



EVIDÊNCIAS DE VALIDADE E PRECISÃO DO TESTE FLUÊNCIA DA BATERIA TSP

EVIDENCIA DE VALIDEZ Y EXACTITUD DE LA PRUEBA DE FLUENCIA DE BATERÍAS TSP

EVIDENCE OF VALIDITY AND ACCURACY OF THE FLUENCE TEST THE TSP BATTERY

Kellyane Madureira Figueiredo¹
Elizabeth do Nascimento²

RESUMO: A Bateria TSP é um instrumento utilizado para avaliação de pessoas no contexto organizacional e do trabalho, composto por nove testes objetivos para avaliar diferentes aptidões cognitivas. Entre eles, o teste Fluência busca avaliar a capacidade de vocabulário simples do candidato, representando um aspecto de aptidão verbal importante para diferentes funções no trabalho. O objetivo deste estudo foi verificar as propriedades psicométricas do teste Fluência, para isto, participaram 1352 pessoas, entre 15 e 76 anos, 65,6% do sexo feminino. Os resultados de precisão foram satisfatórios para o Alfa de Cronbach ($\alpha = 0,79$), ômega de McDonald ($\omega = 0,81$) e o teste e reteste com correlação de Pearson de magnitude forte ($r = 0,77$; $p < 0,001$). Para os estudos de validade realizou-se análise fatorial confirmatória (AFC) na qual indicou adequados índices de ajuste ao modelo unidimensional testado ($\chi^2 = 2,97$; $gl = 2$; $p < 0,227$; TLI = 0,998; CFI = 0,999; RMSEA = 0,02; SRMR = 0,006). Um estudo de validade convergente foi feito com a prova de Raciocínio Verbal (RV) da BPR-5. O resultado de correlação para os 83 sujeitos submetidos ao teste de Fluência e à prova RV foi moderado e estatisticamente significativo ($r = 0,46$; $p < 0,001$). Por fim, estudos de evidências de validade de critério com as variáveis socio-demográficas encontraram resultados condizentes com a literatura, ou seja, não houve uma diferença importante no desempenho em função do sexo; houve uma diferença significativa entre os grupos etários extremos, houve uma diferença significativa em função da escolaridade, na qual os menores desempenhos médios foram observados para os participantes com menor grau de escolaridade. O teste de Fluência indicou adequadas evidências de validade e de precisão, podendo ser utilizado para avaliação da fluência de vocabulário em amostras semelhantes à do presente estudo.

PALAVRAS-CHAVE: Psicometria; Teste Psicológico; Inteligência; Fluência Verbal; Vocabulário.

RESUMEN: La Bateria TSP es un instrumento utilizado para evaluar a las personas en el contexto organizacional y laboral, compuesto por nueve pruebas objetivas para evaluar diferentes habilidades cognitivas. Entre ellos, la prueba de Fluidez busca evaluar la capacidad de vocabulario simple del candidato, representando un aspecto importante de la aptitud verbal para diferentes funciones laborales. El objetivo de este estudio fue verificar las propiedades psicométricas del test de Fluidez, para ello participaron 1352 personas, entre 15 y 76 años, 65,6% mujeres. Los resultados de precisión fueron satisfactorios para el alfa de Cronbach ($\alpha = 0,79$), el omega de McDonald ($\omega = 0,81$) y el test y retest con fuerte correlación de Pearson ($r = 0,77$; $p < 0,001$). Para los estudios de validez se realizó análisis factorial confirmatorio (AFC), que indicó índices de ajuste adecuados al modelo unidimensional probado ($\chi^2 = 2,97$; $df = 2$; $p < 0,227$; TLI = 0,998; CFI = 0,999; RMSEA = 0,02; SRMR = 0,006). Se realizó un estudio de validez convergente mediante la prueba de Razonamiento Verbal (RV) BPR-5. El resultado de correlación para los 83 sujetos sometidos al test de Fluidez y al test de RV fue moderado y estadísticamente significativo ($r = 0,46$; $p < 0,001$). Finalmente, los estudios de evidencia de validez de criterio con variables sociodemográficas encontraron resultados consistentes con la literatura, es decir, no hubo diferencia importante en el desempeño según el género; hubo una diferencia significativa entre los grupos de edades extremas, hubo una diferencia significativa según la educación, en la que los rendimientos promedio más bajos se observaron para los participantes con un nivel de educación más bajo. La prueba de fluidez indicó evidencia adecuada de validez y precisión y puede usarse para evaluar la fluidez del vocabulario en muestras similares a la del presente estudio.

PALABRAS CLAVE: Psicometría; Examen psicológico; Inteligencia; Fluidez verbal; Vocabulario.

¹ kemf_int@yahoo.com.br

² bethdonascimento@gmail.com

ABSTRACT: The TSP Battery is an instrument used to assess people in the organizational and work context, consisting of nine objective tests to assess different cognitive skills. Among them, the Fluency test seeks to assess the candidate's simple vocabulary capacity, representing an important aspect of verbal aptitude for different job functions. The objective of this study was to verify the psychometric properties of the Fluency test, for this, 1352 people participated, between 15 and 76 years old, 65,6% female. Accuracy results were satisfactory for Cronbach's alpha ($\alpha = 0,79$), McDonald's omega ($\omega = 0,81$) and the test and retest with strong Pearson correlation ($r = 0,77$; $p < 0,001$). For validity studies, confirmatory factor analysis (CFA) was carried out, which indicated adequate fit indices to the tested unidimensional model ($\chi^2 = 2,97$; $df = 2$; $p < 0,227$; TLI = 0,998; CFI = 0,999; RMSEA = 0,02; SRMR = 0,006). A convergent validity study was carried out using the BPR-5 Verbal Reasoning (RV) test. The correlation result for the 83 subjects submitted to the Fluency test and the RV test was moderate and statistically significant ($r = 0,46$; $p < 0,001$). Finally, studies of evidence of criterion validity with sociodemographic variables found results consistent with the literature, that is, there was no important difference in performance depending on gender; there was a significant difference between the extreme age groups, there was a significant difference depending on education, in which the lowest average performances were observed for participants with a lower level of education. The Fluency test indicated adequate evidence of validity and precision and can be used to assess vocabulary fluency in samples similar to that of the present study.

KEYWORDS: Psychometrics; Psychological test; Intelligence; Verbal Fluency; Vocabulary.

1 INTRODUÇÃO

A inteligência e, de maneira mais ampliada, as habilidades cognitivas, ao longo do tempo tornaram-se um dos construtos mais estudados e pesquisados no contexto educacional, na psicologia industrial, do trabalho e organizacional (Chiodi, 2012; Larson, 2019). Como resultado, foi demonstrado o poder da inteligência para prever o desempenho humano em várias situações, entre elas, no local de trabalho, na capacidade de produção e desenvoltura laboral. Esses achados trazem implicações para a importância de se ter instrumentos para mensuração da inteligência que podem ser utilizados nos processos de recrutamento e seleção a fim de gerar melhores resultados individuais e coletivos à organização (Hunter; Schmidt, 1996; Hunter; Schmidt; Judiesch, 1990).

Uma empresa que busca selecionar funcionários deve considerar o tipo de trabalho a ser desempenhado pois isso poderá influenciar sobre qual seria a melhor característica preditora, uma habilidade mental geral ou uma habilidade específica, para avaliar o desempenho laboral. Por exemplo, uma vaga para engenheiro de aviação pode considerar que a habilidade espacial seja mais importante neste contexto de trabalho e ter maior validade ecológica do que uma habilidade mental geral (Kell; Lang, 2017). Com o reconhecimento das habilidades mentais específicas e a sua relação com o desempenho dos indivíduos, como exemplo, a aptidão numérica para a atividade de um contador, capacidade percepto-motora e de atenção para o operador de máquinas, memória visual para seguranças, rapidez perceptual para o investigador de incêndio, fica demonstrada a relevância dos instrumentos para mensuração das habilidades amplas, mas também das específicas (Nascimento; Flores-Mendoza, 2006; Robbins; Judge; Sobral, 2010).

Desde as primeiras investigações sobre a inteligência diferentes esforços são empreendidos para compreender como a inteligência funciona, quais são as aptidões e capacidades relacionadas às diferenças individuais de desempenho cognitivo e como a inteligência se constitui e evolui (Andrés-Pueyo, 2006). Essa busca está intimamente relacionada à criação de instrumentos que pretendem a avaliação psicológica da inteligência. Associado aos diferentes modelos de inteligência que foram construídos ao longo do tempo estão os testes psicológicos que fazem parte das investigações e maturação sobre o conceito de inteligência. A começar com os trabalhos de Spearman, em 1904, sobre o modelo bifatorial da inteligência (fator *g*) e o modelo das aptidões mentais primárias do psicólogo norte-americano Thurstone, em 1938, muitos dos testes criados neste período tinham como base as teorias psicométricas (fatoriais) da inteligência desenvolvidas por estes autores (Laros; Jesus; Karino, 2013).

O Teste de Seleção Profissional foi criado em 1947, nos Estados Unidos, pelo norte americano Joseph E. King (Nascimento; Flores-Mendoza, 2006). Os testes da Bateria TSP foram baseados no modelo de inteligência proposto por Thurstone e um dos princípios básicos postulados por ele é que existem diversas capacidades e aptidões intelectuais especializadas e independentes. Dessa maneira, quando os indivíduos precisam resolver algum problema ou responder a uma tarefa, como um teste psicológico, é exigido apenas a aptidão intelectual relacionada à natureza da situação exposta, ou seja, as capacidades intelectuais podem atuar de maneira autônoma. O modelo de Thurstone identificou sete aptidões mentais primárias que correspondem a diferentes tipos de raciocínio: compreensão verbal (V), fluência verbal (W), raciocínio (indutivo) (R), capacidade numérica (N), velocidade perceptual (P), capacidade espacial (S) e memória (M), dando ênfase em uma estrutura básica de fatores de inteligência independentes e aplicados a tarefas com especificidades diferentes (Andrés-Pueyo, 2006).

A Bateria TSP é composta por nove testes objetivos para avaliar diferentes aptidões cognitivas. Foi criada para uso na indústria e comércio e, é historicamente utilizada no contexto organizacional e do trabalho. Um dos testes que compõe a Bateria é o teste Fluência que tem como objetivo avaliar a capacidade de vocabulário simples do examinando, representando um aspecto da aptidão verbal, pois indica a facilidade em utilizar palavras (Nascimento; Flores-Mendoza, 2006). Este tipo de tarefa verifica a amplitude de palavras adquiridas pelo examinando relacionadas às habilidades básicas de escrever palavras isoladas com características fonéticas específicas (como demonstrando pelo item exemplo). O teste Fluência demanda a rapidez e a capacidade de evocação de informações estabelecidas na memória de longo prazo, relacionada ao vocabulário adquirido pela pessoa. As pessoas diferem quanto ao armazenamento de informações na memória de longo prazo e, este aspecto da habilidade tem se

tornado cada vez mais reconhecido devido ao seu papel na compreensão da leitura, da escrita e da comunicação (Schneider; McGrew, 2012).

O teste Fluência pode ser associado ao fator de primeira ordem Fluência de palavras (FW) no modelo Cattell-Horn-Carroll (CHC) de inteligência, definida como a capacidade de rapidamente pensar e produzir palavras isoladas que compartilham uma característica fonêmica, estrutural ou ortográfica específica (McGrew, 2005; Schneider; McGrew, 2012). Esta aptidão pertencia originalmente à habilidade ampla, Armazenamento e recuperação de longo prazo (Glr), no entanto a partir da evolução da teoria CHC e de suas melhorias graduais ao longo do tempo, à habilidade cognitiva Glr foi separada formalmente em 2018, formando duas categorias amplas: Eficiência de aprendizagem (Gl) e Fluência de recuperação (Gr). Esta última, relacionada à velocidade de recuperação seletiva e estratégica de informação verbais e não verbais armazenadas na memória de longo prazo. Em Gr ficaram presentes as habilidades estreitas relacionadas à fluência em relação à produção de ideias e figuras, recordação fluente de palavras, como exemplo: Fluência ideacional (FI), Fluência de palavras (FW), Fluência figural (FF), entre outras (Schneider; McGrew, 2018). Dessa maneira, supõe-se que o teste Fluência da Bateria TSP, esteja relacionado ao fator amplo do modelo CHC de inteligência, Fluência de recuperação (Gr), na habilidade estreita FW.

O teste Fluência pode também aparecer associado a outra aptidão ampla do modelo CHC, Inteligência cristalizada (Gc), como no estudo da Bateria de provas de Raciocínio (BPR-5), em que o teste de Raciocínio verbal (RV) apresentou cargas fatoriais interpretadas para Gc (Almeida *et al.*, 2010; Primi; Correia; Almeida, 2018). No entanto, a hipótese de trabalho é que o teste Fluência da bateria TSP será associado a Gr em função dessa habilidade estar teoricamente mais relacionada ao estímulo proposto pela tarefa, pois o mesmo requer do respondente velocidade para produzir palavras a partir de um recurso fonológico solicitado, condizente com FW, não sendo necessária a ortografia correta da palavra (o que poderia ser esperado em Gc).

A Bateria TSP tornou-se um instrumento desfavorável para uso no Brasil em dezembro de 2022. Com a autorização da Edites – Empresa Distribuidora de Testes Ltda., responsável pela comercialização no Brasil, a presente pesquisa pretende contribuir com estudos recentes que possam atestar a qualidade desse instrumento e, com isso, continuar a ser utilizado pela comunidade de psicólogos em sua atuação profissional.

2 MÉTODO

2.1 Participantes

Participaram desse estudo 1352 sujeitos, a amostragem foi por conveniência, residentes nos estados de Minas Gerais (N=809), Distrito Federal (N=299), Espírito Santo (N=150) e Ceará (N=94). Foram 887 mulheres (65,6%), a idade mínima foi a estabelecida de 15 anos e a máxima 76 anos (M=26,4; DP=11,45). Sobre a escolaridade a maioria foram sujeitos com Ensino superior completo ou incompleto (N=703), seguido pelo Ensino Médio completo ou incompleto (N=392) e Ensino Fundamental completo ou incompleto (N=257). O presente estudo foi submetido à avaliação do Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais (COEP/UFMG) e, recebeu parecer favorável sob o número 3.358.718.

2.2 Instrumentos para a coleta de dados

Cada participante foi submetido à aplicação do teste de Fluência, instrumento foco da pesquisa. O teste Fluência é subdividido em quatro seções de tarefas com estímulos verbais (letra, sufixo, prefixo, por exemplo), que envolvem a evocação e a escrita de um maior número de palavras em um tempo limite de dois minutos para cada seção. No item de treinamento do teste apresentado abaixo é solicitado ao examinando que escreva todas as palavras terminadas em TAR, ele poderá escrever quantas palavras recordar e em qualquer ordem (Nascimento; Flores-Mendoza, 2006).

Escreva todas as palavras que conhece terminadas em: TAR

ajuntar

O escore é formado sendo atribuído 0 para as respostas erradas e 1 para as palavras que obedeceram ao comando solicitado. Uma observação importante é que erros gramaticais não são considerados, então, desde que o examinando tenha atendido ao comando do item, será atribuído um ponto mesmo à escrita incorreta da palavra. No entanto, não poderá usar variações de uma mesma palavra já escrita, como aumentativo, diminutivo, palavras ou tempos verbais, nem mesmo nomes próprios. Neste teste não há limite máximo de pontuação (Nascimento; Flores-Mendoza, 2006).

Responderam também a um questionário sociodemográfica para caracterização da

amostra com itens referentes à idade, sexo, escolaridade, por exemplo. Para realizar estudos de evidências de validade baseada na relação com outras variáveis foi utilizado outro instrumento de inteligência, a Bateria de Provas de Raciocínio (BPR-5) em uma subamostra. A BPR-5 é composta por cinco subtestes de raciocínio. Fundamenta-se em aspectos teóricos consistentes e mais recentes da inteligência (modelo CHC de inteligência). De acordo com a necessidade de investigação no estudo foi utilizado apenas a prova de Raciocínio Verbal (RV). A última atualização das propriedades psicométricas foi no estudo aprovado pelo CFP em 11 de abril de 2003 e a atualização das normas foi aprovada na Plenária de 19 de maio de 2018 (Primi; Almeida, 2018; Primi; Correia; Almeida, 2018). O subteste RV associa-se com a inteligência cristalizada ($G_c - K0$), relacionado ao conhecimento verbal e conceitos aprendidos previamente (Campos; Zaia; Primi, 2019; Primi; Nakano; Wechsler, 2012). A precisão evidenciada pelos coeficientes de consistência interna avaliados pela fórmula de Lâmbda 6 de Guttman foi apresentado por idade e variaram entre 0,72 a 0,95 (RV) entre as formas A e B estudadas, com idade mínima de 12 e máxima de 23 anos. Os coeficientes apresentados demonstram uma alta consistência dos escores (Primi; Almeida, 2018). Primi, Correia e Almeida (2018) apresentam um capítulo sobre a BPR-5 e agrupam diferentes resultados para a precisão, evidências de validade baseadas no conteúdo, na estrutura interna e na relação com outras variáveis com resultados favoráveis quanto à qualidade do instrumento. Desde sua adaptação para o Brasil, a BPR-5 é um dos instrumentos mais utilizados na avaliação das capacidades cognitivas e diversos estudos foram realizados apontando as suas qualidades psicométricas (Júnior; Nascimento; Roazzi, 2019; Primi; Almeida, 2000, 2018).

2.3 Procedimentos para a coleta de dados

A coleta de dados foi realizada em duas etapas, sendo a primeira voltada exclusivamente para pessoas com ensino fundamental, com início em 2015 e se estendeu em razão da dificuldade de acesso a adultos com baixa escolaridade. A segunda etapa incluiu pessoas com diferentes níveis de escolaridade e ocorreu entre dezembro de 2018 e maio de 2023. Durante a pandemia COVID-19, a coleta foi interrompida em razão das barreiras sanitárias.

Para o desenvolvimento da coleta de dados a pesquisadora responsável formou uma equipe de pesquisa composta por estudantes de Psicologia e psicólogas formadas que foram treinadas para a aplicação, correção dos testes e lançamento dados em um banco. Essa equipe foi responsável por entrar em contato com a coordenação ou diretor acadêmico das escolas com ensino fundamental, médio, escolas técnicas e cursos de graduação, a fim de obter auto-

rização formal para a realização da pesquisa. Uma vez autorizado o contato com os estudantes das instituições, aqueles que concordaram assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) quando maiores de 18 anos e, responderam aos instrumentos. No caso de participantes menores de idade, estes foram submetidos à pesquisa somente após a autorização dos pais por meio da assinatura TCLE e do Termo de Assentimento Livre e Esclarecido (TALE) para participação de adolescentes em pesquisa. Ressalta-se que toda a participação na pesquisa foi voluntária e podia ser interrompida a qualquer momento pelo participante.

A aplicação dos instrumentos foi coletiva, realizada em uma ou duas sessões de acordo com a necessidade e disponibilidade oferecida pela instituição e participantes. Os procedimentos de aplicação seguiram as recomendações internacionais para testes educacionais e psicológicos (AERA; APA; NCME, 2014) e as orientações de Muñiz e Fonseca-Pedrero (2019), entende-se que seguir um procedimento padrão contribui para que todos os participantes tenham iguais condições para demonstrar suas competências.

2.4 Procedimentos para a análise dos dados

Os dados coletados foram organizados em planilha e, posteriormente, foram realizadas análises utilizando o software JAMOV – versão 2.3 (2022). Em função dos dados terem sido digitados por diferentes pessoas da equipe, para verificação da qualidade do lançamento utilizou-se da estratégia de extrair aleatoriamente uma determinada quantidade de participantes e examinou a correspondência entre as pontuações anotadas no teste físico e o dado que foi digitado (Muñiz; Fonseca-Pedrero, 2019). O resultado demonstrou a qualidade do banco de dados, em função da correspondência exata das pontuações investigadas dos participantes.

Para avaliar a forma como os dados se distribuem, foi realizado o teste de regressão/correlação Shapiro-Wilk (W). Espera-se um valor de $p > 0,05$ e o valor de W próximo de 1,0 para afirmar o pressuposto paramétrico. Foi feito também um teste gráfico pelo diagrama quantil-quantil (Q-Q), nele é feita uma comparação das proporções dos dados ordenados da amostra e os quantis para uma distribuição normal padrão. Observa-se uma distribuição normal pelo diagrama Q-Q quando os pontos representantes da amostra situam-se ao longo de uma linha diagonal reta. Dados não paramétricos são detectados quando for observada uma distribuição em curva S ou em formas de banana. Ressalta-se a importância de se conciliar estratégias de julgamento diferentes para a verificação da suposição da distribuição dos dados, isso se faz relevante em função do teste estatístico mostrar-se sensível ao erro tipo II na situação de amostras grandes, como na presente pesquisa. A escolha do teste estatístico Shapiro-

Wilk foi por demonstrar, entre os testes formais, um alto grau de discriminação e por ser bem apoiado na literatura. Os diagramas quantil-quantil, em comparação com outros métodos de inspeção (histogramas e gráficos) demonstram maior acurácia e são mais indicados para amostras grandes, contrapondo os testes estatísticas de normalidade mais suscetíveis ao erro tipo II (Bishara; Li; Conley, 2021; Henderson, 2006; Miot, 2017; Vetter, 2017).

Para a investigação de evidências de validade baseadas na estrutura interna do instrumento, inicialmente, foram conduzidas provas estatísticas para verificar a adequação dos dados à fatoração pelo coeficiente de *Kaiser-Meyer-Olkin* (KMO) e pelo teste de Esfericidade de *Bartlett*. O valor da medida de adequação amostral KMO varia entre zero e um. Maiores são as chances de reduzir os dados e encontrar as dimensões subjacentes quanto mais próximo de um for o KMO, indicando uma amostragem adequada. Foi adotado o valor do índice KMO de 0,60 ou mais, desde que também o teste de *Bartlett* seja significativo ($p < 0,05$) indicando que a hipótese nula foi rejeitada e fatorabilidade da matriz de dados foi aceita (Tabachnick; Fidell, 2019).

Análises fatoriais exploratórias (AFE) foram utilizadas em função de se ter poucas pesquisas sobre a estrutura do construto e da medida de interesse. Posteriormente, foram conduzidas análises fatoriais confirmatórias (AFC) para comparação dos resultados da estrutura mais parcimoniosa. Os índices de ajuste para os modelos estruturais foram testados por meio da verificação dos índices absolutos, relativos e parcimoniosos para avaliação da adequação dos modelos (Brown, 2006). As referências utilizadas para comparar e verificar a qualidade dos modelos foram os índices: 1) índices absolutos: qui-quadrado / graus de liberdade, que possibilita a percepção da discrepância entre o modelo hipotético e o modelo revelado na amostra do estudo. O valor adequado de referência é $p < 0,05$; a estatística raiz quadrada média do erro de aproximação (RMSEA - *Root Mean Square Error of Approximation*) e estatística raiz quadrática média Residual Estandartizada (SRMR - *Standardized Root Mean Square Residual*) apontam a qualidade de ajuste ao estimar os parâmetros do modelo hipotético à matriz de covariância (Pilati; Laros, 2007). Valores menores que 0,05 indicam um bom ajuste, entre 0,06 e 0,08 ajuste razoável dos dados (Marsh; Hau; Wen, 2004); 2) índices relativos: índice comparativo de ajuste (CFI - *Comparative Fit Index*) e índice de Tucker-Lewis (TLI - *Tucker-Lewis Index*) são índices comparativos entre o modelo hipotético e os dados observados e, tem aspectos que compensam a complexidade do modelo. Valores acima de 0,95 indicam um excelente ajuste para TLI e CFI, sendo o índice de 0,90 aceitável (Hu; Bentler, 1998); 3) índices de parcimônia: critérios de informação de Akaike (AIC - *Akaike Information Criterion*) e critério de informação Bayesiano (BIC - *Bayesian Information Criterion*) sendo consi-

derado o melhor modelo àquele que apresenta um menor valor (Morin; Arens; Marsh, 2016).

Foi utilizado o método de Estimativa Máxima Verossimilhança (MLE – *Maximum Likelihood Estimator*) em função dos resultados das análises descritivas revelarem a normalidade dos dados. O método de rotação para AFE foi oblíquo (*oblimin*) definido de acordo com o modelo hipotético a ser estudado em função de sua especificidade e por permitir que os fatores se correlacionem entre si (Sass; Schmitt, 2010). Para as análises foi adotado o $p = 0,05$ como referência de significância (Muthén; Muthén, 2012) e ponto de corte de carga fatorial igual ou superior a 0,30 (Hair *et al.*, 2009). A magnitude das correlações (covariâncias) foi avaliada segundo os critérios de Dancey e Reidy (2019).

Uma abordagem complementar pôde testar o instrumento fazendo uso da validade convergente entre o teste de Fluência e o de Raciocínio Verbal (RV) da BPR-5, espera-se resultados superiores a 0,40 para a evidência de validade convergente. Foram feitas análises de diferenças por meio do *Test-t* de *Student* para amostras independentes, ANOVA na comparação em mais de três variáveis e índice *d* (tamanho do efeito), para verificação do grau de sobreposição dos grupos para os critérios sexo, escolaridade e faixa etária (validade de critério). As estatísticas paramétricas indicadas foram usadas em função de uma distribuição normal dos dados encontrada.

A verificação das evidências de precisão foi realizada pelos métodos de consistência interna coeficiente Alfa de Cronbach e estimativa ômega (ω) de McDonald por ser uma escolha viável quando a condição de *tau-equivalência* não é alcançada (Hauck-Filho; Valentini, 2020; Hayes; Coutts, 2020; McNeish, 2018; Taber, 2018). Resultados superiores a 0,70 serão considerados adequados para alfa e ômega (Gadermann; Guhn; Zumbo, 2012). Outra evidência será pelo estudo da estabilidade temporal (teste-reteste), verificado pelo coeficiente de correlação de *Pearson*, para comparação entre o desempenho dos sujeitos nas duas aplicações. A magnitude das correlações foi avaliada segundo os critérios de Dancey e Reidy (2019).

3 RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os dados descritivos para o teste Fluência. Nota-se uma diferença do tamanho amostral entre as seções 1 a 4 do teste e o resultado Fluência Total. Na digitação de dados com pessoas com ensino fundamental alguns protocolos foram digitados somente o escore bruto do sujeito, ou seja, os resultados parciais não foram lançados. Por este motivo há uma quantidade maior de sujeitos para o resultado total quando comparada às seções.

Para avaliação da normalidade dos dados foram considerados dois parâmetros, o W de Shapiro-Wilk e diagrama quantil-quantil. Pelos resultados demonstrados na Tabela 1, pode-se pressupor uma distribuição normal dos dados para o teste Fluência ao verificar que o teste de Shapiro-Wilk ($W = 0,998$; $p = 0,173$) com resultado não significativo ($p > 0,05$) demonstra que esta amostra não é diferente de uma distribuição normal (Miot, 2017; Vetter, 2017).

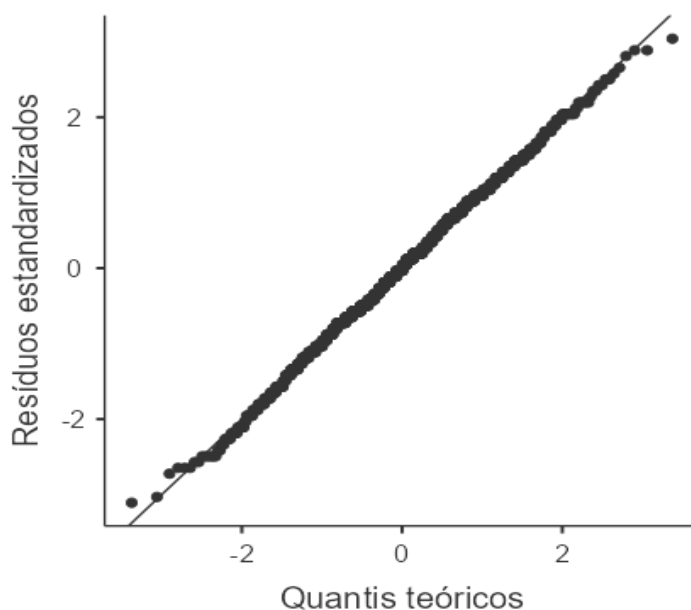
Tabela 1 - Estatísticas descritivas do Teste Fluência

	Flu1	Flu2	Flu3	Flu4	Teste Fluência
N	1251	1251	1251	1251	1352
Média	9,10	7,06	6,70	18,6	40,4
Mediana	9	7	7	19	40,0
Moda	8,00	7,00	5,00	23,0	43,0
Desvio-padrão	3,69	3,10	3,47	5,30	13,0
Amplitude	22	17	19	30	80
Mínimo	0	0	0	0	0
Máximo	22	17	19	30	80
W de Shapiro-Wilk	0,988	0,979	0,981	0,980	0,998
p Shapiro-Wilk	< 0,001	< 0,001	< 0,001	< 0,001	0,173

Nota. Flu1: Fluência seção 1 do teste (o mesmo é válido para Flu2, 3 e 4).

Conforme observado no diagrama Q-Q (Figura 1), a distribuição dos quantis da amostra de dados estão distribuídos de maneira linear sobre a reta esperada para a distribuição normal. Sendo assim, o diagrama Q-Q confirma o resultado do este estatístico W (Bishara; Li; Conley, 2021; Henderson, 2006).

Figura 1 - Diagrama quantil-quantil do teste Fluência



3.1 Evidências de Confiabilidade

Evidências de confiabilidade foram obtidas pelos coeficientes de consistência interna Alfa de Cronbach e ômega de McDonald, os resultados foram α 0,79 e ω 0,81, respectivamente. Os resultados apontam adequada consistência interna do teste Fluência (Gadernann, Guhn; Zumbo, 2012), condizente com as atuais diretrizes normativas (Conselho Federal de Psicologia, 2022). A evidência de estabilidade temporal (Tabela 2) demonstrou coeficiente de correlação de *Pearson* de magnitude forte ($r = 0,77$) e estatisticamente significativo, demonstrando também uma satisfatória estabilidade temporal.

Tabela 2 - Estatísticas do Estudo de Estabilidade Temporal do teste Fluência (N=78).

Teste	1ª aplicação		2ª aplicação		CORRELAÇÃO	
	M	DP	M	DP	r_{12}	Valor p
Fluência	48,06	12,97	53,85	11,59	0,77	<0,001

Nota. M: Média, DP: Desvio-Padrão, r_{12} : correlação bivariada, valor significativo se $p < 0,05$.

3.2 Evidências de Validade Baseadas na Estrutura Interna

Foi realizada a análise fatorial exploratória (AFE), o Método de extração 'Máxima verossimilhança' foi usado em combinação com uma rotação 'oblimin'. Os resultados apontaram que a matriz de dados é passível de redução fatorial [$KMO = 0,799$, Teste de Esfericidade de Barlett $\chi^2(6) = 1566$, $p < 0,001$] (Tabachnick; Fidell, 2019). A AFE indicou que o teste de Fluência apresenta um único fator, com *eigenvalue* (autovalor) de 2,06, e 51,6% de variância total explicada, demonstrando boa qualidade sobre o quanto o fator explica a covariância

compartilhada entre os itens. Conforme a Tabela 3, os itens são fortemente explicados pelo fator com cargas fatoriais maiores que 0,638. Ao realizar análises de correlação entre as seções do teste de Fluência, foram identificadas correlações positivas e moderadas, entre $r = 0,44$ e $0,57$, para todos os itens. Identifica-se, assim, uma relação importante entre os itens constituintes do teste.

Tabela 3 - Cargas Fatoriais dos itens do teste Fluência

Itens	Fator	Comunalidade (h^2)
Flu1	0,770	0,407
Flu2	0,638	0,593
Flu3	0,744	0,447
Flu4	0,715	0,488

A Tabela 4, apresenta os indicadores da qualidade de ajuste do modelo testado pela AFE e observa-se um bom resultado para a estatística RMSEA ($0,019 / IC90\% = 0,0 - 0,06$) e o índice TLI apontou uma excelente qualidade de ajuste dos dados ($>0,95$) (Hair *et al.*, 2009; Morin; Arens; Marsh, 2016; Pilati; Laros, 2007).

Tabela 4 - Medidas de Ajustamento do Modelo pela AFE

RMSEA	IC 90% RMSEA		TLI	BIC	Teste do Modelo		
	Limite Inferior	Limite Superior			χ^2	gl	p
0,019	0	0,063	0,998	-11,3	2,96	2	0,228

Nota. Método de extração 'Máxima verosimilhança' foi usado em combinação com uma rotação 'oblimin'

Realizou-se a AFC para verificar se o modelo é unidimensional conforme previamente estabelecido no instrumento original e apontado pela AFE. O teste qui-quadrado de ajuste apresentou $\chi^2 = 2,97$, $gl = 2$; $p < 0,227$, indicando que o modelo previsto teoricamente é observado nos dados, ou seja, a unidimensionalidade do teste Fluência. Conforme Tabela 5, apresenta os outros índices de ajuste da AFC que revelaram um bom ajuste ao modelo previamente estabelecido, já que produziu índices de ajustamento com os valores recomendados pela literatura CFI/ TLI $> 0,95$; RMSEA/ SRMR $< 0,05$ (Hair *et al.*, 2009; Morin; Arens; Marsh, 2016; Pilati; Laros, 2007).

Tabela 5 - Medidas de Ajustamento do Modelo pela AFC

CFI	TLI	SRMR	RMSEA	IC 90% RMSEA		AIC	BIC
				Limite			

				Inferior	Superior		
0,999	0,998	0,006	0,020	0,00	0,063	26039	26100

3.3 Evidências de validade baseadas nas relações com variáveis externas

3.3.1 Evidência de validade convergente

Foram 83 sujeitos submetidos ao teste Fluência e a prova de Raciocínio Verbal da BPR-5. Participaram 57 mulheres (68,7%), a idade variou entre 18 e 43 anos ($M=21,47$; $DP = 4,79$), com ensino superior incompleto e completo. Os resultados analisados apontaram evidências de validade convergente para o teste Fluência com correlação moderada e estatisticamente significativa ($r = 0,46$; $p < 0,001$). Ao correlacionar os resultados do teste objeto de estudo a um outro instrumento de qualidade estabelecida que avalia o mesmo construto, tem-se uma evidência sobre a capacidade de o instrumento avaliar o construto pretendido (Kalkbrenner, 2021).

3.3.2 Evidências de validade de critério (variáveis sociodemográficas)

O estudo sobre a associação entre o desempenho no teste e as variáveis sociodemográficas (sexo, faixa etária e escolaridade) se baseou nos achados sobre habilidades cognitivas. Consoante com a literatura, a suposição foi que: 1) as médias de desempenhos de homens e de mulheres seriam semelhantes; 2) as médias de desempenho difeririam entre os grupos etários; 3) as médias de desempenho entre os grupos difeririam segundo a escolaridade. As estatísticas usadas permitiram a comparação entre médias (Teste *t* de *Student* e ANOVA), com nível de significância ($p < 0,05$) e tamanho do efeito moderado ou forte ($d > 0,50$, $p^2 > 0,02$).

Com relação à variável sexo, a Tabela 6, apresenta que as diferenças entre as médias de desempenhos de homens e mulheres foram estatisticamente significativas no teste Fluência ($p = 0,002$), no entanto, o tamanho do efeito foi inferior ao ponto de corte adotado ($d = 0,17$). Deste modo, a grande sobreposição da distribuição dos desempenhos das mulheres e dos homens no teste apontou que a diferença estatisticamente significativas na variável sexo têm pouco impacto prático sobre os resultados nos testes.

Tabela 6 - Desempenho dos participantes no teste Fluência de acordo com o sexo

Teste	Sexo	N	Média	Desvio-padrão	Teste <i>t</i>	Valor <i>p</i>	<i>d</i> de Cohen
Fluência	Feminino	887	41,22	12,39	3,05	0,002	0,17
	Masculino	465	38,95	14,01			

A variável idade foi agrupada em cinco faixas etárias (15 a 16 anos; 17 a 18 anos; 19 a 25 anos; 26 a 39 anos e ≥ 40 anos). Os resultados das análises, Tabela 7, indicaram a presença de diferenças estatisticamente significativas entre os grupos etários ($p < 0,001$). Considerando o tamanho do efeito, as diferenças entre as médias foram moderadas ($\eta^2 = 0,05$). De modo geral, as comparações *post hoc* indicaram haver diferenças significativas entre os desempenhos médios dos grupos etários extremos, menor e maior faixa etária estudadas.

Tabela 7 - Desempenho dos participantes no teste Fluência segundo a faixa etária.

Teste	Faixa etária	N	Média	DP	F	Valor <i>p</i>	η^2
Fluência	15-16	88	31,79	11,41	19,59	<0,001	0,05
	17-18	273	39,38	11,42			
	19-25	524	43,36	12,23			
	26-39	250	40,50	12,62			
	≥ 40 anos	216	38,07	15,46			

Nota. DP: Desvio padrão; F = resultado das diferenças entre grupos ANOVA; Valor *p* = significativas caso $p < 0,05$; η^2 = tamanho do efeito.

A terceira variável investigada foi escolaridade (Tabela 8), considerando o nível de ensino (fundamental, médio e superior). As diferenças nas médias de desempenho dos grupos foram estatisticamente significativas ($p < 0,001$) e com tamanho de efeito forte ($\eta^2 = 0,25$). As comparações *post hoc* apontaram que as médias dos desempenhos dos participantes com ensino fundamental diferiram daqueles com ensino médio e superior e, as comparações entre os grupos do ensino médio e superior também foram estatisticamente significativas. Observou-se que as médias foram menores para as pessoas com ensino fundamental e as maiores foram obtidas por participantes com ensino superior completo ou incompleto.

Tabela 8 - Desempenho dos participantes no teste Fluência segundo o nível de escolaridade.

Teste	Escolaridade	N	Média	DP	F	Valor <i>p</i>	η^2
Fluência	Fundamental	257	28,56	11,24	230,27	<0,001	0,25

Médio	392	38,53	10,79
Superior	703	45,84	11,49

Nota. DP: Desvio padrão; F = resultado das diferenças entre grupos ANOVA; Valor p = significativas caso $p < 0,05$; n^2 = tamanho do efeito.

4 CONSIDERAÇÕES FINAIS

Neste estudo, buscou-se atualizar as evidências sobre as propriedades psicométricas do teste Fluência da Bateria TSP. Destaca-se um avanço importante em comparação com os estudos que sustentaram o manual anterior, entre eles: amostra obtida em quatro estados e três diferentes regiões brasileiras e ampliação do nível de escolaridade para contemplar sujeitos com ensino fundamental incompleto ou completo.

As análises de confiabilidade pelos índices de consistência interna foram acima do que é preconizado na literatura e o resultado da estabilidade temporal também foi considerado satisfatório. As evidências de validade baseadas na estrutura interna (AFE e AFC) apresentaram empiricamente bons índices para a relação dos itens do instrumento e a representação teórica do construto, apontando a estrutura unidimensional do teste Fluência. As evidências de validade baseadas nas relações com variáveis externas demonstrou desempenhos semelhantes dos respondentes na prova de Raciocínio Verbal da BPR-5 e o teste Fluência. A análise dos resultados dos testes comparada às variáveis sociodemográficas foi capaz de discriminar grupos diferentes em função da idade e escolaridade, conforme previsto na literatura especializada. Tais aspectos indicam que o instrumento mostra-se adequado teórica e metricamente, podendo ser utilizado para avaliação de fluência de vocabulário para sujeitos a partir de 15 anos, com ensino fundamental incompleto.

Embora as contribuições observadas, reconhece-se que a presente pesquisa possui limitações. A pandemia do Covid-19 no qual a pesquisa foi atravessada acarretou na diferença e extensão de tempo entre as coletas, recomenda-se também contar com amostras mais diversificadas com o intuito de assegurar maior heterogeneidade, contribuindo para melhor representatividade para os diferentes grupos estudados considerando também a diversificação do seu contexto social e geográfico da população testada. Outra relevante sugestão é de um investimento em estudos de equivalência para versão informatizada e/ou *online* do instrumento, tanto para condições de aplicação, quanto para a correção e, se a equivalência não for verificada, devem ser feitas novas pesquisas de validação para a sua autorização neste formato.

REFERÊNCIAS

ALMEIDA, Leandro S. *et al.* Bateria de provas de raciocínio (BPR-5): estudo exploratório em alunos universitários. **Avaliação Psicológica**, v. 9, n. 2, p. 155-162, ago. 2010. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712010000200002&lng=pt&tlng=pt. Acesso em: 10 dez. 2023.

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION [AERA]; AMERICAN PSYCHOLOGICAL ASSOCIATION [APA]; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION [NCME]. **Standards for educational and psychological testing**. Washington, DC: American Educational Research Association, 2014.

ANDRÉS-PUEYO, Antônio. Modelos psicométricos da inteligência. In: Flores-Mendoza, Carmen *et al.* (Orgs.). **Introdução à psicologia das diferenças individuais**. Porto Alegre: Artmed, 2006. p. 73-100.

BISHARA, Anthony J.; LI, Jiexiang; CONLEY, Christian. Informal versus formal judgment of statistical models: The case of normality assumptions. **Psychonomic Bulletin & Review**, v. 28, p. 1164-1182, 2021. Disponível em: <https://doi.org/10.3758/s13423-021-01879-z>. Acesso em: 01 jan. 2024.

BROWN, Timothy A. **Confirmatory factor analysis for applied research**. New York: The Guilford Press, 2006.

CAMPOS, Carolina Rosa; ZAIA, Priscila; PRIMI, Ricardo. Avaliação psicológica da inteligência. In: BAPTISTA, Makilim Nunes *et al.* (Orgs.). **Compêndio de Avaliação Psicológica**. Petrópolis: Vozes, 2019. p. 349-363.

CHIODI, Marcelo Gulini. **Estudo de adaptação e validação da bateria de habilidades cognitivas Woodcock-Johnson-III versão ampliada**. Tese (Doutorado em Psicologia) - Pontifícia Universidade Católica de Campinas, Campinas, 2012.

CONSELHO FEDERAL DE PSICOLOGIA. **Resolução CFP nº 31/2022**. Estabelece diretrizes para a realização de Avaliação Psicológica no exercício profissional da psicóloga e do psicólogo, regulamenta o Sistema de Avaliação de Testes Psicológicos – SATEPSI e revoga a Resolução CFP nº 09/2018. Brasília, 2022. Disponível em: <https://atosoficiais.com.br/cfp/resolucao-do-exercicio-profissional-n-31-2022-estabelece-diretrizes-para-a-realizacao-de-avaliacao-psicologica-no-exercicio-profissional-da-psicologa-e-do-psicologo-regulamenta-o-sistema-de-avaliacao-de-testes-psicologicos-satepsi-e-revoga-a-resolucao-cfp-no-09-2018?origin=instituicao>. Acesso em: 01 dez. 2023.

DANCEY, Christine P.; REIDY, John. **Estatística sem matemática para psicologia**. Tradução de Lori Viali. 7. ed. Porto Alegre: Penso, 2019.

GADERMANN, Anne M.; GUHN, Martin; ZUMBO, Bruno D. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. **Practical Assessment, Research & Evaluation**, v. 17, n. 3, p.1-13, 2012. Disponível em: <https://openpublishing.library.umass.edu/pare/article/id/1351/>. Acesso em: 01 dez. 2023.

HAIR, Joseph F. *et al.* **Análise multivariada de dados**. Tradução de Adonai Schlup Sant'Anna. 6. ed. Porto Alegre: Bookman, 2009.

HAUCK-FILHO, Nelson; VALENTINI, Felipe. Coeficientes de fidedignidade e violações dos pressupostos essencialmente tau-equivalentes. **Avaliação psicológica**, Campinas, v. 19, n. 3, p. a-b, set. 2020. Disponível em: <http://dx.doi.org/10.15689/ap.2020.1903.ed>. Acesso em: 14 dez. 2023.

HAYES, Andrew F.; COUTTS, Jacob J. Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But. **Communication Methods and Measures**, v. 14, n.1, p. 1-24, 2020. Disponível em: <https://doi.org/10.1080/19312458.2020.1718629>. Acesso em: 14 dez. 2023.

HENDERSON, Ralph A. Testing experimental data for univariate normality. **Clinica Chimica Acta**, v. 366, n. 1-2, p. 112-129, abr. 2006. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0009898105006790?via%3Dihub>. Acesso em: 12 jan. 2024.

HU, Li-tze; BENTLER, Peter M. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. **Psychological Methods**, v. 3, n. 4, p. 424-453, 1998. Disponível em: <https://sci-hub.se/10.1037/1082-989X.3.4.424>. Acesso em: 8 fev. 2024.

HUNTER, John E.; SCHMIDT, Frank L. Intelligence and job performance: Economic and social implications. **Psychology, Public Policy, and Law**, v. 2, n. 3-4, p. 447-472, 1996. Disponível em: <https://doi.org/10.1037/1076-8971.2.3-4.447>. Acesso em: 8 fev. 2024.

HUNTER, John E.; SCHMIDT, Frank L.; JUDIESCH, Michael K. Individual differences in output variability as a function of job complexity. **Journal of Applied Psychology**, v. 75, n. 1, p. 28-42, 1990. Disponível em: gwern.net/doc/iq/ses/1990-hunter.pdf. Acesso em: 8 fev. 2024.

JÚNIOR, Renê Marcelino da Silva; NASCIMENTO, Alexsandro Medeiros; ROAZZI, Antonio. Bateria de Provas de Raciocínio (BPR-5): Avaliação das qualidades psicométricas em adolescentes do Nordeste. **Amazônica - Revista de Psicopedagogia, Psicologia Escolar e Educação**, v. 23, n.1, p. 264-288, jan-jun. 2019. Disponível em: <https://periodicos.ufam.edu.br/index.php/amazonica/article/view/5174>. Acesso em: 8 dez. 2023.

KALKBRENNER, Michael T. A Practical Guide to Instrument Development and Score Validation in the Social Sciences: The MEASURE Approach. **Practical Assessment, Research, and Evaluation**, v. 26, n. 1, p. 1-18, 2021. Disponível em: <https://doi.org/10.7275/svg4-e671>. Acesso em: 20 dez. 2023.

KELL, Harrison J.; LANG, Jonas W. B. Specific Abilities in the Workplace: More Important Than g? **Journal of Intelligence**, v. 5, n. 2, p.13, 2017. Disponível em: <https://www.mdpi.com/2079-3200/5/2/13>. Acesso em: 20 fev. 2024.

LAROS, Jacob Arie; JESUS, Girlene Ribeiro de; KARINO, Camila Akemi. Validação brasileira do teste não-verbal de inteligência SON-R 2½-7[a]. **Avaliação psicológica**, Itatiba, v. 12, n. 2, p. 233-242, ago. 2013. Disponível em:

http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712013000200014. Acesso em: 10 fev. 2024.

LARSON, Elliott. C. A Meta-Analysis of Information Processing Measures of Intelligence, Performance, and Group Score Differences. **CUNY Academic Works**, 2019. Disponível em: https://academicworks.cuny.edu/gc_etds/3040/. Acesso em: 10 fev. 2024.

MARSH, Herbert W.; HAU, Kit-Tai; WEN, Zhonglin. In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. **Structural Equation Modeling: A Multidisciplinary Journal**, v. 11, n. 3, p. 320-341, 2004. Disponível em: https://doi.org/10.1207/s15328007sem1103_2. Acesso em: 02 jan. 2024.

McGREW, Kevin S. The Cattell-Horn-Carroll theory of cognitive abilities: past, presente, and future. In: FLANAGAN, Dawn P.; HARRISON, Patti L. (Orgs.), **Contemporary intellectual assessment: Theories, tests and issues**. 2. ed. New York: Guilford Press, 2005. p. 136-181.

McNEISH, Daniel. Thanks coefficient alpha, we'll take it from here. **Psychological Methods**, v. 23, n. 3, p. 412-433, 2018. Disponível em: <https://doi.org/10.1037/met0000144>. Acesso em: 12 dez. 2023.

MIOT, Hélio Amante. Avaliação da normalidade dos dados em estudos clínicos e experimentais. **Jornal Vascular Brasileiro**, v. 16, n.2, p. 88-91, abr-jun. 2017. Disponível em: <https://www.scielo.br/j/vb/a/FPW5hwZ6DTH4gvj5mJYpt6B/>. Acesso em: 10 jan. 2024.

MORIN, Alexandre J. S.; ARENS, A. Katrin; MARSH, Herbert W. A Bifactor Exploratory Structural Equation Modeling Framework for the Identification of Distinct Sources of Construct-Relevant Psychometric Multidimensionality. **Structural Equation Modeling: A Multidisciplinary**, v. 23, n. 1, p. 116-139, 2016. Disponível em: <https://doi.org/10.1080/10705511.2014.961800>. Acesso em: 20 dez. 2023.

MUÑIZ, José; FONSECA-PEDRERO, Edudardo. Diez pasos para la construcción de un test. **Psicothema**, v. 31, n. 1, p. 7-16, 2019. Disponível em: [doi: 10.7334/psicothema2018.291](https://doi.org/10.7334/psicothema2018.291). Acesso em: 20 dez. 2023.

MUTHÉN, Linda K.; MUTHÉN, Bengt O. **Mplus user's guide**. 7. ed. Los Angeles: Muthén & Muthén, 2012. Disponível em: https://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf. Acesso em: 20 dez. 2023.

NASCIMENTO, Elizabeth do; FLORES-MENDOZA, Carmen E. **Manual da Bateria TSP** (revisão e reavaliação). São Paulo: Edites, 2006.

PILATI, Rornaldo; LAROS, Jacob Arie. Modelos de Equações Estruturais em Psicologia: conceitos e aplicações. **Psicologia: Teoria e Pesquisa**, v. 23, n.2, p. 205-216, jun. 2007. Disponível em: <https://www.scielo.br/j/ptp/a/7dtRLCKQVY5fhtpCFBYPXHB/>. Acesso em: 15 jan. 2024.

PRIMI, Ricardo; ALMEIDA, Leandro S. Estudo de validação da bateria de provas de raciocínio (BPR-5). **Psicologia: Teoria e Pesquisa**, v. 16, n. 2, p. 165-173, ago. 2000. Disponível em:

<https://www.scielo.br/j/ptp/a/6rvvQsNBT9ZmgqQz7SyJ5TL/?lang=pt>. Acesso em: 10 dez. 2023.

PRIMI, Ricardo; ALMEIDA, Leandro S. **BPR-5 – Bateria de Provas de Raciocínio: Estudos de Revisão das Normas**. Pearson Clinical Brasil, 2018.

PRIMI, Ricardo; CORREIA, Tatiane Alves; ALMEIDA, Leandro S. Bateria de Provas de Raciocínio (BPR-5). *In*: HUTZ, Claudio Simon; BANDEIRA, Denise Ruschel; TRENTINI, Clarissa Marcelli (Orgs.). **Avaliação psicológica da inteligência e da personalidade**. Porto Alegre: Artmed, 2018. p. 109-122.

PRIMI, Ricardo; NAKANO, Tatiana de Cássia; WECHSLER, Solange Muglia. Cross-battery factor analysis of the Battery of Reasoning Abilities (BPR-5) and Woodcock-Johnson Tests of Cognitive Ability (WJ-III): Bateria de Habilidades de Raciocínio (BPR-5) e Bateria de Habilidades Cognitivas Woodcock-Johnson III. **Temas em Psicologia**, Ribeirão Preto, v. 20, n.1, p. 121-132, jun. 2012. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1413-389X2012000100010&lng=pt&nrm=iso. Acesso em: 10 dez. 2023.

ROBBINS, P. Stephen; JUDGE, Timothy A.; SOBRAL, Filipe. **Comportamento Organizacional: teoria e prática no contexto brasileiro**. Tradução Rita de Cássia Gomes. 14. ed. São Paulo: Pearson Prentice Hall, 2010.

SASS, Daniel. A.; SCHMITT, Thomas A. A Comparative Investigation of Rotation Criteria Within Exploratory Factor Analysis. **Multivariate Behavioral Research**, v. 45, n. 1, p. 73-103, feb. 2010. Disponível em: <https://doi.org/10.1080/00273170903504810>. Acesso em: 20 dez. 2023.

SCHNEIDER, W. Joel; MCGREW, Kevin S. The Cattell-Horn-Carroll model of intelligence. *In*: FLANAGAN, Dawn P.; HARRISON, Patti L. **Contemporary intellectual assessment: Theories, tests, and issues**. 3. ed. New York: The Guilford Press, 2012. p. 99–144.

SCHNEIDER, W. Joel; MCGREW, Kevin S. The Cattell-Horn-Carroll Theory of Cognitive Abilities. *In*: FLANAGAN, Dawn P.; McDONOUGH, Erin M. **Contemporary intellectual assessment: Theories, tests, and issues**. 4. ed. New York: The Guilford Press, 2018. p. 73-163.

TABACHNICK, Barbara G.; FIDELL, Linda S. **Using multivariate statistics**. 7. ed. Boston: Pearson, 2019. Disponível em: <https://www.pearsonhighered.com/assets/preface/0/1/3/4/0134790545.pdf>. Acesso em: 20 jan. 2024.

TABER, Keith S. The use of Cronbach's alpha when developing and reporting research instruments in science education. **Research in Science Education**, v. 48, n. 6, p. 1273-1296, 2018. Disponível em: <https://doi.org/10.1007/s11165-016-9602-2>. Acesso em: 20 dez. 2023.

THE JAMOVİ PROJECT (2022). Jamovi. (Version 2.3) [Computer Software]. Disponível em: <https://www.jamovi.org>. Acesso em: 20 dez. 2023

VETTER, Thomas R. Fundamentals of Research Data and Variables: The Devil Is in the Details. **Anesthesia & analgesia**, v. 125, n. 4, p. 1375-1380, oct. 2017. Disponível em: https://journals.lww.com/anesthesia-analgesia/Fulltext/2017/10000/Fundamentals_of_Research_Data_and_Variables__The.45.aspx. Acesso em: 20 jan. 2024.