

## Uso de Medidas Quantitativas de Eficácia em Reabilitação Neuropsicológica

### *Use of Effectiveness Quantitative Measures to Neuropsychological Rehabilitation*

Thiago da Silva Gusmão Cardoso<sup>I</sup>

Claudia Berlim de Mello<sup>II</sup>

Patrícia Martins de Freitas<sup>III</sup>

#### Resumo

O objetivo deste estudo foi oferecer orientações quanto ao uso de medidas quantitativas para verificar a eficácia da reabilitação neuropsicológica, apresentando os pressupostos para emprego de medidas de magnitude de efeito. O método utilizado foi a revisão bibliográfica, baseada na escolha arbitrária de artigos em periódicos científicos e livros que tratem do tema. A partir desta revisão, recomendamos o uso de medidas de magnitude de efeito na mensuração da eficácia de intervenções em reabilitação neuropsicológica, com especial atenção para o uso do  $d$  de Cohen e delta de Cliff, bem como a descrição do poder estatístico dos testes utilizados e seus respectivos intervalos de confiança.

**Palavras-chave:** Avaliação da eficácia terapêutica; medidas estatísticas; reabilitação neuropsicológica.

#### Abstract

The aim of this study was to offer guidelines on the use of quantitative measures to verify the neuropsychological rehabilitation effectiveness, presenting the assumptions for the use of effect size measures. The method was literature review, based on the arbitrary choice of articles in scientific journals and books that addressed the topic. From this review, we recommend the use of effect size measures in the mensuration of the treatment effectiveness in neuropsychological rehabilitation, with special attention to Cohen's  $d$  and Cliff's delta usage, as well as the description of the statistical power of the applied tests and their respective confidence intervals.

**Keywords:** Treatment effectiveness evaluation; statistical measurement; neuropsychological rehabilitation.

<sup>I</sup>Escola de Filosofia, Letras e Ciências Humanas da Universidade Federal de São Paulo (São Paulo), Brasil

<sup>II</sup>Centro Paulista de Neuropsicologia da Universidade Federal de São Paulo (São Paulo), Brasil

<sup>III</sup>Instituto Multidisciplinar de Saúde da Universidade Federal da Bahia (Vitória da Conquista), Brasil

A preocupação constante com a busca de medidas quantitativas para a avaliação dos programas de reabilitação neuropsicológica deve-se a inter-relação de várias tendências (Wilson, 1991; 2003). Dentre essas tendências, Wilson (1991) destaca: (a) dificuldades em desenvolver estratégias padronizadas de intervenção para cada distúrbio neuropsicológico; (b) diversidade da população clínica, os diferentes tipos de lesões e as características particulares dos déficits observados fazem com que existam dificuldades na avaliação dos programas de reabilitação; (c) a criação de estratégias em reabilitação nem sempre estão fundamentadas no conhecimento teórico que se tem sobre determinada função ou distúrbio; (d) as estratégias bem fundamentadas teoricamente nem sempre apresentam subsídios de estudos empíricos quanto a sua efetividade.

Entre os fatores que contribuíram para a intensificação dessa tendência à quantificação da eficácia das intervenções em reabilitação neuropsicológica pode-se citar: 1) a necessidade de organização dos serviços de saúde com foco na

reabilitação de pessoas com transtornos neurológicos e psiquiátricos; 2) a racionalização dos gastos públicos com procedimentos que se revelem mais eficazes; 3) a preocupação quanto à generalização dos resultados de pesquisas para a prática de reabilitação; 4) a análise mais profunda dos processos envolvidos, através da aplicação de técnicas estatísticas confiáveis.

A mensuração é, essencialmente, um processo descrito. Envolve geralmente a atribuição de um número para expressar em termos quantitativos o grau em que uma pessoa ou espécie possui determinada característica (Pasquali, 2004). A quantificação permite aumentar a precisão e a objetividade da descrição, de modo que tenha um significado padronizado, ou seja, um mesmo significado em diferentes momentos e para um grupo de pessoas (Pasquali, 2004). As medidas descritivas auxiliam a análise do padrão dos dados e quando devidamente refinadas por testes estatísticos permitem a confirmação ou a refutação de hipóteses sobre o fenômeno estudado.

A eficácia remete às condições controladas e aos resultados desejados de experimentos, com foco na

validade interna dos resultados (Lutz, 2003). Cientistas clínicos concentram-se na questão da eficácia dos procedimentos clínicos, através de experimentos que visam responder se uma nova intervenção produz melhores resultados do que uma intervenção tradicional ou uma condição de controle (Howard, Moras, Brill, Martinovich & Lutz, 1996, Loschiavo-Alvares et al., 2011). Desta forma, a avaliação da eficácia de uma intervenção perpassa a questão do desenho metodológico dos estudos, verificando a qualidade da pesquisa a partir de vários critérios, tais como: a randomização da amostra; a presença de grupo controle; a aplicação de instrumentos válidos e confiáveis; avaliações cegas; mascaramentos; técnicas estatísticas utilizadas; resultados compatíveis com os objetivos; protocolo de seguimento, entre outros (Seligman, 1995).

A neuropsicologia utiliza-se de uma série de instrumentos de medidas e procedimentos: testes psicológicos, escalas, desenho, tarefas de desempenho, observação, entrevista e paradigmas experimentais. O seu objetivo enquanto ciência é investigar a relação entre o cérebro e o comportamento humano em condições normais ou patológicas (Teeter & Clikeman, 1997). Na prática clínica, o exame neuropsicológico investiga o funcionamento cognitivo e comportamental, o que permite inferir sobre o funcionamento cerebral, compreendendo as disfunções cognitivas e comportamentais decorrentes de condições neurológicas e psiquiátricas específicas e seu impacto na vida das pessoas (Teeter & Clikeman, 1997). Enquanto que a reabilitação, seja de orientação cognitiva ou neuropsicológica, pretende minimizar os efeitos de déficits cognitivos e alterações de comportamento que se constituem como obstáculos ao desempenho adequado em tarefas do cotidiano, bem como estimular a aquisição de novas habilidades e a adaptação às perdas permanentes, por meio do treino de estratégias compensatórias (Solberg & Mateer, 2009).

Na abordagem cognitiva, o foco da reabilitação está no prejuízo neuropsicológico e na restauração da função cognitiva por meio de treinos específicos, ao passo que a reabilitação neuropsicológica é orientada a metas funcionais significativas, aplicadas às diversas áreas de vida, buscando a integração da pessoa com dano cerebral ao seu ambiente da forma mais autônomo possível (Wilson, 2011; Prigatano, 1997). Independentemente da abordagem, o desenvolvimento e manejo dos processos pelo clínico demanda amplo conhecimento sobre modelos em

cognição, emoção e aprendizagem, entre outros. Conforme Wilson (2002) enfatiza, ao discorrer sobre o modelo compreensivo da reabilitação cognitiva, é necessário combinar metodologia científica, teoria e relevância clínica neste contexto.

Diversas pesquisas têm sido realizadas com o objetivo de demonstrar a adequação dos instrumentos neuropsicológicos, assim como a eficácia de seus programas de intervenção (Feinberg & Farah, 1997; Teeter & Clikeman, 1997). A intervenção tem sido denominada de reabilitação, pois historicamente os procedimentos de intervenção na neuropsicologia surgiram no intuito de recuperar funções intactas e que foram danificadas por inúmeros fatores, como por exemplo, em consequência de um acidente vascular encefálico.

A reabilitação é a área da saúde que trata da independência funcional, da qualidade de vida e da adaptação psicossocial (Solberg & Mateer, 2009). A reabilitação neuropsicológica é a área focada na elaboração de procedimentos preventivos e adaptativos para indivíduos que apresentem comprometimentos cognitivos e comportamentais que tenham impactos para o desempenho das funções diárias (Santos, 2006; Solberg & Mateer, 2009).

Existem dois modelos teóricos principais de reabilitação neuropsicológica: a restituição funcional e a compensação (Solberg & Mateer, 2009). A abordagem de restituição funcional se baseia no conceito de neuroplasticidade e o objetivo é o restabelecimento ou desenvolvimento do funcionamento cognitivo. O modelo de compensação utiliza-se de estratégias cognitivas e comportamentais visando suprir de maneiras alternativas os papéis funcionais dos sistemas neuronais lesionados. As técnicas de reabilitação podem ser efetivas do ponto de vista da recuperação funcional em crianças e adultos (Ginarte-Arias, 2002; Kessels & Haan, 2003).

A literatura tem demonstrado resultados consistentes nos estudos de reabilitação funcional para alguns transtornos de aprendizagem em crianças (Flecher, 2009; Fadini & Capellini, 2011, Fukuda & Capellini, 2012). No estudo de Schneider, Roth e Ennemoser (2000), os resultados da reabilitação de crianças com Dislexia demonstram que através da aplicação de técnicas comportamentais associadas com tarefas de desenvolvimento fonológico os ganhos para as habilidades de leitura são significativos. O estudo de Fukuda e Capellini (2012) teve por objetivo verificar a eficácia de um programa de treinamento

fonológico e habilidades de correspondência grafema-fonema aplicado a 60 escolares com sinais de risco para dislexia, os resultados do estudo, analisados a partir do teste de Wilcoxon, indicaram a melhora significativa nas habilidades fonológicas dos participantes, superando, em algumas provas, a média de desempenho dos escolares não submetidos ao programa.

Um exemplo extremamente bem sucedido de recuperação funcional é o treinamento da discriminação de sequências temporais rápidas que promove o desenvolvimento da linguagem em algumas crianças com transtorno específico do desenvolvimento da linguagem oral (Flecher, 2009).

Outros estudos desenvolvidos para demonstrar a eficácia de programas de intervenção em crianças com déficits linguísticos, envolvendo a compreensão e produção oral e textual apresentam resultados bem sucedidos aplicando técnicas cognitivo-comportamentais associadas com atividades estruturadas no princípio de contar e recontar histórias. O foco do treinamento tem sido na estimulação do desenvolvimento de redes conceituais e consciência fonológica (Fonseca & Parente, 2007, Hein, Macedo, Hyppolito, Guadanin & Lukasova, 2007; Spinillo & Brandão, 2001). Os resultados para a reabilitação de crianças com alterações comportamentais são ainda mais expressivos (Ylvisaker & Feney, 1998; Muñoz-Céspedes & Tirapu-Ustároz, 2004; Fabiano et al., 2009). Todavia, os estudos citados apenas utilizaram testes de hipóteses baseados no valor de  $p$  ao analisar a eficácia dos seus programas de intervenção o que inviabiliza a análise da magnitude com essas intervenções promovem mudanças no funcionamento cognitivo e comportamental dos participantes.

No caso da Discalculia do Desenvolvimento, outro transtorno específico da aprendizagem que interfere diretamente no processamento matemático, os resultados de pesquisas com procedimentos de reabilitação são, em alguma medida, inconsistentes (Kaufmann & Aster, 2012). A maioria dos estudos empíricos sobre as intervenções para melhorar a capacidade de cálculo foram realizados em países de língua inglesa por profissionais da educação especial e com crianças com dificuldades na matemática, não necessariamente com diagnóstico de Discalculia do Desenvolvimento (Kaufmann & Aster, 2012). Uma meta-análise de Kroesbergen e Van Luit (2003), envolvendo 58 estudos de intervenção primária para escolares, revelou que as

intervenções para promover habilidades numéricas básicas foram eficazes, com tamanhos de efeitos variando de  $-0,44$  para acima de  $3$ .

Diversos estudos de prevenção e intervenção precoce para dificuldades na matemática apresentaram bons resultados com base no relato de valores de  $p$  estatisticamente significativos (Fuchs et al., 2005; Fuchs, Compton, Fuchs & Davis, 2008). Fuchs et al. (2005) identificaram 169 alunos em 41 salas de aula como apresentando risco para dificuldades de matemática com base em seus desempenhos acadêmicos. Estas crianças foram aleatoriamente alocadas em um grupo controle ou em grupos pequenos de intervenção, sendo que estes últimos receberiam tutoria três vezes por semana durante 16 semanas. Os resultados mostraram que o desenvolvimento da matemática foi significativamente superior para os grupos que receberam tutoria em relação ao grupo controle. Além disso, a incidência de alunos com dificuldades de matemática foi substancialmente reduzida e esta redução se manteve na série escolar subsequente (Fuchs et al., 2008). Novamente, os estudos citados se restringiram ao uso de testes de hipóteses baseados no valores de  $p$  ao analisarem a eficácia dos seus programas de intervenção.

Em relação à reabilitação neuropsicológica de adultos com diferentes condições neurológicas, a eficácia de programas de intervenção com treinamento cognitivo tem sido um tema bastante polêmico e controverso (Cicerone et al., 2000). Três estudos de revisão e meta-análise indicam a eficácia de técnicas de reabilitação cognitiva para pacientes com comprometimento cognitivo leve, amnésia e traumatismo cranioencefálico (Li et al., 2011; Kessels & Haan, 2003; Cicerone et al., 2000). Enquanto que outro estudo de meta-análise foi mais modesto nas suas conclusões, descrevendo apenas tamanhos de efeito positivo e significativos (média do  $g$  de Hedge= $0,41$ , IC95%  $0,31-0,50$ ,  $p<0,05$ ) para o uso de estratégias de instruções metacognitivas na intervenção para funções executivas após traumatismo cranioencefálico em adultos, porém com a ressalva de que as provas são insuficientes para fazer recomendações clínicas de uso dessas estratégias até o momento (Kennedy et al., 2008).

Na revisão de Rosti-Otajärvi e Hämäläinen (2011) foram encontrados tamanhos de efeito pequenos para o treinamento cognitivo em adultos com esclerose múltipla, com diferença média padronizada de  $0,54$  para o span de memória (IC95%  $0,2-0,88$ ,  $p=0,002$ ), de  $0,33$  para memória

de trabalho (IC95% 0,09–0,57,  $p=0,006$ ), e de 0,33 para memória de curto prazo visual (IC95% 0,04–0,6,  $p=0,02$ ). No mesmo caminho, o estudo de revisão sistemática de Clare e Woods (2003) que avaliou a eficácia e o impacto do treinamento cognitivo e reabilitação cognitiva na melhoria do funcionamento da memória episódica verbal para pessoas nos estágios iniciais da doença de Alzheimer e demência vascular, relatou tamanhos de efeito insignificantes para memória verbal de curto prazo (tamanho de efeito de 0,07, IC95% -0.26–0.41) e memória verbal de longo prazo (tamanho de efeito de -0.05, IC95% -3.47–3.38). Os autores não encontraram evidências que apoiem a eficácia desta modalidade de tratamento e destacaram a ausência de ensaios clínicos randomizados sobre tratamentos individuais em reabilitação cognitiva.

Esses exemplos servem para demonstrar o potencial e os desafios da reabilitação neuropsicológica infantil e no adulto, chamando a atenção para o crescente número de pesquisas sobre o tema e a necessidade de mais estudos bem planejados com foco na demonstração da eficácia desses procedimentos (Ylvisaker, Hanks & Johnson-Green, 2003). Portanto, a preocupação quanto à metodologia adotada por esses estudos no que se refere aos parâmetros estatísticos escolhidos para avaliar e descrever os resultados levaram ao desenvolvimento do presente estudo.

O estudo em questão tratou de revisar os procedimentos estatísticos em relação a medidas de eficácia em reabilitação neuropsicológica, principalmente as técnicas estatísticas recomendadas pela literatura para a descrição e avaliação dos resultados de pesquisa na área, considerando os pressupostos para emprego de cada uma dessas estatísticas, suas vantagens, limites e contextos de aplicação. Nesse sentido, recomendamos o uso de métodos de magnitude de efeito, intervalos de confiança e poder do teste.

Este artigo está organizado da seguinte forma. A primeira seção define o que são medidas de magnitude de efeito e discute a necessidade do relato dessas medidas de forma complementar ao emprego de testes de hipótese. A segunda seção descreve as medidas de efeito do tipo *d*, os pressupostos para emprego, cálculo e interpretação dessas medidas. A terceira centra-se nos pressupostos para emprego, cálculo e interpretação de uma medida de efeito não paramétrica, o delta de Cliff. Por último, são oferecidos alguns comentários finais sobre o tema.

## Magnitude de Efeito

A utilização de medidas de efeito é recomendada pela American Psychological Association (APA), que considera a descrição de medidas de magnitudes de efeito como fundamental na grande maioria dos estudos em psicologia (APA, 2001). Entretanto, pouca informação sobre esse tema está disponível na maioria dos livros didáticos de estatística, os quais costumam dedicar pouca atenção às orientações quanto à escolha, uso e interpretação de medidas de magnitude de efeito ou simplesmente medidas de efeito (*Effect Sizes*) (Capraro & Capraro, 2002). Como resultado, muitos pesquisadores não se sentem seguros ou mesmo reconhecem a importância da incorporação de medidas de efeito em seus trabalhos e muitos dos relatórios de pesquisa em psicologia carecem do emprego destas estatísticas (Ward, 2002; Volker, 2006).

Cohen (1988) foi um dos primeiros teóricos a defender a utilização dessas medidas como alternativa aos testes baseados no *valor p*, que embora informem sobre se um resultado ou diferença é estatisticamente significativo, ou seja, que as diferenças observadas são superiores aqueles que se esperaria encontrar por mero acaso, pouco ou nada nos dizem da dimensão desta diferença.

O uso de testes de hipótese, embora, tenha sua importância costumam fornecer apenas informação limitada e dicotômica do tipo: rejeita-se a hipótese nula ou falha-se em rejeitá-la; existe ou não existe associação entre os fatores (Conboy, 2003). Enquanto que as medidas de magnitude de efeito podem informar sobre a direção dessas diferenças e o grau em que diferem, bem como a força e porcentagem de variação explicativa entre os fatores.

Um dos principais problemas dos testes de hipótese que as medidas de magnitude de efeito podem ajudar a resolver se refere ao uso do valor de *p*, pois ele apresenta fortes relações com o tamanho da amostra (Dancey & Reidy, 2006). Por exemplo, ao compararmos dois estudos de intervenção neuropsicológica, podemos observar que o primeiro apresenta um valor de *p* muito pequeno (p. ex., 0,001) e o segundo um valor grande (p. ex., 0,21), entretanto, somente com os valores de *p* em mãos não podemos dizer que o primeiro estudo mostra um grande efeito, e o segundo, um pequeno. Talvez o primeiro estudo tenha utilizado uma amostra grande, enquanto que o segundo uma amostra pequena, além disso, afirmações sobre o efeito com base nesses

dados continuariam mascaradas pelo tamanho da amostra, pois efeitos pequenos em grandes amostras podem apresentar significância estatística. A melhor abordagem para enfrentar esses problemas é o recurso das medidas de magnitude do efeito, pois estas se mostram na sua grande maioria pouco influenciáveis pelo tamanho da amostra (Maroco, 2010).

Embora as medidas de magnitude de efeito via de regra não dependam do tamanho amostral, o seu uso é circunscrito a satisfação de alguns pressupostos que apresentaremos a fim de que haja uma utilização adequada por conta dos pesquisadores interessados no assunto. Primeiro, existem numerosas medidas de magnitude de efeito ou famílias de medidas de efeito como alguns teóricos preferem referir. Os dois principais tipos de família de medidas de efeito são a família *r* e a família *d* (Morgan, Leech, Gloeckner & Barrett, 2004).

A família de medidas de efeito *r* são métodos para expressar medidas de efeito em termos da força de associação, a mais conhecida dessas abordagens é o coeficiente de correlação de Pearson, *r* (Morgan et al., 2004). Esta família inclui muitas outras estatísticas de associação tais como rho ( $\rho$ ), phi ( $\varphi$ ), eta ( $\eta$ ), e uma medida de correlação múltipla (*R*). Entretanto, não trataremos neste artigo da família *r*, por fugir ao objetivo que é apresentar medidas de efeito para diferenças entre grupos e não de associação.

### Medidas de Efeito *d*

A família de medidas de efeito *d* são métodos que focam a magnitude das diferenças em vez da força da associação. São especialmente indicados para medir a diferença padronizada entre dois grupos, por exemplo, um grupo experimental e controle (Morgan et al., 2004). Na avaliação da eficácia da reabilitação neuropsicológica são medidas que possibilitam comparar a diferenças nos resultados de testes psicológicos para um mesmo grupo no pré e pós-teste ou ainda para o grupo experimental e controle. O modelo proposto por Cohen (1988) é sem dúvida o mais influente, geralmente utilizado para medir o efeito para diferenças de médias de dois grupos e é dado por  $d = (X_1 - X_2) / s'$ , sendo  $s' = \sqrt{(s_1^2 + s_2^2) / 2}$ , onde  $s_1$  representa o desvio padrão do primeiro grupo e  $s_2$  o desvio padrão do segundo grupo. Nota-se que a medida de magnitude de efeito *d* é uma medida da diferença entre médias em termos de unidades de desvio-padrão, sendo que a melhor estimativa para o

parâmetro populacional é medido através do desvio-padrão combinado dos dois grupos, presente no denominador da fórmula (Cohen, 1988).

A fórmula do *d* de Cohen assume o pressuposto de que os *N*s dos dois grupos são iguais, existindo assim homogeneidade das variâncias e garantindo que o desvio padrão seja a média aritmética dos desvios padrões das amostras. Nos casos em que não seja possível dispor de grupos de mesmo tamanho amostral o *g* de Hedges é a versão mais indicado e acurada para o cálculo do *d* de Cohen e é dada pela fórmula da Figura 1. Observe que a fórmula possui duas partes, a primeira que é o *g* de Hedges, propriamente dito, e a segunda localizada após o sinal de multiplicação que é o ajustamento para pequenas amostras. Todavia, ressaltados que embora possamos ajustar a fórmula para cálculo do tamanho do efeito, a família *d* não depende do tamanho amostral.

$$\hat{g} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{(n_{total} - 2)}}} \times 1 - \left( \frac{3}{4(n_1 - n_2) - 9} \right)$$

**Figura 1.** Fórmula do *g* de Hedges

Fonte: Conboy (2003)

Uma das recomendações dos estatísticos para garantir a qualidade na descrição dos resultados de pesquisa é o relato dos intervalos de confiança, sejam estes para médias, variâncias e proporções (Dancey & Reidy, 2006). No caso, das medidas de magnitude do efeito consideramos também fundamental o cálculo de intervalos de confiança. Como o próprio nome já diz os intervalos de confiança fornecem ao pesquisador e aos futuros leitores do estudo maior confiança para analisar os resultados. Neste sentido, o manual da APA 6ª Edição recomenda que além de medidas de efeito, sejam relatados nos estudos intervalos de confiança e poder estatístico (APA, 2009).

O intervalo de confiança é a escala de certeza ou área de confiança para os dados obtidos, fornecendo os limites superiores e inferiores para aceitarmos a estatística em função do erro padrão e nível de significância (Dancey & Reidy, 2006). No caso, do *d* de Cohen (1988) o intervalo de confiança (IC) pode ser calculado através da expressão matemática  $IC = d \pm Z_{\alpha/2} \sigma[d]$ , sendo o  $\sigma[d]$  obtido pela fórmula da Figura 2.

$$\sigma[d] = \sqrt{\frac{N_E + N_C}{N_E \times N_C} + \frac{d^2}{2(N_E \times N_C)}}$$

**Figura 2.** Fórmula para o intervalo de confiança do *d* de Cohen

Fonte: Cohen (1988)

Onde  $N_E$  e  $N_C$  são o tamanho da amostra para o grupo experimental e controle, e  $d$ , o valor da magnitude de efeito.

A interpretação quanto aos valores variam conforme a medida de efeito utilizada, no caso do  $d$  de Cohen (1988) é possível se obter medidas que variam de zero a +1 ou -1, ou até mesmo valores superiores ou inferiores a esses números. Os resultados negativos da magnitude do efeito  $d$  indicam desempenho inferior para o grupo 1 da fórmula. Para uma interpretação da força do efeito nas famílias pode-se visualizar a Tabela 1.

O segundo pressuposto que devemos assumir ao escolhermos a medida de efeito que utilizaremos é se a amostra obedece ou não a distribuição normal. A família de medidas de efeito  $d$  se baseia em parâmetros da distribuição normal para calcular as diferenças padronizadas (Ledesma, Macbeth & Kohan, 2008). Todavia, como na área de pesquisa em reabilitação neuropsicológica é difícil obtermos grandes amostras, especialmente para grupos com doenças raras e distúrbios neurológicos, as amostras obtidas geralmente são pequenas e violam os pressupostos da distribuição normal. Assim, o  $d$  de Cohen (1992) nem sempre é a estatística mais adequada para utilizarmos para avaliação da eficácia de intervenções em reabilitação neuropsicológica. Nestes casos, em que as amostras não apresentam distribuição normal os estatísticos indicam a utilização de outra medida de magnitude de efeito bem menos conhecida, o delta de Cliff ( $\delta$ ) (Cliff, 1993).

**Tabela 1.** Interpretação da força da relação ou da magnitude do efeito

Interpretação	Família <i>d</i>		Família <i>r</i>	
	<i>d</i>	<i>r e φ</i>	<i>R</i>	<i>η</i>
Muito Grande	≥1,00	≥0,70	0,70	0,45
Grande	0,80	0,50	0,51	0,35
Médio	0,50	0,30	0,36	0,24
Pequeno	0,20	0,10	0,14	0,10

Fonte: Landau e Everitt (2004)

## Delta de Cliff

O delta de Cliff ( $\delta$ ) é a medida não paramétrica que permite quantificar a magnitude da diferença entre dois grupos de observação que não atendem aos pressupostos da normalidade, sendo que a sua descrição permite complementar a interpretação do valor de  $p$  associado ao correspondente teste de hipóteses utilizado (MacBeth, Razumiejczyk & Ledesma, 2011). A recomendação quanto ao uso do delta de Cliff está ligada aos casos em que as distribuições são assimétricas e violam a homogeneidade das variâncias. O delta de Cliff não depende da média, mas sim de um conceito de dominância, ou seja, ele considera as propriedades ordinais ao invés das propriedades intervalares dos dados (Cliff, 1993).

O delta de Cliff é a mais direta e simples das medidas de efeito ordinal, mais poderosa e robusta do que o  $d$  de Cohen sob certas condições como quando as distribuições são marginais e enviesadas e, especialmente para a análise de escalas do tipo Likert. O  $\delta$  de Cliff pode ser calculado através da seguinte fórmula:

Nesta fórmula  $X_1$  e  $X_2$ , são escores respectivamente do grupo 1 e grupo 2, e  $n_1$  e  $n_2$  são o tamanho da amostra de cada grupo. O símbolo # indica contagem. Nesta medida de efeito estimasse a probabilidade que um valor selecionado de um dos grupos seja maior do que o valor selecionado do outro grupo, subtraindo a probabilidade reversa. Cliff (1993) argumenta que o delta de Cliff analisa o grau de sobreposição entre as duas distribuições, o que ele chama de medida de dominância. O delta varia de -1 a +1, podendo assumir qualquer valor para este intervalo. Um delta de -1 ou +1 indica uma ausência de sobreposição, enquanto que um delta de 0.0 indica que os grupos estão completamente sobrepostos (Kromrey & Hogarty, 1998). Além disso, (+1) indica que o grupo 1 é maior do que o grupo 2, e (-1) significa que o grupo 1 é menor do que o grupo 2, enquanto que (0) se traduz por grupo 1 igual a grupo 2. A estatística gera uma matriz de  $i$  linhas e  $j$  colunas, assumindo três valores possíveis +1, -1 e 0 (Macbeth, Razumiejczyk & Ledesma, 2011). Um exemplo de matriz de dominância pode ser visualizado a seguir (Tabela 2):

Cliff (1996) sugere uma variedade de métodos para inferências sobre o  $\delta$ , mas para o cálculo do intervalo de confiança, uma estimativa consistente da variância do delta,  $\sigma^2[\delta]$  é dada por (figura 4):

**Tabela 2.** *Matriz de Dominância*

	1	2	3	4	4	5	<i>Di</i>
1	0	-1	-1	-1	-1	-1	-0.833
1	0	-1	-1	-1	-1	-1	-0.833
2	1	0	-1	-1	-1	-1	-0.500
2	1	0	-1	-1	-1	-1	-0.500
2	1	0	-1	-1	-1	-1	-0.500
3	1	1	0	-1	-1	-1	-0.167
3	1	1	0	-1	-1	-1	-0.167
3	1	1	0	-1	-1	-1	-0.167
4	1	1	1	0	0	-1	0.333
5	1	1	1	1	1	0	0.833
<i>d.j.</i>	0.8	0.3	-0.3	-0.7	-0.7	-0.9	-0.250

Fonte: Cliff (1996)

$$\hat{\delta} = \frac{\#(x_{i1} > x_{j2}) - \#(x_{i1} > x_{j2})}{n_1 - n_2}$$

**Figura 3.** *Fórmula para o delta de Cliff*

Fonte: Cliff (1993)

$$S_{dij}^2 = \frac{\sum \sum (d_{ij} - \hat{\delta})^2}{(n_1 - 1) + (n_2 - 1)}$$

**Figura 5.** *Cálculo de  $S_{dij}^2$*

Fonte: Cliff (1996)

$$\hat{\sigma}_{dc}^2 = \frac{(n_2 - 1)S_{di}^2 + (n_1 - 1)S_{dj}^2 + S_{dij}^2}{n_1 n_2}$$

**Figura 4.** *Fórmula para a variância do  $\delta$  de Cliff*

Fonte: Cliff (1996)

$$\frac{\hat{\delta} - \hat{\delta}^3 \pm Z_{\alpha/2} \hat{\sigma} \sqrt{(1 - \hat{\delta}^2)^2 + Z^2_{\alpha/2} \hat{\sigma}_{dc}^2}}{1 - \hat{\delta}^2 + Z^2_{\alpha/2} \hat{\sigma}_{dc}^2}$$

**Figura 6.** *IC para o  $\delta$  de Cliff*

Fonte: Cliff (1996)

Note que na fórmula que  $d_i$  é o valor marginal da linha  $i$ ,  $d_j$  é o valor marginal da coluna  $j$ ,  $d_{ij}$  é o valor de elementos  $ij$  na matriz de dominância (ver Tabela 2). O  $S_{di}^2$  é igual ao somatório dos quadrados da diferença ( $d_i - \delta$ ), dividido por  $n_1 - 1$ ; O  $S_{dj}^2$  é igual ao somatório dos quadrados da diferença ( $d_j - \delta$ ), dividido por  $n_2 - 1$ ; e  $S_{dij}^2$  é obtido pela fórmula abaixo (Figura 5):

Os limites para o IC para o delta de Cliff podem ser calculados através da expressão matemática da Figura 6.

O delta de Cliff também pode ser calculado através de software estatístico de interface amigável como o Cliff's Delta Calculator (Macbeth, Razumiejczyk & Ledesma, 2011) ou através do Statistical Analysis Software (SAS) (Kromrey & Hogarty, 1998).

## Poder Estatístico

Ward (2002) examinou artigos em três importantes revistas de psicologia para avaliar a frequência com que são relatadas medidas de efeito e poder estatístico. A autora encontrou que apenas 7% dos estudos estimam ou discutem poder estatístico e 30% calculam medidas de magnitude de efeito e que frequentemente os estudos não têm poder estatístico suficiente para detectar medidas de efeito.

O poder estatístico pode ser entendido como a probabilidade que possui uma prova de obter resultados significativos, sendo comumente associado ao erro do tipo II (a probabilidade de aceitar erroneamente a hipótese nula). O cálculo do poder estatístico é uma função matemática que depende do tamanho amostral, nível de significância

assumido e tamanho do efeito. Quanto maior o tamanho do efeito, maior é o poder do teste e menor a probabilidade de um erro do tipo II (Cohen, 1992).

O valor do tamanho do efeito deve vir acompanhado do valor correspondente do poder estatístico para que os dados da pesquisa não sejam falseados, refletindo medidas de efeito que não poderiam ter sido obtidas. Para o cálculo do poder estatístico existem softwares estatísticos como o *Gpower 3.0.10*.

### Considerações Finais

Os cuidados quanto à análise estatística da eficácia da reabilitação neuropsicológica são necessários, pois o que se procura é a obtenção de dados confiáveis que venham confirmar, ou não, as hipóteses formuladas nos trabalhos e subsidiar a leitura crítica dos resultados. Um conjunto de resultados ou observações duvidosos não permite uma transposição e generalização dos resultados de pesquisas científicas para a prática dos profissionais que trabalham com reabilitação neuropsicológica e que geralmente buscam nos artigos da área contribuições para o seu próprio trabalho. A aplicação de medidas de magnitude de efeitos como métodos estatísticos de eficácia em reabilitação neuropsicológica depende da maneira pela qual as observações ou os dados foram obtidos, ou seja, do delineamento do experimento e do tamanho amostral.

O conhecimento necessário para o cálculo de medidas de magnitude de efeito não é a única questão premente na área da reabilitação neuropsicológica para garantir o adequado relato e interpretação dos resultados de pesquisas. Os produtores e consumidores dos resultados de pesquisa na área se deparam com uma questão fundamental em torno do uso dessas medidas, a saber, qual significado atribuir aos tamanhos dos efeitos relatados. Um *d* de Cohen de 0,45 é simplesmente um número e seu significado e importância devem ser explicados pelo pesquisador e compreendidos pelos interessados nas pesquisas (Durlak, 2009). Segundo Hedges & Hedberg (2007) com base no que foi alcançado em muitos tipos diferentes de intervenções, os pesquisadores educacionais têm indicado que magnitudes em torno de 0,20 são de interesse prático até mesmo político quando se referem a medidas de desempenho acadêmico. O que estes dados revelam é que para a área da reabilitação

neuropsicológica, os próprios pesquisadores e profissionais precisam definir qual o valor prático ou clínico das medidas de magnitude de efeito, traçando padrões específicos para cada domínio avaliado com base no corpo teórico e empírico disponível na literatura, em vez de apenas adotarem as diretrizes arbitrárias presentes nos livros textos de estatística (Volker, 2006).

Este estudo focalizou o uso de métodos de magnitude de efeito especialmente úteis na área da psicologia como o *d* de Cohen (1988) e delta de Cliff (1999), seus respectivos intervalos de confiança e poder do teste, considerando os pressupostos para emprego de cada uma dessas estatísticas e as recomendações da APA para a sua utilização em artigos científicos. Não se pretendia com esse trabalho revisar sistematicamente toda a literatura das áreas da Estatística e/ou Psicologia relacionada ao empregado dessas medidas, mas apresentar aos pesquisadores informações que permitissem a segurança necessária para a utilização dessas medidas. Desta forma, buscou-se evitar a superficialidade de um artigo meramente descritivo sobre o tema bem como a profundidade dos procedimentos matemáticos técnicos de um artigo voltado para o público de ciências exatas.

### Referências

- American Psychological Association [APA]. (2001). *Publication manual of the American Psychological Association*. 5ª edição. Washington, DC: APA.
- APA. (2009). *Publication manual of the American Psychological Association*. 6ª edição. Washington, DC: APA.
- Capraro, R. M., & Capraro, M. (2002). Treatments of effect sizes and statistical significance tests in textbooks. *Education and Psychological Measurement*, 62, 771-782.
- Cicerone, K. D., Dahlberg, C., Kalmar, K., Langenbahn, D. M., Malec, J. F., Berquist, T. F., Felicetti, T., Giacino, J. T., Harley, J. P., Harrington, D. E., Herzog, J., Kneipp, S., Laatsch, L., & Morse, P.A. (2000). Evidence-based cognitive rehabilitation: Recommendation for clinical practice. *Archives of Physical Medicine and Rehabilitation*, 81, 1596-1615.
- Clare, L., & Woods, R. T (2003). Cognitive rehabilitation and cognitive training for early-stage Alzheimer's disease and vascular dementia.

- Cochrane Database Systematic Review*, (4), CD003260.
- Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, 114: 494-509 .
- Cliff, N. (1996). *Ordinal Methods for Behavioral Data Analysis*. New Jersey: Lawrence Erlbaum Associates.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1), 155-159.
- Conboy, J. E. (2003). Algumas medidas típicas univariadas da magnitude do efeito. *Análise Psicológica*, 21(2), 145-158.
- Dancey, C. P., & Reidy, J. (2006). *Estatística sem matemática para psicologia: usando SPSS para Windows*. Porto Alegre: Artmed.
- Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal Pediatric Psychology*, 34: 917-928.
- Fabiano, G. A., Pelham, W. E., Coles, E. K., Gnagy, E. M., Chronis, A. M., & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, 29, 129-140.
- Fadini, C. C., & Capellini, S. A. (2011). Eficácia do treinamento de habilidades fonológicas em crianças de risco para dislexia. *Revista CEFAC*, 13(5), 856-865.
- Feinberg, T. E., & Farah, M. J. (1997). *Behavioral neurology and neuropsychology*. New York: McGraw-Hill.
- Flecher, J. M., Lyons, G. R., Fuchs, L. S., & Barnes, M. A. (2009). *Transtornos de aprendizagem: da identificação à intervenção*. Porto Alegre: Artmed.
- Fonseca, R. P., & Parente, M. A. (2007). Avaliação e reabilitação neuropsicológicas da atenção aos aspectos funcionais da linguagem: O caso da lesão do hemisfério direito. In A. L. Sennyey, L. I. Z. Mendonça, B. B. G. Schlecht, E. F. Santos, E. C. Macedo, *Neuropsicologia e inclusão: Tecnologias em (re)habilitação cognitiva*. São Paulo: Artes Médicas, pp. 219-227
- Fuchs, D., Compton, D. L., Fuchs, L. S., & Davis, G. C. (2008). Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal study at the National Research Center on Learning Disabilities. *Reading and Writing: A Contemporary Journal*, 21:413-436.
- Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, 97,493-513.
- Fukuda, M. T. M., & Capellini, S. A. (2012). Programa de intervenção fonológica associado à correspondência grafema-fonema em escolares de risco para a dislexia. *Psicologia Reflexão Crítica, Crit.*, 25(4), 783-790.
- Ginarte-Arias, Y. (2002). Rehabilitación cognitiva. Aspectos teóricos y metodológicos. *Revista de Neurologia*, 34(9), 870-876.
- Grissom, R. J., & Kim, J. J. (2005). *Effect Sizes for Research: A Broad Practical Approach*. New Jersey: Lawrence Erlbaum Associates.
- Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29, 60-87.
- Hein, J. M., Macedo, E. C., Hyppolito, C. R. S., Guadanin, F., & Lukasova, K. (2007). Leitura e consciência fonológica na deficiência mental: Intervenções com o programa alfabetização fônica computadorizada. In A. L. Sennyey, L. I. Z. Mendonça, B. B. G. Schlecht, E. F. Santos, E. C. Macedo, *Neuropsicologia e inclusão: Tecnologias em (re)habilitação cognitiva*. São Paulo: Artes Médicas, pp. 247-256.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). (2012) The evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, 51, 1059-1064.
- Kaufmann, L., & Aster, M. V. The diagnosis and management of dyscalculia. *Deutsches International Arzteblatt* 109(45), 767-778.
- Kennedy, M. R., Coelho, C., Turkstra, L., Ylvisaker, M., Moore, Sohlberg, M., Yorkston, K., Chiou, H. H., & Kan, P. F. (2008). Intervention for executive functions after traumatic brain injury: A systematic review, meta-analysis and clinical recommendations. *Neuropsychological Rehabilitation*, 18(3), 257-99.
- Kessels, R. P., & De Haan, E. H. (2003). Implicit learning in memory rehabilitation: A meta-analysis on errorless learning and vanishing cues methods. *J Clin Exp Neuropsychol*, 25(6), 805-14.

- Kroesbergen E., Van Luit, J. E. H. (2003). Mathematics Intervention for Children with Special Educational Needs. *Remedial and Special Education, 24*, 97-114.
- Kromrey, J. D., & Hogarty, K. Y. (1998). Analysis options for testing group differences on ordered categorical variables: an empirical investigation of type I error control and statistical power. *Multiple Linear Regression Viewpoints, 25*, 70-82.
- Landau, S., & Everitt, B. S. (2004). *A Handbook of Statistical Analyses using SPSS*. Chapman, & Hall: Washington, D. C.
- Ledesma, R., Macbeth, G., & Kohan N. C. (2008). Tamaño del efecto: Revisión teórica y aplicación con el sistema estadístico vista. *Revista Latinoamericana de Psicología, 40*(3), 425-439.
- Li, H., Li, J., Li, N., Li B., Wang, P., & Zhou, T. (2011). Cognitive intervention for persons with mild cognitive impairment: A meta-analysis. *Ageing Research Reviews, 10*(2), 285-96.
- Loschiavo-Alvares, F. Q., Sediya, C. Y. N., Rivero, T. S., Nicolato, R., Neves, F., Bueno, O. F. A., Corrêa, H., & Malloy-Diniz, L. F. (2011). Tools of efficacy's assessment of neuropsychological rehabilitation programs: A systematic review. *Clinical Neuropsychiatry, 8* (3), 1-8.
- Lutz, W. (2003). Efficacy, effectiveness, and expected treatment response in psychotherapy. *Journal of Clinical Psychology, 59* (7), 745-750.
- Macbeth, G., Razumiejczyk, E., & Ledesma, R. D. (2011). Cliff's Delta Calculator: A non-parametric effect size program for two groups of observations. *Universitas Psychologica, 10*(2), 545-555.
- Maroco, J. (2010). *Análise estatística com utilização do SPSS*. 3ª Ed. Lisboa: Sílabo.
- Morgan, G. A., Leech, N. L., Gloeckner, G. W., & Barrett, K. C. (2004). *SPSS for introductory statistics: use and interpretation*. 2ª Edition. LEA: New Jersey.
- Muñoz-Céspedes, J. M., & Tirapu-Ustároz, C. (2004). Rehabilitación de las funciones ejecutivas. *Revista de Neurología, 38*: 656-663.
- Pasquali, L. (2004). *Psicometria: Teoria dos testes na psicologia e na educação*. Petrópolis, RJ: Vozes.
- Prigatano, G. P. (1997). Learning from our Successes and Failures: Reflections and Comments on "Cognitive Rehabilitation: How it is and How it Might Be". *Journal of the International Neuropsychological Society, 3*, 497-499.
- Rosti-Otajärvi, E. M., & Hämäläinen, P. I. (2011). Neuropsychological rehabilitation for multiple sclerosis. *Cochrane Database Syst Rev.*, 9(11).
- Salles, J. F. (2007). Intervenção nas dificuldades de leitura em crianças: Uma abordagem neuropsicológica cognitiva. In A. L. Sennyey, L. I. Z. Mendonça, B. B. G. Schlecht, E. F. Santos, E. C. Macedo. *Neuropsicologia e inclusão: tecnologias em (re)habilitação cognitiva*. São Paulo: Artes Médicas, pp. 239-246.
- Santos, F. H. (2006). Reabilitação neuropsicológica infanto-juvenil. In J. Abrisqueta-Gomez, & F. H. Santos, *Reabilitação Neuropsicológica: da teoria à prática*. São Paulo: Artes Médicas, pp. 17-33
- Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Education Psychology 92*(2), 284-95.
- Seligman, M. E. P. (1995). The effectiveness of Psychotherapy: The consumer reports study. *American Psychologist, 50* (12), 965-974.
- Solberg, M. M., & Mateer, C. A. (2009). *Reabilitação cognitiva: uma abordagem neuropsicológica integrativa*. São Paulo: LSantos.
- Spinillo, A. C. P., & Brandão, A. G. (2001). Produção e compreensão de textos em uma perspectiva de desenvolvimento. *Estudos de Psicologia (Natal), 6*(1), 51-62.
- Teeter, A. P., & Clikeman, M. S. (1997). *Child Neuropsychology*. Boston: Allyn and Bacon.
- Volker, M. A. (2006). Reporting effect sizes in school psychology research. *Psychology in the Schools, 43*, 653-672.
- Ward, R. M. (2002). *Highly significant findings in psychology: A power and effect size survey*. ETD Collection for University of Rhode Island: Rhode Island, USA.
- Wilson, B. A. (2011). *Reabilitação da Memória: Integrando teoria e prática*. Artmed: Porto Alegre.
- Wilson, B. A. (2003). The theory and practice of neuropsychological rehabilitation: An overview. In B. A. Wilson (Org.), *Neuropsychological rehabilitation: Theory and practice* (pp. 1-10). Lisse: Swets, & Zeitlinger.
- Wilson, B. A. (2002) Towards a comprehensive model of cognitive rehabilitation. *Neuropsychological Rehabilitation, 12*(2), 97-110.

- Wilson, B. A. (1991). Theory, assessment and treatment in neuropsychological rehabilitation. *Neuropsychology*, 5, 281-291.
- Ylvisaker, M., & Feney, T. J. (1998). *Collaborative Brain Injury Intervention: Positive Everyday Routines*. San Diego: Singular Publishing.
- Ylvisaker, M., Hanks, R., & Johnson-Green, D. (2003). Rehabilitation of children and adults with cognitive-communication disorders after brain injury. *ASHA Supplement*, 23, 59-72.

**Endereço para correspondência:**

Thiago da Silva Gusmão Cardoso  
Centro Paulista de Neuropsicologia  
Rua Embaú, 54, Vila Clementino  
CEP 04039-060 – São Paulo/SP  
E-mail: thiago\_gusmao1@hotmail.com

Recebido em 15/03/13

Revisto em 04/04/2013

Aceito em 27/04/2013