

## Aspectos psicométricos de instrumentos neuropsicológicos: revisão conceitual, proposta de interpretação de percentis e classificações

Luis Anunciação<sup>1</sup>

Anna Carolina Portugal

J. Landeira-Fernandez

### RESUMO

A avaliação neuropsicológica é um procedimento fundamental especialmente a pacientes que possam apresentar suspeita de transtornos neurológicos, psicológicos ou psiquiátricos. Uma de suas finalidades é mapear aspectos do perfil psicológico e comportamental do examinando e, com isso, auxiliar o diagnóstico em saúde mental. Frequentemente, esse processo integra os resultados obtidos por instrumentos de medida e aspectos dinâmicos relacionados à interpretação clínica. No entanto é possível identificar algumas dificuldades relacionadas a esse processo, tais como 1) limitações na compreensão de aspectos psicométricos e estatísticos por parte dos profissionais, 2) a utilização de instrumentos, em contextos clínicos que originalmente foram desenvolvidos para avaliação de trânsito e seleção organizacional e 3) um baixo consenso da relação entre percentil e classificação, o que pode gerar inconsistências entre resultados e avaliações. Posto isso, o presente trabalho é um estudo teórico que revisita conceitos fundamentais em Estatística e Psicometria associados à Neuropsicologia, discute condições em que testes voltados para trânsito e organizações são utilizados em Neuropsicologia e apresenta uma proposta de classificação.

*Palavras-chave:* Neuropsicologia; Psicometria; Estatística; Classificação; Percentil.

### ABSTRACT

#### Título em inglês\*\*\*\*

Neuropsychological assessment is a fundamental procedure, especially in cases of suspected neurological, psychological, or psychiatric disorders. One of its purposes is to map aspects of the patient's psychological and behavioral profiles and, with that, contribute to mental health's diagnostic process. This process combines the results obtained by psychometric instruments and dynamic aspects related to clinical evaluation. However, we note some difficulties related to this process, such as 1) limitations in the understanding of psychometric and statistical aspects by the professionals or clinicians; 2) use of instruments developed for Traffic Psychology assessment and personnel selection in industrial and organizational settings; 3) a low consensus on the relationship between percentile and classification or descriptors to scores that represent the performance obtained in the test, which can lead to inconsistencies between results and evaluations. That said, this theoretical study revisits fundamental concepts in Psychometrics and Statistics related to Neuropsychology, discusses the conditions under which contextual-unrelated tests are used in Neuropsychology, and provides a proposal for classification.

*Keywords:* Neuropsychology; Psychometrics; Statistics; Classification; Percentile.

### Sobre os autores

L. A.  
<http://orcid.org/0000-0001-5303-5782>  
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) – Rio de Janeiro, RJ  
[luisfca@puc-rio.br](mailto:luisfca@puc-rio.br)

A. C. P.  
<http://orcid.org/0000-0003-3693-9979>  
Instituto Brasileiro de Neuropsicologia e Ciências Cognitivas – Brasília, DF  
[portugal.aca@gmail.com](mailto:portugal.aca@gmail.com)

J. L. F.  
<http://orcid.org/0000-0002-8395-8008>  
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) – Rio de Janeiro, RJ  
[landeira@puc-rio.br](mailto:landeira@puc-rio.br)

### Direitos Autorais

Este é um artigo de acesso aberto e pode ser reproduzido livremente, distribuído, transmitido ou modificado, por qualquer pessoa desde que usado sem fins comerciais. O trabalho é disponibilizado sob a licença Creative Commons CC-BY-NC.



<sup>1</sup> Trabalho submetido ainda com vinculação de Professor do Departamento de Psicometria da Universidade Federal do Rio de Janeiro. E-mail de submissão: [luisfca@ufrj.br](mailto:luisfca@ufrj.br)

*There is no known way to present test results  
without confusing someone.*

*John O. Willis*

A Psicologia, no geral, e a Neuropsicologia, em específico, tradicionalmente utilizam instrumentos de medida em processos de avaliação de pessoas. Testes, escalas e inventários buscam operacionalizar dimensões abstratas e complexas de maneira objetiva e padronizada (Anunciação, 2018). Uma condição particular dessas medidas é que elas produzem resultados numéricos que, indiretamente, representam o fenômeno psicológico que está sendo avaliado, seja uma habilidade (como inteligência e memória), característica ou traço de personalidade ou condições afetivas.

Comumente, os resultados brutos obtidos nesses instrumentos neuropsicológicos são relacionados com indicadores estatísticos como percentis e classificações textuais, o que auxilia o profissional em diversos aspectos. Entre eles, tais resultados permitem levantar e explorar o perfil neuropsicológico do avaliado, auxiliar o processo diagnóstico em saúde mental, investigar as possíveis relações entre cérebro-comportamento de disfunções observadas, estimando também a sua localização e amplitude/tamanho, bem como monitorar os efeitos de algum tratamento em específico, entre outros (Lezak et al., 2012; Schoenberg & Scott, 2014).

No entanto há três condições em Neuropsicologia que frequentemente despertam o interesse acadêmico. Primeiro, a intensa relação que testes neuropsicológicos apresentam com estatística e psicometria não costuma vir acompanhada do conhecimento de termos técnicos nessas duas áreas (Bowman, 2002; Stigler, 1992); segundo, parte considerável da avaliação neuropsicológica feita no Brasil conta com instrumentos desenvolvidos, ou no contexto da Psicologia do Trânsito, ou na Psicologia Organizacional, o que gera um grande desafio durante a interpretação dos resultados da avaliação neuropsicológica (Guerra et al., 2020; Hazin et al., 2019; Mäder, 1996); e, finalmente, há um grande hibridismo em relação aos valores de percentil que um avaliando alcança e as categorias de classificação proposta pelos diferentes instrumentos (Guilmette et al., 2008, 2020).

Com isso posto, este trabalho teórico congrega esforços para revisar conceitos fundamentais em Psicometria e Estatística aplicados à Neuropsicologia, apresentar os possíveis problemas da utilização de testes desenvolvidos para finalidades diversas à Neuropsicologia e também apresentar uma referência que possa auxiliar na classificação dos resultados obtidos em avaliações neuropsicológicas.

Dessa maneira, este estudo visa diminuir parte da confusão que costuma ocorrer quando é necessário interpretar os

resultados obtidos por instrumentos psicológicos e, consequentemente, ampliar a comunicabilidade entre os profissionais. Para tanto, este trabalho está subdividido em 7 seções, que discorrem sobre o processo de aplicação dos testes em avaliações neuropsicológicas sob um panorama psicométrico e estatístico, expõem algumas dificuldades de quando instrumentos construídos em outros contextos são utilizados na prática neuropsicológica e, finalmente, introduz uma proposta de classificação que possa auxiliar na interpretação e tomada de decisão de psicólogos e neuropsicólogos clínicos.

## O PROCESSO DE APLICAÇÃO DE TESTES NEUROPSICOLÓGICOS

Há uma forte relação entre um processo de avaliação neuropsicológica e a administração de testes neuropsicológicos. No geral, testes psicológicos podem e são utilizados em um processo de avaliação neuropsicológica (Goldstein et al., 2019). Alguns autores, no entanto, tentam organizar e diferenciar entre ambos os tipos de testes. Assim, apesar de haver definições estritas que entendem que um teste só deva ser classificado como neuropsicológico quando ele é "sensível à presença de compromissos funcionais no sistema nervoso central" (Hebben & Milberg, 2009, p. 44), pragmaticamente, entende-se que os testes neuropsicológicos são parte específica da instrumentação psicológica, composta por atividades baseadas em performance, em que os resultados têm utilidade clínica e permitem não apenas identificar eventuais disfunções psicológicas e psiquiátricas, mas também associar tais condições clínicas com aspectos estruturais e funcionais do funcionamento cerebral. Quando uma série de testes neuropsicológicos é utilizado em conjunto para finalidades clínicas, essa coleção de atividades e instrumentos recebe o nome de bateria (Harvey, 2012; Russell, 2012).

Após a aplicação de testes neuropsicológicos, é necessário que sejam executadas algumas ações para que seja possível utilizar os resultados da testagem como suporte em decisões clínicas. Isto é realizado por um processo de correção, que pode ser dividido em quatro etapas sucessivas, que são: 1) correção do instrumento, pela implementação de procedimentos técnicos e matemáticos que possam apurar os resultados brutos; 2) identificação e posicionamento desse resultado apurado em tabelas específicas, quase sempre relacionadas a normas intragrupos e geradas por processos de tratamento e/ou transformação dos dados, como percentil, Escore Z, Escore T ou Escala QI/QI de Desvio; 3) classificação desse percentil (ou outra medida) em categorias textuais específicas associadas a um critério de performance (médio, médio superior etc.) e 4) interpretação dos resultados à luz de hipóteses clínicas e outras questões específicas. Apesar desses instrumentos não terem capacidade diagnóstica, seus resultados são bastante úteis nesse processo (Martinelli et al., 2014; Schoenberg & Scott, 2014)

A primeira etapa é relacionada à apuração. Etimologicamente, essa palavra significa tornar mais puro e talvez isso indique justamente o que se faz aqui. Nesse momento, o profissional deve ter a possibilidade de obter ou derivar um resultado que possa caracterizar a performance que o examinando teve no instrumento e que, indiretamente, representa ou reflete o fenômeno psicológico avaliado. Apesar dessa métrica ter importância, a interpretação dela é limitada, o que será exposto detalhadamente na próxima seção.

Na apuração, os procedimentos matemáticos necessários costumam ser simples. Como exemplo, somar pontos dos itens que foram respondidos corretamente (como é feito no Mini Exame do Estado Mental), cronometrar o tempo total que foi gasto para executar a atividade (como é feito na Figura Complexa de Rey) ou implementar fórmulas ou equações específicas (como é feito no Teste de Aprendizagem Auditivo-Verbal de Rey; Spreen & Strauss, 1998).

Na segunda e terceira etapas (normatização e classificação), o profissional deve utilizar tabelas específicas. Nelas, é possível localizar o resultado apurado obtido por um examinando, identificar medidas estatísticas em que transformações foram realizadas e uma classificação textual relacionada a essa medida. Com isso, é possível comparar e posicionar o resultado bruto obtido pelo examinando com uma amostra normativa.

Por suas propriedades descritivas, com muita frequência, o percentil é a medida utilizada nessas tabelas (Witz et al., 1990). A interpretação dos resultados é dinâmica, necessariamente varia em função da finalidade da avaliação e deve se relacionar com a história clínica do examinando.

## A RELAÇÃO ENTRE DADOS BRUTOS E SUAS TRANSFORMAÇÕES

Tal como exposto, os resultados obtidos por um instrumento psicológico são chamados de dados brutos e são pouco úteis ou informativos sobre a variável psicológica (chamada também de traço latente) que ele visa medir (Anastasi, 1976; Cermak, 1989). Por exemplo, se um examinando tiver obtido 80 pontos em um teste de atenção seletiva ou 40 pontos em um instrumento de inteligência, isso não teria nenhum significado fora do próprio sistema de pontuação do instrumento, tampouco auxiliaria na interpretação dos resultados e eventual decisão a ser tomada futuramente. Dessa maneira, quando o interesse se volta em posicionar e comparar o desempenho de um examinando nessas funções em uma métrica que seja interpretável, comunicável e lógica, esses dados brutos precisam passar por conversões, comumente chamadas de transformações.

Como em todas as ciências, métodos estatísticos fornecem ferramentas importantes para a organização dos dados e isso não é diferente em Neuropsicologia. Em síntese, essas transformações costumam ser agrupadas em lineares e não lineares. Por via de regra, transformações lineares preservam todas as relações matemáticas dos dados brutos, enquanto as não lineares ajustam os dados para que sejam compatíveis a uma distribuição de interesse. A Figura 1 apresenta algumas transformações e métodos utilizados em Neuropsicologia, que serão seletivamente discutidos em seguida.

De forma geral, dois conceitos são necessários para entender esse processo: 1) A distribuição normal e os Escores Padrão; e o 2) Percentil. Ambos os conceitos são inter-relacionados e as seções abaixo servem para apresentá-los e também apresentar como eles podem auxiliar na comunicação de resultados em processos de avaliação neuropsicológica.

## A DISTRIBUIÇÃO NORMAL E ESCORES PADRÃO

Os resultados obtidos (escores apurados) em um teste neuropsicológico são assumidos como fenômenos ou variáveis aleatórias (Belhekar, 2016). Apesar da palavra "aleatória(o)" ter diferentes interpretações, em estatística, significa que existe um processo probabilístico, ou uma distribuição, subjacente aos dados. Dessa maneira, os valores dessa variável ocorrem em função de uma determinada distribuição de probabilidade (Everitt, 2002). Assim, ao saber a distribuição de probabilidade que caracteriza os resultados, além de ser possível descrevê-los, também se torna viável desenvolver um modelo matemático que possa estimar ou prever como e quais serão os futuros dados.

Como em diferentes áreas da Psicologia, em Neuropsicologia se assume que os fenômenos sigam uma distribuição normal (Koziol et al., 2016, p. 52; Russell, 2012, p. 62). As características dessa distribuição são as seguintes: 1) ela integra a família das distribuições contínuas, em que 2) as observações podem assumir qualquer valor no conjunto dos reais (de  $-\infty$  a  $+\infty$ ) e 3) sua área total acumulada igual a 1. Além disso, ela apresenta 4) formato de sino (duas caudas e um pico), com os 5) valores da média, moda e mediana iguais e localizados ao centro em que 6) todas as outras observações se afastam ou aproximam deles de maneira simétrica. Finalmente, a distribuição normal é 7) definida por dois parâmetros (média populacional -  $\mu$  e variância populacional -  $\sigma^2$ ), além de ter 8) Coeficiente de curtose igual a 3 e Coeficiente de assimetria igual a 0. É importante ressaltar que, como todas as distribuições estudadas em estatística, a distribuição normal é um conceito teórico útil à modelagem de dados e não um resultado empírico observado a todo tempo ao se medir fenômenos psicológicos.

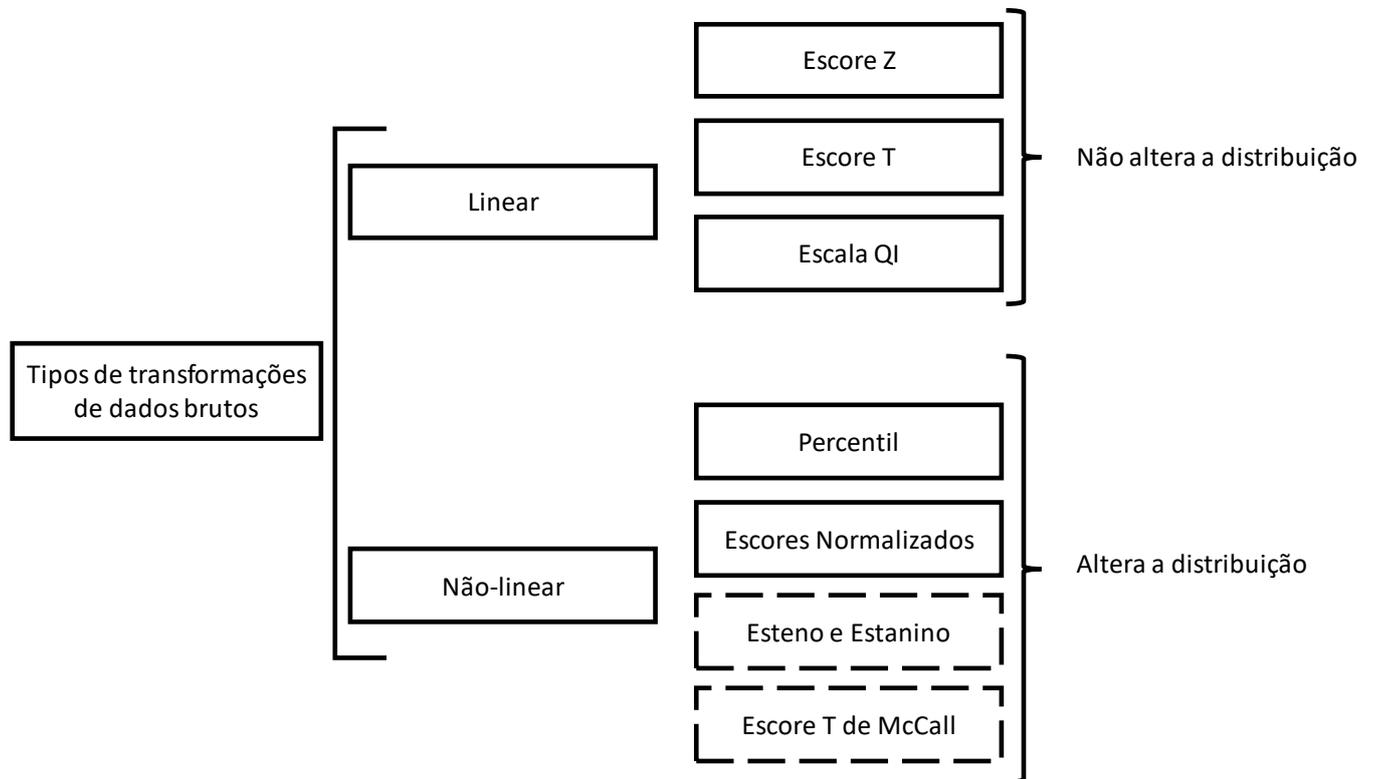


Figura 1. Tipos de transformações de dados brutos.  
Nota: Retângulos com linha pontilhada referem-se às transformações menos usadas atualmente.

Uma vez que as médias e os desvios-padrão dos escores brutos (que foram calculados na etapa de apuração) obtidos podem variar, é possível perceber que infinitas distribuições normais poderão coexistir. Trazendo isso à Neuropsicologia, esse fenômeno torna a interpretação dos resultados inteiramente dependente do sistema de correção desenvolvido pelo instrumento. Ou seja, é possível que um teste neuropsicológico tenha valores entre 0 e 36 (por exemplo, a Figura Completa de Rey), enquanto outro possa gerar valores entre 0 e 10 (por exemplo, o Teste do Desenho dos Relógios).

Diferentes procedimentos estatísticos são propostos para auxiliar na construção de uma métrica que sirva para interpretação desses resultados e as comparações intragrupos são as mais frequentemente solicitadas. O termo Escore Padrão é utilizado para identificar toda classe de procedimentos estatísticos em que se fixa a média e o desvio-padrão dos dados em valores pré-determinados. Três métricas são tradicionalmente utilizadas em Neuropsicologia, que são o Escore Z, o Escore T e o QI de Desvio.

O Escore Z é um tipo de transformação linear que indica, em unidades de desvio-padrão, a direção e a intensidade com que determinado resultado bruto se afasta da média da distribuição à qual pertence. A fórmula de seu cálculo amostral

está abaixo e, como um cânone ou uma convenção matemática, não pode ser alterada:

$$Z_i = \frac{x_i - \bar{X}}{s}$$

Onde:

$Z_i$  = Escore Z do examinando i

$x_i$  = Escore bruto obtido pelo examinando i

$\bar{X}$  = Média da amostra normativa

$s$  = Desvio-padrão da amostra normativa

Com o Escore Z calculado para cada participante, é possível identificar o quanto o seu resultado está, em unidades de desvios-padrão, distante da média de seu grupo. Essa padronização permite três condições vantajosas, que são: 1) a maior clareza na interpretação dos resultados, permitindo identificar pontos discrepantes ou anômalos, 2) o cálculo da probabilidade esperada a partir da área sob a curva de densidade das observações e 3) a comparação entre resultados originalmente medidos em escalas diferentes. Por exemplo, uma pessoa com o Escore Z igual a 0 está situada exatamente na média de seu grupo, enquanto uma outra com o Escore Z de -1.5 está a -1.5 desvios-padrão da média de seu grupo normativo.

Muitos diagnósticos em saúde mental tem a participação do Escore Z, como o caso do Transtorno Específico da Aprendizagem, que pede Escore Z de -1.5 ou do Transtorno do Desenvolvimento Intelectual, "em que o Escore Z deve ser de -2" (American Psychiatric Association, 2013, p. 37). Torna-se importante ressaltar que essas métricas não são diagnósticas, mas servem tão somente para auxílio clínico.

O Escore T foi originalmente desenvolvido em 1922 por William A. McCall. À época, seu cálculo precisava tanto de transformações lineares como não lineares dos dados. Por isso, o Escore T de McCall é entendido como um tipo de transformação mista. Uma das motivações para escolha da letra "T" foi homenagear os psicólogos Lewis Terman e L. Thorndike, que foi seu orientador de doutorado. A escolha do valor do desvio-padrão tentou evitar que resultados negativos fossem presentes, mas também por McCall julgar que esse procedimento "gerava um zero matemático perto do suposto zero absoluto das funções mentais" (McCall, 1922, p. 294). Atualmente, o Escore T teve sua operacionalização alterada e hoje é uma transformação linear dos Escores Z obtidos, em que se fixa média teórica em 50 e desvio-padrão igual a 10 (Krus & Krus, 1977). É importante atentar que não há qualquer relação entre Escore T e distribuição ou estatística T de *Student* (Belhekar, 2016).

Finalmente, o QI de desvio tem relação histórica próxima ao que aconteceu com o Escore T. Essa métrica tem média teórica de 100 e desvio-padrão em 15, mas seu cálculo pode ser feito tanto por uma transformação linear do Escore Z, como por uma transformação não linear dos dados. Nessa última, conveniou-se chamar de Escore Padrão Normalizado, já que sua operacionalização envolve calcular frequência acumulada percentual dos escores brutos (ou seja, o percentil) e, em seguida, verificar o valor do Escore Z em uma distribuição normal teórica equivalente àquela proporção para, finalmente, atribuí-lo ao escore bruto originalmente apurado (Burns, 1988).

Frequentemente, a métrica QI é utilizada comumente em testes de inteligência. Uma nota de atenção é importante aqui. O termo QI sofreu por muitas modificações estatísticas desde seu surgimento, cunhado por Wilhelm Stern (Goldstein et al., 2019) em 1912, e, por vício de linguagem, frequentemente o termo "testes de QI" é utilizado para se referir a todos os instrumentos de inteligência.

Ao calcular o Escore Z para todos os dados ou escores de um teste, uma nova distribuição será formada, agora com a média igual a 0 e o desvio-padrão igual a 1. Essa distribuição é chamada de distribuição normal padronizada e não altera o formato original dos dados. Conseqüentemente, a aplicação do Escore Z não altera o formato da distribuição e, portanto, não normaliza os dados (Howell, 2011).

Uma das vantagens operacionais de trabalhar com dados assumindo uma distribuição normal subjacente é a implementação da regra empírica. Essa regra assegura que, em variáveis normalmente distribuídas, cerca de 68% das observações estarão entre  $\pm 1$  desvio-padrão, cerca de 95% estarão entre  $\pm$  dois desvios-padrão da média e cerca de 99.7% das observações estarão entre  $\pm$  três desvios-padrão da média (Witz et al., 1990). Em muitas convenções em Estatística descritiva, um intervalo médio é visto entre -1 e +1 desvio-padrão (Guilmette et al., 2020).

É importante ressaltar que uma crítica frequentemente realizada argumenta que as variáveis psicológicas não são normalmente distribuídas. Em outras palavras, quando se constrói histogramas ou gráficos de densidade para apresentar a distribuição dos dados obtidos por instrumentos psicológicos, o seu formato é assimétrico. Por exemplo, é bem aceito que medidas de depressão obtidas por inventários costumam ter formato assimétrico positivo (à direita), enquanto medidas de bem-estar possuem formato assimétrico negativo (à esquerda) (Micceri, 1989; Stahel & Wilcox, 1998).

No entanto, para modelagem dos dados e inferência estatística, a estatística frequentista considera que os dados obtidos por uma pesquisa são apenas uma realização particular e independente de uma variável que, em teoria, pode se realizar infinitamente. Como os processos inferenciais são baseados no Teorema Central do Limite (TCL), tecnicamente não é inadequado contar com procedimentos chamados de inferência paramétrica para esse tipo de dado (Kwak & Kim, 2017). Nesse sentido, utilizar médias e desvios-padrão, bem como toda classe de transformações lineares e construção de intervalos de confiança tem justificativa.

O TCL assegura que a distribuição das médias de amostras aleatórias tomadas de uma população de distribuição qualquer, desde que com variância definida e não infinita, converge à distribuição normal quando o tamanho dessas amostras ( $n$ ) se torna suficientemente grande (Equação 2). Em simulação computacional, 30 já é considerado grande para verificar o processo em alguns casos (Howell, 2011).

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} Z \sim N(0,1) \text{ quando } n \rightarrow \infty$$

Onde:

$\mu$  = Média populacional

$\sigma$  = Desvio-padrão populacional

Com isso posto, torna-se mais claro que a apresentação apenas dos resultados brutos obtidos em uma avaliação neuropsicológica é pouco informativa. Unir essas informações

com Escores padrão (como o Escores Z ou T) é fundamental à comunicação dos resultados em uma avaliação neuropsicológica e também permite que análises inferenciais sejam desenvolvidas (Schoenberg & Scott, 2014).

## OS PERCENTIS (OU POSTOS PERCENTÍLICOS)

O percentil é uma medida de posição relativa. Ele é calculado pelo ranqueamento de todas as observações, seguida pela divisão da distribuição dos dados (agora ordenados) em 100 partes iguais. Assim, por sua própria definição, o percentil indica a posição de um escore em relação aos demais, indicando a porcentagem de resultados abaixo ou iguais àquele. Existem vantagens e desvantagens na utilização do percentil para interpretação dos resultados de um teste neuropsicológico. Entre as vantagens, está 1) a interpretação intuitiva mesmo às pessoas sem conhecimento técnico. Por exemplo, se um estudante obteve um resultado bruto qualquer (exemplo, 33) e esse valor indica o percentil de 70, isso indica que 70% do grupo que compôs a amostra normativa tiveram resultados abaixo ou iguais a esse. Pragmaticamente, pode-se interpretar que a performance obtida está acima de 70% dos outros estudantes que fizeram o teste (e, conseqüentemente, abaixo de 30% desse grupo) (Witz et al., 1990).

Além disso, é 2) fácil de computar, 3) independe da distribuição subjacente dos dados, tendo a 4) a mesma interpretação em quaisquer distribuições, sendo 5) possível também de se aproximar a partir do Escore Z, se algumas características forem assumidas, e 6) aplicável a todos os instrumentos psicológicos e neuropsicológicos. Dessa maneira, muitos pesquisadores consideram o percentil uma medida descritiva vantajosa quando comparada aos Escores padrão, especialmente quando a distribuição dos dados não é aproximadamente normal (Crawford et al., 2009).

No entanto algumas características limitadoras são importantes. Como o cálculo do percentil realiza uma transformação não linear (monotônica retangular) dos escores brutos, ele tem nível de medida ordinal, em que as propriedades intervalares se perdem. Por exemplo, embora haja distância constante entre dois escores Z, isso gera uma diferença variável na quantidade de observações em cada intervalo percentil. O percentil de 16 a 50, contem 34% da distribuição e equivale ao Escore Z de -1 a 0. No entanto o percentil de 2 até 16, que equivale ao Escore Z de -2 a -1, contém apenas 14% das observações.

Essa perda dessa propriedade impossibilita que 1) operações matemáticas sejam feitas pelos percentil (por exemplo, não se deve somar ou calcular a média de dois percentis), 2) haja uma amplificação artificial das diferenças entre escores mais ao centro da distribuição original dos dados e 3) haja uma redução das diferenças de escores extremos (Bowman, 2002; Witz et al., 1990). Por exemplo, em um teste de

inteligência que conte com uma distribuição normal subjacente aos dados, uma diferença de apenas 3 pontos ao centro da distribuição (100 vs 103) gera uma distância aproximada de 10 pontos percentis, enquanto seriam necessários mais de 20 pontos de diferença ao fim da distribuição para a mesma distância percentilica (120 versus 145 pontos) (Crawford & Garthwaite, 2009; Wechsler, 2008).

Uma condição ainda importante em relação ao percentil é que existem diferentes maneiras de realizar o cálculo, o que pode explicar resultados diferentes entre programas (Hyndman & Fan, 1996). Neste trabalho, por exemplo, a Tabela 1 conta com a apresentação do percentil assumindo uma distribuição normal subjacente, tal como é tipicamente utilizado em manuais de testes no país. Finalmente, a despeito da parte matemática associada ao percentil, existem recomendações matematicamente inconsistentes e incorretas de buscar o percentil mais alto quando os resultados brutos não apresentam valores percentílicos diretamente associados (Anastasi & Urbina, 1997).

Isto posto, na apresentação dos resultados de uma avaliação neuropsicológica, é necessário apresentar o Escore Z obtido, mas também o percentil associado a ele. Essa inclusão é alinhada às recomendações internacionais e torna a comunicação mais efetiva (Guilmette et al., 2020; Schoenberg & Scott, 2014). Isso se dá pois, enquanto os percentis oferecem descrições fáceis de serem entendidas, os Escores padrão auxiliam em eventuais inferências feitas (Crawford et al., 2009; Schoenberg & Scott, 2014). Uma vez que agora torna-se claro ver a relação entre os conceitos expostos, a figura 2 traz à cena as convergências explicitadas.

## AS CLASSIFICAÇÕES DOS RESULTADOS OBTIDOS

As classificações são textos que sintetizam a performance obtida em um instrumento de maneira clara. O objetivo aqui é que elas possam ser facilmente compreendidas mesmo por profissionais sem devido treinamento técnico e sirvam também para embasar tomadas de decisão. Por definição, esses textos são feitos por convenção e as categorias criadas são exaustivas e mutuamente excludentes. Ou seja, elas devem abranger todos os níveis do fenômeno estudado e, se um determinado resultado for alocado a uma classificação específica, ele não poderá estar em outra (Guilmette et al., 2008).

O processo de criação/nomeação das categorias costuma ter uma associação direta com o percentil que um examinado obteve. Em testes que foram desenvolvidos para avaliação neuropsicológica, a iniciativa de criar classificações gradativas associadas à performance é recente, tendo iniciado por volta de 1970 com Russell, Neuriger e Goldstein, "que utilizaram os descritores leve, moderado e severo para graduar a incapacidade neuropsicológica" (Hebben & Milberg, 2009, p. 161).

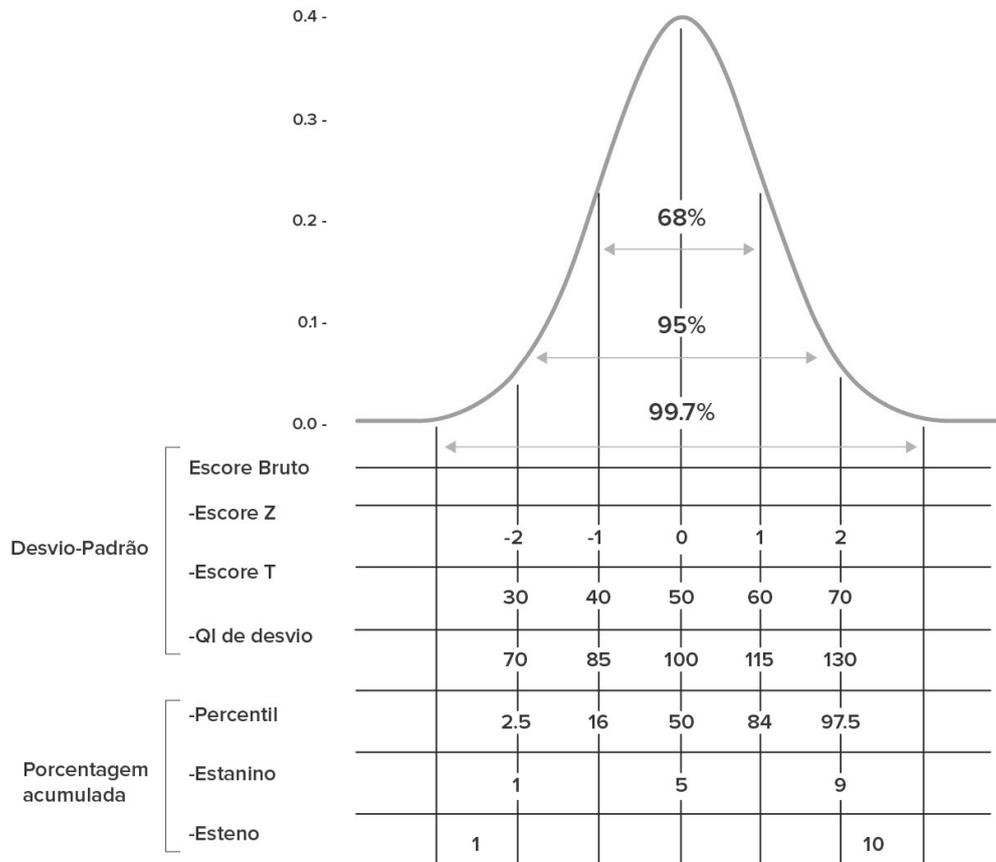


Figura 2. Convergências da distribuição normal.

Apesar das categorias terem uma aplicação direta como fontes que embasam decisões clínicas, existem divergências na literatura sobre 1) quais devem ser os rótulos ou descritores textuais para sintetizar a performance obtida, 2) quantas categorias devem ser consideradas para análise, seguido por 3) como ancorar cada uma das categorias ao valor percentil ou outra medida normativa e, finalmente, 4) ao se considerar o percentil, qual deve ser o percentil limite para inclusão em categorias (Horton Jr. & Wedding, 2008, p. 297).

Em relação aos rótulos e à quantidade, há um consenso na literatura brasileira e internacional de que os rótulos devem possibilitar a comparação da performance de um examinando em relação à média de seu grupo (Guilmette et al., 2008). Apesar da média ser um único valor (pontual), com frequência, ela é dividida também em subgrupos, como “média inferior” e “média superior” e aparecem seguidas da categoria “superior” e “inferior”, totalizando cinco categorias (Swinson, 2013).

Entretanto, dentro dessa parcial convenção, o sistema Wechsler adiciona também a classificação “muito superior”, “limítrofe” e “extremamente baixo” e retira a categoria “inferior”. Por esse sistema, há o total de sete categorias (Wechsler, 2008).

Em oposição a esse sistema, o sistema Heaton oferece oito categorias, em que uma categoria intitulada “comprometido” é inserida. Eventualmente, há também um hibridismo entre os termos “médio inferior” e “abaixo da média”, com significados diferentes (Schoenberg et al., 2018).

No que diz respeito à ancoragem de cada uma das categorias ao valor percentil ou outra medida normativa, a convenção estatística exposta na regra empírica não costuma ser utilizada e divergências ocorrem entre os sistemas. Por exemplo, o Sistema Wechsler classifica como médio os participantes com percentil 25 (Escore Z = -0.67), mas o sistema Heaton considera essa uma performance abaixo da média. Já o Escore Z -0,69 corresponde ao percentil 25 e o sistema Wechsler considera médio, enquanto o Heaton acima da média ou no Escore Z -0.6, que é classificado como médio pelo sistema Wechsler e abaixo da média pelo Heaton. Parte dessas divergências ocorre justamente por uma transformação ser linear (Escore Z) e outra não linear (Percentil), bem como por critérios dos limites percentílicos necessários para sair de uma categoria e ir para outra (Willis, 1998).

Em instrumentos psicológicos considerados favoráveis para uso, quase sempre os próprios manuais e textos-técnicos apresentam as informações necessárias para que o profissional consiga realizar todo o procedimento. Entretanto, mesmo assim, a relação entre valor percentílico e classificação frequentemente é diferente de um manual ou texto-técnico para outro (Sedó et al., 2015). Como exemplos, alguns testes cognitivos consideram que todo o conjunto composto por resultados entre médio inferior e médio superior está no intervalo em 10 e 90 postos percentis. Como exemplo, o Teste Atenção Concentrada (AC – Cambraia, 2009) e o Teste de Habilidade para o Trabalho Mental (HTM – Santarosa et al., 2011). De maneira distinta, outros testes consideram que essas mesmas classificações estão compreendidas entre os percentis 25 e 75 (como exemplo, a Bateria Psicológica para Avaliação da Atenção (BPA; Rueda, 2013).

## UMA DIFICULDADE: QUANDO TESTES PARA AVALIAÇÃO NOS CONTEXTOS DO TRÂNSITO E ORGANIZACIONAL SÃO UTILIZADOS EM NEUROPSICOLOGIA

No Brasil, a utilização de instrumentos psicológicos é restrita a psicólogas e psicólogos regularmente inscritos no Conselho Regional de Psicologia (CRP). Por sua vez, esses profissionais só podem utilizar instrumentos com parecer “Favorável” na lista oferecida pelo Sistema de Avaliação de Testes Psicológicos (SATEPSI), excetuado os casos de pesquisa. Atualmente (2020), essa lista tem cerca de 155 instrumentos privativos distintos, que foram desenvolvidos para avaliar características psicológicas heterogêneas, incluindo desde fenômenos tradicionalmente intitulados como cognitivos (atenção, inteligência e memória, por exemplo), como traços e características de personalidade (que contam também com métodos projetivos) e funções menos específicas (por exemplo, orientação profissional e empregabilidade; CFP, 2018).

Evidências prévias têm sistematicamente demonstrado que a maior parte desse conjunto de testes foi desenvolvida para uma finalidade diferente da avaliação neuropsicológica (Guerra et al., 2020; Hazin et al., 2019; Mäder, 1996) e que o contexto da avaliação psicológica do trânsito e organização figuram como principais canais para criação de instrumentos de avaliação psicológica (CFP, 2016; Cruz et al., 2017; Mansur-Alves et al., 2016; Sampaio & Nakano, 2011; Silva, 2012).

Nesse sentido, as características subjacentes ao desenvolvimento de instrumentos psicológicos para serem utilizados no trânsito ou em organizações são bastante distintas daqueles presentes no desenvolvimento de testes com finalidades neuropsicológicas em, ao menos, seis aspectos, descritos a seguir.

Inicialmente, o público-alvo e sua característica clínica em ambos os procedimentos são diferentes. No contexto do trânsito ou organizacional, assume-se que examinandos tenham idade superior a 18 anos e inferior à idade legal de aposentadoria, bem como não apresentem disfunções psicológicas e/ou transtornos mentais. Já na avaliação neuropsicológica, por oposição, o público é majoritariamente formado ou por crianças ou idosos e o procedimento costuma fazer parte dos exames clínicos utilizados se chegar a um diagnóstico de saúde mental. Essa última característica é importante, pois estudos voltados à utilidade clínica de instrumentos (sensibilidade, especificidade, curva ROC) são pouco presentes nesses primeiros exames psicológicos.

Além disso, em procedimentos de avaliação psicológica, especialmente para o trânsito ou para organizações, os testes utilizados são majoritariamente desenvolvidos para avaliar características psicológicas aderentes ao contexto específico. Assim, é esperado e plenamente justificável que os instrumentos cognitivos apresentem estímulos como placas automotivas ou de trânsito, o que tem baixa aderência em avaliações neuropsicológicas (Cruz et al., 2017).

Ainda por características que fazem interface entre aspectos legislativos e operacionais, muitos testes utilizados no contexto do trânsito ou organizacional são rápidos e aplicados com finalidade exclusiva de sondagem ou rastreamento cognitivo, o que marca o quarto aspecto. Como em avaliações neuropsicológicas é esperado que haja alguma condição clínica de saúde a investigar, tais testes com características de sondagem ou rastreios poderiam gerar resultados incorretos para grande parte das avaliações em que o avaliando ainda não apresenta grande comprometimento cognitivo.

Em relação à forma de aplicação, o quinto aspecto pode ser visto na tentativa de controle de variáveis. Muitos instrumentos desenvolvidos para o contexto de avaliação psicológica, mas especialmente para utilização em avaliação compulsórias, tentam controlar o máximo possível das variáveis e ter estímulos distantes do cotidiano do avaliado. Em Neuropsicologia, com frequência, tenta-se justamente dispor de instrumentos que repliquem ou mimetizem características do cotidiano e aumentem a validade ecológica do instrumento.

Finalmente, características normativas merecem atenção e marcam o sexto aspecto previamente mencionado. Da mesma forma como exposto anteriormente, a relação entre classificação e percentil em testes psicológicos utilizados no contexto do trânsito e organizacional também é heterogênea. Uma parte considerável dos instrumentos entende que a classificação média envolve um intervalo percentílico entre 10 e 90, conforme previamente exposto. No entanto há sugestão preliminar de que isso possa ser associado também às características legislativas. Uma vez que a avaliação psicológica

nesse contexto tem caráter eliminatório e conclusivo (CFP, 2019), candidatos considerados inferiores seriam eliminados do processo e essa classificação ocorre àqueles com percentil inferior a 10 (Brasil, 1998).

É importante notar que essas características não desencorajam ou proíbem que instrumentos psicológicos desenvolvidos para o contexto do trânsito ou organizacional sejam utilizados em avaliações neuropsicológicas. Como exposto anteriormente, os testes neuropsicológicos fazem parte da instrumentação psicológica. Ainda nesse sentido, como poucos instrumentos foram desenvolvidos inicialmente para uso em neuropsicologia clínica, é inclusive recomendado que tais testes sejam incorporados na avaliação neuropsicológica e sirvam como auxílio ao procedimento. No entanto, isso deve ser realizado de maneira cautelosa.

## UMA PROPOSTA DE UNIFICAÇÃO DA CLASSIFICAÇÃO

Tendo em vista que há uma divergência em como os autores realizam as conversões entre aspectos estatísticos e categorias de classificação, bem como a particular condição de uso feita por neuropsicólogos brasileiros aos instrumentos criados para diferentes contextos, a tabela a seguir apresenta uma recomendação de utilização clínica. Ela é útil nos casos em que se utiliza instrumentos não desenvolvidos primariamente para avaliação neuropsicológica, bem como naqueles com divergência de tabelas.

Por esse modelo, aproveita-se o parcial consenso acadêmico em considerar valores abaixo de 1.5 Escore Z como ponto de corte (American Psychiatric Association, 2013; Green, 2000), bem como torna a classificação mais próxima às tabelas frequentemente dispostas em manuais técnicos de testes psicológicos. Em direção similar, essa tabela permite que o profissional tenha disponível uma métrica menos dependente da tradicional classificação do Sistema Wechsler, criado para avaliação da inteligência e também mais próxima à instrumentação psicológica utilizada no Brasil.

## CONSIDERAÇÕES FINAIS

Neste trabalho teórico, três objetivos foram trabalhados. Inicialmente, a relação entre avaliação neuropsicológica e aspectos psicométricos e estatísticos foi apresentada, o que permitiu que conceitos analíticos, como padronização de resultados, fossem expostos e discutidos. Em um segundo momento, discutiu-se algumas das características potencialmente problemáticas que ocorrem quando instrumentos psicológicos desenvolvidos para o contexto do trânsito ou organizacional são utilizados em avaliações neuropsicológicas. Finalmente, a relação entre postos percentílicos e classificação dos resultados foi debatida, possibilitando o oferecimento de uma possibilidade de classificação.

Tabela 1. *Recomendação de unificação da classificação a partir do intervalo percentílico*

Percentil	Escore Z $\mu=0 \sigma=1$	Escore T $\mu=50 \sigma=10$	QI de desvio $\mu=100 \sigma=15$	Escore Ponderado $\mu=10 \sigma=3$	Classificação
<5	<-1.64	<34	<75	<5	Risco clínico
<10	<-1.28	<37	<81	<6	Inferior
10-29	<-0.52	37-44	81-91	6-8	Médio inferior
30-70	<0.52	45-55	92-107	9-11	Média
71-90	<1.28	56-63	108-119	12-14	Médio superior
>90	>1.28	>63	>119	>14	Superior

Como toda proposta, há pontos positivos e questões que podem trazer problemas. Em relação aos pontos positivos, essa proposta converge o percentil do intervalo médio para, aproximadamente, entre -0.5 e 0.5 Escore Z. Similar ao este-no, cerca de 40% das observações estarão dentro desse intervalo. Valores médios inferiores podem ser considerados como indicadores clínicos, mas apenas os valores < 10 (Escore Z  $\cong$  -1.3) são vistos como inferiores, enquanto os < 5 (Escore Z  $\cong$  -1.6) são vistos como Risco clínico.

Esse último ponto acarreta diversas consequências, tais como: problemas na interpretação dos resultados, impactando assim decisões clínicas diferentes, baixa comunicação entre profissionais, dificuldades na comparação de resultados obtidos por diferentes instrumentos e certo risco jurídico quando as avaliações fazem parte de procedimentos na justiça. De forma a auxiliar na resolução de, ao menos, parte desses problemas, uma proposta de interpretação de resultados foi apresentada e discutida.

Em síntese, pretendeu-se tornar mais clara a relação intrínseca que a avaliação neuropsicológica apresenta tanto com a Psicometria como com a Estatística. Entretanto é importante também destacar que todos os instrumentos utilizados na avaliação trazem resultados importantes, mas somente complementares e que não podem ser utilizados diretamente para diagnóstico.

Por definição, todo diagnóstico faz parte da escala de medida nominal, em que qualquer número alocado (seja da CID-10 ou outro manual de classificação), permite apenas uma identificação. No entanto, para que esse diagnóstico seja válido, é importante que seu procedimento tenha sido embasado por evidências empíricas, em que se conjugue condições clínicas e resultados de instrumentos psicológicos.

Espera-se que o estudo em questão sirva para auxiliar no entendimento de aspectos psicométricos e estatísticos desses instrumentos. De maneira colateral, espera-se que este trabalho tenha tornado ainda mais clara a vívida e intrínseca relação que a Psicologia e a Neuropsicologia apresentam com a Psicometria e a Estatística.

## DECLARAÇÃO DA CONTRIBUIÇÃO DOS AUTORES

Certificamos que todos os autores participaram suficientemente do trabalho para tornar pública sua responsabilidade pelo conteúdo. A contribuição de cada autor pode ser atribuída como se segue:

L.A. contribuiu com a conceitualização, escrita do trabalho e revisão de conceitos. A.P. contribuiu com a revisão do trabalho. J.L.F. contribuiu com a conceitualização e revisão do texto. Todos os autores são igualmente responsáveis pela redação final (revisão e edição).

## DECLARAÇÃO DE CONFLITOS DE INTERESSE

Os autores declaram que não há conflitos de interesse no manuscrito submetido.

## REFERÊNCIAS

- American Psychiatric Association. (2013). DSM-5. In *Diagnostic and statistical manual of mental disorders*.
- Anastasi, A. (1976). *Psychological Testing* (4th ed.). New York: Macmillan Publishing.
- Anastasi, Anne, & Urbina, S. (1997). Psychological testing. In *Journal of sport & exercise psychology*. (Vol. 20, Issue 1).
- Anunciação, L. (2018). An Overview of the History and Methodological Aspects of Psychometrics-History and Methodological aspects of Psychometrics. *Journal for ReAttach Therapy and Developmental Diversities*. <https://doi.org/10.26407/2018jrtd.1.6>
- Belhekar, V. M. (2016). Probability: Basic Concepts and Random Variables. In *Statistics for Psychology Using R* (pp. 1-37). SAGE Publications, Inc. <https://doi.org/10.4135/9789353282493.n1>
- Bowman, M. (2002). The perfidy of percentiles. *Archives of Clinical Neuropsychology*, 17(3), 295-303. [https://doi.org/10.1016/S0887-6177\(01\)00116-0](https://doi.org/10.1016/S0887-6177(01)00116-0)
- Brasil, Conselho Nacional de Trânsito (1998). Resolução CONTRAN nº 51 de 21/05/1998. Dispõe sobre os exames de aptidão física e mental e os exames de avaliação psicológica a que se refere o inciso I, do artigo 147 do Código de Trânsito Brasileiro e os §§ 3º e 4º do artigo 2º da Lei nº 9.602/98.
- Burns, E. (1988). Normalized standard score spreadsheet and norm table generator. *Journal of School Psychology*. [https://doi.org/10.1016/0022-4405\(88\)90038-6](https://doi.org/10.1016/0022-4405(88)90038-6)
- Cambraia SV. (2009). *O Teste de Atenção Concentrada AC*. Manual. São Paulo: Vetor Editora.
- Cermak, S. (1989). Chapter 5: Norms and Scores. *Physical & Occupational Therapy In Pediatrics*, 9(1), 91-120. [https://doi.org/10.1300/J006v09n01\\_05](https://doi.org/10.1300/J006v09n01_05)
- Conselho Federal de Psicologia. (2016). *Psicologia do Tráfego: Características e desafios no contexto do MERCOSUL*. Brasília: CFP
- Conselho Federal de Psicologia. (2018). Resolução CFP nº 009/2018. Institui o Manual de Elaboração de Documentos Escritos produzidos pelo psicólogo, 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 as Resoluções nº 002/2003, nº 006/2004 e nº 005/2012 e Notas Técnicas nº 01/2017 e 02/2017.
- Crawford, J. R., & Garthwaite, P. H. (2009). Percentiles Please: The Case for Expressing Neuropsychological Test Scores and Accompanying Confidence Limits as Percentile Ranks. *The Clinical Neuropsychologist*, 23(2), 193-204. <https://doi.org/10.1080/13854040801968450>
- Crawford, J. R., Garthwaite, P. H., & Slick, D. J. (2009). On percentile norms in neuropsychology: Proposed reporting standards and methods for quantifying the uncertainty over the percentile ranks of test scores. In *Clinical Neuropsychologist*. <https://doi.org/10.1080/13854040902795018>
- Cruz, R. M., Wit, P. A. J. M. de, & Souza, C. Z. de. (2017). *Manual de psicologia do trânsito*. Nila Press Livraria e Editora Ltda.
- Everitt, B.S. (2002). *The Cambridge Dictionary of Statistics*. 2nd edition. Cambridge University Press, Cambridge, UK.
- Goldstein, G., Allen, D. N., & DeLuca, J. (2019). Historical perspectives. In *Handbook of Psychological Assessment* (pp. 3-27). Elsevier. <https://doi.org/10.1016/B978-0-12-802203-0.00001-8>
- Green, J. (2000). Neuropsychological evaluation of the older adult: A clinician's guidebook. [https://doi.org/10.1164/ajrcm-conference.2013.187.1\\_MeetingAbstracts.A2574](https://doi.org/10.1164/ajrcm-conference.2013.187.1_MeetingAbstracts.A2574)

- Guerra, A., Hazin, I., Siebra, C., Rezende, M., Silvestre, I., Le Gall, D., & Roy, A. (2020). Assessing executive functions in Brazilian children: A critical review of available tools. *Applied Neuropsychology: Child*, 1-13. <https://doi.org/10.1080/21622965.2020.1775598>
- Guilmette, T. J., Hagan, L. D., & Giuliano, A. J. (2008). Assigning Qualitative Descriptions to Test Scores in Neuropsychology: Forensic Implications. *The Clinical Neuropsychologist*, 22(1), 122-139. <https://doi.org/10.1080/13854040601064559>
- Guilmette, T. J., Sweet, J. J., Hebben, N., Koltai, D., Mahone, E. M., Spiegler, B. J., Stucky, K., & Westerveld, M. (2020). American Academy of Clinical Neuropsychology consensus conference statement on uniform labeling of performance test scores. *The Clinical Neuropsychologist*, 34(3), 437-453. <https://doi.org/10.1080/13854046.2020.1722244>
- Harvey, P. D. (2012). Clinical applications of neuropsychological assessment. *Dialogues in Clinical Neuroscience*.
- Hazin, I., Fernandes, I., Gomes, E., & Garcia, D. (2019). Neuropsicologia no Brasil: passado, presente e futuro. *Estudos e Pesquisas Em Psicologia*. <https://doi.org/10.12957/epp.2018.42228>
- Hebben, N., & Milberg, W. (2009). Essentials of Neuropsychological Assessment. *Neuropsychological Rehabilitation*. <https://doi.org/10.1017/CBO9781107415324.004>
- Horton Jr., A. M., & Wedding, D. (2008). The Neuropsychology Handbook. In *Review Literature And Arts Of The Americas*. <https://doi.org/10.1007/s00122-002-0929-x>
- Howell, D. C. (2011). *Fundamental Statistics for the Behavioral Sciences*. Wadsworth Cengage Learning.
- Hyndman, R. J., & Fan, Y. (1996). Sample Quantiles in Statistical Packages. *The American Statistician*, 50(4), 361. <https://doi.org/10.2307/2684934>
- Kozioł, L. F., Beljan, P., Bree, K., Mather, J., & Barker, L. (2016). Large-scale brain systems and neuropsychological testing: An effort to move forward. In *Large-Scale Brain Systems and Neuropsychological Testing: An Effort to Move Forward*. <https://doi.org/10.1007/978-3-319-28222-0>
- Krus, D. J., & Krus, P. H. (1977). Lost: McCall's T Scores: Why? *Educational and Psychological Measurement*, 37(1), 257-261. <https://doi.org/10.1177/001316447703700134>
- Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. *Korean Journal of Anesthesiology*, 70(2), 144. <https://doi.org/10.4097/kjae.2017.70.2.144>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment (5th ed.)*. Oxford University Press.
- Mäder, M. J. (1996). Avaliação neuropsicológica: aspectos históricos e situação atual. *Psicologia: Ciência e Profissão*, 16(3), 12-18. <https://doi.org/10.1590/S1414-98931996000300003>
- Mansur-Alves, M., Silva, R. S., & Fernandes, S. C. de Á. (2016). Impact of the Psychological Testing Assessment System (SATEPSI) for Scientific Publications in Psychological Assessment. *Psico-USF*, 21(1), 179-188. <https://doi.org/10.1590/1413-82712016210115>
- Martinelli, J. E., Cecato, J. F., Bartholomeu, D., & Montiel, J. M. (2014). Comparison of the Diagnostic Accuracy of Neuropsychological Tests in Differentiating Alzheimer's Disease from Mild Cognitive Impairment: Can the Montreal Cognitive Assessment Be Better than the Cambridge Cognitive Examination. *Dementia and Geriatric Cognitive Disorders Extra*, 4(2), 113-121. <https://doi.org/10.1159/000360279>
- McCall, W. A. (1922). How to Measure in Education. *Nature*, 110(2766), 601-601. <https://doi.org/10.1038/110601a0>
- Micceri, T. (1989). The Unicorn, The Normal Curve, and Other Improbable Creatures. *Psychological Bulletin*. <https://doi.org/10.1037/0033-2909.105.1.156>
- Rueda, F. J. M. (2013). Bateria Psicológica para Avaliação da Atenção - BPA. São Paulo: Vetor.
- Russell, E. (2012). The Scientific Foundation of Neuropsychological Assessment. In *The Scientific Foundation of Neuropsychological Assessment*. <https://doi.org/10.1016/C2011-0-04279-5>
- Sampaio, M. H. de L., & Nakano, T. de C. (2011). Avaliação psicológica no contexto do trânsito: revisão de pesquisas brasileiras. *Psicol. Teor. Prát.*
- Santarosa, L.M.C., Wwinstein, O., Prado, Z.R. (2011). Manual de Teste H.T.M.. São Paulo: Vetor.
- Schoenberg, M. R., Osborn, K. E., Mahone, E. M., Feigon, M., Roth, R. M., & Pliskin, N. H. (2018). Physician Preferences to Communicate Neuropsychological Results: Comparison of Qualitative Descriptors and a Proposal to Reduce Communication Errors. *Archives of Clinical Neuropsychology*, 33(5), 631-643. <https://doi.org/10.1093/arclin/acx106>
- Schoenberg, M. R., & Scott, J. G. (2014). The Little Black Book of Neuropsychology. A Syndrome-Based Approach. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Sedó, M., Paula, J. J. de, & Malloy-Diniz, L. F. (2015). *FDT - Teste Dos Cinco Dígitos - Manual* (Hogrefe Cetep Ed.).
- Silva, F. H. V. de C. e. (2012). A Psicologia do trânsito e os 50 anos de profissão no Brasil. *Psicologia: Ciência e Profissão*, 32(spe), 176-193. <https://doi.org/10.1590/S1414-98932012000500013>
- Spren, O., & Strauss, E. (1998). A Compendium of Neuropsychological Tests. In *Administration Norms And Commentary* (p. 1216). <https://doi.org/10.3390/nu3080735>
- Stahel, W., & Wilcox, R. R. (1998). Introduction to Robust Estimation and Hypothesis Testing. *Journal of the American Statistical Association*, 93(443), 1240. <https://doi.org/10.2307/2669876>

- Stigler, S. M. (1992). A Historical View of Statistical Concepts in Psychology and Educational Research. *American Journal of Education*. <https://doi.org/10.1086/444032>
- Swinson, J. (2013). British Ability Scales 3. *Educational Psychology in Practice*. <https://doi.org/10.1080/02667363.2013.853380>
- Wechsler, D. (2008). WAIS-IV: Wechsler adult intelligence scale. San Antonio, TX: Psychological Corporation. [https://doi.org/10.1016/S0043-1354\(02\)00019-2](https://doi.org/10.1016/S0043-1354(02)00019-2)
- Willis, J. O. (1998). Guide to identification of learning disabilities. Copley Custom Printing Group.
- Witz, K., Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1990). Applied Statistics for the Behavioral Sciences. *Journal of Educational Statistics*. <https://doi.org/10.2307/1164825>

Data de submissão: 28/01/2020  
Primeira decisão editorial: 22/07/2020  
Aceite em: 06/10/2020