas de busca lexical para completar frase. De modo geral, a junção das duas etapas do teste permite avaliar os componentes de iniciação e inibição.

Em relação à amostra total, os dados presentes na Tabela 26 indicaram correlações estatisticamente significativas entre os tempos de execução entre as variáveis do Stroop e do Hayling, variando de valores entre  $\rho=0,24$  ( $p=0,008$ ) e  $\rho=0,40$  ( $p=0,001$ ). Este último valor, foi identificado na correlação entre a etapa B do Hayling e etapa cor-palavra do Stroop, ambas

consideradas tarefas de supressão de resposta. Em relação ao escore geral do Hayling - que compreende a subtração do tempo de execução da etapa B menos o tempo de execução da etapa A – e o escore de interferência do Stroop (que também se refere à diferença de tempo de execução do cartão cor e palavra menos a média aritmética da etapa cor e da etapa palavra) o coeficiente de correlação se mostrou em menor magnitude, com valor de  $\rho$  0,31 ( $p=0,001$ ).

**Tabela 26**

*Coefficientes de correlação de Spearman entre os tempos de execução do Teste Stroop e do Teste Hayling para amostra total*

		Dados de correlação	Hayling Etapa A	Hayling Etapa B	Hayling Escore Geral
<b>Stroop – cartão pontos (etapa 1)</b>		coeficiente de correlação	0,26**	0,28**	0,23*
		$p$	0,003	0,001	0,011
		$n$	122	122	122
<b>Stroop – cartão palavras (etapa 2)</b>		coeficiente de correlação	0,27**	0,32**	0,29**
		$p$	0,002	0,001	0,001
		$n$	122	122	122
<b>Stroop – cartão cor-palavra (etapa 3)</b>		coeficiente de correlação	0,30**	0,40**	0,37**
		$p$	0,001	0,001	0,001
		$n$	122	122	122
<b>Stroop – média etapas 1 e 2</b>		coeficiente de correlação	0,31**	0,36**	0,30**
		$p$	0,001	0,001	0,001
		$n$	122	122	122
<b>Stroop escore de interferência</b>		coeficiente de correlação	0,24**	0,33**	0,31**
		$p$	0,008	0,001	0,001
		$n$	122	122	122

\*\* A correlação é significativa no nível 0,01

\* A correlação é significativa no nível 0,05

Quando selecionados os casos da amostra de participantes com idade igual ou superior a 49 anos, os valores apresentaram ausência de relacionamento estatístico entre as mesmas variáveis, como pode ser observado na Tabela 27.

**Tabela 27**

*Coefficientes de correlação de Spearman entre os tempos de execução do Teste Stroop e do Teste Hayling: amostra de pessoas com 49 anos de idade ou mais.*

		Dados de correlação	Hayling Etapa A	Hayling Etapa B	Hayling Escore Geral
<b>Stroop – cartão pontos (etapa 1)</b>		coeficiente de correlação	0,12	0,38	0,38
		<i>p</i>	0,590	0,079	0,076
		<i>n</i>	22	22	22
<b>Stroop – cartão palavras (etapa 2)</b>		coeficiente de correlação	-0,12	0,10	0,17
		<i>p</i>	0,589	0,640	0,449
		<i>n</i>	22	22	22
<b>Stroop – cartão cor-palavra (etapa 3)</b>		coeficiente de correlação	-0,13	0,06	0,16
		<i>p</i>	0,561	0,778	0,476
		<i>n</i>	22	22	22
<b>Stroop – média etapas 1 e 2</b>		coeficiente de correlação	-0,06	0,27	0,34
		<i>p</i>	0,764	0,219	0,118
		<i>n</i>	22	22	22
<b>Stroop      escore      de interferência</b>		coeficiente de correlação	0,24	-0,04	0,09
		<i>p</i>	0,275	0,859	0,584
		<i>n</i>	22	22	22

#### 5.2.3.5 Estudos de correlação entre o Wisconsin Teste de Classificação de Cartas - WCST e o Teste Stroop de Cores e Palavras

O Teste Wisconsin de Classificação de Cartas (WCST) é um instrumento que foi desenvolvido em 1948 para avaliar a capacidade do indivíduo em raciocinar abstratamente e modificar suas estratégias cognitivas como resposta a alterações nas contingências ambientais. Ao longo dos anos, foi adaptado em diferentes versões, sendo um instrumento aprovado pelo Conselho Federal de Psicologia para uso profissional. Como medidas derivadas, o WCST requer do respondente planejamento estratégico e exploração organizada utilizando *feedback* ambiental para mudar contextos cognitivos, direcionando o comportamento para alcançar um objetivo e modulando a responsividade impulsiva (Heaton, *et al.* 2005).

As hipóteses iniciais desta tese previam moderadas correlações entre algumas das variáveis do WCST e o teste Stroop, considerando que ambos são referenciados, na literatura, como medidas para avaliação de funções executivas ou de seus subcomponentes.

A Tabela 28 apresenta os coeficientes de correlação de Spearman entre a variável escore de interferência e as variáveis do WCST. Levando em conta a amostra total (n=108) os valores

que apontaram relacionamentos estatisticamente significativos foram entre o escore de interferência do Stroop e o número de categorias completadas ( $r=0,23$ ,  $p=0,017$ ); escore de interferência e acertos ( $r=0,22$ ,  $p=0,021$ ); escore de interferência e erros ( $r=-0,24$ ,  $p=0,009$ ); e escore de interferência e erros perseverativos ( $r=0,25$ ,  $p=0,008$ ). Diferentemente do esperado, as correlações são consideradas de fraca magnitude para a amostra total. Quando selecionados os casos de participantes com idade igual ou superior a 49 anos, as correlações estatisticamente significativas foram entre o escore de interferência do Stroop e categorias completadas do WCST ( $r=0,55$ ,  $p=0,017$ ) e escore de interferência e erros não perseverativos ( $r=0,51$ ,  $p=0,028$ ).

**Tabela 28**

*Coefficientes de correlação de Spearman entre o Teste Stroop e o Teste Wisconsin*

			<b>Stroop – Escore de Interferência Log</b>	
			Amostra total	49 anos ou +
<b>WCST – categorias completadas</b>	coeficiente de correlação		-0,23*	-0,55*
	<i>p</i>		0,017	0,017
	n		108	18
<b>WCST - acertos</b>	coeficiente de correlação		-0,22*	-0,27
	<i>p</i>		0,021	0,26
	n		109	19
<b>WCST - erros</b>	coeficiente de correlação		-0,24*	0,46
	<i>p</i>		0,009	0,055
	n		108	18
<b>WCST – ensaios administrados</b>	coeficiente de correlação		-0,06	-
	<i>p</i>		0,477	
	n		108	
<b>WCST – rupturas</b>	coeficiente de correlação		-0,09	0,22
	<i>p</i>		0,330	0,371
	n		108	18
<b>WCST – erros perseverativos</b>	coeficiente de correlação		0,25*	0,34
	<i>p</i>		0,008	0,157
	n		108	18
<b>WCST – erros não perseverativos</b>	coeficiente de correlação		0,13	0,51*
	<i>p</i>		0,162	0,028
	n		108	18

\* A correlação é significativa no nível 0.05

Foram analisados, também, os relacionamentos entre os escores de acertos e erros entre ambos os testes e, a partir destas, foi possível verificar que somente no grupo com idade superior a 49 anos os valores foram estatisticamente significativos e indicaram a presença de correlações (Tabela 29). Neste recorte da amostra, os erros da etapa 3 (cartão cor e palavra) do Stroop apresentaram-se correlacionados tanto com os erros perseverativos do WCST ( $\rho=0,48$ ,  $p=0,040$ ) quanto com os erros não perseverativos ( $\rho=0,50$ ,  $p=0,034$ ).

**Tabela 29**

*Coefficientes de correlação de Spearman entre as variáveis de acertos e erros do Teste Stroop e do Teste Wisconsin*

		Stroop - acertos etapa 3	Stroop - erros etapa 3	Stroop - acertos etapa 3	Stroop - erros etapa 3
		Amostra total		49 anos ou +	
<b>WCST - acertos</b>	coef. correlação	0,08	-0,11	0,31	-0,31
	<i>p</i>	0,361	0,242	0,190	0,190
	n	114	114	19	19
<b>WCST - erros</b>	coef. correlação	-0,12	0,14	-0,46	0,46
	<i>p</i>	0,187	0,128	0,053	0,053
	n	113	113	18	18
<b>WCST - erros perseverativos</b>	coef. correlação	-0,12	0,15	-0,48*	0,48*
	<i>p</i>	0,175	0,110	0,040	0,040
	n	113	113	18	18
<b>WCST - erros não perseverativos</b>	coef. correlação	-0,14	0,14	-0,50*	0,50*
	<i>p</i>	0,128	0,115	0,034	0,034
	n	113	113	18	18

\* A correlação é significativa no nível 0.05

**Tabela 30**

*Síntese dos maiores coeficientes de correlação (estatisticamente significativos)*

<b>Correlação bivariada</b>	<b>Coefficientes</b>	
Stroop cartão pontos x FDT leitura	<i>rho</i>	0,44
Stroop cartão pontos x FDT contagem	<i>rho</i>	0,40
Stroop cartão palavras x FDT leitura	<i>rho</i>	0,53
Stroop cartão palavras x FDT contagem	<i>rho</i>	0,55
Stroop cartão palavras x FDT escolha	<i>rho</i>	0,58
Stroop cartão palavras x FDT alternância	<i>rho</i>	0,52
Stroop cartão cor-palavra x FDT leitura	<i>rho</i>	0,42
Stroop cartão cor-palavra x FDT contagem	<i>rho</i>	0,44
Stroop cartão cor-palavra x FDT escolha	<i>rho</i>	0,53
Stroop cartão cor-palavra x FDT alternância	<i>rho</i>	0,52
Stroop escore de interferência x FDT inibição	<i>rho</i>	0,24
Stroop escore de interferência x FDT flexibilidade	<i>rho</i>	0,23
Stroop cartão pontos x BETA - III (códigos)	<i>rho</i>	0,53
Stroop cartão palavras x BETA - III (códigos)	<i>rho</i>	0,65
Stroop cartão cor-palavra x BETA - III (códigos)	<i>rho</i>	0,54
Stroop média aritmética entre as etapas 1 e 2 x BETA - III (códigos)	<i>rho</i>	0,64
Stroop escore de interferência x BETA - III (códigos)	<i>rho</i>	0,30
Stroop cartão cor-palavra x EASV pontuação total	<i>rho</i>	0,19
Stroop cartão palavras x EASV pontuação total	<i>rho</i>	0,19
Stroop escore de interferência x EASV - pontuação básica	<i>r</i>	0,21
Stroop cartão cor-palavra x Hayling B	<i>rho</i>	0,40
Stroop escore de interferência x Hayling Escore Geral	<i>rho</i>	0,31
Stroop escore de Interferência x WCST - categorias completadas	<i>r</i>	0,23
Stroop escore de Interferência x WCST erros	<i>r</i>	0,24

### 5.3 DISCUSSÃO

### 5.3.1 Análises e discussões dos estudos entre o Teste Stroop de Cores e Palavras e as variáveis sociodemográficas

A influência da idade no desempenho do teste Stroop figura como uma das variáveis recorrentemente analisadas nas pesquisas que utilizam o instrumento. Marrow (2013), um dos pesquisadores que se propôs a descrever sobre esta relação, desenvolveu um estudo com os objetivos de gerar dados normativos para o Teste Stroop de Cores e Palavras para a população canadense e norte-americana, identificar a relação entre testes que mensuram velocidade de processamento e o Stroop e verificar a influência entre dados demográficos como idade, escolaridade e gênero e escores no teste. Os testes foram administrados em 146 adultos saudáveis, com idade média de 37,5 anos ( $dp=10,9$ ). Em relação à descrição da amostra, 68,5% era do sexo feminino e a idade média de escolarização foi de 14,3 anos escolares. A versão utilizada nesta pesquisa foi de Golden. O resultado final representou que altos escores no teste indicaram melhor desempenho na capacidade cognitiva avaliada, indicada como atenção seletiva. Dentre os resultados, a variável idade foi significativamente associada com os escores do Stroop ( $r=-0,403$ ,  $p<0,001$ ), indicando uma correlação negativa moderada. Tal resultado apontou que à medida que a idade do participante aumenta, o desempenho, em termos de escores no Stroop, diminui. Em relação às demais variáveis demográficas, gênero e educação, nenhuma apresentou correlações com as pontuações do Stroop nesta população.

Tais resultados apontados por Marrow (2013) correspondem com os dados obtidos nesta tese em relação à associação entre o Stroop e medidas sociodemográficas. Neste trabalho, em relação à idade, os dados apresentaram uma correlação positiva entre a idade dos participantes e o escore geral do teste ( $rho=0,44$   $p<0,001$ ). O sentido inverso da correlação - em comparação ao estudo supracitado - justifica-se pelo método utilizado na formação do escore de acordo com a versão do teste. Na versão de Victoria - utilizada aqui - quanto menor é o resultado na medida, melhor é o desempenho na capacidade cognitiva. Deste modo, a correlação entre idade e os escores no Stroop (versão Victoria) demonstra que à medida que a idade do participante aumenta, os escores do Stroop também aumentam, indicando assim, altos escores como associados a piores desempenhos. Em relação ao gênero, nesta pesquisa, também não foram encontradas diferenças entre homens e mulheres em relação ao desempenho no teste Stroop. O mesmo é identificado em relação à escolaridade dos participantes. Os dados obtidos pela análise

da ANOVA indicaram que os grupos não possuem médias diferentes em relação ao nível de escolarização.

Ainda em relação ao fator idade, Zimmermann e colegas (2015), apresentaram diferenças estatisticamente significativas nos escores do teste Stroop em relação à progressão da idade dos participantes. O artigo contemplou uma amostra composta por 158 adultos com idades entre 19 e 75 anos, divididos em três grupos, adultos jovens (19-39), adultos de meia-idade (40-59), e adultos idosos (60-75). Como resultados, os autores indicaram que o grupo mais jovem teve melhor desempenho que o grupo de idosos nas tarefas que envolvem a nomeação de cores, e também, na etapa cor-palavra. Outros estudos reportam, também, que a idade é um forte preditor de performance no Stroop, como mencionado nos trabalhos de Zalonis, et al. (2008), Moering, Schinka, Mortimer e Graves (2004), Van der Elst, Boxtel, Breukelen e Jolles (2006) e Peña-Casanova et al. (2009).

Previamente já descrita na literatura, a lentidão cognitiva é um importante fator associado ao envelhecimento, e por isso, é importante levá-la em consideração ao interpretar os efeitos de interferência do teste em relação ao aumento da idade. Por isso, as variáveis de interesse principal com o uso do teste Stroop na população idosa devem incluir, além do escore de interferência, as pontuações de erros. Entretanto, escores que indicam os erros são esporadicamente relatados em estudos envolvendo tal população, todavia, a literatura aponta evidências de que adultos mais velhos cometem mais erros do que adultos mais jovens nas tarefas Stroop (Delis et al., 2001; West, 1999; Whelihan & Leshner, 1985).

Considerando este raciocínio, Troyer, Leach e Strauss (2006) verificaram a associação entre o efeito da idade nas medidas do Stroop, examinando além dos escores de efeito de interferência, os erros cometidos pelos participantes. Foram avaliados 272 adultos saudáveis, de ambos os gêneros, com idades entre 18 e 94 anos. Os dados apresentaram correlações entre a idade e o tempo requerido para completar as três tarefas do Stroop Teste versão de Victoria, com valores de  $r=0,44$  para idade e etapa pontos ( $p < 0,01$ ),  $r= 0.58$  para etapa palavras neutras ( $p < 0,01$ ), e  $r=0,62$  na etapa cor e palavra ( $p < 0,01$ ). Os resultados são congruentes com os obtidos nesta pesquisa, que também indicaram associações entre a idade e o tempo de execução em cada uma das etapas do Stroop. Os coeficientes de correlação de Spearman foram iguais a 0,26, 0,46 e 0,49, todos estatisticamente significativos ( $p < 0,01$ ), para as respectivas etapas (etapa cor, etapa palavra e etapa cor-palavra).

O segundo objetivo avaliado por Troyer, Leach e Strauss (2006) se referiu aos valores de erros e a associação com a idade dos participantes. Os resultados apresentados em tabelas de frequências demonstram que a média de erros foi maior nos participantes idosos (por ex. média de erros entre 18 a 39 anos foi igual à 0,8, quanto a média de erros para participantes entre 80 e 94 anos foi de 2.1). No presente trabalho, os dados apresentaram correlações positivas entre a idade e o número de erros obtidos na etapa 3, cor-palavra ( $\rho=0,18$   $p<0,05$ ) e entre a idade e o total de erros obtidos nas três tarefas do teste ( $\rho=0,21$   $p<0,05$ ), indicando que participantes mais velhos apresentam maior número de erros na execução total da tarefa. Tais dados convergem também com o estudo de Delis et al. (2001), que sinaliza que participantes mais jovens cometem menos erros em comparação aos mais velhos, cometem mais erros. Pondera-se que as correlações possuem relacionamento considerado fraco em termos de magnitude, entretanto, tal fator pode sofrer influências de elementos como, por exemplo, o formato do teste. A versão Victoria possui um número reduzido de itens para cada cartão, o que também pode estar associado ao menor número de erros. Ainda, a amostra é composta em sua maioria por adultos jovens, o que, do mesmo modo, pode interferir nos dados, já que não possui um valor expressivo de participantes idosos.

Tais resultados são importantes de serem descritos e analisados já que o fator velocidade trata-se de uma medida frequentemente associada ao teste Stroop. Alguns autores defendem a hipótese de que a dificuldade sentida pelos participantes mais velhos é reflexo do processo de lentidão na sua velocidade de execução. No entanto, o fator precisão, que se refere a realização do teste com acurácia, com ausência ou poucos erros, também se mostrou, nesta tese, diminuído de acordo com a progressão idade, o que reforça a conjectura de que a lentidão cognitiva não pode ser considerada como o único fator associado ao desempenho na tarefa Stroop, neste grupo. Entretanto, não é possível determinar até que ponto a velocidade *versus* a precisão são fatores que exercem maior influência nos resultados no teste visto que no formato da versão Victoria as duas pontuações estão relacionadas, já que a ocorrência dos erros aumenta o escore de tempo, pois o participante é obrigado a corrigi-los à medida que eles são cometidos.

Outras variáveis demográficas têm se mostrado divergentes com a performance no Stroop. Nas análises descritas por Troyer, Leach e Strauss (2006) não foram identificadas associações entre escolaridade e as medidas examinadas do teste. Diferenças entre os grupos de participantes masculino e feminino, do mesmo modo, não foram estatisticamente significativas para nenhuma das variáveis medidas pelo teste, dados que convergem com os obtidos nesta

tese. Todavia, não se pode estabelecer um consenso a respeito da associação, ou ausência dela, entre as diferenças de gênero ou escolarização e os resultados das medidas do Teste Stroop de Cores e Palavras em suas diferentes versões.

Ao longo dos aproximados 90 anos de investigação do fenômeno, publicado por John Ridley Stroop em 1935, os estudos divergem sobre as diferenças entre sexos no seu desempenho. Dentre as publicações, merece destaque o estudo de meta-análise de Sjöberg, Baguely e Cole (2022) que relacionou todos os experimentos publicados entre os anos 1935 até 2013 que propuseram avaliar tais diferenças entre os grupos. Dos 60 artigos selecionados após os critérios de inclusão e exclusão, 74% dos estudos não apresentaram diferenças entre gênero. 22% apresentaram resultados favoráveis de desempenho para a amostra feminina, e somente 4% indicaram resultados favoráveis para a amostra masculina. Entretanto, ao olharem para os estudos que não relataram diferenças, foi possível calcular o tamanho do efeito entre os grupos, que resultou em um dado de que 77,5% destes apontam um tamanho de efeito a favor das mulheres ( $d=0,12$ ,  $p<0,001$ ).

Em relação à influência da escolarização e o desempenho no teste Stroop, Zimmermann e colegas (2015) apresentaram dados que contrastam com os obtidos nesta pesquisa de doutoramento. Os autores relataram que indivíduos com alta escolaridade tiveram melhor desempenho do que seus pares com baixa escolaridade, e que, as habilidades menos automáticas, presentes na condição incongruente cor-palavra, são influenciadas tanto pela idade ( $F=4.30$ ,  $p<0,001$ ) quando pela educação ( $F=13.84$ ,  $p<0,001$ ), separadamente. Pode-se utilizar como fator de análise para a divergência entre os estudos o número amostral de pessoas com baixa escolaridade. Na pesquisa de Zimmermann e colegas, 57 participantes foram descritos como pertencentes ao grupo com baixa escolaridade, composto por pessoas que possuíam entre 5 e 8 anos de escolarização, entendido como equivalente ao Ensino Fundamental. Já nesta pesquisa, o número de participantes equivalentes ao Ensino Fundamental foi inexpressivo, com somente 5 pessoas, o que tende a justificar a ausência de diferenças estatisticamente significativas. A distribuição dos participantes em relação a variável categorias de escolarização dividiu-se de forma equivalente somente nas demais divisões, categorizadas como Ensino Médio ( $n=30$ ), Ensino Superior Incompleto ( $n=39$ ) e Ensino Superior Completo ( $n=36$ ). Ressalta-se que mesmo com número amostral considerado satisfatório para as análises de dados, não houve, do mesmo modo, diferenças entre as médias do teste Stroop entre os demais grupos.

Dados consonantes com os obtidos nesta pesquisa podem ser observados no trabalho de Brandelero e de Toni (2015). Em um estudo de validade do teste Stroop para controle inibitório, os autores não encontraram resultados estatisticamente significativos entre o desempenho dos participantes no teste com relação à variável escolaridade ( $p=0,154$ ), indicando que tal variável não foi correlacionada com os resultados do Stroop. Em relação à idade é possível, também, identificar convergência com os dados desta tese. Sobre tal, os autores apontaram uma correlação significativa entre idade e o desempenho no teste Stroop ( $r=-0,224$ ,  $p=0,003$ ) demonstrando que a performance no Stroop é prejudicada com o incremento da idade dos participantes.

### **5.3.2 Análises e discussões dos estudos entre o Teste Stroop de Cores e Palavras e os demais testes e tarefas cognitivas**

#### **5.3.2.1 Estudos correlacionais entre o teste FDT – Teste dos Cinco Dígitos e o Teste Stroop de Cores e Palavras**

O Teste dos Cinco Dígitos – FDT se caracteriza como uma medida neuropsicológica destinada à avaliação da velocidade de processamento e da eficiência do controle cognitivo, aprovada para fins profissionais pelo Sistema de Avaliação de Testes Psicológicos – SATEPSI do Conselho Federal de Psicologia - CFP. Desenvolvido por Manuel Sedó, se apresenta como um teste multilíngue de funções cognitivas baseado em conhecimentos linguísticos mínimos, que envolvem a contagem de dígitos de 1 a 5, compreensão de quantidades de 1 a 5 e a leitura de dígitos numéricos, também, de 1 a 5. Sua origem é fundamentada no Teste Stroop de Cores e Palavras - TSCP, tendo seu formato como uma adaptação ao paradigma proposto pelo TSCP. A adaptação consiste na utilização de símbolos praticamente universais (substituição de palavras, no caso do Stroop, por dígitos numéricos, no FDT) e menos suscetíveis à variáveis confundidoras, como a capacidade, ou não, do indivíduo identificar cores, ou ainda, a alfabetização ou destreza na leitura de palavras ou fluência no idioma corrente. Abreviadamente, o teste possui quatro etapas e em cada uma delas são anotados os tempos de execução e erros cometidos. Ademais, a partir destes escores são obtidas duas pontuações complementares, denominadas inibição e flexibilidade (Sedó, Paula & Malloy-Diniz, 2015).

Considerando a origem do FDT como um teste derivado do TSCP, hipotetizou-se que o mesmo se caracterizaria como uma medida altamente correlacionada, em termos de magnitude, com o teste Stroop, sugerindo evidências de validade de teste-critério convergente. Por validade convergente se entendem as relações significativas entre os escores do teste alvo e outras medidas que já detenham evidências consistentes de validade. Quando a validade convergente utiliza como método a correlação entre dois testes que se destinam ao alcance do mesmo construto, espera-se que estes dois instrumentos apresentem empiricamente altas correlações entre si. Neste caso, os coeficientes de correlação esperados devem ser de elevada magnitude, acima de 0,509 (Nunes & Primi, 2010). Entretanto, quando o teste que será utilizado como medida comparativa ao teste alvo não se destina a avaliação do mesmo construto, denomina-se evidências de validade por construtos relacionados. Nesta proposta, os coeficientes de correlação esperados são de moderada magnitude, variando entre 0,20 a 0,50 (Nunes & Primi, 2010).

Um conjunto de dados foram analisados, nesta pesquisa, com objetivo de verificar a existência de correlações estatisticamente significativas e suas magnitudes. Inicialmente, utilizando a amostra total foram comparados os tempos de execução das etapas do FDT e do teste Stroop. Considerando todas os escores parciais possíveis de serem relacionados, destaca-se a correlação positiva ( $\rho=0,53$ ,  $p<0,001$ ), estatisticamente significativa, entre o tempo de execução do cartão palavra (etapa 2) do Stroop e o tempo de leitura do FDT, ambas tarefas de leitura relacionadas à velocidade de processamento. Segundo Oliveira et al. (2014), no FDT, as duas primeiras etapas dependem primariamente da velocidade de processamento, e nelas, todos os coeficientes de correlação se mostraram positivos e significativos com as duas primeiras etapas do Stroop. As correlações entre tempo de leitura do FDT e cartão pontos do TSCP, tempo de leitura do FDT e cartão palavra, tempo de contagem do FDT e cartão pontos do TSCP, e tempo de contagem do FDT e cartão palavra foram de  $\rho=0,44$ ,  $\rho=0,53$ ,  $\rho=0,40$ , e  $\rho=0,55$  respectivamente, todos com valores de  $p<0,001$ . Tais dados mostram convergência no sentido de que as etapas 1 e 2 do FDT, que requerem a função de velocidade de processamento, estão moderadamente correlacionadas à etapa de leitura do Stroop, o que reforça a hipótese de que ambas convergem na avaliação deste construto, seguindo o parâmetro definido por Nunes e Primi (2010) para valores maiores que 0,509.

Nas duas últimas etapas do FDT, escolha e alternância, que envolvem processos controlados descritos como dependentes das funções executivas, as medidas de tempo de

execução novamente foram significativamente correlacionadas com o tempo de execução da etapa cor-palavra do Stroop ( $\rho=0,53$  e  $\rho=0,52$ ,  $p<0,001$ ). Entretanto, os escores complementares de inibição e de flexibilidade do FDT mostraram correlações fracas com o escore de interferência do Stroop ( $\rho=0,24$  e  $\rho=0,23$ , respectivamente,  $p=0,001$ ), diferentemente do que era esperado como hipótese inicial.

Na sequência, procedeu-se a divisão da amostra selecionando somente os indivíduos com idade superior a 49 anos ( $n=22$ ). Nesta análise, foi utilizada a variável logaritimizada do escore de interferência do Stroop, considerada como uma pontuação geral do teste, e, portanto, utilizou-se o coeficiente de correlação de Spearman. Para esta parcela da amostra, foi possível identificar o considerável aumento dos coeficientes de correlação.

O escore de interferência do Stroop e o escore de Inibição do FDT apresentaram um coeficiente de  $\rho=0,56$  ( $p=0,007$ ), enquanto o escore de interferência do Stroop e o escore de flexibilidade do FDT indicaram correlação igual a  $\rho=0,42$  ( $p=0,052$ ), apontando um maior relacionamento no que tange o construto de inibição de respostas.

Do mesmo modo, houve um aumento nas correlações entre os tempos de execução dos testes quando selecionados os adultos com idade maior que 49 anos. O tempo necessário para completar a etapa escolha do FDT foi altamente correlacionado com o tempo da etapa cor-palavra do Stroop ( $\rho=0,73$ ,  $p<0,001$ ), e o tempo gasto na etapa de alternância do FDT também obteve maior correlação - em comparação à mesma medida com a amostra total - com a etapa cor-palavra do Stroop ( $\rho=0,69$ ,  $p<0,001$ ) quando separados os grupos de indivíduos.

Tal resultado está de acordo com a dissertação de DeCristoforo (2000 apud Sedó & DeCristoforo, 2001), que administrou ambos os testes a uma amostra de 91 adultos idosos com idades superior a 60 anos. Todos os coeficientes de correlação apresentaram valores de  $r$  entre 0,50 e 0,71 ( $p=0,01$ ), similares aos obtidos nesta pesquisa. Para os autores, os altos índices correlacionais sugerem que ambos os instrumentos são análogos, e, portanto, mensuram construtos similares.

Em um estudo com 81 participantes dependentes químicos de diversas substâncias e 37 pessoas que compuseram o grupo controle, Verdejo-García e Pérez-García (2007) realizaram a aplicação de uma bateria de testes neuropsicológicos com objetivo de medir uma lista de habilidades cognitivas, e entre os testes figuravam o teste Stroop e o FDT. Após a utilização da

Análise dos Componentes Principais (rotação ortogonal), os autores identificaram a extração de 4 fatores: processos mentais, inibição, flexibilidade e tomada de decisão. O fator 2 foi composto por três medidas de interferência e controle atencional, que foram: escore de interferência do FDT, escore de alternância do FDT (nomenclaturas traduzidas conforme o artigo original, mas que são compatíveis em relação à tarefa executada com os escores de escolha e alternância de acordo com o manual técnico do teste utilizado nesta tese) e escore de interferência do Stroop. Na seção discussão, os autores indicaram que este fator extraído corresponde ao componente de inibição, tal como descrito por Miyake et al (2000). O cálculo para obter o escore das variáveis utilizadas no teste FDT foram a diferença entre o tempo de execução na parte 3 e a média das partes 1 e 2 (escore de interferência), e a diferença entre o tempo de execução na parte 4 e as médias de partes 1 e 2 (escore de alternância).

Outro estudo que traz resultados correlacionados entre os testes é de Paula et al. (2001). Com o objetivo de analisar a validade do FDT em pacientes com transtorno bipolar, os pesquisadores avaliaram 35 pacientes utilizando uma bateria de testes que contemplava o teste Stroop como uma medida de validade de “critério-relacionado”. Nas análises, encontraram correlações significativas entre todas as medidas de tempo de execução do FDT, tanto nas etapas automáticas, quanto controladas, e as medidas de tempo dos três cartões do teste Stroop. O coeficiente de correlação entre o cartão cor-palavra do Stroop e etapa escolha do FDT foi de 0,57, e entre o cartão cor-palavra do Stroop e etapa alternância do FDT foi de 0,52. Na pesquisa, os erros também se mostraram correlacionados entre as mesmas medidas, com coeficientes 0,83 e 0,69, entretanto, os autores não indicam a utilização dos escores de interferência como variáveis analisadas. Os resultados apresentados pelos autores convergem, portanto, com os obtidos nesta tese, indicando coeficientes de correlação estatisticamente significativos e de moderada magnitude entre as medidas.

#### 5.3.2.2 Estudos correlacionais entre o Teste não verbal de inteligência geral - BETA-III – Subteste Códigos e o Teste Stroop de Cores e Palavras

Como hipóteses iniciais, eram esperadas moderadas correlações entre o Teste Stroop (especificamente entre os tempos de execução dos cartões que compreendem as três etapas do teste) e o BETA-III – Subteste Códigos considerando que, na literatura, ambos os testes se

propõem a avaliar velocidade de processamento. A velocidade de processamento é definida como a capacidade que o indivíduo tem de realizar rapidamente tarefas simples, automatizadas, mantendo o foco atencional. Pode ser entendida, portanto, como uma medida de velocidade de conclusão de tarefas com precisão considerável. Usualmente, relaciona-se a tarefas de intervalor fixo, em que a pessoa executa o maior número possível de atividades simples e repetitivas (Primi, 2013).

Os resultados obtidos nesta pesquisa apontam para a confirmação das hipóteses iniciais que previam correlações, no mínimo, moderadas entre os tempos de execução das três etapas do teste Stroop, que são: cartão pontos, cartão palavra e cartão cor-palavra, e o escore geral do Subteste Códigos do BETA-III. Utilizando o coeficiente não paramétrico de Spearman, os valores foram de  $\rho = -0,53$ ,  $-0,65$  e  $-0,54$ , respectivamente (valores de  $p < 0,001$ ), para a amostra total. A média de tempo das duas primeiras etapas também foi correlacionada com o escore geral do BETA, com coeficiente igual a  $-0,64$  ( $p < 0,001$ ). Os resultados indicam um relacionamento negativo, o que significa que altas pontuações no BETA-III foram relacionadas a menores tempos de execução do Stroop.

Quando selecionados os adultos com idade igual ou superior a 49 anos, os dados mantiveram-se estatisticamente significativos, aumentando os valores dos coeficientes de correlação. Na nova análise, foram obtidos coeficientes entre  $-0,62$  e  $-0,75$  entre as medidas de tempo de execução das etapas do Stroop e o escore geral do Subteste Códigos do BETA ( $p < 0,001$ ).

Nota-se, contudo, um importante dado oriundo da comparação entre o escore de interferência e o escore geral do Subteste Códigos. Na amostra total, a correlação entre tais valores foi de  $-0,33$ , coeficiente menor do que os demais obtidos nas análises desta população. A mesma proporção pôde ser verificada, também, com o grupo de adultos com 49 anos ou mais. O coeficiente entre o escore de interferência do Stroop e o escore geral do BETA sofreu um acréscimo de valor, obtendo um  $\rho = -0,51$  ( $p < 0,001$ ), contudo, ainda é de magnitude menor do que os demais valores obtidos pelas correlações que envolvem somente a variável tempo, nestes indivíduos.

Tais resultados direcionam para a interpretação de que os maiores valores de correlação, entre os testes, se ligam às medidas que envolvem, somente, a variável tempo de execução do

Stroop. Como o escore de interferência, no Stroop, acrescenta a interferência entre a leitura automática em relação a nomeação da cor em que a palavra está impressa, a variável de inibição da resposta parece diminuir a magnitude da relação entre os testes. Em suma, os dados permitem inferir que as medidas de execução, em segundos, das três etapas do Stroop possuem maior relacionamento com o BETA-III, e que quando adicionada a variável de inibição/supressão das respostas, tal relação diminui. Quando esta variável é adicionada, portanto, percebe-se uma diminuição do coeficiente, prejudicando a correlação. Assim, sugere-se que os tempos de execução das etapas do Stroop possuem boas evidências de validade convergente a partir dos resultados do Subteste Códigos do BETA, contudo, a sensibilidade desta relação está na estimativa de velocidade de processamento, dada pelos tempos simples de execução das 3 tarefas do Stroop.

#### 5.3.2.3 Estudos correlacionais entre a Escala de Atenção Seletiva Visual – EASV e o Teste Stroop de Cores e Palavras

A atenção é considerada uma função cognitiva essencial para o adequado funcionamento das demais funções cognitivas. Na literatura neuropsicológica e cognitivista são definidos três tipos de atenção - em termos de funcionalidade - mais consensualmente utilizados, sendo eles atenção sustentada, alternada e seletiva. A atenção seletiva, frequentemente indicada como um dos construtos avaliados pelo teste Stroop, é conceituada como a capacidade de responder a estímulos específicos. Refere-se à habilidade de selecionar, em meio a uma gama de outros estímulos distratores, aqueles considerados relevantes (Sternberg, 2000).

A Escala de Atenção Seletiva Visual – EASV propõe-se avaliar a atenção seletiva por meio da busca de elementos semelhantes em conjuntos de itens que possuem vários elementos distratores. Quatro medidas são fornecidas para a interpretação dos resultados. A primeira, “capacidade seletiva visual básica” compreende itens compostos por menos elementos, seu nível de dificuldade é o mais baixo e a capacidade para resolver esses itens demanda uma habilidade atencional seletiva visual mais simples. A segunda “capacidade seletiva visual mediana” trabalha com uma exigência intermediária da atenção seletiva visual, e inclui grupos com seis elementos, considerados como um nível de exigência intermediária. A terceira

“capacidade seletiva visual superior” exige maior capacidade atencional e demanda que o respondente trabalhe com mais estímulos-alvos e mais distratores (nove elementos por item, com seis e sete distratores). Por fim, a “pontuação total” se caracteriza pela soma de pontos obtidos (total de acertos), independentemente do nível de exigência envolvido nos diferentes itens do teste.

Os escores de todas as medidas da EASV foram correlacionadas com todas as medidas do Stroop, e entre os resultados foi identificado que, na população geral, as correlações foram muito próximas de zero ou fracas, o que indicou pouco relacionamento entre as variáveis, hipótese contrária ao esperado por esta tese, tendo em vista o referencial teórico que associa o teste Stroop como uma medida de atenção seletiva (Klojčnik, Kavcic, & Bakracevic Vukman, 2017; Acera et al, 2019; Yilmaz, & Kafadar, 2020; Toyoshima, et al., 2019; Jongstra et al., 2017; MacLeod, 1991; MacLeod & MacDonald, 2000; Bench et al., 1993). Entretanto, contrariamente aos estudos que teoricamente indicam o Stroop como uma medida de atenção seletiva, percebe-se a carência de estudos que relatam esta relação direta de modo empírico.

Utilizando o teste d2 como uma medida de atenção seletiva, Ferreira (2016) correlacionou as variáveis deste teste com as variáveis do teste Stroop (utilizado com o propósito da avaliação de inibição) em uma amostra de adolescentes na cidade de Porto-Portugal. Dentre as variáveis de interesse, os resultados indicam ausências de correlações e correlações fracas, não ultrapassando coeficientes de correlação maiores do que  $r=0,47$ . Em relação ao indicador de concentração (IC) do teste d2, medida que indica a capacidade de concentração na tarefa, houve uma correlação significativamente positiva entre IC e o número de cores e palavras lidas no Stroop igual a  $r=0,44$  ( $p<0,001$ ). Entretanto, o mesmo indicador quando correlacionado ao escore de interferência do Stroop resulta em uma correlação significativa de  $r=0,14$  ( $p=0,032$ ), valor de baixa magnitude, que não indica convergência em relação ao construto de concentração e o escore de interferência do Stroop.

Posteriormente à separação da amostra, que excluiu os participantes com idade inferior a 49 anos, os dados apresentaram mudanças nos coeficientes de correlação. Os novos resultados apontaram o relacionamento estatisticamente significativo entre o tempo de leitura do cartão palavras e o escore total na EASV ( $\rho=0,66$ ;  $p=0,001$ ), tempo de leitura do cartão cor e palavra e escore total to EASV ( $\rho=0,58$ ;  $p=0,006$ ), e escore de interferência e escore total do EASV ( $\rho=0,55$ ;  $p=0,012$ ).

#### 5.3.2.4 Estudos correlacionais entre o teste Hayling e o Teste Stroop de Cores e Palavras

As hipóteses iniciais sobre os possíveis correlações entre os testes Stroop e Hayling foram refutadas a partir da análise dos coeficientes de correlação entre as tarefas. O esperado, presumia que os tempos de execução entre as etapas 1 e 2 do Stroop (cartão pontos e cartão palavras, respectivamente) fossem moderadamente correlacionadas com a etapa A do Hayling. Nesta, solicita-se ao participante que ele complete a frase fornecendo uma palavra com adequado sentido a sentença, o que, semelhantemente às etapas 1 e 2 do Stroop, não exigem supressão de resposta, mas fornecem uma medida de tempo para a execução de uma tarefa simples, tempo este, que será utilizado como base para o cálculo do índice de interferência de ambos os testes. Contrariamente ao previsto, os coeficientes de correlação se mostraram estatisticamente significativos, porém de baixa magnitude ( $\rho=0,28$ ,  $p=0,001$  entre cartão pontos do Stroop e etapa A do Hayling e  $\rho=0,32$ ,  $p=0,001$  entre cartão palavras do Stroop e etapa A do Hayling).

As etapas subsequentes de ambos os testes compreendem, então, uma tarefa que envolveria a supressão da resposta automática, habitual, em favor da menos automática, não habitual. No Stroop, denominada como etapa cor-palavra, compreende a nomeação da cor em que a palavra está impressa em detrimento da leitura da palavra. No Hayling, a etapa B, requer a complementação de uma frase com uma palavra que não possua nenhuma relação com o sentido da frase. Nestas, esperava-se, novamente, moderados coeficientes de correlação, tendo em vista a natureza da tarefa, o que foi identificado em termos das análises estatísticas, que apontaram um coeficiente de  $\rho=0,40$  ( $p=0,001$ ), indicando um relacionamento moderado entre ambas as tarefas, atendendo a hipótese de evidências de validade convergente por construtos relacionados.

Jantscher et al. (2011) realizaram um estudo utilizando ambos os testes com o intuito de analisar uma versão alemã do teste Hayling utilizando o teste Stroop como medida de correlação. Dentre os resultados, identificaram resultados estatisticamente significativos entre os cartões 1 e 2 do Stroop (cartão cor e cartão palavra) e a primeira etapa do Hayling (nomeada como etapa A). Para estas variáveis, encontraram um coeficiente que indica uma moderada correlação ( $r = 0.378$ ;  $p = 0.001$ ), valores similares aos apontados nesta tese, sugerindo

consonância entre os resultados. Entretanto, identificaram correlações sem significância estatística entre o cartão cor-palavra do Stroop e a etapa B do Hayling ( $r = 0.070$ ;  $p = 0.560$ ).

Coefficientes de correlação fracos ou nulos entre estes testes, também, foram obtidos por Borella et al. (2009) quando compararam os índices de interferência/inibição das medidas. Ao analisar uma amostra de adultos jovens e idosos, com ambas versões tradicionais e sem o controle de outras variáveis, identificaram a ausência de correlações estatisticamente significativas entre o Stroop e o Hayling. Tais estudos apresentam resultados, em termos dos coeficientes de correlação, díspares dos encontrados nesta tese. Jantscher et al. (2011) concluíram seus achados indicando a necessidade de investigações adicionais considerando a avaliação de funções executivas. No mesmo sentido, Borella et al. (2009) mencionam que processos inibitórios distintos e específicos parecem estar envolvidos na execução de cada uma das tarefas, logo, a inibição, não poderia ser considerada como um construto unitário e, portanto, os pesquisadores devem analisar mais profundamente as tarefas utilizadas e o tipo de inibição exigida. Sobre tais conclusões, permite-se concordar com a análise sobre as diferentes habilidades inibitórias exigidas em ambos os testes, o que faz jus ao moderado coeficiente de correlação encontrado entre eles, neste trabalho. Em termos de variância compartilhada os testes representam somente 16% de covariância, o que aponta que o teste Stroop e o teste Hayling podem ser medidas associadas, entretanto, se direcionariam a avaliação de diferentes atributos dentro do conceito de inibição, para a amostra total.

#### 5.3.2.5 Estudos correlacionais entre o Teste Wisconsin de Classificação de Cartas Modificado (WCSTM-48) e o Teste Stroop de Cores e Palavras

O Teste Wisconsin de Classificação de Cartas Modificado (WCSTM-48) é considerado um paradigma de avaliação das funções executivas (FEs) que é composto por escores em sete variáveis, que são: categorias completadas, considerada como uma medida geral do teste, acertos, ensaios administrados, erros totais, erros do tipo perseverativo, erros não perseverativos e rupturas.

Juntamente com o teste Stroop, o Wisconsin (WCST) é amplamente utilizado na avaliação das FEs em pacientes com diferentes quadros clínicos. Tradicionalmente, o WCST é utilizado para avaliar a flexibilidade cognitiva. Kosmidis et al. (2006) citam que o teste é

utilizado substancialmente em estudos com pacientes com esquizofrenia, demonstrando que tais pessoas possuem um desempenho pior na tarefa quando comparados a controles saudáveis.

Rossi et al. (1997) se propuseram a investigar os construtos compartilhados entre o Stroop e WCST por considerar que ambos os testes empregados na avaliação de funções executivas compartilham da necessidade de inibir uma determinada resposta e empregar uma estratégia diferente. No caso do Stroop, a inibição se daria na tendência de inibir o significado lexical do estímulo, e no WCST, na mudança de estratégia quando o respondente é informado de que o cartão anterior foi classificado incorretamente. A pesquisa proposta pelos autores contou com uma amostra de 25 pacientes com diagnóstico de esquizofrenia e entre os resultados obtidos tiveram correlações estatisticamente significativas entre número de erros perseverativos no WCST e o tempo gasto na etapa de cor-palavra do Stroop ( $r=0.34$ ,  $p=0.09$ ), bem como no escore de interferência do Stroop e na porcentagem de erros perseverativos para a amostra total ( $r=0.44$ ,  $p=0.08$ ) e para a amostra masculina ( $r=0,61$ ,  $p<0,01$ ). Com tais resultados, os autores apontaram que a hipótese de que um fator subjacente comum poderia influenciar os índices de desempenho do teste WCST e do Stroop em indivíduos esquizofrênicos parece ser confirmada pela correlação significativa encontrada entre os escores. Concluem, portanto, que o construto compartilhado seria denominado como *cognitive set shifting and maintenance*, traduzidos aqui como mudança e manutenção, considerados como homogêneos entre as tarefas.

Os dados de Rossi et al. (1997) apresentam convergência com os obtidos nesta pesquisa. Em relação à população geral, obteve-se correlações significativas entre algumas variáveis medidas por ambos os testes, como por exemplo, entre o escore de interferência do Stroop e os erros perseverativos no WCST, com valores de  $r=0,25$  ( $p=0,008$ ), sugerindo que pessoas com menores valores no escore de interferência (quão menor é o escore de interferência, melhor é o desempenho no teste, indicando maior habilidade na supressão da resposta automática em favor da menos automática) também apresentaram menor quantidade de erros perseverativos no WCST. Ainda, o escore de interferência no Stroop foi correlacionado com a variável categorias completadas, considerada como uma medida geral do WCST, com valores de  $r= -0,23$  ( $p=0,017$ ), o que indica que escores baixos de interferência no Stroop se relacionam a um bom desempenho geral no Wisconsin.

Entretanto, se seguirmos os valores indicativos de acordo com Nunes e Primi (2010), tais coeficientes, em relação à magnitude, indicam evidências de validade por construtos

relacionados. Tendo em vista que a medida de erros perseverativos no WCST, segundo Zimmermann e Fonseca (2017), se relaciona a dificuldade de inibir uma condição prévia (inibição) para iniciar uma nova estratégia de combinação (flexibilidade), e o escore de interferência do Stroop também se destina a capacidade de inibição, ambos, de acordo com as hipóteses desta tese, deveriam possuir maior percentual de variância compartilhada, o que não foi identificado. Entretanto, este dado remonta a discussão pontuada por Friedman e Miyake (2004) no Capítulo I, em relação ao equívoco de considerar inibição como um conceito unitário. Assim, mesmo indicados como medidas de tal função entende-se que a exigência entre as tarefas não pode ser considerada como análoga, o que justifica os coeficientes de correlação obtidos.

A partir dos achados de Rossi e colegas (1997), Kosmidis et al. (2006) realizaram nova pesquisa com o intuito de atualizar e ampliar o trabalho anterior no sentido de verificar se o WCST e o Stroop medem, de fato, processos cognitivos semelhantes em pacientes com esquizofrenia, removendo, contudo, o fator de velocidade de resposta, com objetivo de obter uma medida mais limpa de controle inibitório na tarefa Stroop. Em uma amostra com 55 pacientes com esquizofrenia, por meio de análises de regressões lineares, removeram a influência da velocidade na nomeação de cores e então, identificaram que o índice de interferência na tarefa Stroop não pode ser previsto por nenhuma das variáveis do WCST, incluindo erros perseverativos. Dentre os resultados, identificaram que a velocidade na nomeação de cores é responsável por 63% da variância na performance da etapa cor-palavra. Como conclusão adicional, apontam que os testes parecem medir componentes ou habilidades diferentes do controle inibitório, e reforçam que tal estudo demonstra que o controle inibitório não é um construto unitário e precisa ser estudado com o objetivo de diferenciar e isolar seus componentes, para então, ter uma medida mais limpa da função avaliada.

#### 5.3.2.6 Síntese dos coeficientes de correlação entre as tarefas neuropsicológicas e o teste Stroop.

A Tabela 30, apresentada na sessão resultados, sintetiza os maiores coeficientes de correlação, estatisticamente significativos, entre os testes/tarefas utilizados como fontes de evidência de validade do teste Stroop na população geral. Iniciando com os valores de maior magnitude, é possível observar as correlações obtidas entre o Stroop e o subteste códigos do BETA – III (aprovado pelo SATEPSI para avaliação do construto de velocidade de

processamento). Percebe-se que os maiores coeficientes obtidos se relacionam às variáveis de tempo de execução das três etapas do Stroop, ou ainda, da média aritmética entre as duas primeiras etapas do Stroop e o escore geral do BETA – III. Nota-se, ademais, que o coeficiente de correlação diminui – em termos de magnitude – quando este é correlacionado com o escore de interferência do Stroop. Para a medida de escore de interferência, é subtraído o tempo de execução da etapa 3, da média aritmética das etapas 1 e 2, ou seja, retira-se o efeito da velocidade/tempo gasto. Nesta operação, pode-se perceber que o relacionamento estatístico entre o escore de interferência (resultado residual, retirando a variável tempo de nomeação) e o BETA – III diminui em termos de grau de relacionamento. Tais resultados dão suporte à hipótese de que as evidências de validade se direcionam para a afirmação de que os escores de tempo do Stroop (em ambas etapas ou na média aritmética do cartão cor e do cartão palavra) e o construto de velocidade de processamento, indicado pelo manual do BETA – III, podem ser considerados convergentes, seguindo o critério indicado por Nunes e Primi (2010).

Na sequência, um conjunto de dados dão suporte a hipótese anterior, quando analisados os coeficientes de correlação entre o Stroop e o FDT. O FDT é descrito em seu manual de aplicação como um instrumento indicado para avaliação da velocidade de processamento, da atenção e das funções executivas (subcomponentes controle inibitório e flexibilidade cognitiva). As análises entre as diversas variáveis do Stroop e o FDT indicaram, novamente, maiores coeficientes de correlação entre as medidas de tempo de execução entre os testes, o que reforça a hipótese de evidências de validade convergente entre o Stroop e as medidas de velocidade de processamento do FDT. Adicionalmente, ainda com valores de coeficientes de correlação maiores do que 0,509, tem-se o relacionamento entre o cartão cor-palavra do Stroop e as etapas controladas de escolha e alternância (ambas medidas pelo tempo de execução) do FDT. Diferentemente do esperado, os escores derivados de inibição e flexibilidade indicaram fraco relacionamento com o escore de interferência do Stroop. Esperava-se que o escore de interferência do Stroop fosse altamente correlacionado com o escore de inibição do FDT, já que ambas as tarefas possuem o mesmo modo de funcionamento, e derivam de um sistema de pontuação semelhante. Contudo, como tal resultado não ocorreu, pontuam-se as discussões trazidas por Sedó, Paula e Malloy-Diniz (2015) sobre o tipo de estímulos utilizados (palavras *versus* numerais), e sobre as habilidades subjacentes na execução dos testes, como, por exemplo, o nível de destreza na identificação de palavras no caso do Stroop que pode afetar a

validade do teste, tópico que emerge outras discussões, como, por exemplo, a impureza das tarefas.

Dando sucessão, observa-se os valores correlacionais entre o Stroop e o teste Hayling. Este teste é indicado como uma medida de avaliação do controle inibitório, mais precisamente da iniciação e inibição verbal. A parte B, relacionada ao construto de inibição, obteve um coeficiente de correlação moderado com o teste Stroop. O escore de interferência do Stroop e o escore geral do Hayling também apresentaram um resultado moderado em termos de magnitude. Tais valores sugerem evidências de validade por construtos relacionados, pois se enquadram entre 0,20 e 0,50 (Nunes & Primi, 2010). Magalhães (2013) cita, em seu trabalho, o debate entre a definição e a circunscrição do fenômeno relativo ao controle inibitório, mencionando como promissora a ideia de questionar a validade unitária deste construto, já que, atualmente, há o entendimento de que o controle inibitório é formado por várias habilidades subjacentes e que se faz necessário abdicar do termo inibição como um conceito guarda-chuva, para então, especificar, definir e investigar os subtipos distintos desses processos, pois fenômenos diversos, como, conflito, hesitação, interferência e apreensão têm sido, não especificadamente, designados como inibição.

Avançando, aponta-se para as análises entre o Teste Wisconsin de Classificação de Cartas (WCST) e o Stroop. O WCST é um teste utilizado na avaliação de funções executivas, baseado na flexibilidade de pensamento do sujeito. Propõe-se mensurar a capacidade do indivíduo em raciocinar abstratamente e modificar suas estratégias cognitivas como resposta a alterações nas contingências ambientais. A variável “categorias completadas”, do WCST, pode ser indicativa de diferentes funções cognitivas, como inibição, atenção e flexibilidade cognitiva. Tal variável, quando correlacionada com o Stroop, apresentou baixos coeficientes de correlação. As variáveis “erros” e “erros perseverativos” que, também, apresentaram correlações estatisticamente significativas com o escore de interferência do Stroop, porém de baixa magnitude, são referenciadas, por Zimmermann e colegas (2015), como relacionadas a dificuldades de inibir uma condição prévia (inibição) para iniciar uma nova estratégia de combinação (flexibilidade). Considerando tais resultados, nesta tese, indica-se que os dados apontam evidências de validade por construtos relacionados (0,20 e 0,50 segundo Nunes e Primi, 2010), contudo, ressalta-se que os valores foram limítrofes, o que sugere cautela em tal apontamento, estando próximos dos indicadores de evidências de validade divergente.

Por fim, os menores coeficientes foram entre o teste Stroop e a Escala de Atenção Seletiva Visual (EASV). Tanto no cartão palavras, quanto no cartão cor-palavra do Stroop, os coeficientes de correlação com a pontuação do EASV mostraram ser de fracas magnitudes, indicando evidências de validade divergentes, pois foram inferiores a 0,20. O escore de interferência do Stroop e a pontuação básica do EASV, novamente, se mostrou estatisticamente correlacionado, porém, em valor limítrofe. Considerando o volume de publicações que relacionam o uso do teste Stroop como uma medida de atenção seletiva, os resultados ora obtidos foram antagônicos dos esperados como hipóteses, já que era estimado que a atenção seletiva fosse uma variável com maior correlação em alguma das etapas do Stroop ou com algum dos escores possíveis.

#### 5.3.2.7 Testes neuropsicológicos e o problema da impureza das tarefas

Um aspecto que merece destaque na observação dos resultados é o problema da impureza das tarefas. Strauss et al. (2006) indicam, por exemplo, a existência de mecanismos psicológicos subjacentes à realização da tarefa Stroop incluindo, memória de trabalho, velocidade de processamento de informação, ativação semântica e habilidade para resistir a uma resposta característica, além dos construtos relacionados às funções executivas, como atenção seletiva, flexibilidade cognitiva e suscetibilidade a interferência (Strauss, Sherman, & Spreen, 2006). Friedman e Gustavson (2022) adicionam, ainda, habilidades que envolvem a percepção de cores, leitura de palavras e velocidade vocal. O problema da impureza estaria, portanto, associado ao fato de que para a execução de uma tarefa neuropsicológica outras funções seriam recrutadas para o seu desempenho, não sendo possível identificar, por conseguinte, somente o conteúdo cognitivo almejado.

Van der Sluis et al. (2006) relacionam as dificuldades fundamentais encontradas nos estudos que mensuram funções executivas ao problema da impureza das tarefas. Como as FEs demandam de uma estrutura de tarefas para serem manifestas, elas sempre envolvem outras habilidades cognitivas não executivas, como, por exemplo, habilidade verbal, velocidade motora ou habilidade visuo-espacial. Mariano (2017) cita, por exemplo, a linguagem como uma função mediadora com papel significativo nas tarefas neuropsicológicas e, portanto, esta função se torna uma espécie de impureza na mensuração da performance em diversas tarefas

cognitivas. Assim, o desempenho em uma tarefa executiva não pode ser atribuído, simplesmente, a presença ou a ausência de uma determinada capacidade executiva.

Friedman e Miyake (2004), citados anteriormente no Capítulo II, ao se referirem ao construto de inibição, realçam que nenhuma tarefa das utilizadas em seu trabalho poderia ser considerada uma medida pura para tal construto. Exemplificam que o processo de inibição é sempre a inibição de algo e por tal, o resultado - em termos de performance - envolve outros processos. Por este motivo, não se poderia justificar que baixas pontuações em uma tarefa seriam derivadas, somente, de uma inabilidade na função, ou ainda, que baixas correlações entre diferentes testes podem não necessariamente ser decorrentes de diferentes processos inibitórios, mas sim, justificadas por outras variáveis que podem mascarar as semelhanças atribuíveis (Miyake, Friedman, et al., 2000; Shilling et al., 2002).

A impureza das tarefas é considerada uma das justificativas para pesquisas que utilizam análises correlacionais entre testes e que encontram baixos coeficientes de correlações entre as medidas, mesmo quando estas são consideradas razoavelmente confiáveis Friedman e Miyake (2017). Uma possível solução para este problema seria a administração de diferentes testes direcionados ao mesmo processo, mas que diferem em seus requisitos básicos de execução, e a posterior extração da variância comum pelo modelo de variáveis latentes. Como as variáveis latentes capturam apenas a variação que se correlaciona entre as tarefas, elas se tornam mais puras e removem erros de medição aleatórios (Friedman & Gustavson, 2022).

### 5.3.2.8 Levantamento de hipóteses a respeito das análises correlacionais entre o grupo de pessoas com idade igual ou superior a 49 anos

Com o processo natural do envelhecimento, algumas alterações em diferentes áreas da cognição e do comportamento podem ocorrer. De acordo com Malloy-Diniz, Fuentes e Cosenza (2013) as FEs atingem maturidade funcional na idade adulta, iniciando, um declínio natural a partir do fim da terceira década de vida. Vários mecanismos neurobiológicos são mencionados como associados a este processo, como, por exemplo, a diminuição do peso cerebral, que se inicia por volta dos 30 anos de idade, processo que se acentua com o envelhecimento. Destacam, também, os circuitos pré-frontais, principalmente a substância branca, que parece ter maior

precocidade no processo natural de degeneração em comparação com outras regiões encefálicas.

Em termos cognitivos, a diminuição da velocidade de processamento, apontada por Span, Ridderinkof e van der Molen, (2004), é identificada como um fator que influencia outras tarefas relacionadas às FEs. Segundo os autores, há a predominância da hipótese de que as mudanças na velocidade de processamento das informações – relacionadas ao incremento da idade – respondem pela maioria das variações relevantes no desempenho em uma ampla gama de tarefas cognitivas, independentemente da composição dessas tarefas. Green (2000) acrescenta que no envelhecimento, os déficits observados nas tarefas neuropsicológicas relacionadas ao controle executivo e a memória episódica ocorrem por estarem associados à diminuição no processamento da informação, nos processos atencionais, inibitórios e na flexibilidade cognitiva.

No presente trabalho, com exceção do teste Hayling, os coeficientes de correlação entre o teste Stroop e as demais tarefas sofreram um incremento significativo, em termos da força dos relacionamentos, quando selecionados os casos de respondentes com idade igual ou superior a 49 anos. A respeito destes achados, levantam-se algumas hipóteses, indicadas como inquietações para pesquisas futuras.

A primeira delas, refere-se à capacidade de discriminação do teste Stroop em diferentes amostras/grupos. O Stroop é reconhecido como um instrumento útil na avaliação de idosos, ou ainda, em pacientes clínicos em diferentes estágios de quadros demenciais (Malloy-Diniz, Fuentes e Cosenza, 2013). Do mesmo modo, testes como o Wisconsin e o teste dos 5 dígitos - FDT (Sedó; Paula & Malloy-Diniz, 2015) também são indicados na literatura como ferramentas úteis para estas especificidades. Considerando isto, é possível questionar se a sensibilidade destes testes não estaria relacionada à identificação de déficits cognitivos, que pouco ocorrem em pacientes mais jovens, ou ainda, se em pacientes mais jovens, os testes teriam menor sensibilidade na identificação de seus construtos, tendo em vista um aporte cognitivo dito como saudável, ou dentro do esperado.

A respeito desta hipótese, Ben-David et al. (2014) investigaram o impacto da Doença de Alzheimer (DA) no desempenho do teste Stroop. Como delineamento, realizaram um estudo de meta-análise e modelagem matemática de 18 trabalhos, comparando 637 pacientes com DA

e 977 pacientes saudáveis da mesma faixa etária. Dentre os resultados, encontraram um aumento significativo em todos os estudos no escore de efeito do teste Stroop nos pacientes com DA. Ademais, observaram que esta diferença entre os grupos foi associada a lentidão na velocidade generalizada no processamento, que explicou uma porção moderada da variância total, equivalente à 25%.

Os dados apresentados por Ben-David et al. (2014) destacam a importância da velocidade generalizada do processamento na interpretação dos resultados do Stroop, constatação que adiciona outra hipótese sobre os dados desta presente pesquisa. Considerando que os testes Stroop, Códigos do BETA-III e FDT são instrumentos em que os escores são apurados pelos tempos de execução, poderiam eles ser mais suscetíveis ao avanço da idade. Contudo, tal avanço, não necessariamente indicaria maior prejuízo nas demais funções cognitivas avaliadas, mas sim, que o modo de obtenção de escores seria um viés na análise dos dados. Para tal indagação, sugere-se que pesquisas futuras analisem o formato de obtenção de escores, com o intuito de demarcar quais são os melhores indicadores, considerando os diferentes grupos amostrais.

#### 5.4 CONCLUSÃO

O Teste Stroop de Cores e Palavras consiste em uma tarefa neuropsicológica utilizada tanto para finalidades clínicas, quanto para pesquisas, nacional e internacionalmente. Ao longo de sua existência, que remonta ao ano de 1935, somam-se inúmeras traduções para diferentes idiomas, versões que introduziram modificações em termos do formato de apresentação, distintos métodos para obtenção dos escores, além de pesquisas com diversos grupos clínicos, não clínicos, com participantes de variadas faixas etárias e, ainda, com diversificados propósitos avaliativos, abarcando uma variedade de construtos ditos como alvos de mensuração.

A utilização de um teste no campo da avaliação psicológica ou neuropsicológica implica na escolha de ferramentas confiáveis para a observação de construtos ou traços latentes. Para tal, é imprescindível que a escolha destas ferramentas seja pautada nos fundamentos teóricos que as sustentam, bem como, no apanhado de evidências de validade, que legitimam as interpretações dos resultados obtidos por meio dela. Tais pressupostos, são considerados parâmetros mínimos para a confiabilidade das interpretações realizadas a partir dos escores

obtidos, quando ponderados dentro dos usos e propósitos pré-avaliados pelas pesquisas anteriores.

Reconhecendo a importância destes parâmetros, esta pesquisa de doutorado teve por objetivo realizar a busca de evidências de validade do Teste Stroop de Cores e Palavras para adultos. Para tal, inicialmente, levantou-se, por meio de uma revisão integrativa de literatura, quais construtos psicológicos eram mencionados no escopo da utilização do Stroop. Posteriormente, foram selecionados os construtos com maior frequência de citações e as respectivas medidas neuropsicológicas que reuniam evidências prévias de validade para tais indicações. Por fim, por meio da análise estatística dos dados, foram desenvolvidos estudos de evidências de validade convergente, divergente e por construtos relacionados entre o Stroop e as medidas utilizadas, que foram, o Teste dos Cinco Dígitos – FDT, Wisconsin Teste de Classificação de Cartas – WCST, Teste de Avaliação da Atenção Seletiva – EASV, BETA-III – Subteste Códigos e *Hayling Sentence Completion Test* (HSCT). Tais medidas foram eleitas por avaliarem alguns dos subcomponentes das funções executivas, como inibição, flexibilidade cognitiva, atenção seletiva, e também, a velocidade de processamento.

Dentre as possíveis variáveis de análise correlacionadas e avaliadas, alguns dados se sobressaíram e foram discutidos de acordo com a literatura científica. Em relação às variáveis sociodemográficas, os escores obtidos pelos participantes da amostra indicaram haver um relacionamento entre o escore de interferência do Stroop e a idade dos participantes, e, também, no número de erros cometidos, possibilitando identificar diferenças estatisticamente significativas entre o grupo de adultos com idade igual ou superior a 49 anos e o grupo de adultos jovens, indicando melhor desempenho dentre os mais jovens, e pior desempenho entre os mais velhos. Não foram identificadas diferenças entre as variáveis de escolaridade e gênero dos respondentes. É importante ressaltar que a literatura aponta o fator escolaridade como uma variável que influencia o desempenho do Stroop, porém, nesta pesquisa, a amostra não obteve participação de um número considerável de pessoas com baixa escolaridade, o que pode ter limitado a capacidade de replicar as diferenças indicadas na literatura. Sugere-se, que pesquisas futuras incluam um número mais expressivo de pessoas com ensino básico e fundamental para dirimir dúvidas em relação a esta questão.

A respeito dos estudos de evidências de validade, o montante de dados permitiu um número relevante de análises. Destas, sintetizam-se alguns apontamentos. As principais fontes

de evidências de validade convergentes indicaram relacionamento entre o Stroop e as medidas de velocidade de processamento dos testes FDT e Códigos do BETA-III. Por evidências de construtos relacionados, foram obtidos coeficientes de correlação moderados entre o Stroop e a medida de inibição (Hayling). Com valores limítrofes entre evidências de validade por construtos relacionados e divergentes, foram observadas as correlações entre o Stroop e as medidas do teste Wisconsin que avaliam inibição e flexibilidade, bem como, entre estes mesmos indicadores do FDT. E, por fim, os dados apontaram evidências de validade divergente entre o Stroop e os escores da EASV, utilizada na avaliação da atenção seletiva.

Faz-se necessária uma ponderação na interpretação dos dados, que diz respeito ao problema da impureza das tarefas. Já devidamente mencionado no presente trabalho, merece destaque o fato que nenhuma tarefa neuropsicológica é considerada como pura na mensuração de seus construtos, pois se entende que para execução de uma ação outras funções também serão recrutadas, o que gera um desafio a mais na análise e interpretação dos resultados, tanto no campo da pesquisa, quanto no campo clínico. É recomendado, para próximas pesquisas, a utilização de diferentes modelos de análise dados, estratégia que pode reduzir tal problemática.

Outro resultado sobressaliente diz respeito à mudança na magnitude dos relacionamentos observados entre os coeficientes de correlação quando selecionados os participantes da amostra com idade igual ou superior a 49 anos. De modo geral, neste recorte, os valores assumiram maiores magnitudes, levantando a hipótese de maior sensibilidade das medidas utilizadas para esta parcela da população estudada. Ademais, outra hipótese sugerida para pesquisas futuras, se trata da verificação da influência da velocidade generalizada do processamento no aumento da idade, e se este indicador pode ser uma variável significativa que influencia o desempenho nas demais medidas neuropsicológicas, mesmo naquelas em que o construto almejado seja diferente deste.

Por fim, este estudo apresenta evidências de validade do Teste Stroop de Cores e Palavras. Pesquisas com esse propósito são essenciais para o avanço nessa área, pois fornecem fundamentos que conferem maior legitimidade às interpretações dos escores dos instrumentos e justificam seu uso para finalidades práticas e, também, em pesquisas. Investigações sobre a validade dos testes contribuem para a melhoria e progresso dos instrumentos de avaliação, além de ajudar a identificar lacunas que ressaltam a importância de pesquisas adicionais e atualizações do conhecimento.

## REFERÊNCIAS

- Acera, M., Molano, A., Tijero, B., Bilbao, G., Lambarri, I., Villoria, R., ... & Gomez-Esteban, J. C. (2019). Impacto de la estimulación subtalámica a largo plazo sobre la situación cognitiva de los pacientes con enfermedad de Parkinson avanzada. *Neurología*, 34(9), 573-581. doi: [10.1016/j.nrl.2017.05.009](https://doi.org/10.1016/j.nrl.2017.05.009)
- Adólfssdóttir, S., Wollschlaeger, D., Wehling, E., & Lundervold, A. J. (2017). Inhibition and switching in healthy aging: a longitudinal study. *Journal of the International Neuropsychological Society*, 23(1), 90-97. doi: 10.1017/S1355617716000898
- Albein-Urios, N., Martínez-González, J. M., Lozano, Ó., Clark, L., & Verdejo-García, A. (2012). Comparison of impulsivity and working memory in cocaine addiction and pathological gambling: Implications for cocaine-induced neurotoxicity. *Drug and alcohol dependence*, 126(1-2), 1-6. doi: [10.1016/j.drugalcdep.2012.03.008](https://doi.org/10.1016/j.drugalcdep.2012.03.008)
- Alizadeh, F., Tavakkoly-Bazzaz, J., Bozorgmehr, A., Azarnejhad, A. A., Tabrizi, M., & Shahsavand Ananloo, E. (2017). Association of transcription factor 4 (TCF4) gene mRNA level with schizophrenia, its psychopathology, intelligence and cognitive impairments. *Journal of Neurogenetics*, 31(4), 344-351. doi: [10.1080/01677063.2017.1396330](https://doi.org/10.1080/01677063.2017.1396330)
- American Educational Research Association AERA, American Psychological Association APA, National Council on Measurement in Education NCME (1999). *Standards for educational and psychological testing*. New York: American Educational Research Association.
- American Educational Research Association AERA, American Psychological Association APA, National Council on Measurement in Education NCME (2014). *Standards for educational and psychological testing*. New York: American Educational Research Association.
- Arhan, E., Gücüyener, K., Soysal, Ş., Şalvarlı, Ş., Gürses, M. A., Serdaroğlu, A., ... & Atalay, Y. (2017). Regional brain volume reduction and cognitive outcomes in preterm children at

- low risk at 9 years of age. *Child's Nervous System*, 33, 1317-1326. doi: 10.1007/s00381-017-3421-2
- Arslan, F., Tasdemir, S., Durmaz, A., & Tosun, F. (2018). The effect of nasal polyposis related nasal obstruction on cognitive functions. *Cognitive neurodynamics*, 12(4), 385-390. doi: [10.1007/s11571-018-9482-4](https://doi.org/10.1007/s11571-018-9482-4)
- Bayard, S., Gély-Nargeot, M. C., Raffard, S., et al. French Version of the Hayling Sentence Completion Test, Part I: Normative Data and Guidelines for Error Scoring. *Arch Clin Neuropsychol*. 2017;32(5):585–591. doi:10.1093/arclin/acx010
- Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind, Brain, and Education*, 10(1), 10-33. doi: [10.1111/mbe.12100](https://doi.org/10.1111/mbe.12100)
- Ben-David, B. M., Tewari, A., Shakuf, V., & Van Lieshout, P. H. (2014). Stroop effects in Alzheimer's disease: selective attention speed of processing, or color-naming? A meta-analysis. *Journal of Alzheimer's Disease*, 38(4), 923-938. doi: 10.3233/JAD-131244
- Bench, C., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, 31(9), 907-922. doi: [10.1016/0028-3932\(93\)90147-R](https://doi.org/10.1016/0028-3932(93)90147-R)
- Brandelero, V., & de Toni, P. M. (2015). Estudo de validade do Teste Stroop de Cores e Palavras para controle inibitório. *Revista Psicologia e Argumento*, 33 (80), 282-297. doi: [10.7213/psicol.argum.33.080.AO05](https://doi.org/10.7213/psicol.argum.33.080.AO05)
- Borella, E., Delaloye, C., Lecerf, T., Renaud, O., & De Ribaupierre, A. (2009). Do age differences between young and older adults in inhibitory tasks depend on the degree of activation of information?. *European Journal of Cognitive Psychology*, 21(2-3), 445-472. doi: [10.1080/09541440802613997](https://doi.org/10.1080/09541440802613997)
- Borella, E., Carretti, B., Mitolo, M., Zavagnin, M., Caffarra, P., Mammarella, N., ... & Piras, F. (2017). Characterizing cognitive inhibitory deficits in mild cognitive impairment. *Psychiatry Research*, 251, 342-348. doi: [10.1016/j.psychres.2016.12.037](https://doi.org/10.1016/j.psychres.2016.12.037)

- Cammisuli, D. M., & Timpano Sportiello, M. (2017). Conscious attention defect and inhibitory control deficit in Parkinson's disease-mild cognitive impairment: A comparison study with amnesic mild cognitive impairment multiple domain. *Psychiatria Danubina*, 29(4), 437-455. doi: [10.24869/psyd.2017.437](https://doi.org/10.24869/psyd.2017.437)
- Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological assessment: Past and future. *Journal of the International Neuropsychological Society*, 23(9-10), 778-790. doi:10.1017/S1355617717001060
- Casarin, F. S. (2010). *Bateria Montreal de avaliação da comunicação breve – Bateria Mac Breve: Estudos de Adaptação*. (Dissertação de mestrado). Recuperado de <http://repositorio.pucrs.br/dspace/handle/10923/4905>
- Caixeta, L., Soares, V. L., Vieira, R. T., Soares, C. D., Caixeta, V., Ferreira, S. B., & Aversif-Ferreira, T. A. (2017). Executive Function Is Selectively Impaired in Old Age Bipolar Depression. *Frontiers in psychology*, 8, 194. doi: 10.3389/fpsyg.2017.00194
- Cao, Y., Chen, X., Xie, H., Zou, L., Hu, L. J., & Zhou, X. J. (2017). Correlation between Electroencephalogram Alterations and Frontal Cognitive Impairment in Esophageal Cancer Patients Complicated with Depression. *Chinese medical journal*, 130(15), 1785–1790. doi:10.4103/0366-6999.211552
- Clerici, F., Ghiretti, R., Di Pucchio, A., Pomati, S., Cucumo, V., Marcone, A., ... & Cappa, S. F. (2017). Construct validity of the Free and Cued Selective Reminding Test in older adults with memory complaints. *Journal of neuropsychology*, 11(2), 238-251. doi: [doi.org/10.1111/jnp.12087](https://doi.org/10.1111/jnp.12087)
- Conselho Federal de Psicologia (2018). Resolução nº 09/2018 que estabelece diretrizes para a realização e Avaliação Psicológica no exercício profissional da psicóloga e do psicólogo, regulamenta o Sistema de Avaliação de Testes Psicológico – SATEPSI e revoga as Resoluções nº 002/2003 , nº 006/2004 e nº 005/2012 e Notas Técnicas nº 01/2017 e 02/2017.
- Cotrena, C., Branco, L. D., Ponsoni, A., Shansis, F. M., Kochhann, R., & Fonseca, R. P. (2017). The predictive role of daily cognitive stimulation on executive functions in bipolar disorder. *Psychiatry Research*, 252, 256-261. doi: [10.1016/j.psychres.2017.03.011](https://doi.org/10.1016/j.psychres.2017.03.011)

- Cunha, J. A., Trentini, C. M., Argimon, I. L. Oliveira, M. S. Werlang, B. G., & Prieb, R. G (2005). *Teste Wisconsin de classificação de cartas: Manual*. São Paulo: Casa do Psicólogo.
- Cunha, P. J., & Novaes, M. A. (2004). Avaliação neurocognitiva no abuso e dependência do álcool: implicações para o tratamento. *Revista Brasileira de Psiquiatria*, 26, 23-27. doi: [10.1590/S1516-44462004000500007](https://doi.org/10.1590/S1516-44462004000500007).
- Davison, S., Deeprase, C., & Terbeck, S. (2018). A comparison of immersive virtual reality with traditional neuropsychological measures in the assessment of executive functions. *Acta neuropsychiatrica*, 30(2), 79–89. doi: [10.1017/neu.2017.14](https://doi.org/10.1017/neu.2017.14)
- de Moura Kolling, N., da Silva, C. R., Carvalho, J. C. N., da Cunha, S. M., & Kristensen, C. H. (2007). Avaliação neuropsicológica em alcoolistas e dependentes de cocaína. *Avaliação Psicológica: Interamerican Journal of Psychological Assessment*, 6(2), 127-137. Recuperado em 11 de abril de 2018, de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712007000200003&lng=pt&tlng=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712007000200003&lng=pt&tlng=pt).
- Della Corte, et. al. (2018). A simple measure of cognitive reserve is relevant for cognitive performance in MS patients. *Neurological Sciences*, 39(7), 1267-1273. doi: [10.1007/s10072-018-3422-2](https://doi.org/10.1007/s10072-018-3422-2)
- Del-Bem, C. M. (2005). Neurobiologia do transtorno de personalidade anti-social. *Revista de Psiquiatria Clínica*, 32 (1), 27-36. doi: [10.1590/S0101-60832005000100004](https://doi.org/10.1590/S0101-60832005000100004)
- de Paula, J. J., Abrantes, S., Silva Neves, F., & Fernandes Malloy-Diniz, L. (2014). The Five Digits Test on the assessment of psychiatric patients with heterogeneous educational backgrounds: Evidences of validity on the assessment of bipolar disorder. *Clinical Neuropsychiatry*, 11 (3), 103-107.
- Dias, N. M. (2009). Avaliação neuropsicológica das funções executivas: tendências desenvolvimentais e evidências de validade de instrumentos (Dissertação de mestrado). Recuperado de <http://tede.mackenzie.br/jspui/handle/tede/1746>

- Dong, G., Zhou, H., & Zhao, X. (2011). Male Internet addicts show impaired executive control ability: evidence from a color-word Stroop task. *Neuroscience letters*, 499(2), 114–118. <https://doi.org/10.1016/j.neulet.2011.05.047>
- Diamond A. (2013). Executive functions. *Annual review of psychology*, 64, 135–168. [doi: 10.1146/annurev-psych-113011-143750](https://doi.org/10.1146/annurev-psych-113011-143750)
- Duchesne, M., Mattos, P., Fontenelle, L. F., Veiga, H., Rizo, L., & Appolinario, J. C. (2004). Neuropsicologia dos transtornos alimentares: revisão sistemática da literatura. *Brazilian Journal of Psychiatry*, 26(2), 107-117. doi: 10.1590/S1516-44462004000200008
- Duncan, M. T. (2006). Obtenção de dados normativos para desempenho no teste de Stroop num grupo de estudantes do ensino fundamental em Niterói. *Jornal Brasileiro de Psiquiatria*. doi: [10.1590/S0047-20852006000100006](https://doi.org/10.1590/S0047-20852006000100006).
- Eggermont, L. H., Aly, M. F., Vuijk, P. J., de Boer, K., Kamp, O., van Rossum, A. C., & Scherder, E. J. (2017). Cardiac function and cognition in older community-dwelling cardiac patients. *Psychogeriatrics*, 17(6), 356-363. doi: [10.1111/psyg.12245](https://doi.org/10.1111/psyg.12245)
- Faria, C. D. A., Alves, H. V. D., & Charchat-Fichman, H. (2015). The most frequently used tests for assessing executive functions in aging. *Dementia & neuropsychologia*, 9(2), 149-155. doi: [10.1590/1980-57642015DN92000009](https://doi.org/10.1590/1980-57642015DN92000009)
- Ferreira, M. B., Pereira, P. A., Parreira, M., Sousa, I., Figueiredo, J., Cerqueira, J. J., & Macedo, A. F. (2018). Relationships between neuropsychological and antisaccade measures in multiple sclerosis patients. *PeerJ*, 6, e5737. doi:10.7717/peerj.5737.
- Fernández-Serrano, M. J., Perales, J. C., Moreno-López, L., Pérez-García, M., & Verdejo-García, A. (2012). Neuropsychological profiling of impulsivity and compulsivity in cocaine dependent individuals. *Psychopharmacology*, 219(2), 673-683. doi: [10.1007/s00213-011-2485-z](https://doi.org/10.1007/s00213-011-2485-z)
- Ferreira, M. F. B. N. (2016). *Atenção, inibição e resistência psicológica à mudança*. (Dissertação de mestrado). Recuperado de <http://repositorio.ulusiada.pt/bitstream/11067/2570/1/Dissertacao%20de%20mestrado.pdf>

- Friedman, N. P., & Gustavson, D. E. (2022). Do rating and task measures of control abilities assess the same thing?. *Current directions in psychological science*, 31(3), 262-271. doi: 10.1177/09637214221091824
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of experimental psychology: General*, 133(1), 101. doi: 10.1037/0096-3445.133.1.101
- Gindri, G., Zibetti, M. R., & Fonseca, R. P. (2008). Funções executivas pós-lesão de hemisfério direito: estudo comparativo e frequência de déficits. *Psico*, 39(3), 8. Recuperado em 05 de junho de 2018, de <https://dialnet.unirioja.es/servlet/articulo?codigo=5161432>
- Golden, C. J., & Freshwater, S. M. (2002). *Stroop Color and Word Test Adult Version. A manual for clinical and experimental uses*. (2<sup>a</sup> ed.). Wood Dale, Illinois: Stoelting.
- Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A history of executive functioning as a theoretical and clinical construct. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (p. 3–12). Springer Science + Business Media. doi: [10.1007/978-1-4614-8106-5\\_1](https://doi.org/10.1007/978-1-4614-8106-5_1)
- Grande, P. H. A. (2013). Avaliação neuropsicológica das funções executivas em idosos com comprometimento cognitivo leve e demência do tipo Alzheimer. (Dissertação de mestrado). Recuperado de <https://acervodigital.ufpr.br/handle/1884/30646>
- Guevara, M. L. C, Fernández-Serrano M. J, Del Paso R. G. A, & Duschek S. Executive function impairments in fibromyalgia syndrome: Relevance of clinical variables and body mass index. *PLoS One*. 2018;13(4):e0196329. Published 2018 Apr 25. doi:10.1371/journal.pone.0196329
- Haldane, M., & Frangou, S. (2005). Maudsley Bipolar Disorder Project: insights sobre o papel do córtex préfrontal em pacientes com transtorno de humor bipolar tipo I. *Revista de Psiquiatria do Rio Grande do Sul*, 27 (3), 241-250. doi: [10.1590/S0101-81082005000300003](https://doi.org/10.1590/S0101-81082005000300003)

- Hamdan, A. C., & Pereira, A. P. D. A. (2009). Avaliação neuropsicológica das funções executivas: considerações metodológicas. *Psicologia: Reflexão e Crítica*, 22(3), 386-393. doi: [10.1590/S0102-79722009000300009](https://doi.org/10.1590/S0102-79722009000300009)
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (2005). *Teste Wisconsin de Classificação de Cartas: manual revisado e ampliado. Adaptação e padronização brasileira*. Cunha JÁ, Trentini CM, Argimon IL, Oliveira MS, Werlang BG, Prieb RG. São Paulo: Casa do Psicólogo.
- Huang, M., Lu, S., Yu, L., Li, L., Zhang, P., Hu, J., ...Weng, J. (2017). Altered fractional amplitude of low frequency fluctuation associated with cognitive dysfunction in first-episode drug-naïve major depressive disorder patients. *BMC psychiatry*, 17(1), 11. doi: [10.1186/s12888-016-1190-1](https://doi.org/10.1186/s12888-016-1190-1).
- Hutz, C. S. (2015). O que é avaliação psicológica – métodos, técnicas e testes. In: Hutz, C. S., Bandeira, D. R. & Trentini, C. M. *Psicometria*. Porto Alegre: 2015.
- Instituto brasileiro de avaliação psicológica (2003). *Diretrizes para o uso de testes: international test commission*. Recuperado em 05 de junho de 2019 de [https://www.intestcom.org/files/guideline\\_test\\_use\\_portuguese\\_brazil](https://www.intestcom.org/files/guideline_test_use_portuguese_brazil)
- International Test Commission. (2005). ITC guidelines on computer-based and internet delivered testing. *Downloaded electronically on*, 8, 2014.
- Jantscher, S., Willinger, U., Schmoeger, M., Mueller, C., & Auff, E. (2011). Validation of the Hayling sentence completion test-German version & Stroop-test. *European Psychiatry*, 26(S2), 420-420. doi: [10.1016/S0924-9338\(11\)72128-9](https://doi.org/10.1016/S0924-9338(11)72128-9)
- Jongstra, S., Wijsman, L. W., Cachucho, R., Hoevenaer-Blom, M. P., Mooijaart, S. P., & Richard, E. (2017). Cognitive testing in people at increased risk of dementia using a smartphone app: the Vitality proof-of-principle study. *JMIR Health and Health*, 5(5), e6939. doi:[10.2196/mhealth.6939](https://doi.org/10.2196/mhealth.6939)
- Karlsson, L., Gezelius, A., Nordenström, A., Hirvikoski, T., & Lajic, S. (2017). Cognitive impairment in adolescents and adults with congenital adrenal hyperplasia. *Clinical endocrinology*, 87(6), 651-659. doi: [10.1111/cen.13441](https://doi.org/10.1111/cen.13441)

- Kiernan, J. A. (2002). *Neuroanatomia Humana de Barr* (7<sup>a</sup> ed). Barueri: Manole.
- Kim, H., Lee, Y. N., Jo, E. M., & Lee, E. Y. (2017). Reliability and validity of culturally adapted executive function performance test for Koreans with stroke. *Journal of Stroke and Cerebrovascular Diseases*, 26(5), 1033-1040. doi: [10.1016/j.jstrokecerebrovasdis.2016.12.013](https://doi.org/10.1016/j.jstrokecerebrovasdis.2016.12.013)
- Klojčnik, M., Kavcic, V., & Bakracevic Vukman, K. (2017). Relationship of depression with executive functions and visuospatial memory in elderly. *The International Journal of Aging and Human Development*, 85(4), 490-503. doi: [10.1177/0091415017712186](https://doi.org/10.1177/0091415017712186)
- Kosmidis, M. H., Bozikas, V. P., Zafiri, M., & Karavatos, A. (2006). Shared cognitive processes underlying performance on the Wisconsin Card Sorting Test and the Stroop Test in patients with schizophrenia: A measurement artifact?. *Neuroscience letters*, 409(3), 234-238. doi: [10.1016/j.neulet.2006.09.049](https://doi.org/10.1016/j.neulet.2006.09.049)
- Kulaif, T. (2005). *O teste de cores e palavras de Stroop modificado para analfabetos*. (Dissertação de mestrado). Recuperado de <https://repositorio.usp.br/item/001490875>
- Lezak, M. D. (1995). *Neuropsychological Assessment* (3<sup>a</sup> ed.). England: Oxford University Press.
- Lins, M. R. C. & Borsa, J. C. (Orgs.). (2017). *Avaliação psicológica: aspectos teóricos e práticos*. Petrópolis: Vozes. (477 p.)
- Logan, G. D. (1978). Attention in character-classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General*, 107(1), 32. doi: [10.1037/0096-3445.107.1.32](https://doi.org/10.1037/0096-3445.107.1.32)
- Malloy-Diniz, L. F., Fuentes, D., & Cosenza, R. M. (2013). *Neuropsicologia do envelhecimento: uma abordagem multidimensional*. Artmed Editora.
- Malloy-Diniz, L. F., Paula, J. J., Loschiavo-Alvares, F. Q., Fuentes, D., & Leite, W. B. (2010). Exame das Funções Executivas. In: Malloy-Diniz, L. F., Fuentes, D., Mattos, P., & Abreu, N. *Avaliação Neuropsicológica*. Porto Alegre: Artmed.

- Malloy-Diniz, L. F., Sedo, M., Fuentes, D. & Leite, W. B. (2008). Neuropsicologia das Funções Executivas. In: Fuentes, D., Malloy-Diniz, L. F., Camargo, C. H. P., & Consenza, R. M. *Neuropsicologia: teoria e prática*. Porto Alegre: Artmed.
- Malloy-Diniz, L. F., Costa, D. S., Querino, E. H. G., de Paula, J. J. (2014). Propriedades psicométricas do Teste dos Cinco Dígitos para o contexto brasileiro: estudo preliminar com a população adulta. Relato de pesquisa apresentado no I Congresso da Sociedade Brasileira de Neuropsicologia Jovem.
- Mana, J., Vaneckova, M., Klempíř, J., Lišková, I., Brožová, H., Poláková, K., ...Maréchal, B. (2019). Methanol poisoning as an acute toxicological basal ganglia lesion model: evidence from brain volumetry and cognition. *Alcoholism: Clinical and Experimental Research*, 43(7), 1486-1497. doi: [10.1111/acer.14077](https://doi.org/10.1111/acer.14077)
- Mariano, L. I. (2017). Impulsividade e controle inibitório: avaliação cognitiva, comportamental e diagnóstica na demência frontotemporal. (Dissertação de mestrado). Recuperado de [https://repositorio.ufmg.br/bitstream/1843/BUOS-B3VN95/1/luciano\\_inacio\\_mariano\\_dissertacao\\_final\\_impulsividade\\_e\\_controle\\_inibitorio....pdf](https://repositorio.ufmg.br/bitstream/1843/BUOS-B3VN95/1/luciano_inacio_mariano_dissertacao_final_impulsividade_e_controle_inibitorio....pdf)
- Marín-Navarrete, R., Toledo-Fernández, A., Villalobos-Gallegos, L., Pérez-López, A., & Medina-Mora, M. E. (2018). Neuropsychiatric characterization of individuals with inhalant use disorder and polysubstance use according to latent profiles of executive functioning. *Drug and alcohol dependence*, 190, 104-111. doi: [10.1016/j.drugalcdep.2018.06.005](https://doi.org/10.1016/j.drugalcdep.2018.06.005)
- Masiliūnas, R., Vitkutė, D., Stankevičius, E., Matijošaitis, V., & Petrikonis, K. (2017). Response inhibition, set shifting, and complex executive function in patients with chronic lower back pain. *Medicina*, 53(1), 26-33. doi: [10.1016/j.medici.2016.12.001](https://doi.org/10.1016/j.medici.2016.12.001)
- McLeod, C. M. (1991). Half a century of research on the Stroop Test effect: an integrative review. *Psychological Bulletin*, 2 (109), 163-203. doi: [10.1037/0033-2909.109.2.163](https://doi.org/10.1037/0033-2909.109.2.163)
- MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(1), 126. doi: [10.1037/0278-7393.14.1.126](https://doi.org/10.1037/0278-7393.14.1.126)

- MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in cognitive sciences*, 4(10), 383-391. doi: [10.1016/S1364-6613\(00\)01530-8](https://doi.org/10.1016/S1364-6613(00)01530-8)
- Morrow, S. A. (2013). Normative data for the Stroop color word test for a North American population. *Canadian Journal of Neurological Sciences*, 40(6), 842-847.
- Mesquita, C., Coutinho, G. & Mattos, P. (2010). Perfil neuropsicológico de adultos com queixas de desatenção: diferenças entre portadores de TDAH e controles clínicos. *Revista de Psiquiatria Clínica*, 37 (5), 212-215. doi: [10.1590/S0101-60832010000500005](https://doi.org/10.1590/S0101-60832010000500005)
- Milioni, A. L. V., Chaim, T. M., Cavallet, M., de Oliveira, N. M., Annes, M., Dos Santos, B., ... & Cunha, P. J. (2017). High IQ may “mask” the diagnosis of ADHD by compensating for deficits in executive functions in treatment-naïve adults with ADHD. *Journal of attention disorders*, 21(6), 455-464. doi: 10.1177/1087054714554933
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, 41(1), 49-100. doi: [10.1006/cogp.1999.0734](https://doi.org/10.1006/cogp.1999.0734)
- Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current directions in psychological science*, 21(1), 8–14. doi: 10.1177/0963721411429458
- Moering, R. G., Schinka, J. A., Mortimer, J. A., & Graves, A. B. (2004). Normative data for elderly African Americans for the Stroop color and word test. *Archives of Clinical Neuropsychology*, 19(1), 61-71. doi: [10.1093/arclin/19.1.61](https://doi.org/10.1093/arclin/19.1.61)
- Montagnero, A. V., & Esteves, M. A. S. (2013). Atenção seletiva e memória implícita em sujeitos com diferentes IMCs. *Revista Brasileira de Terapias Cognitivas*, 9(1), 02-09. doi: [dx.doi.org/10.5935/1808-5687.20130002](https://doi.org/10.5935/1808-5687.20130002)
- Moreira, Helena S., Ana S. Costa, São L. Castro, César F. Lima, and Selene G. Vicente. "Assessing executive dysfunction in neurodegenerative disorders: a critical review of brief

- neuropsychological tools." *Frontiers in aging neuroscience* 9 (2017): 369. doi: [10.3389/fnagi.2017.00369](https://doi.org/10.3389/fnagi.2017.00369)
- Morrow, S. A. (2013). Normative data for the Stroop color word test for a North American population. *Canadian Journal of Neurological Sciences*, 40(6), 842-847. doi: [10.1017/S0317167100015997](https://doi.org/10.1017/S0317167100015997)
- Nigg J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological bulletin*, 126(2), 220–246. doi:10.1037/0033-2909.126.2.220
- Nowakowska-Domagala, K., Jabłkowska-Górecka, K., Mokros, Ł., Koprowicz, J., & Pietras, T. (2017). Differences in the verbal fluency, working memory and executive functions in alcoholics: Short-term vs. long-term abstainers. *Psychiatry research*, 249, 1–8. doi: [10.1016](https://doi.org/10.1016)
- Nunes, C. H. S. (2011). Importância da especificação dos contextos de aplicação e propósitos nos manuais de testes psicológicos. In: Conselho Federal de Psicologia. *Ano da Avaliação Psicológica – Textos geradores* - Brasília: Conselho Federal de Psicologia.
- Nunes, C. H. S., Primi, R. (2010). Aspectos técnicos e conceituais da ficha de avaliação dos testes psicológicos. In: *Avaliação Psicológica: Diretrizes na regulamentação da profissão*. Brasília: Conselho Federal de Psicologia.
- Nyongesa, M. K., Ssewanyana, D., Mutua, A. M., Chongwo, E., Scerif, G., Newton, C. R. J. C., & Abubakar, A. (2019). Assessing Executive Function in Adolescence: A Scoping Review of Existing Measures and Their Psychometric Robustness. *Frontiers in Psychology*, 10, 311. <https://doi.org/10.3389/fpsyg.2019.00311>
- Oliveira, T. D. O., Malloy-Diniz, L. F., Magalhães, S., Costa, D. S., Lacerda, S. R., Querino, E. H. G., ... & De Paula, J. J. (2014, May). Propriedades psicométricas do Teste dos Cinco Dígitos para o contexto brasileiro: estudo preliminar com a população adulta. In *I Congresso da Sociedade Brasileira de Neuropsicologia Jovem*. Disponível em: < [https://www.researchgate.net/publication/277873709\\_Propriedades\\_psicomtricas\\_do\\_Testes\\_dos\\_Cinco\\_Digitos\\_para\\_o\\_contexto\\_brasileiro\\_estudo\\_preliminar\\_com\\_a\\_populacao\\_adulta](https://www.researchgate.net/publication/277873709_Propriedades_psicomtricas_do_Testes_dos_Cinco_Digitos_para_o_contexto_brasileiro_estudo_preliminar_com_a_populacao_adulta).

- Pasquali, L. (2017) *Psicometria: teoria dos testes na psicologia e na educação*. Petrópolis: Vozes.
- Pawlowski, J., Trentini, C. M., & Bandeira, D. R. (2007). Discutindo procedimentos psicométricos a partir da análise de um instrumento de avaliação neuropsicológica breve. *PsicoUSF*, 12 (1), 211-219. doi: [10.1590/S1413-82712007000200009](https://doi.org/10.1590/S1413-82712007000200009)
- Peña-Casanova, J., Quiñones-Úbeda, S., Gramunt-Fombuena, N., Quintana, M., Aguilar, M., Molinuevo, J. L., ... & Blesa, R. (2009). Spanish Multicenter Normative Studies (NEURONORMA Project): norms for the Stroop color-word interference test and the Tower of London-Drexel. *Archives of Clinical Neuropsychology*, 24(4), 413-429. doi: [10.1093/arclin/acp043](https://doi.org/10.1093/arclin/acp043)
- Primi, R. (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica*, 2(1), 67-77. Recuperado em 26 de outubro de 2019, de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712003000100008&lng=pt&tlng=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712003000100008&lng=pt&tlng=pt).
- Primi, R., Muniz, M., & Nunes, C. H. S. S. (2009). Definições contemporâneas de validade de testes psicológicos. In Hutz, C. S. (Org.) *Avanços e polêmicas em avaliação psicológica* (pp. 243-265). São Paulo: Casa do Psicólogo.
- Rabelo, I. S., Pacanaro, S. V., Leme, I. F. A. S., Ambiel, R. A. M., & Alves, G. A. S. (2011). *Teste não verbal de inteligência geral – BETA III subtestes raciocínio matricial e cógiso – manual técnico*. São Paulo: Casa do Psicólogo.
- Ramos, A. A., & Hamdan, A. C. (2016). O crescimento da avaliação neuropsicológica no Brasil: uma revisão sistemática. *Psicologia: Ciência e Profissão*, 36(2), 471-485. doi: [10.1590/1982-3703001792013](https://doi.org/10.1590/1982-3703001792013)
- Reeves, T. D., & Marbach-Ad, G. (2016). Contemporary test validity in theory and practice: A primer for discipline-based education researchers. *CBE—Life Sciences Education*, 15(1), rm1. doi: [10.1187/cbe.15-08-0183](https://doi.org/10.1187/cbe.15-08-0183)

- Reppold, C. T., Gomes, C. M. A., Seabra, A. G., Muniz, M., Valentini, F., & Laros, J. A. (2015). (Artigo teórico) Contribuições da psicométrica para os estudos em neuropsicologia cognitiva. *Revista Psicologia: Teoria e Prática*, 17(2)
- Rezaei, F., Hosseini Ramaghani, N. A., & Fazio, R. L. (2017). The effect of a third party observer and trait anxiety on neuropsychological performance: the Attentional Control Theory (ACT) perspective. *The Clinical Neuropsychologist*, 31(3), 632-643. doi: [10.1080/13854046.2016.1266031](https://doi.org/10.1080/13854046.2016.1266031)
- Rossi, A., Daneluzzo, E., Mattei, P., Bustini, M., Casacchia, M., & Stratta, P. (1997). Wisconsin card sorting test and Stroop test performances in Schizophrenia: a shared construct. *Neuroscience letters*, 226(2), 87-90. doi: [10.1016/S0304-3940\(97\)00253-X](https://doi.org/10.1016/S0304-3940(97)00253-X)
- Sampieri, R. H., Callado, C. F., & Lucio, M. D. P. B. (2013). *Metodologia da pesquisa* (5ed.). Porto Alegre: Penso.
- Sedó, M. A., & DeCristoforo, L. (2001). All-language verbal tests free from linguistic barriers. *Revista Española de Neuropsicología*, 3(3), 68-82.
- Sedó, M. A., Paula, J. J., & Malloy-Diniz, L. (2015). *Five Digit Test (FDT) –Teste dos Cinco Dígitos*. São Paulo, Brazil: Hogrefe CETEPP.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, 84(2), 127. doi: [10.1037/0033-295X.84.2.127](https://doi.org/10.1037/0033-295X.84.2.127)
- Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia*, 40(6), 605-619.
- Shoghi Javan, S., & Ghonsooly, B. (2018). Learning a foreign language: A new path to enhancement of cognitive functions. *Journal of Psycholinguistic Research*, 47, 125-138. doi: [10.1007/s10936-017-9518-7](https://doi.org/10.1007/s10936-017-9518-7)

- Siqueira, L. D. S., Scherer, L. C., Reppold, C. T., & Fonseca, R. P. (2010). Hayling Test-adult version: applicability in the assessment of executive functions in children. *Psychology & Neuroscience*, 3, 189-194. doi: 10.3922/j.psns.2010.2.008
- Sisto, F. F., & Castro, N. R. (2010). *EASV: Escala de atenção seletiva visual*. São Paulo: Casa do Psicólogo.
- Sjoberg, E. A., Baguely, T., & Cole, G. (2017). Gender differences in the Stroop Colour-Word Test: a meta-analysis. In *Boise, Idaho, Human Behavior and Evolution Conference, HBES*.
- Sousa, V. D., Driessnack, M., & Mendes, I. A. C. (2007). Revisão dos desenhos de pesquisa relevantes para enfermagem: Parte 1: desenhos de pesquisa quantitativa. *Revista Latino-Americana de Enfermagem*, 15(3), 502-507. doi: [10.1590/S0104-11692007000300022](https://doi.org/10.1590/S0104-11692007000300022)
- Span, M. M., Ridderinkhof, K. R., & van der Molen, M. W. (2004). Age-related changes in the efficiency of cognitive processing across the life span. *Acta psychologica*, 117(2), 155-183. doi: [10.1016/j.actpsy.2004.05.005](https://doi.org/10.1016/j.actpsy.2004.05.005)
- Spreen, O.; Strauss, E. (1998). *A Compendium of Neuropsychological tests - administration, norms, and commentary* (2a ed.). New York: Oxford University Press.
- Sternberg, R. J. (2000). *Psicologia Cognitiva* (pp. 87-90). Porto Alegre: Artmed Editora.
- Strauss, E., Spreen, O., & Sherman, E.M.S.. (2006). *A Compendium of neuropsychological tests: Administration, norms, and commentary* (3ª ed.). Nova York: Oxford University Press.
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales: A practical guide to their development and use* (5 ed.). New York: Oxford University Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6), 643. doi: [10.1037/h0054651](https://doi.org/10.1037/h0054651)
- Sweet, J. J. (1999). *Forensic Neuropsychology: Fundamentals and practice*. Nova York: Psychology Press.
- Talarico, M. K., Lynall, R. C., Mauntel, T. C., Weinhold, P. S., Padua, D. A., & Mihalik, J. P. (2017). Static and dynamic single leg postural control performance during dual-task

- paradigms. *Journal of sports sciences*, 35(11), 1118-1124. doi: [10.1080/02640414.2016.1211307](https://doi.org/10.1080/02640414.2016.1211307)
- Toyoshima, K., Kako, Y., Toyomaki, A., Shimizu, Y., Tanaka, T., Nakagawa, S., ... & Kusumi, I. (2019). Associations between cognitive impairment and quality of life in euthymic bipolar patients. *Psychiatry research*, 271, 510-515. doi: [10.1016/j.psychres.2018.11.061](https://doi.org/10.1016/j.psychres.2018.11.061)
- Troyer, A. K., Leach, L., & Strauss, E. (2006). Aging and response inhibition: Normative data for the Victoria Stroop Test. *Aging, Neuropsychology, and Cognition*, 13(1), 20-35. doi: [10.1080/138255890968187](https://doi.org/10.1080/138255890968187)
- Tuokko, H., Griffith, L. E., Simard, M., & Taler, V. (2017). Cognitive measures in the Canadian longitudinal study on aging. *The Clinical Neuropsychologist*, 31(1), 233-250. doi: [10.1080/13854046.2016.1254279](https://doi.org/10.1080/13854046.2016.1254279)
- Uebel-von Sandersleben, H., Albrecht, B., Rothenberger, A., Fillmer-Heise, A., Roessner, V., Sergeant, J., ...Banaschewski, T. (2017). Revisiting the co-existence of Attention-Deficit/Hyperactivity Disorder and Chronic Tic Disorder in childhood-The case of colour discrimination, sustained attention and interference control. *PloS one*, 12(6), e0178866. doi: [10.1371/journal.pone.0178866](https://doi.org/10.1371/journal.pone.0178866)
- Uehara, E., Charchat-Fichman, H., & Landeira-Fernandez, J. (2013). Funções executivas: Um retrato integrativo dos principais modelos e teorias desse conceito. *Neuropsicologia Latinoamericana*, 5(3). doi: [10.5579/rnl.2013.145](https://doi.org/10.5579/rnl.2013.145)
- Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.
- Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2006). The Stroop color-word test: influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, 13(1), 62-79. doi: [10.1177/1073191105283427](https://doi.org/10.1177/1073191105283427)
- Van der Sluis, S., De Jong, P. F., & Van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, 35(5), 427-449. doi: [10.1016/j.intell.2006.09.001](https://doi.org/10.1016/j.intell.2006.09.001)

- Verdejo-García A, Pérez-García M. Profile of executive deficits in cocaine and heroin polysubstance users: common and differential effects on separate executive components. *Psychopharmacology (Berl)*. 2007 Mar;190(4):517-30. doi: 10.1007/s00213-006-0632-8. Epub 2006 Nov 29. PMID: 17136401.
- Westerhausen, R., Kompus, K., & Hugdahl, K. (2011). Impaired cognitive inhibition in schizophrenia: a meta-analysis of the Stroop interference effect. *Schizophrenia research*, 133(1-3), 172-181. doi: [doi.org/10.1016/j.schres.2011.08.025](https://doi.org/10.1016/j.schres.2011.08.025)
- Yilmaz, S., & Kafadar, H. (2020). Investigating the relationship between decision-making processes and cognitive processes, personality traits, and affect via the structural equation model in young adults. *Applied Neuropsychology: Adult*, 27(6), 558-569. doi: [10.1080/23279095.2019.1576690](https://doi.org/10.1080/23279095.2019.1576690)
- Yu, F., Vock, D. M., & Barclay, T. R. (2018). Executive function: Responses to aerobic exercise in Alzheimer's disease. *Geriatric nursing*, 39(2), 219-224. doi: 10.1016/j.gerinurse.2017.09.005
- Zalonis, I., Christidi, F., Bonakis, A., Kararizou, E., Triantafyllou, N. I., Paraskevas, G., ... & Vasilopoulos, D. (2009). The stroop effect in Greek healthy population: normative data for the Stroop Neuropsychological Screening Test. *Archives of Clinical Neuropsychology*, 24(1), 81-88. doi: [10.1093/arclin/acp011](https://doi.org/10.1093/arclin/acp011)
- Zimmermann, N., Cardoso, C. D. O., Trentini, C. M., Grassi-Oliveira, R., & Fonseca, R. P. (2015). Brazilian preliminary norms and investigation of age and education effects on the Modified Wisconsin Card Sorting Test, Stroop Color and Word test and Digit Span test in adults. *Dementia & Neuropsychologia*, 9, 120-127. doi:[10.1590/1980-57642015DN92000006](https://doi.org/10.1590/1980-57642015DN92000006)
- Zysset, S., Schroeter, M. L., Neumann, J., & von Cramon, D. Y. (2007). Stroop interference, hemodynamic response and aging: an event-related fMRI study. *Neurobiology of aging*, 28(6), 937-946. doi: [10.1016/j.neurobiolaging.2006.05.008](https://doi.org/10.1016/j.neurobiolaging.2006.05.008)

**APÊNDICES**

**APÊNDICE 1****CLINITRAN – CLÍNICA DE AVALIAÇÃO MÉDICA E PSICOLÓGICA****DECLARAÇÃO DE AUTORIZAÇÃO DE PESQUISA**

Declaro para os devidos fins e efeitos legais que tomei conhecimento da pesquisa "Evidências de validade do Testes Stroop de Cores e Palavras para adultos", sob responsabilidade de Vanessa Brandelero Abi-Abib, aluna de doutorado do Programa de Pós-Graduação em Psicologia do Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina, sob orientação do professor Dr. Carlos Henrique Sancineto da Silva Nunes e coorientação do professor Dr. Roberto Moraes Cruz. Como sócia-proprietária da instituição, autorizo a sua execução e declaro que acompanharei o seu desenvolvimento para garantir que será realizada dentro do que preconiza a Resolução CNS 466/12, de 12/09/2012 e complementares.



---

Luciana Ballardin  
Sócia-proprietária  
CPF: 027.405.439-64

Florianópolis, 18 de junho de 2020.

## APÊNDICE 2

### HABILITA TRAN - CLÍNICA DE AVALIAÇÃO MÉDICA E PSICOLÓGICA

#### DECLARAÇÃO DE AUTORIZAÇÃO DE PESQUISA

Declaro para os devidos fins e efeitos legais que tomei conhecimento da pesquisa "Evidências de validade do Testes Stroop de Cores e Palavras para adultos", sob responsabilidade de Vanessa Brandelero Abi-Abib, aluna de doutorado do Programa de Pós-Graduação em Psicologia do Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina, sob orientação do professor Dr. Carlos Henrique Sancineto da Silva Nunes e coorientação do professor Dr. Roberto Moraes Cruz. Como sócia-proprietária da instituição, autorizo a sua execução e declaro que acompanharei o seu desenvolvimento para garantir que será realizada dentro do que preconiza a Resolução CNS 466/12, de 12/09/2012 e complementares.



---

Luciana Ballardin  
Sócia-proprietária  
CPF: 027.405.439-64

Florianópolis, 18 de junho de 2020.

## APÊNDICE 3



## DECLARAÇÃO DE AUTORIZAÇÃO DE PESQUISA

Declaro para os devidos fins e efeitos legais que tomei conhecimento da pesquisa "Estudos de validade do Testes Stroop de Cores e Palavras para adultos", sob responsabilidade de Vanessa Brandeleiro Abi-Abib, aluna de doutorado do Programa de Pós-Graduação em Psicologia do Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina, sob orientação do professor Dr. Carlos Henrique Sancineto da Silva Nunes e coorientação do professor Dr. Roberto Moraes Cruz. Como coordenadora do curso de graduação em Psicologia, autorizo a sua execução e declaro que acompanharei o seu desenvolvimento para garantir que será realizada dentro do que preconiza a Resolução CNS 466/12, de 12/09/2012 e complementares.

*Andréa de Fátima Rua Estácio*  
 Andréa de Fátima Rua Estácio  
 Coordenação do curso de Psicologia

Andréa de Fátima Rua Estácio  
 Coordenadora do Curso  
 UniCesumar

☎ 79.265.617/0001-9  
 CENTRO DE ENSINO SUPERIOR  
 DE MARINGÁ  
 RUA ITAJUBÁ, 673 - PORTÃO  
 CEP 81070-190 - CURITIBA - PR

## APÊNDICE 4



UNIVERSIDADE FEDERAL DE SANTA CATARINA

Centro de Filosofia e Ciências Humanas

Programa de Pós-Graduação em Psicologia

### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Prezado(a) Colaborador(a),

Você está sendo convidado(a) a participar da coleta de dados Tese de Doutorado intitulada “*Evidências de validade do Teste Stroop de Cores e Palavras para adultos*” sob a responsabilidade de Vanessa Brandelero Abi-Abib, aluna de doutorado do Programa de Pós-Graduação em Psicologia do Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina, sob orientação do professor Dr. Carlos Henrique Sancineto da Silva Nunes e coorientação do professor Dr. Roberto Moraes Cruz, que tem por objetivo realizar a busca de evidências de validade do Teste Stroop de Cores e Palavras, em versão impressa e informatizada, para adultos.

O Teste Stroop de Cores e Palavras é um teste neuropsicológico desenvolvido em 1935 por John Ridley Stroop. Tradicionalmente, o Teste Stroop é utilizado na avaliação das funções executivas (FE), e também, de alguns de seus componentes, como atenção seletiva, inibição, flexibilidade, memória de trabalho e velocidade de processamento. Entretanto, até o momento, não há na literatura referências de evidências de validade para a população adulta brasileira. Por este motivo, o presente projeto tem por objetivo realizar a busca de evidências de validade do Teste Stroop de Cores e Palavras.

O presente projeto de pesquisa foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Santa Catarina e cumpre os termos da Resolução CNS 446/12 e suas complementares (item IV.5.a).

#### **DADOS DO PARECER DE APROVAÇÃO**

emitido Pelo Comitê de Ética em Pesquisa, CEPESH-UFSC

Número do parecer: 4.137.183

Data da relatoria: 06/06/2020

**1. PARTICIPAÇÃO NA PESQUISA:** Ao participar desta pesquisa você realizará seis testes/tarefas neuropsicológicas. Os testes a serem realizados serão: Teste Stroop de Cores e Palavras; Teste dos Cinco Dígitos - FDT; Wisconsin Teste de Classificação de Cartas - WCST; Teste de Avaliação da Atenção Seletiva - EASV; BETA-III – Subteste Códigos; e *Hayling Sentence Completion Test* (HSCT). A aplicação ocorrerá em um único encontro e será individual. A duração média prevista para a participação será de 50 minutos. A sua participação é voluntária e você tem a liberdade de não querer participar. Ainda, reforçamos que é permitido desistir desta coleta, em qualquer momento, mesmo após ter iniciado o questionário sem nenhum prejuízo.

**2. RISCOS E DESCONFORTOS:** O procedimento utilizado nesta pesquisa obedece aos critérios da Ética na Pesquisa com Seres Humanos conforme a Resolução 466/12 de 12/06/2012

do Conselho Nacional de Saúde, de modo que não oferecem riscos graves ao participante. Informamos que mesmo pouco provável, há a possibilidade em sentir desconfortável ou mesmo experienciar sentimentos de fadiga, ansiedade, cansaço ou aborrecimento ao responder aos testes/tarefas neuropsicológicas, devido à extensão de testes que serão realizados e de suas eventuais complexidades. Os pesquisadores garantirão esforços no intuito de reduzir tais riscos, adotando medidas que proporcionem conforto e assistência durante toda a aplicação. O (a) senhor (a) poderá alimentar-se, bem como beber líquidos, descansar, ir ao banheiro, amamentar, ou quaisquer outras necessidades durante a participação. Caso você apresente qualquer desconforto, será disponibilizada assistência imediata e/ou integral, com encaminhamento ao serviço de atenção psicológica da UFSC. Porém, vale a informação de que os danos previsíveis serão evitados com a realização cautelosa da pesquisa e o estabelecimento de uma relação de confiança com o pesquisador, respeitando-se as condições, tempo e características de cada indivíduo. Você poderá expressar suas dúvidas e apreensões, em qualquer momento, e será acompanhado durante todo o procedimento de participação, por um dos pesquisadores, que prestará toda a assistência e tomará as providências e cautelas necessárias para evitar e reduzir efeitos e condições adversas. Sendo assim, no caso de sentir qualquer desconforto decorrente da pesquisa, pedimos que informe ao pesquisador e assim, a tarefa será suspensa. Deste modo, informamos que o tipo de procedimento apresenta um risco mínimo, que será reduzido pelos pesquisadores, ficando sob responsabilidade destas o encaminhamento aos serviços de apoio sem custo algum para o participante.

**3. BENEFÍCIOS:** Os benefícios esperados com o estudo se destinam à comunidade que poderá se beneficiar de uma nova ferramenta neuropsicológica para avaliação de funções executivas com adequadas evidências de validade.

**4. CONFIDENCIALIDADE:** Todas as informações que o (a) Sr. (a) nos fornecer ou que sejam conseguidas pelo questionário serão utilizadas somente para esta pesquisa. Seus(Suas) respostas ficarão em sigilo e o seu nome não aparecerá em lugar nenhum do questionário nem quando os resultados forem apresentados. Entretanto, torna-se necessário alerta-lo sobre a quebra de sigilo. Se por qualquer situação ocorrer quebra de sigilo, esta somente dar-se-á de forma involuntária e não intencional. As informações obtidas serão empregadas com ética na elaboração do trabalho científico, que poderá ser publicado em meios acadêmicos e científicos, garantindo a manutenção do sigilo e da privacidade dos participantes.

**5. RESSARCIMENTO DAS DESPESAS:** Caso o (a) Sr. (a) aceite participar da pesquisa, a pesquisadora se compromete ao ressarcimento de possíveis gastos decorrentes de vossa participação. Sendo assim, custos referentes ao deslocamento e alimentação para a participação poderão ser ressarcidas pela pesquisadora mediante envio dos comprovantes via e-mail. Caso venha a sofrer qualquer prejuízo, material ou imaterial, decorrente de sua participação nesta pesquisa, você poderá solicitar indenização de acordo com a legislação vigente. De acordo com a legislação brasileira, sua participação é absolutamente voluntária e não remunerada.

**6. ESCLARECIMENTOS:** Se tiver alguma dúvida a respeito da pesquisa e/ou dos métodos utilizados na mesma, pode procurar a qualquer momento o pesquisador responsável.

Nome da pesquisadora: Vanessa Brandelero Abi-Abib

E-mail: [vanessa.brandelero@hotmail.com](mailto:vanessa.brandelero@hotmail.com)

Nome do pesquisador responsável (orientador): Carlos Henrique Sancineto da Silva Nunes.

E-mail: [carlos.sancineto@pesquisador.cnpq.br](mailto:carlos.sancineto@pesquisador.cnpq.br)

Endereço institucional dos pesquisadores: Laboratório de Pesquisa em Avaliação Psicológica. Sala 15A. Universidade Federal de Santa Catarina – Campus Universitário- Trindade. Centro de Filosofia e Ciências Humanas – Departamento de Psicologia. CEP 88040-900 – Florianópolis, SC – Brasil Telefone: (48) 3721-9904.

**7. INFORMAÇÕES SOBRE O COMITÊ DE ÉTICA EM PESQUISA COM SERES HUMANOS-CEPSH-UFSC:** O CEPESH é um órgão colegiado interdisciplinar, deliberativo, consultivo e educativo, vinculado à Universidade Federal de Santa Catarina, mas independente na tomada de decisões, criado para defender os interesses dos participantes da pesquisa em sua integridade e dignidade e para contribuir no desenvolvimento da pesquisa dentro de padrões éticos.

**8. ENDEREÇO CEPESH-UFSC:** Prédio Reitoria II, R: Desembargador Vitor Lima, nº 222, sala 401, Trindade, Florianópolis/SC, CEP 88.040-400, Contato: (48) 3721-6094, [cep.propesq@contato.ufsc.br](mailto:cep.propesq@contato.ufsc.br).

**9. DISPOSIÇÕES FINAIS:** A pesquisadora responsável, que também assina este documento, declara que cumprirá os termos da Resolução CNS 466/12 e suas complementares (item IV.5.a) e compromete-se a conduzir a pesquisa de acordo com o que preconiza a Resolução 466/12 de 12/06/2012, que trata dos preceitos éticos e da proteção aos participantes da pesquisa. Esclarecemos, ainda, que poderá ser realizada uma devolutiva dos resultados da pesquisa de forma geral a você. Para isso, basta informar seu interesse ao pesquisador que estiver conduzindo a coleta dos dados. Após a leitura do presente termo e do aceite em participar do estudo, solicito sua assinatura, no presente termo, em duas vias, sendo que uma delas permanecerá em seu poder.

**10. CONCORDÂNCIA NA PARTICIPAÇÃO:** Se o (a) Sr. (a) estiver de acordo em participar deverá preencher e assinar o Termo de Consentimento Pós-esclarecido que se segue, em duas vias, sendo que uma via ficará com você.

Pelo presente instrumento que atende às exigências legais, eu, \_\_\_\_\_, portador (a) do RG \_\_\_\_\_, declaro por meio do presente documento o meu consentimento em participar desta pesquisa. Fui devidamente informado e esclarecido sobre os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes da participação da mesma. Foi-me garantida a confidencialidade das informações pessoais e que posso retirar meu consentimento a qualquer momento, sem que isso leve a qualquer penalidade.

Cidade: \_\_\_\_\_ Data: \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

\_\_\_\_\_  
Assinatura do participante

Assinatura do Pesquisador Responsável  
Prof. Dr. Carlos Henrique Sancineto da Silva Nunes  
[carlos.nunes@ufsc.br](mailto:carlos.nunes@ufsc.br)

Assinatura da Pesquisadora  
Doutoranda: Vanessa Brandelero Abi-Abib  
[vanessa.brandelero@hotmail.com](mailto:vanessa.brandelero@hotmail.com)

**ANEXOS**

## ANEXO 1

UNIVERSIDADE FEDERAL DE  
SANTA CATARINA - UFSC



**PARECER CONSUBSTANCIADO DO CEP**

**DADOS DO PROJETO DE PESQUISA**

**Título da Pesquisa:** Evidências de Validade do Teste Stroop de Cores e Palavras para Adultos

**Pesquisador:** Carlos Henrique Sancineto da Silva Nunes

**Área Temática:**

**Versão:** 1

**CAAE:** 33926120.4.0000.0121

**Instituição Proponente:** Universidade Federal de Santa Catarina

**Patrocinador Principal:** Financiamento Próprio

**DADOS DO PARECER**

**Número do Parecer:** 4.137.183

**Apresentação do Projeto:**

Tese de doutorado de Vanessa Brandelero Abi-Abib, aluna do Programa de Pós-Graduação em Psicologia do Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina, sob orientação do professor Dr. Carlos Henrique Sancineto da Silva Nunes e coorientação do professor Dr. Roberto Moraes Cruz.

Estudo quantitativo, com previsão de 500 participantes, que serão submetidos à seis testes/tarefas neuropsicológicas. Os testes a serem realizados serão: Teste Stroop de Cores e Palavras; Teste dos Cinco Dígitos - FDT; Wisconsin Teste de Classificação de Cartas - WCST; Teste de Avaliação da Atenção Seletiva - EASV; BETA-III – Subteste Códigos; e Hayling Sentence Completion Test (HSCT). A aplicação ocorrerá em um único encontro e será individual. A duração média prevista para a participação será de 50 minutos.

**Objetivo da Pesquisa:**

**Objetivo Primário:**

Realizar a busca de evidências de validade do Teste Stroop de Cores e Palavras, em versão impressa e informatizada, para adultos.

**Objetivo Secundário:**

a) Levantar, por meio de revisão integrativa de literatura, quais construtos psicológicos são

**Endereço:** Universidade Federal de Santa Catarina, Prédio Reitoria II, R: Desembargador Vitor Lima, nº 222, sala 401  
**Bairro:** Trindade **CEP:** 88.040-400  
**UF:** SC **Município:** FLORIANOPOLIS  
**Telefone:** (48)3721-6094 **E-mail:** cep.propesq@contato.ufsc.br

Continuação do Parecer: 4.137.183

avaliados pelo Teste Stroop de Cores e Palavras;b) Desenvolver estudos de evidências de validade convergente do instrumento;c) Desenvolver estudos de evidências de validade discriminante do instrumento;d) Desenvolver estudos de evidências de validade por construtos relacionados do instrumento;

**Avaliação dos Riscos e Benefícios:**

A análise de riscos e benefícios está adequada no projeto, no formulário da Plataforma Brasil e no TCLE.

**Comentários e Considerações sobre a Pesquisa:**

Sem comentários.

**Considerações sobre os Termos de apresentação obrigatória:**

As autorizações das Clínicas (CLINITRAN e HABILITA TRAN) são da mesma pessoa, Luciana Ballardin, que consta como sócia-proprietária de ambas as Clínicas de Avaliação Médica e Psicológica.

A folha de rosto vem assinada pelo orientador, constando como pesquisador responsável e pela autoridade institucional competente.

O cronograma informa que a coleta de dados acontecerá a partir de 01/11/2020.

O orçamento informa despesas de R\$ 2.500,00, com financiamento próprio.

O TCLE é esclarecedor a respeito de objetivos, procedimentos, riscos e direitos dos participantes, e cumpre as exigências da res. 466/12.

**Recomendações:**

Observar as restrições sanitárias impostas pela Pandemia de COVID-19 na execução da pesquisa.

**Conclusões ou Pendências e Lista de Inadequações:**

Sem pendências.

**Considerações Finais a critério do CEP:**

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações	PB_INFORMAÇÕES_BÁSICAS_DO_P	22/06/2020		Aceito

Endereço: Universidade Federal de Santa Catarina, Prédio Reitoria II, R: Desembargador Vitor Lima, nº 222, sala 401  
 Bairro: Trindade CEP: 88.040-400  
 UF: SC Município: FLORIANOPOLIS  
 Telefone: (48)3721-6094 E-mail: cep.propesq@contato.ufsc.br

UNIVERSIDADE FEDERAL DE  
SANTA CATARINA - UFSC



Continuação do Parecer: 4.137.183

Básicas do Projeto	ETO_1579822.pdf	09:17:53		Aceito
Declaração de Instituição e Infraestrutura	Declaracao_Graduacao_Marcela.pdf	22/06/2020 09:17:06	Vanessa Brandelero	Aceito
Folha de Rosto	FolhaDeRosto_assinada.pdf	22/06/2020 09:09:23	Vanessa Brandelero	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_COMITE_DE_ETICA.pdf	19/06/2020 15:12:51	Vanessa Brandelero	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_PPGP_UFSC.pdf	19/06/2020 15:03:51	Vanessa Brandelero	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_Laboratorio_LPAP.pdf	19/06/2020 15:03:14	Vanessa Brandelero	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_Habilitatran.pdf	19/06/2020 15:02:20	Vanessa Brandelero	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_Clinitran.pdf	19/06/2020 15:01:49	Vanessa Brandelero	Aceito
Cronograma	Cronograma.pdf	19/06/2020 14:56:24	Vanessa Brandelero	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	19/06/2020 14:49:30	Vanessa Brandelero	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

FLORIANOPOLIS, 06 de Julho de 2020

Assinado por:  
Nelson Canzian da Silva  
(Coordenador(a))

Endereço: Universidade Federal de Santa Catarina, Prédio Reitoria II, R: Desembargador Vitor Lima, nº 222, sala 401  
Bairro: Trindade CEP: 88.040-400  
UF: SC Município: FLORIANOPOLIS  
Telefone: (48)3721-6094 E-mail: cep.propesq@contato.ufsc.br