

Contribuições da psicometria para os estudos em neuropsicologia cognitiva

Caroline Tozzi Reppold¹

Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre – RS – Brasil

Cristiano Mauro Assis Gomes

Universidade Federal de Minas Gerais, Belo Horizonte – MG – Brasil

Alessandra Gotuzo Seabra

Universidade Presbiteriana Mackenzie, São Paulo – SP – Brasil

Monalisa Muniz

Universidade Federal de São Carlos, São Carlos – SP – Brasil

Felipe Valentini

Universidade Salgado de Oliveira, Niterói – RJ – Brasil

Jacob Arie Laros

Universidade de Brasília, Brasília – DF – Brasil

Resumo: A psicometria e a neuropsicologia cognitiva, embora com tradições independentes, apresentam diversos aspectos que podem ser utilizados de maneira interdisciplinar para a construção do conhecimento em psicologia. O presente texto: 1. discute como essas duas disciplinas podem interagir, 2. apresenta as contribuições da psicometria para os estudos em neuropsicologia cognitiva e 3. destaca os avanços ocorridos na neuropsicologia cognitiva ao fazer uso da psicometria. Discute-se também que os modelos de testagem da neuropsicologia estiveram, em grande parte, embasados em estudos de validade convergente e de critério. Contudo, tendências atuais da psicometria apontam a fragilidade desses modelos, bem como indicam a necessidade de estudos referentes à estrutura dos escores, processamento de resposta, validade consequencial e, principalmente, estudos que avaliem e comparem os modelos teóricos e seus construtos postulados. Uma relação recíproca entre psicometria e neuropsicologia cognitiva é importante para o avanço substancial nessas duas áreas e na psicologia geral.

Palavras-chave: psicometria; neuropsicologia cognitiva; avaliação psicológica; validade; testes neuropsicológicos.

CONTRIBUTIONS OF PSYCHOMETRICS TO STUDIES IN COGNITIVE NEUROPSYCHOLOGY

Abstract: Although they have their own independent traditions, psychometrics and cognitive neuropsychology present various aspects that may be used in an interdisciplinary approach to widen knowledge in the psychology field. This paper aims to: 1. discuss how these two areas of knowledge might interact, 2. present contributions of psychometrics to studies in the field of cognitive neuropsychology, and 3. highlight advances in the field of cognitive neuropsychology using psychometric theories. This paper also discourses that neuropsychological assessment models are often based on convergent and criterion validity. However, new psychometrics trends point to the weakness of these models, and indicate the need for research related to the structure

¹ **Endereço para correspondência:** Caroline Tozzi Reppold, Universidade Federal de Ciências da Saúde de Porto Alegre, Rua Sarmento Leite, 245, Centro Histórico – Porto Alegre – RS – Brasil. CEP: 90050-170. E-mail: carolinereppold@yahoo.com.br.

of scores, response processes, consequential validity, and, principally, the fit between theoretical models and their latent constructs. A reciprocal relation between psychometrics and cognitive neuropsychology is important for the development of both these areas, as well as for psychology in general.

Keywords: psychometrics; cognitive neuropsychology; psychological assessment; validity; neuropsychological tests.

CONTRIBUCIONES PSICOMÉTRICAS A LOS ESTUDIOS DE LA NEUROPSICOLOGÍA COGNITIVA

Resumen: La psicometría y neuropsicología cognitiva, aunque con tradiciones independientes, presentan diversos aspectos que se pueden utilizar de forma interdisciplinaria para la construcción del conocimiento en psicología. El presente texto: 1. analiza cómo estas dos disciplinas pueden interactuar; 2. presenta las contribuciones de la psicometría para los estudios de neuropsicología cognitiva, y 3. destaca los progresos de esta última al hacer uso de la psicometría. A su vez, sostiene que los modelos de las pruebas neuropsicológicas han tenido, en gran parte, base en estudios de validez convergente y de criterio. Sin embargo, las tendencias actuales de la psicometría apuntan a la fragilidad de estos modelos, así como indican la necesidad de realizar estudios sobre la estructura de las puntuaciones, procesamiento de respuesta, validez consecuencial y, especialmente, estudios que evalúen y comparen los modelos teóricos y sus constructos postulados. La relación recíproca entre la psicometría y la neuropsicología cognitiva es importante para el progreso sustancial de estas dos áreas y de la psicología en general.

Palabras clave: psicometría; neuropsicología cognitiva; evaluación psicológica; validez; tests neuropsicológicos.

Duas grandes tradições em psicologia, a psicometria e a neuropsicologia cognitiva, têm se mostrado profícuas na geração de novas evidências e *insights* para o conhecimento científico psicológico. A psicometria tem passado por grandes transformações em função do desenvolvimento dos microcomputadores e dos *softwares* e da geração de novas técnicas de análise de dados a cada ano. Novos métodos, estratégias e perspectivas trazem um ar revigorado à psicometria, abrindo caminho para novas agendas de estudo e campos de investigação. Por sua vez, a neuropsicologia vivencia um momento único de impulsionamento no campo das neurociências, estudos sobre o cérebro, o sistema nervoso, suas redes e conexões com o comportamento (Coltheart, 2002). Em função de sua relevância e crescimento, não é incomum que essas tradições em psicologia interajam e que pesquisadores reflitam sobre como esses diferentes campos podem se articular para a melhoria da psicologia.

Compreender a relação entre a psicometria e a neuropsicologia cognitiva implica a necessidade da reflexão sobre como esses dois campos da psicologia podem atuar de forma interdisciplinar e contribuir para o aprimoramento da ciência psicológica. Talvez o maior dos desafios presentes na articulação entre esses dois campos envolva não se apropriar dessas tradições em psicologia em benefício específico de uma delas, em detrimento das outras. Dentre os vários enfoques possíveis ao estudar essa relação, o presente artigo abordará especificamente as contribuições que a psicometria pode oferecer à construção de evidências em neuropsicologia cognitiva. Nessa linha de raciocínio, a psicometria poderia subsidiar ferramentas ou meio hábeis para a construção de novas evidências. De fato, o campo da neuropsicologia brasileira, especificamente

da neuropsicologia cognitiva, tem avançado substancialmente na construção e adaptação de testes neuropsicológicos e se utilizado de estratégias de análise de dados provenientes da psicometria para investigar a validade dos escores derivados de seus instrumentos. Pesquisadores em neuropsicologia reconhecem a necessidade de aprimorar a qualidade dos testes utilizados e têm buscado avançar nesse quesito.

Dessa forma, a psicometria tem sido usada como uma ferramenta para a neuropsicologia, o que tem gerado frutos positivos. Contudo, tal uso ainda tem sido insuficiente porque não aproveita, da própria psicometria, o que ela pode oferecer de melhor, ou seja, os seus construtos e modelos teóricos, assim como sua tradição de investigação. Para compreender como a psicometria pode auxiliar de forma mais efetiva a neuropsicologia, primeiramente é importante resgatar a própria definição de psicometria. Ela não é apenas um conjunto de técnicas e procedimentos para a construção e validação de testes e instrumentos psicológicos. É bem mais que isso. A psicometria é um campo do conhecimento em psicologia voltado para a construção de evidências por meio da modelagem dos construtos psicológicos. Esse aspecto fundamental da psicometria muitas vezes é esquecido. A psicometria não pode ser confundida com a estatística. O que a torna um campo especial e sempre jovem não é seu aspecto técnico para a testagem (neuro)psicológica, mas sua busca pela inovação e pela construção de evidências que permitam ao psicólogo investigar seus modelos teóricos, compará-los e refutá-los por meio de métodos quantitativos. A psicometria pode ser uma importante ferramenta para auxiliar as diferentes teorias no questionamento empírico de seus próprios construtos e objetos de investigação. Nessa linha de raciocínio, incorporar de fato a psicometria envolve questionar *ad infinitum* os próprios construtos e modelos da psicologia. No caso em questão, envolve investigar cuidadosamente a validade dos próprios construtos em neuropsicologia, assim como construir e analisar novos modelos, capazes de integrar os construtos dessas duas grandes tradições.

Tendências da psicometria que contribuem para a neuropsicologia

No cerne dessa discussão, várias tendências contemporâneas da psicometria são importantes para refletirmos o seu potencial de relacionamento com outras áreas do conhecimento. Uma primeira é a crescente importância atribuída ao desenvolvimento e aprimoramento de teorias e modelos teóricos. Nesse contexto, as teorias exercem um papel cada vez mais valorizado no processo de análise de dados. Na análise multivariada, por exemplo, a existência de uma teoria sobre a relação de variáveis associadas com a variável de resposta é de suma importância. Na ausência de uma teoria, a escolha entre modelos será fundada unicamente com base nas considerações estatísticas: o modelo com o melhor ajuste será escolhido. Essa escolha nem sempre será a mais adequada. Nesse sentido, Kreft e De Leeuw (1998) afirmou que é uma falácia esperar que os próprios números revelem a verdade. Tal tendência é interessante, uma vez que aproxima a psicometria de outras disciplinas, como a neuropsicologia, que têm forte ênfase no substrato teórico dos achados empíricos.

Outra evidência da crescente importância atribuída pela psicometria à teoria está presente no campo de estudos de validade. No decorrer dos últimos 50 anos, é dado cada vez mais importância ao modelo de validade de construto em detrimento do modelo de validade de critério. O modelo de validade de critério é principalmente baseado em dados empíricos. Um dos maiores problemas com esse modelo de validade foi a falta de disponibilidade de critérios adequados e o fato de que alguns critérios poderiam ser avaliados somente depois de um longo período de tempo (Zumbo & Chan, 2011). Além do mais, o modelo de validade de critério não forneceu uma base teórica suficiente para definir se um critério poderia ser de fato adequado ou não. De acordo com Cronbach (1971), a testagem de personalidade criou a necessidade de desenvolver o modelo de validade de construto, uma vez que, nesse campo de pesquisa, não existem critérios adequados para tipos de personalidade. No modelo de validade de construto, a validade é avaliada com base em considerações teóricas, o que, como anteriormente descrito, é algo que vai ao encontro dos fundamentos da neuropsicologia e aproxima a psicometria.

Uma segunda tendência na psicometria é a crescente importância dada à replicação de estudos. Nos últimos anos, a atividade de replicar estudos na psicologia está sendo cada vez mais valorizada, parcialmente em razão da ocorrência de vários grandes casos de fraude na análise de dados e na publicação de resultados, e, em outros casos, de práticas questionáveis de pesquisa. Estudos de replicação são especialmente importantes em pesquisas aplicadas que não utilizam um delineamento experimental ou amostras aleatórias, uma vez que a generalizabilidade dos resultados desse tipo de estudos é mais questionável. Entretanto, também para pesquisas experimentais, a replicação de estudos é importante. Nesse contexto, Wagenmakers, Wetzels, Borsboom e Maas (2011) argumentam que os psicólogos experimentais precisam mudar a forma de conduzir experimentos e analisar os dados, pois atualmente sustentam, de forma demasiada, a construção de suas evidências no teste de significância, havendo pouco interesse em estudos estritamente confirmatórios.

Uma terceira tendência na psicometria é a crescente crítica contra os testes de significância e a importância dada ao tamanho de efeito e intervalos de confiança. O debate sobre os testes de significância gerou muita controvérsia nos últimos 20 anos. Por exemplo, segundo Schmidt e Hunter (1997), a utilização de testes de significância atrasou o crescimento do conhecimento científico. A maior influência no movimento contra o uso de testes de significância foi de Jacob Cohen (1994). Entretanto, não obstante o longo período de críticas em relação ao uso de testes de significância, ainda há vários pesquisadores que não entendem o que os testes de significância podem e não podem fazer. Thompson (1999) descreve várias crenças equivocadas de pesquisadores em relação aos testes estatísticos de significância. São essas: 1. a significância estatística indica a *importância* dos resultados encontrados; 2. a significância estatística indica *replicabilidade* dos resultados; e 3. a significância estatística é uma indicação do *tamanho de efeito* do resultado encontrado. De acordo com Cumming (2012), o tamanho de efeito e os intervalos de confiança são muito mais informativos do que os testes

estatísticos de significância. Uma razão adicional para calcular o tamanho de efeito está na realização de uma metanálise. Na existência de vários estudos de replicação, uma metanálise pode ser utilizada para sintetizar os resultados desses estudos. Para realizar uma metanálise, é essencial ter acesso a uma medida de tamanho de efeito de cada estudo de replicação. Assim, muitas revistas científicas exigem atualmente que pesquisadores sempre informem o tamanho de efeito do estudo realizado e não apenas relatem os resultados de um teste de significância.

As tendências descritas anteriormente podem favorecer o desenvolvimento de uma relação mais produtiva entre a psicometria e a neuropsicologia, pois aproxima tais abordagens (por exemplo, diante da tendência de fortalecer as bases teóricas dos achados psicométricos) e oferece ferramentas à neuropsicologia (por exemplo, ao enfatizar a importância da replicação de estudos e analisar informações como tamanho de efeito e intervalos de confiança). Borsboom (2006) distingue três desenvolvimentos na psicometria contemporânea que podem ajudar a melhorar a qualidade de comunicação entre a psicologia e a psicometria, com o propósito de diminuir a brecha entre ambas: 1. o número crescente de livros introdutórios sobre modelagem psicométrica; 2. o aumento de pacotes de *softwares* amigáveis para o usuário, baseados nas ideias psicométricas, como Lisrel e Mplus, que possibilitam modelagem psicométrica com variáveis latentes; e 3. o número crescente de artigos de conteúdo psicométrico que foram publicados nos últimos dez anos. Esses três desenvolvimentos podem facilitar o uso da psicometria por parte de pesquisadores de outras abordagens, como os da neuropsicologia.

Definição da neuropsicologia e suas especificidades na relação com a psicometria

Para a compreensão de como a psicometria pode contribuir para a neuropsicologia, é fundamental conhecer essa disciplina científica e, mais especificamente, a neuropsicologia cognitiva. Apesar de a neuropsicologia ser uma disciplina relativamente nova, já passou por várias transformações. Desde seu início, aproximadamente na década de 1960, até meados de 1980, seu foco principal era a identificação de comprometimentos funcionais associados a lesões encefálicas. As avaliações clínicas frequentemente empregavam testes com limitações psicométricas importantes e sem dados normativos adequados (Bilder, 2011). A partir da década de 1980, entretanto, com a ampliação das técnicas de neuroimagem, o objetivo de localização encefálica tornou-se menos evidente, e a neuropsicologia, fortemente influenciada pela psicologia cognitiva, tem mudado o seu foco para a compreensão das operações mentais envolvidas em determinadas tarefas e, em menor grau, sua relação com o funcionamento neurológico. Nessa perspectiva, surge a neuropsicologia cognitiva que pode ser definida como o ramo da psicologia cognitiva que utiliza dados de pessoas com distúrbios cognitivos adquiridos ou de desenvolvimento para melhor compreensão sobre os processos envolvidos na cognição normal (Lezak, Howieson, & Loring, 2004).

De acordo com Ellis e Young (1988, p. 4), que estão entre os primeiros pesquisadores em neuropsicologia cognitiva, essa disciplina tem dois objetivos principais: 1. explicar “modelos de desempenho cognitivo, preservado e alterado, de pacientes com lesão encefálica em termos de prejuízos de um ou mais componentes de uma teoria ou modelo de funcionamento cognitivo normal” e 2. “derivar conclusões sobre os processos cognitivos normais, intactos, a partir de modelos de capacidades deficitárias e intactas vistos em pacientes com lesão encefálica”. Assim, a neuropsicologia cognitiva, com maior número de pesquisas em países da Europa e da América do Norte, tem buscado testar modelos teóricos integrando evidências provenientes tanto de instrumentos psicológicos, que permitem inferências sobre construtos cognitivos, quanto de técnicas de neuroimagem estrutural e funcional, que permitem inferências sobre o processamento neurológico (Bilder, 2011).

Nessa linha de investigação, a neuropsicologia cognitiva tem ressaltado, por meio do estudo de grupos e dos estudos de caso único, a necessidade de instrumentos de avaliação mais adequados, que sejam suficientes para permitir uma análise detalhada de processos cognitivos. Assim, maior atenção foi dada à psicometria, com desenvolvimento de testes embasados teoricamente e com investigação de seus parâmetros psicométricos.

Conforme descrito por Wood, Carvalho, Rothe-Neves e Haase (2001), os passos no desenvolvimento de um instrumento de avaliação neuropsicológica devem seguir os mesmos critérios para desenvolvimento de instrumentos de avaliação psicológica, o que envolve a definição do construto psicológico a ser examinado; a operacionalização desse construto de forma a possibilitar a sua mensuração experimental e/ou psicométrica; a verificação das características psicométricas do instrumento de avaliação neuropsicológica, que poderá envolver a análise dos itens; e análise da precisão e validade dos escores derivados do instrumento, bem como sua normatização. Porém, na avaliação neuropsicológica, o teste neuropsicológico pode ser usado para comparação normativa (ou seja, derivado de uma população apropriada) ou individual (derivado da história prévia do paciente e de suas características). Na normativa, é essencial atentar às habilidades que sofrem grande influência de nível de escolaridade ou nível socioeconômico, de modo a considerar, para comparação, o grupo específico ao qual o paciente pertence. Em tais casos, pode ser preferível conduzir uma avaliação padronizada individual, e não normativa, para comparar as habilidades atuais do paciente neurológico com suas características anteriores à alteração neurológica, como no curso de uma demência ou no acompanhamento de um processo cirúrgico encefálico. A flexibilidade na administração dos instrumentos é um aspecto central da avaliação neuropsicológica (Lezak *et al.*, 2004).

Outra situação destacada por Rusell, Rusell e Hill (2005), além da necessidade de verificação dos parâmetros psicométricos dos testes usados na neuropsicologia, é a investigação das baterias de testes. Tais autores definem como bateria não um simples agrupamento de testes, mas um conjunto que permita uma interpretação acerca do

funcionamento do indivíduo que seja fruto da relação entre os desempenhos dos testes. Esse é um ponto forte da neuropsicologia, que busca a integração entre diferentes medidas e a análise da integridade *versus* prejuízo em diferentes módulos cognitivos. Logo, a bateria provê informações que não poderiam ser obtidas a partir do uso de testes individuais. Sendo esse o caso, faz-se necessária a análise da precisão e da validade da bateria, como conjunto de interpretação, ainda que os testes isolados já obtenham evidências de fidedignidade e validade. Entretanto, poucas baterias são objeto de estudos psicométricos.

Paralelamente à administração de provas standardizadas, uma avaliação neuropsicológica deve realizar também uma observação detalhada das respostas gerais do paciente diante da prova e da situação de avaliação. Para tanto, concomitantemente ao registro quantitativo das respostas, são feitos registros qualitativos da responsividade do paciente, de seu reconhecimento dos próprios erros, de respostas emocionais e características de execução das tarefas (Ardila & Ostrosky-Solís, 1996). Tais informações podem, por exemplo, consistir em um padrão irregular de acertos e erros ao longo de um teste em que, por ter grau de dificuldade crescente, seria esperado um padrão de diminuição do número de acertos. Nesse caso, muito mais relevante do que a análise comparativa do escore total com o grupo de referência é a análise qualitativa do padrão de desempenho, que pode sugerir, por exemplo, a presença de uma negligência ou de um déficit atencional.

Dessa forma, a neuropsicologia tem especificidades em relação à avaliação psicológica na abordagem psicométrica que podem contribuir para a compreensão dos processos cognitivos. Entretanto, carece de recursos que a psicometria pode oferecer de modo a investigar parâmetros dos instrumentos, tornando a avaliação mais precisa. No Brasil, especialmente, ainda há grande carência de instrumentos adequados que sirvam para fundamentar uma avaliação neuropsicológica ou a proposição de novos modelos teóricos de funcionamento cognitivo. No atual panorama nacional, a maioria dos instrumentos que são foco de estudos de adaptação ou busca de evidências de validade destina-se ao rastreamento da identificação de prejuízos cognitivos, sendo funções executivas, memória e linguagem os construtos mais frequentemente cobertos por esses instrumentos. Em sua maioria, os idosos são a população-alvo dos estudos, o que é coerente com o dado anterior, uma vez que essa faixa etária é mais frequentemente acometida por declínios cognitivos. Em termos neuropsicológicos, esse dado mostra o potencial interesse dos pesquisadores em compreender os processos cognitivos à luz da psicologia do desenvolvimento ou, mais especificamente, compreender, aprofundar ou modificar modelos teóricos que descrevem o desenvolvimento cognitivo ou as áreas envolvidas nesse processo. Nesse sentido, estudos com crianças e idosos são relevantes para incremento dos modelos teóricos sobre cognição, na medida em que permitem verificar como a maturação de determinadas áreas cerebrais e o refinamento das funções cognitivas, no caso infantil, impactam o desempenho do sujeito em testes neuropsicológicos. Isso ocorre especialmente em estudos cujo delineamento possibilita a avaliação do mesmo sujeito em momentos distintos.

Em relação às investigações psicométricas, muitos dos estudos neuropsicológicos brasileiros limitam-se à busca de evidências baseadas no conteúdo e na relação com variáveis externas, e, com menos frequência, na estrutura interna. A busca de evidências de validade baseada no conteúdo é apresentada, por exemplo, em instrumentos de rastreio não específicos, como o miniexame do estado mental, e naqueles que avaliam processos comunicativos ou mnemônicos. De fato, na avaliação desses construtos, diferenças culturais têm potencialmente grande impacto no desempenho dos sujeitos nos testes. Por exemplo, em um teste de memória ou atenção, um item que, no instrumento original, é uma palavra formada por uma única sílaba não pode ser traduzido por outra palavra formada por três sílabas (por exemplo: *red versus vermelho*), nem mesmo por uma palavra que seja pouco usual no contexto local. Assim, a preocupação com a extensão da palavra (morfologia) faz mais sentido na adaptação de um teste neuropsicológico do que na adaptação de um teste de personalidade, por exemplo, no qual a apreciação da paridade semântica é imprescindível. No caso dos testes neuropsicológicos, a validade de conteúdo é importante não apenas por “garantir” a abrangência e representatividade dos itens, mas por viabilizar que comparações de resultados internacionais possam ser realizadas com maior segurança.

Já sobre a busca de evidências de validade baseadas nas relações com variáveis externas, observa-se que os estudos mais frequentemente desenvolvidos pelos neuropsicólogos são os de validade convergente ou discriminante. No caso da validade convergente, os autores buscam um padrão de correlação entre os escores do teste-alvo e de outras variáveis que medem o mesmo construto ou construtos relacionados (por exemplo: desempenho em um teste de avaliação de síndrome disexecutiva *versus* desempenho em subtestes da escala de inteligência Wechsler para adultos (*Wechsler adult intelligence scale – Wais-III*) ou em um teste de fluência verbal. No entanto, nota-se que, muitas vezes, os testes escolhidos como padrão-ouro para as análises de correlação são testes de rastreio cognitivo, formados por poucos itens específicos de cada domínio e por fatores pouco sensíveis para identificação de algumas situações clínicas (por exemplo: miniexame do estado mental). Além disso, há de se considerar que, pela complexidade do funcionamento cognitivo, ainda existe certa dificuldade para estabelecer que magnitudes mínimas de correlação deveriam ser esperadas em cada avaliação para que se pudesse considerá-las, de fato, como evidências de validade.

Em se tratando das pesquisas que buscam evidências de validade baseadas em critérios para os testes neuropsicológicos, constata-se que alguns pesquisadores erroneamente classificam como validade discriminante estudos que se enquadram na definição de validade concorrente. Esse é o caso de estudos que buscam descrever padrões de resposta típicos de grupos clínicos que se diferenciam do desempenho de grupos controle. De qualquer forma, observa-se que a investigação de grupos clínicos *versus* controle é usual nos estudos neuropsicológicos e pertinente aos objetivos da área, pois a identificação de tais diferenças permite aos pesquisadores a seleção de grupos amostrais com características semelhantes para futuros estudos que venham a investigar

o que há de específico no desenvolvimento cognitivo desse grupo. Há de se considerar, contudo, que uma das dificuldades encontradas nos estudos de busca de evidências de validade concorrente de instrumentos neuropsicológicos é o fato de que muitos desses apresentam efeito de teto quando aplicados em amostras controle, o que reduz a variabilidade dos escores no grupo não clínico e assim reduz o tamanho do coeficiente de validade. Em contrapartida, é frequente, nos estudos que envolvem testes neuropsicológicos, o cálculo da sensibilidade e especificidade dos instrumentos avaliados, o que é um recurso importante para tomada de decisão dos neuropsicólogos clínicos.

Quanto à subutilização da psicometria nos estudos da neuropsicologia cognitiva

Algumas das discussões importantes sobre a contribuição da psicometria aos estudos neuropsicológicos são a importância dada ao contexto de validação dos instrumentos e a subutilização atual de algumas fontes de evidência de validade dos instrumentos, especialmente no que concerne às evidências baseadas no processo de respostas e na consequência das respostas. Os *Standards for Educational and Psychological Testing* (American Educational Research Association [Aera], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1986), ao modificarem o conceito de validade e as formas de obtenção de resultados que as evidenciem, apresentam cinco fontes de evidências de validade. No documento publicado em 1986, validade era definida como “o grau em que as evidências embasam as inferências feitas a partir dos escores dos testes” (Aera, APA, & NCME, 1986, p. 9), sendo possível, na ocasião, tipificar três classes de validade (de conteúdo, construto e critério). Já em 1999, a revisão dos *Standards* (p. 9) passou a definir validade como “o grau em que evidência e teoria sustentam as interpretações dos escores dos testes vinculados aos usos propostos dos testes”. Nesse sentido, ressaltam que o processo de validação é cumulativo e visa validar não o teste em si,

[...] mas as interpretações propostas por ele e as aplicações práticas tentadas. [...] [As alterações propostas nos *Standards*] apontam a fragilidade da generalização das inferências elaboradas a partir de pesquisas realizadas em contextos específicos para outros muito diversos (Primi, Muniz, & Nunes, 2009, p. 246).

Ou seja, tornam questionável a possibilidade da administração, em contexto da clínica neuropsicológica, de testes validados para contexto escolar ou para avaliação de trânsito, por exemplo. Isso coloca em xeque também a pertinência da utilização de determinados instrumentos como padrão-ouro para busca de evidências de validade convergente de testes neuropsicológicos. Além disso, a nova definição enfatiza a questão da relevância teórica da interpretação dos resultados, o que é ponto central no campo da neuropsicologia.

Na sequência das reformulações propostas em 1999, cinco fontes de evidências de validade foram listadas nos *Standards*, a saber: 1. evidências baseadas no conteúdo,

2. baseadas no processo de resposta, 3. baseadas na estrutura interna, 4. baseadas nas relações com variáveis externas e 5. baseadas nas consequências da testagem. As evidências baseadas nos processos de respostas referem-se à descrição dos processos mentais envolvidos na realização das tarefas propostas pelo teste, o que contribui para detalhar o construto que o teste se propõe a medir e, notadamente, para subsidiar o desenvolvimento de modelos teóricos sobre o funcionamento cognitivo. Assim, deveria ser um dos principais objetivos dos pesquisadores interessados em desenvolver instrumentos neuropsicológicos ou teorias cognitivas. No entanto, são raros os estudos atuais da psicologia brasileira que se destinam à investigação de evidências baseadas nos processos de respostas, com exceção das pesquisas sobre avaliação assistida.

Segundo Primi *et al.* (2009, p. 248), na busca de evidências de validade processual,

[...] a partir do referencial teórico do construto, elaboram-se modelos explicativos do processamento mental que ocorre ao responder os[ic] itens do teste. Elaboram-se também previsões sobre parâmetros do comportamento – acertos, tempo de reação, por exemplo – em relação a diferentes propriedades dos itens. O modelo teórico inclui explicações sobre como as diferentes características dos itens afetam o processamento e, especialmente, como as alterações processuais culminam em alterações comportamentais – ocorrências de erros, maior ou menor tempo de reação, etc. Assim, o modelo teórico pode ser empiricamente testado.

Nesse sentido, os modelos teóricos propostos pela psicologia cognitiva poderiam ser testados por um método dedutivo (ou seja, por meio de argumentos nos quais a forma lógica válida garante a verdade da conclusão se as premissas forem verdadeiras). Por outro lado, devem-se utilizar instrumentos neuropsicológicos e premissas psicométricas. Uma vez que o estudo de evidências de validade pelo processo de respostas pode também acontecer por meio do questionamento dos respondentes sobre suas estratégias de resolução dos itens, esse tipo de evidência de validade que parte de um pressuposto indutivo (do respondente para teoria) pode também ser considerado uma forma de verificação de quão plausível é determinada teoria sobre o desenvolvimento ou funcionamento cognitivo.

O processamento de resposta envolve um conjunto de variáveis com forte base nos estudos da psicologia cognitiva. Nesse sentido, os processos estudados por essa ciência podem auxiliar as pesquisas na área de psicometria, principalmente no que se refere ao processo de validação de um instrumento psicológico. Infelizmente, existem poucos estudos nessa área. Cizek, Rosenberg e Koons (2007) investigaram os artigos publicados sobre validação dos escores de testes psicológicos e educacionais e concluíram que apenas 1,8% dos manuscritos reportavam evidências com base no processo de resposta. No entanto, o número de investigações sobre esse tipo de validade parece ter aumentado na literatura internacional, nos últimos anos (Zumbo & Chan, 2011).

Em que pese a importância dos estudos de processamento de resposta, ainda é necessário aprofundar os métodos para obtenção de tais evidências. Padilla e Benítez

(2014) destacam quatro conjuntos de procedimentos utilizados nos estudos de evidências de validade com base no processamento de resposta:

- *Tempo de resposta*: foca a relação entre o tempo para resolver uma tarefa e o seu nível de complexidade.
- *Rastreamento ocular*: busca rastrear os movimentos dos olhos durante a aplicação de um item de um teste. Tais movimentos podem oferecer pistas sobre o processo atencional e cognitivo utilizado na resolução de uma tarefa.
- *Entrevistas*: os participantes da pesquisa são indagados a pensar como e por que responderam à tarefa solicitada.
- *Grupo focal*: os tópicos específicos de uma tarefa de um teste são discutidos em grupo.

Grande parte dos estudos (na área de psicologia) de validade de processamento de resposta utiliza entrevistas para a coleta de dados, bem como métodos qualitativos de análise de dados. Normalmente, os roteiros de entrevistas envolvem perguntas sobre a maneira como o teste foi respondido, como os respondentes entenderam a relevância dos itens, entre outros aspectos. Contudo, ainda que a estratégia de entrevista cognitiva seja importante para a validação com base no processo de resposta, outros métodos podem (e precisam) ser utilizados. Nesse sentido, a utilização de métodos experimentais pode ampliar o leque de ferramentas da psicometria. Por exemplo, na resolução de um item de raciocínio espacial, um participante pode tentar resolver um item por eliminação de alternativas (utilizando o raciocínio lógico) e não por meio da manipulação mental de uma figura. Para investigar essa diferença no processamento da resposta, o pesquisador poderia manipular experimentalmente as alternativas incorretas de resposta e observar o tempo de resposta e a possível alteração na dificuldade do item. Cabe salientar que métodos semelhantes já são utilizados em pesquisas internacionais, nas quais são empregados os modelos mistos de TRI (*mixture IRT models*) para estimar diferentes classes latentes para diferentes estratégias de resolução de itens e velocidade de resposta, entre outros aspectos (cf., por exemplo, Bolt, Cohen, & Wollack, 2002).

Em última instância, há de se considerar o caráter promissor dos estudos, no campo da neuropsicologia, que buscam evidências de validade com base na consequência das respostas. Essa fonte investiga as consequências sociais, sejam intencionais ou não, do uso dos testes e se a sua utilização está produzindo os efeitos *a priori* esperados. Isso é especialmente relevante no caso da avaliação da efetividade e eficácia de uma intervenção de reabilitação neuropsicológica. Nesse caso, a “chancela” de validade não se restringe aos escores do teste em si, mas ao processo avaliativo e terapêutico como um todo, e, por si, traz uma perspectiva ética sobre a pertinência do uso de determinados instrumentos, intervenções ou sobre os desdobramentos deontológicos e metodológicos de um modelo teórico proposto para compreensão do funcionamento cognitivo. No entanto, ainda são raros os estudos baseados nessa fonte de evidência, sejam esses

próprios do campo da avaliação neuropsicológica ou da avaliação psicológica de forma mais ampla.

O presente texto procurou contribuir para a discussão sobre a importância de unir diferentes áreas da psicologia, especificamente a psicometria e a neuropsicologia, que apresentam papel de destaque na evolução da ciência psicológica e, com isso, propiciam um melhor entendimento sobre o ser humano. A utilização em conjunto dessas grandes áreas somente aumentará a probabilidade de descobertas relevantes sobre a mente humana, culminando em proposições mais adequadas de intervenções que visem ao bem-estar do ser humano. Em especial, o texto optou por abordar a contribuição da psicometria para a neuropsicologia. No entanto, encoraja-se que demais textos sejam construídos apontando para as vertentes da neuropsicologia e da psicologia cognitiva, pois, na proposta interdisciplinar, a junção das especificidades das três áreas é que faz a ciência psicológica avançar.

Referências

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1986). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Ardila, A., & Ostrosky-Solís, F. (1996). *Diagnóstico del daño cerebral: enfoque neuropsicológico*. Mexico: Editorial Trillas.
- Bilder, R. M. (2011). Neuropsychology 3.0: evidence-based science and practice. *Journal of the International Neuropsychological Society, 17*, 7-13.
- Bolt, D. M., Cohen, A. S., & Wollack, J. A. (2002). Item parameter estimation under conditions of test speededness: application of a mixture Rasch model with ordinal constraints. *Journal of Educational Measurement, 39*, 331-348. doi: 10.1111/j.1745-3984.2002.tb01146.x.
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*(3), 425-440. doi: 10.1007/s11336-006-1447-6.
- Cizek, G. J., Rosenberg, S. L., & Koons, H. H. (2007). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement, 68*, 397-412. doi: 10.1177/0013164410379323.
- Cohen, J. (1994). The earth is round ($p < 0.05$). *American Psychologist, 49*, 997-1003. doi: 10.1037/0003-066X.49.12.997.

- Coltheart, M. (2002). Cognitive neuropsychology. In J. Wixted (Ed.). *Stevens' handbook of experimental psychology* (Vol. 4, 3rd ed., pp. 139-174). New York: John Wiley & Sons.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.). *Educational measurement* (2nd ed., pp. 443-407). Washington, DC: American Council on Education.
- Cumming, G. (2012). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analyses*. New York: Routledge.
- Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology*. Hillsdale, NJ: Lawrence Erlbaum.
- Kreft, I. G. G., & De Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage.
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.
- Padilla, J. L., & Benítez, I. (2014). Validity evidence based on response processes. *Psicothema*, 26, 136-144. doi: 10.7334/psicothema2013.259.
- Primi, R., Muniz, M., & Nunes, C. H. S. S. (2009). Definições contemporâneas de validade de testes psicológicos. In C. S. Hutz (Ed.). *Avanços e polêmicas em avaliação psicológica* (pp. 243-265). São Paulo: Casa do Psicólogo.
- Rusell, E. W., Rusell, S. L. K., & Hill, B. D. (2005). The fundamental psychometric status of neuropsychological batteries. *Archives of Clinical Neuropsychology*, 20, 785-794.
- Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In L. L. Harlow, S. A. Mulaik & J. H. Steiger (Eds.). *What if there were no significance tests?* (pp. 37-64). London: Lawrence Erlbaum.
- Thompson, B. (1999). Five methodology errors in educational research: a pantheon of statistical significance and other faux pas. In B. Thompson (Ed.). *Advances in social science methodology* (pp. 23-85). London: JAI Press.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & Maas, H. L. J. van der (2011). Why psychologists must change the way they analyze their data: the case of psi. *Journal of Personality and Social Psychology*, 100, 426-432. doi: 10.1037/a0022790.
- Wood, G. M. O., Carvalho, M. R. S., Rothe-Neves, R., & Haase, V. G. (2001). Validação da bateria de avaliação da memória de trabalho. *Psicologia: Reflexão e Crítica*, 14(2), 305-316.
- Zumbo, B. D., & Chan, E. K. H. (2014). *Validity and validation in social, behavioral, and health sciences*. New York: Springer International Publishing.

Submissão: 15.9.2014

Aceitação: 14.4.2015