

Viabilidade da estimativa de QI a partir dos subtestes Vocabulário e Cubos da WAIS-III

Flávia Wagner¹ – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil
 Josiane Pawlowski – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil
 Denise Balem Yates – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil
 Suzi Alves Camey – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil
 Clarissa Marceli Trentini – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

Resumo

Objetivou-se avaliar a viabilidade de estimar o Quociente de Inteligência (QI) de adultos por meio de duas estratégias que utilizam uma forma curta constituída pelos subtestes Vocabulário e Cubos da Escala Wechsler de Inteligência para Adultos (WAIS-III). Participaram do estudo 77 homens e mulheres, com idade média de 39,1 anos ($dp=13,4$). Foi aplicada a WAIS-III e questionários para cumprir os critérios de inclusão dos participantes sem prejuízos neurológicos e psiquiátricos. Para avaliar a concordância entre as duas estratégias de estimar o QI, utilizou-se o método estatístico proposto por Bland e Altman. Como resultado, encontrou-se que as estimativas de QI geradas pelas duas estratégias apresentam concordância com a medida do QI realizada pela aplicação da escala completa apenas em alguns dos casos avaliados. Sugere-se cautela na utilização dessas estratégias para estimar o QI de adultos.

Palavras-chave: Avaliação, Quociente de inteligência, Estimação, Escala de Inteligência Wechsler para Adultos, Subtestes.

Viability of IQ estimates based on Vocabulary and Block Design WAIS-III subtests

Abstract

The aim of this study was to evaluate the feasibility of estimating the Intelligence Quotient (IQ) in adults based on two strategies that use Vocabulary and Block Design subtests of Wechsler Adult Intelligence Scale (WAIS-III). Seventy-seven men and women participated of the study, with mean age of 39.1 ($sd=13.4$). WAIS-III subtests were administered and questionnaires were used to verify inclusion criteria for participants without neurological or psychiatric impairments. A statistical method proposed by Bland and Altman was used to evaluate the agreement between both strategies to estimate IQ. Results suggest that IQ estimates derived from both strategies only agree with the IQ derived from the full battery in a few cases of the sample. Considering these findings, caution is suggested when using these strategies to estimate adults' IQ.

Keywords: Assessment, Intelligence quotient, Estimation, Wechsler Adult Intelligence Scale, Subtests.

Estudos nacionais e internacionais utilizam combinações de subtestes das escalas Wechsler para estimar o nível de inteligência global de adultos (Kenny & cols., 1997; Mantoan & cols., 2006; Thompson & cols., 2001; Tramontina & cols., 2002). Esse método é utilizado quando se deseja avaliar a inteligência de um indivíduo sem a necessidade de utilizar uma bateria completa como as escalas Wechsler de Inteligência, cuja aplicação dura, em média, 90 minutos (Nascimento & Figueiredo, 2002a). O determinante pode ser o fator tempo, por exemplo, na prática de atendimento ambulatorial em hospitais, em que muitas vezes não é possível a utilização de uma escala longa. Outra razão pode ser a necessidade de apenas se estimar o quociente de inteligência (QI), por exemplo, para análises de comparação em uma investigação científica. Segundo Jeyakumar, Warriner, Raval e Ahmad (2004), várias formas curtas foram

desenvolvidas para diferentes edições da escala, com o objetivo de estimar rapidamente o QI de um indivíduo quando a aplicação da bateria completa não é possível ou viável. Dessa forma, o objetivo do presente estudo é avaliar a viabilidade de se estimar o QI de um indivíduo utilizando os subtestes Vocabulário e Cubos da Escala Wechsler de Inteligência para Adultos, Terceira Edição, WAIS-III (Wechsler, 1997, adaptado por Nascimento, 2005), por meio de dois diferentes métodos.

A Escala Wechsler de Inteligência é um instrumento psicométrico tradicionalmente conhecido para a avaliação da inteligência e determinação do QI. Sua primeira versão, desenvolvida por David Wechsler nos Estados Unidos, foi publicada no ano de 1939, com o nome de *Wechsler-Bellevue Scale* (Nascimento & Figueiredo, 2002b). Desde então, a escala sofreu revisões e foi aprimorada durante as décadas subsequentes, dando origem às edições mais recentes, adaptadas para o contexto brasileiro e publicadas no país, entre elas a WAIS-III (Nascimento, 2005) e a terceira edição da Escala Wechsler de Inteligência para Crianças, WISC-III (Figueiredo, 2002).

A WAIS-III é um instrumento clínico desenvolvido para avaliação da capacidade intelectual

¹ Endereço para correspondência:
 Instituto de Psicologia, UFRGS
 Rua Ramiro Barcelos, 2.600, sala 119 – CEP 90035-003 – Porto Alegre-RS
 E-mail: flavia_scs@yahoo.com.br

de adolescentes e adultos. É indicado para pessoas de 16 a 89 anos de idade e é composta por 14 subtestes, sendo sete verbais e sete de execução. Fornece quatro índices fatoriais, denominados compreensão verbal, organização perceptual, memória operacional e velocidade de processamento, e três medidas compostas, o QI verbal, o QI de execução e o QI total. Fazem parte da Escala Verbal os subtestes Vocabulário, Semelhanças, Aritmética, Dígitos, Informação, Compreensão e Sequência de Números e Letras e da Escala de Execução, Completar Figuras, Códigos, Cubos, Raciocínio Matricial, Arranjo de Figuras, Procurar Símbolos e Armar Objetos (Nascimento, 2005).

A necessidade de um método para a rápida avaliação do QI levou pesquisadores e clínicos a desenvolverem formas alternativas para a mensuração da capacidade intelectual. O amplo uso das escalas Wechsler incentivou o desenvolvimento de formas curtas (*short forms*) de avaliação da inteligência, para obtenção do QI total sem a aplicação da bateria completa (Tellegen & Briggs, 1967), ou seja, mediante uma seleção de subtestes ou de itens de todos os subtestes que compõem a bateria. A utilização de formas curtas² e a sua validade como um método para a mensuração do QI, entretanto, não são um consenso entre os pesquisadores. Kaufman e Kaufman (2001) salientam que a literatura está repleta de diferentes propostas de formas curtas e de controvérsias sobre a adequação do seu uso. Os autores defendem o uso de escalas breves ao invés de formas curtas, pois as escalas breves são instrumentos construídos como tal e possuem indicadores próprios de validade e fidedignidade, além de normas específicas. Tellegen e Briggs (1967) consideram que formas curtas compõem um instrumento distinto da escala original, apresentando, portanto, indicadores de validade e fidedignidade próprios. Cyr e Brooker (1984) também apontam a necessidade de considerar coeficientes de validade e fidedignidade na escolha de uma forma curta.

Pela necessidade de avaliação das propriedades psicométricas de formas curtas, Tellegen e Briggs (1967) sugeriram fórmulas apropriadas para avaliar a validade e a fidedignidade de novas combinações de subtestes das escalas Wechsler. A fórmula da fidedignidade, de acordo com os autores, é obtida por meio dos coeficientes de fidedignidade dos subtestes que compõem a forma curta e das intercorrelações entre eles. Já a validade é determinada através de uma fórmula que avalia a correlação entre os resultados da

forma curta e da escala total. Tellegen e Briggs (1967) introduziram uma correção nessa fórmula, considerando que um número significativo dos estudos realizados em relação a formas curtas utiliza os escores de uma única aplicação para determinar tanto o QI estimado quanto o QI total. Sem essa correção, a fórmula poderia inflar de forma espúria a correlação entre os dois escores, pois os erros de medida das formas curtas nesse caso não são independentes dos erros de medida das escalas completas.

A avaliação da validade de uma forma curta também foi foco de discussões por parte de outros pesquisadores. De acordo com Sattler (1988), a validade de uma forma curta geralmente é avaliada por três critérios, propostos pela primeira vez por Rabin, em 1943 (Rabin, 1943): 1) a correlação entre o QI fornecido pela forma curta e o QI fornecido pela escala total, 2) a comparação das médias dos escores de QI fornecidos pela forma curta e pela escala total e 3) a concordância entre as duas formas em relação à classificação da inteligência. Silverstein (1990), entretanto, considera esses critérios necessários para a avaliação de uma forma curta, porém acredita que estes não fornecem muitas informações acerca da validade da mesma. O autor ressalta que uma correlação alta entre a maior parte dos subtestes e a escala total foram relatadas em todas as amostras de normatização das escalas Wechsler, o que possivelmente levaria a uma alta correlação entre o QI fornecido por uma forma curta e o QI da escala total. Em relação ao segundo critério, Silverstein complementa que pequenas diferenças nas médias de QI fornecidas por meio das duas formas mostrar-se-ão estatisticamente significativas em uma amostra grande, o que invalida a interpretação. O terceiro critério também não é considerado útil pelo autor, que considera que um ponto de diferença pode levar a uma discordância na classificação da inteligência, enquanto uma diferença maior pode não levar. A partir dessas críticas, o autor propõe que a validade e a fidedignidade são de importância fundamental para fazer uso de uma forma curta.

Outra questão que merece atenção na utilização de formas curtas da WAIS-III é a conversão dos escores ponderados dos subtestes em escores de QI. De acordo com Tellegen e Briggs (1967), duas formas de conversão são citadas na literatura. A primeira delas, denominada internacionalmente de *prorating*, permite a utilização das tabelas do manual para conversão da soma dos pontos ponderados em escores de QI, assumindo que o indivíduo obterá, nos subtestes não aplicados, a mesma média de escore ponderado que obteve nos subtestes aplicados. A partir disso, é possível fazer uma estimativa da soma de

² Neste texto, o termo “forma curta” será utilizado sempre que se fizer referência ao uso de combinações de subtestes já existentes.

todos os subtestes e converter essa soma em escore de QI mediante tabelas apropriadas. Já a segunda forma utiliza procedimentos de regressão para determinar estimativas de QI com base na amostra de normatização da Escala Wechsler. Por exemplo, a soma dos escores ponderados de Vocabulário e Cubos da amostra de normatização é submetida a procedimentos de regressão para estimar o QI total, e tabelas de conversão são derivadas dos resultados dessas análises. De acordo com Tellegen e Briggs, ambos os procedimentos são inadequados. Os autores desenvolveram uma fórmula para transformação dos escores ponderados dos subtestes utilizados em escores de QI, considerando as intercorrelações entre esses subtestes. Segundo Silverstein (1984), esse é o procedimento mais adequado para o cálculo do QI estimado nas formas curtas.

Jeyakumar e colaboradores (2004), pelas fórmulas propostas por Tellegen e Briggs (1967), conduziram um estudo com o objetivo de investigar se formas curtas da WAIS-III permitem chegar a estimativas válidas e fidedignas do funcionamento intelectual de um indivíduo sem a aplicação da bateria completa. Os resultados sugeriram que o uso de formas curtas é apropriado e útil clinicamente para estimar o funcionamento intelectual. Tomando por base esse estudo, os autores construíram tabelas que indicam o QI total que corresponde à soma de pontos ponderados de subtestes. Quatro combinações de subtestes foram utilizadas: Vocabulário-Cubos, Informação-Completar Figuras, Vocabulário-Semelhanças-Aritmética-Cubos e Semelhanças-Aritmética-Completar Figuras-Códigos. Essas tabelas foram construídas utilizando a amostra de normatização americana da WAIS-III.

De acordo com Jeyakumar e colaboradores (2004), duas estratégias principais são utilizadas para desenvolver formas curtas de avaliação. Na primeira estratégia, diferentes combinações de subtestes são selecionadas para estabelecer uma estimativa do QI total, enquanto na segunda, todos os subtestes são aplicados, mas o número de itens de cada um é reduzido. A primeira forma permite a posterior aplicação dos outros subtestes que compõem a bateria, enquanto a segunda elimina essa possibilidade. Em relação à primeira forma, Kaufman e Kaufman (2001) salientam que a maior parte dos estudos realizados com combinações de subtestes utiliza a aplicação da bateria completa para derivar as normas da nova combinação. Além disso, muitos dos estudos que propõem essas combinações de subtestes são baseados nos dados da amostra de normatização, sem realizar uma aplicação nova dos subtestes. Os autores ressaltam que o desempenho dos participantes pode

ser afetado pela diminuição do número de subtestes, pois uma tarefa mais breve facilitaria a manutenção da atenção, da concentração e da motivação dos mesmos. Dessa forma, as normas desenvolvidas para um teste mais longo não se aplicariam para a versão mais curta.

Kaufman e Kaufman (2001) também apresentam críticas em relação à segunda estratégia. Eles consideram que a redução de itens de um subteste apresenta validade questionável e ressaltam que algumas das tarefas propostas na WAIS são baseadas na aprendizagem, pois o sucesso em itens mais difíceis dependeria da prática com aqueles mais fáceis. Da mesma forma, Jeyakumar e colaboradores (2004) consideram que a redução do número de itens em um subteste pode causar uma desproporção entre os níveis de dificuldade dos itens que compõem o subteste, diminuindo, assim, a consistência interna do mesmo. Silverstein (1990) ressalta que a aplicação de uma seleção de itens de todos os subtestes é menos fidedigna do que a aplicação de uma combinação de subtestes. Em consequência, evidencia-se que a combinação de subtestes tem sido a estratégia mais aceita entre os pesquisadores e será, portanto, a estratégia investigada neste estudo.

Os subtestes que compõem uma forma curta podem ser escolhidos baseados em diferentes critérios. Sattler (1988) ressalta que a seleção de uma combinação geralmente é baseada em critérios de fidedignidade e validade, tempo disponível para a testagem, capacidade de prover dados clinicamente relevantes e capacidades físicas do examinando. De acordo com o autor, uma forma curta popular é a díade Vocabulário e Cubos, visto que ambos apresentam excelentes coeficientes de fidedignidade, altas correlações com a escala total e são uma boa medida do fator *g* de inteligência. Cyr e Brooker (1984), em estudo realizado com a amostra de normatização da versão americana da WAIS-R, encontraram que o conjunto Vocabulário e Cubos apresentou os melhores coeficientes de validade e fidedignidade entre as formas curtas compostas por dois subtestes. De acordo com Kaufman e Kaufman (2001), a díade Vocabulário-Cubos tem sido a forma curta comumente utilizada em estudos epidemiológicos realizados pelos serviços de saúde, como, por exemplo, estudos que têm por objetivo relacionar alguma característica clínica e o QI. No Brasil, observa-se a mesma tendência em estudos que buscam estimar a inteligência para fins de comparação ou diagnóstico (Mantoan & cols., 2006; Tramontina e cols., 2002).

Dentre os subtestes específicos referentes às medidas verbal e de execução da WAIS-III, o subteste Vocabulário é o que mais se correlaciona à Escala

Verbal e os subtestes Cubos e Completar Figuras são os mais correlacionados à Escala de Execução (Nascimento, 2005). Considerando-se a alta correlação com as escalas Verbal e de Execução e a ampla utilização em estudos, entende-se a possibilidade de utilizar a díade Vocabulário-Cubos no presente estudo.

Conseqüentemente, objetivou-se avaliar a viabilidade de duas estratégias para estimar o QI de um indivíduo. Para tanto, foi investigada a correspondência entre os escores de QI obtidos por meio da escala completa e os escores de QI estimado obtidos por meio das estratégias.

Método

Participantes

Os resultados da WAIS-III de 77 casos previamente coletados foram analisados neste estudo. Em relação aos participantes, 67,5% eram do sexo feminino e 32,5% do sexo masculino, com idade, em média, de 39,1 anos ($dp=13,4$; mínimo=16 e máximo=61) e com escolaridade de 12,2 anos, em média, de estudo formal ($dp=3,6$). Foram incluídos apenas brasileiros natos, falantes do português brasileiro, sem histórico de doenças neurológicas ou psiquiátricas autorrelatadas. Todos os participantes negaram história de uso atual ou prévio de drogas ilícitas e benzodiazepínicos, assim como presença de distúrbios sensoriais – visuais e/ou auditivos – sem correção. Os critérios de inclusão foram verificados a partir de autorrelato do participante, através de entrevista e questionário padronizado.

Procedimentos, instrumentos e descrição das estratégias de QI estimado

Este trabalho é um recorte de uma pesquisa quantitativa, incluindo grupos clínico e controle, de delineamento transversal (Frankfort-Nachmias & Nachmias, 1996), elaborada seguindo os princípios éticos da realização de pesquisa com seres humanos e com aprovação prévia do Comitê de Ética do Hospital de Clínicas de Porto Alegre. Os participantes do grupo controle, do qual uma amostra foi selecionada para o presente estudo, haviam sido previamente selecionados pela técnica de amostragem não-aleatória de conveniência. Todos assinaram um Termo de Consentimento Livre e Esclarecido e, com o propósito de observar os critérios de inclusão, responderam a um questionário de dados sociodemográficos. Após, seguiu-se a aplicação da WAIS-III completa, excluindo o subteste de Armar Objetos, que não é utilizado para o cálculo dos QIs ou índices fatoriais. A coleta de dados e a pontuação da escala foram realizadas por

estudantes e profissionais da psicologia devidamente treinados.

O levantamento dos resultados da WAIS-III foi realizado conforme o manual da terceira edição brasileira da escala, com o objetivo de apurar o resultado de QI total. Os resultados brutos dos subtestes Vocabulário e Cubos de cada participante foram convertidos em resultados ponderados de acordo com a idade, conforme tabelas A.1 do manual da WAIS-III edição brasileira. Para o cálculo do QI estimado 1, os escores ponderados dos dois subtestes foram somados e o valor do QI estimado foi estabelecido baseado nas tabelas propostas por Jeyakumar e colaboradores (2004) para os subtestes Vocabulário e Cubos. Para o cálculo do QI Estimado 2, foi calculada a média da soma dos resultados ponderados dos dois subtestes e o resultado foi convertido em escore de QI por meio de correspondência direta ponto ponderado-QI. O resultado ponderado dos subtestes, tendo média 10 e desvio padrão 3, é convertido proporcionalmente nas escalas de escores de QI, com média 100 e desvio-padrão 15, a qual se refere ao QI total (Nascimento, 2005). Segundo a versão brasileira do manual da WAIS-III, a normalização da soma dos resultados ponderados de todos os subtestes corresponde à Escala de QI Total (Nascimento, 2005). Considerando-se a presença de alta correlação dos subtestes Vocabulário e Cubos com o escore total de QI, a conversão da média dos escores ponderados desses subtestes em escores de QI pode ser aplicada como uma estimativa do QI total.

Análise dos dados

Foram realizadas análises descritivas para caracterização da amostra através do Programa SPSS versão 13. O teste *t* pareado³ foi aplicado para comparação da média geral entre os métodos de QI estimado e o QI fornecido pela Escala Total. Foi calculada a porcentagem de casos que apresentaram diferenças de escore acima de 5 pontos e acima de 10 pontos de QI na comparação de cada QI Estimado com o QI da Escala Total, de acordo com o estudo realizado por Ringe, Saine, Lacritz, Hynan e Cullum (2002). Verificou-se a porcentagem de classificações de QI estimado que se mantiveram iguais às fornecidas pelo QI avaliado mediante a aplicação de todos os subtestes.

Para a avaliação da concordância de cada estratégia de exame de QI estimado com o escore de

³ O teste *t* pareado foi realizado com a finalidade de indicar e interpretar os resultados por meio de análise estatística tradicionalmente utilizada em estudos psicométricos e compará-la a uma proposta nova de análise.

QI da Escala Total, foi utilizado o método estatístico proposto por Bland e Altman (1986) e denominado metodologia de Bland-Altman, criado para avaliar a concordância entre dois métodos de medida, um já existente e um novo. Essa análise avalia a relação entre a diferença dos escores dos dois métodos e a média entre eles, considerada por Bland e Altman (1986) como a melhor estimativa na comparação dos dois métodos. Essa relação é verificada para cada participante, além de ser avaliada a média de todas as diferenças da amostra, a qual é denominada viés estimado. Esse viés indica quanto, em média, os dois métodos apresentam de discrepância entre si, tal como pode ser verificado com a utilização do teste *t* pareado. A metodologia de Bland-Altman apresenta a vantagem de comparar graficamente os dados caso a caso, apontando os limites de concordância entre os métodos avaliados. Os limites de concordância são definidos a partir de dois desvios padrão acima e abaixo da diferença média. Assim, se os limites de concordância entre os dois métodos não forem clinicamente importantes, ambos podem ser usados sem distinção. Para a análise do presente estudo, considerar-se-á aceitável a diferença máxima de 5 pontos entre a Escala Total e cada método de estimativa do QI. Esse valor foi adotado porque se entende que uma variação de 5 pontos no QI para cima ou para baixo não seja clinicamente relevante, considerando que o intervalo entre a pontuação de QIs na WAIS-III é de 10 pontos para que a classificação qualitativa do QI do indivíduo mude.

Resultados

No que se refere às características de nível intelectual, a amostra estudada inclui participantes com níveis de QI de 87 a 134, cuja média foi de 112,74 ($dp=10,28$). Esses resultados indicam que a amostra apresentava, em média, classificação de QI em nível médio superior, variando de médio inferior a muito superior. Como resultado das análises inferenciais, a análise de concordância pelo teste *t* para medidas repetidas indicou diferença significativa ($t=-3,11$; $gl=76$; $p=0,03$) entre QI total ($M=112,74$; $dp=10,28$) e QI estimado 1 ($M=110,75$; $dp=11,90$) e diferença significativa ($t=5,96$; $gl=76$; $p<0,001$) entre QI total ($M=112,74$; $dp=10,28$) e QI estimado 2 ($M=109,35$; $dp=10,27$).

Na análise da porcentagem de casos que apresentaram diferenças entre a medida de QI total e o QI estimado, 32,5% dos casos de QI estimado 1 e 39% dos casos de QI estimado 2 diferiram em mais de 5 pontos. Para diferenças acima de 10 pontos, foram encontrados, em ambos os métodos, 9,1% dos casos.

Na avaliação da quantidade de casos que mudaram de classificação quando comparados ao QI da Escala Total, foram encontrados 31,2% de casos de QI estimado 1 e 29,9% de casos de QI estimado 2.

No que se refere aos resultados da análise Bland-Altman em relação à concordância entre o QI total e o QI estimado 1, foi encontrado um viés estatisticamente significativo de 2,0 pontos (IC95%: 0,7-3,3), indicando que QI estimado 1, em média, subestima em 2 pontos o QI total. Os limites de concordância demonstrados na Figura 1 indicam que as diferenças entre os dois métodos podem variar entre -9,0 e 13,0 pontos.

Sobre o QI estimado 2, os resultados da análise Bland-Altman em relação à concordância com o QI total indicam um viés estatisticamente significativo de 3,4 pontos (IC95%: 2,3-4,5), indicando que QI estimado 2, em média, subestima em 3,4 pontos o QI total. Na Figura 2, os limites de concordância apontam que as diferenças entre os dois métodos podem variar de -6,4 a 13,2 pontos.

Discussão

Os resultados do teste *t* para medidas repetidas indicaram diferenças estatisticamente significativas entre cada estratégia e o QI fornecido pela Escala Total, sendo de 2 pontos para o QI estimado 1 e de 3,39 pontos para o QI estimado 2, diferenças que não se mostram clinicamente importantes. A partir desses dados, poder-se-ia considerar viável a utilização das duas estratégias propostas neste estudo. Contudo, a porcentagem de casos que apresentaram diferenças de escore acima de 5 pontos e acima de 10 pontos de QI na comparação QI estimado e QI Escala Total indicam a possibilidade de erros de classificação quando este é apenas estimado. O percentual de mudanças de classificação na amostra avaliada também foi elevado, apesar de a classificação se modificar, em geral, em um nível apenas. É necessário considerar, ainda, que a amostra estudada possuía níveis médios de QI ou maiores, sendo essencial a investigação com uma amostra mais heterogênea nos níveis de QI para avaliar se os resultados se aplicam também à população com QIs distantes da média.

Esses dados são confirmados e mais bem demonstrados pela metodologia de Bland-Altman, que indicou diferenças que podem variar de -9,0 a 13,0 pontos para o QI estimado 1 e de -6,4 a 13,2 para o QI estimado 2 quando considerado um limite de concordância de 95%. Adotando um limite máximo de 5 pontos de diferença como aceitável para fins clínicos, uma parte significativa dos valores obtidos pelos QIs estimados 1 e 2 não é clinicamente válida.

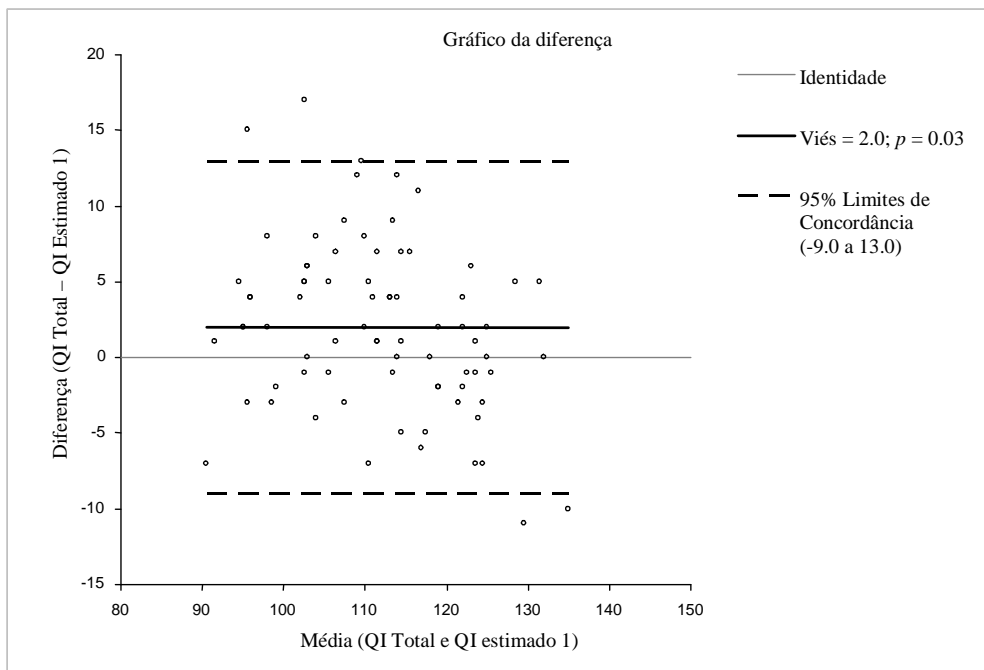


Figura 1 – Gráfico da diferença QI total – QI estimado 1

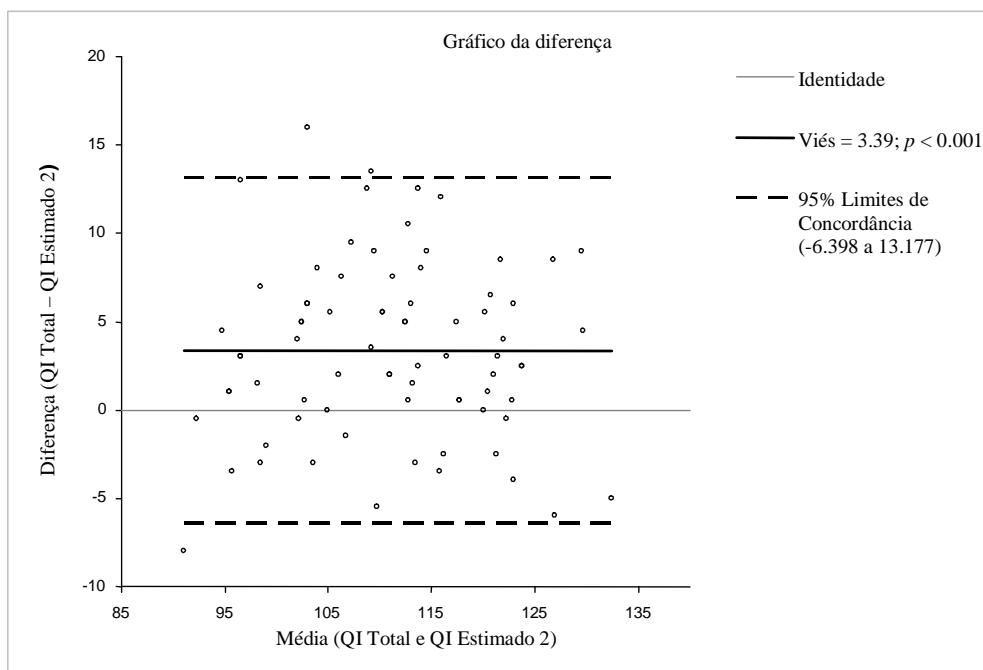


Figura 2 – Gráfico da diferença QI total – QI estimado 2

Assim, embora a média de todos os casos e alguns casos individualmente tenham apresentado níveis de concordância aceitáveis entre os QIs estimados e o QI avaliado por todos os subtestes, cerca de um terço dos casos apresentaram uma diferença clínica relevante.

Em razão disso, recomenda-se cautela na utilização desses métodos de obtenção do QI. O

primeiro método, QI estimado 1, foi desenvolvido considerando os dados da população americana de normatização da WAIS-III (Wechsler, 1997). Dessa forma, não se trata de um método adaptado para a população brasileira e, portanto, apresenta limitações. O segundo método, QI estimado 2, transforma a média dos escores ponderados de dois subtestes que

compõem a forma curta em escores de QI. Considerando que as análises de precisão da Escala Wechsler foram realizadas a partir da aplicação de 11 subtestes e que para a obtenção de um valor de QI total preciso devam ser somados os escores ponderados dos 11 subtestes aplicados, pode-se pensar que esta estimativa de QI, com base em apenas dois subtestes seja pouco precisa.

Sattler (1988) ressalta que os escores fornecidos por formas curtas são menos estáveis do que aqueles fornecidos pela escala total e, além disso, a fidedignidade do QI obtido é reduzida quando subtestes são eliminados. Especificamente em relação à diáde utilizada neste estudo, Thompson (1987), citado por Kaufman e Kaufman (2001), conduziu uma pesquisa com pacientes psiquiátricos, comparando a aplicação da forma curta Vocabulário e Cubos, seguidos dos demais subtestes, e da aplicação da bateria completa, em que os subtestes Vocabulário e Cubos eram o segundo e o quinto na ordem de aplicação, respectivamente. Os resultados apontaram que os pacientes que inicialmente fizeram os subtestes Vocabulário e Cubos obtiveram estimativas de QI significativamente mais altas em relação ao QI fornecido pela bateria completa. Esses resultados podem estar relacionados à diminuição da atenção, concentração e motivação durante a aplicação da bateria completa, o que pode ser mais evidente em grupos que cansam mais rapidamente e/ou apresentam dificuldade de sustentar a atenção, tais como idosos (Valerianni, Ranghi & Giaquinto, 2003). Também pode ocorrer uma redução das funções executivas associadas à fadiga, em especial, das capacidades de planejamento e de tomada de decisões, percebidas entre idosos (Mejia, Pineda, Alvarez & Ardila, 1998; Albert, 1993) e em pacientes com lesões cerebrais, tais como as decorrentes de doença vascular isquêmica subcortical (Alves & cols., 2008).

Outras limitações apontadas sobre a utilização de formas curtas é a perda de informações sobre o padrão cognitivo dos examinandos, o que pode ser especialmente importante em avaliações cognitivas e neuropsicológicas. Da mesma forma, o padrão de resolução de problemas não pode ser observado, limitando a compreensão das potencialidades e dificuldades de cada paciente (Sattler, 1988).

Assim, há um consenso na literatura sobre formas curtas de que o seu uso é adequado apenas em alguns contextos, por exemplo, como triagem, ou em pesquisas, para caracterizar o nível intelectual de um grupo, e quando a avaliação da inteligência é periférica e a administração da forma curta pode ser seguida pela administração dos subtestes que faltaram (Watkins Jr., 1986). Além disso, os autores destacam que formas

curtas não são recomendadas para decisões importantes com base nos seus resultados (Jeyakumar & cols., 2004; Ringe, Saine, Lacritz, Hynan & Cullum, 2002; Sattler, 1988; Silverstein, 1990).

Nesse sentido, Kaufman e Kaufman (2001) recomendam a utilização de escalas breves, tais como a *Wechsler Abbreviated Scale of Intelligence* – WASI (The Psychological Corporation, 1999). A WASI é composta por quatro subtestes similares aos existentes na WAIS-III, Vocabulário, Cubos, Semelhanças e Raciocínio Matricial, porém compostos por itens diferentes. O instrumento, adaptado para a população americana, apresenta normas para pessoas entre 6 e 89 anos e fornece três índices compostos: QI total, QI verbal e QI de execução. São fornecidas tabelas para determinação do QI total a partir dos quatro subtestes que compõem o instrumento e da combinação de dois subtestes (Vocabulário e Raciocínio Matricial). Kaufman e Kaufman (2001) destacam que escalas breves possuem normas próprias, bem como estudos sobre validade e fidedignidade. No entanto, eles não sugerem que escalas breves possam substituir escalas completas. Assim como as formas curtas, as escalas breves devem ser usadas em situações apropriadas, tais como pesquisas e casos em que a medida da inteligência não é o foco central. Ressalta-se que, no Brasil, está ocorrendo o processo de adaptação e validação da Escala WASI para a população brasileira (Yates & cols., 2006).

Alguns autores, contudo, têm questionado o desempenho da WASI. Axelrod (2002) ressalta que a WASI obteve correlações mais baixas que formas curtas em relação à WAIS-III em estudo conduzido com versões americanas dos instrumentos. A partir disso, o autor sugere cautela na utilização desse instrumento, especialmente quando há necessidade de uma estimativa fidedigna do QI. Ringe e colaboradores (2002) também discutem a possibilidade da WASI apresentar estimativas errôneas do QI em pacientes com determinados transtornos, dependendo da sensibilidade dos subtestes para cada caso em específico. Eles complementam que estimativas de QI fornecidas através de formas curtas podem apresentar algumas vantagens se comparadas à WASI.

No que se refere ao presente estudo, ressaltam-se algumas limitações. A amostra poderia ser mais ampla, incluindo diversas faixas etárias, de escolaridade e de níveis de QI, possibilitando verificar a interferência de variáveis sociodemográficas na viabilidade de utilização das estratégias de estimar o QI. Outra questão que merece atenção refere-se ao fato de os subtestes Vocabulário e Cubos terem sido aplicados junto aos demais subtestes. Esses subtestes, quando aplicados sozinhos, podem diminuir o efeito

de fadiga e aumentar a motivação do testando, o que influenciaria o desempenho, melhorando os escores destes. Além disso, o cálculo do QI total incluiu os subtestes que também foram utilizados para a conversão de resultados em QI estimado. Portanto, é possível que a diferença de escores proveniente da aplicação de apenas Vocabulário e Cubos seja ainda maior quando comparado ao QI obtido pela Escala Total. É importante ressaltar, contudo, que esta não é uma limitação apenas deste estudo, pois diversas outras pesquisas com formas curtas, exceto Thompson, Howard e Anderson (1986), utilizaram o mesmo procedimento, ou seja, aplicaram a bateria completa e após estimaram o QI.

Dessa forma, sugere-se que procedimentos mais rígidos sejam aplicados para a replicação das análises realizadas. Por exemplo, poderia ser realizado um estudo com dois tempos de aplicação na mesma amostra, um incluindo a Escala Total e outro só os subtestes Vocabulário e Cubos; entretanto, estaria presente o fator reteste e a possibilidade de aprendizagem dos subtestes. Outra possibilidade seria um estudo transversal, em que uma amostra responderia à Escala Total e outra apenas os subtestes Vocabulário e Cubos. Tais amostras deveriam ser pareadas em relação a diversas variáveis sociodemográficas, para reduzir a limitação de não serem comparados os mesmos sujeitos. Outra possibilidade seria a divisão da amostra em que metade dos casos respondessem a Vocabulário e Cubos primeiro, seguidos dos demais subtestes, e a outra metade respondesse aos subtestes na ordem sugerida no manual, sendo avaliado o possível efeito de fadiga.

Apesar das limitações do presente estudo, destaca-se a utilização da análise proposta por Bland e Altman (1986), que demonstra o quanto o QI Estimado pode variar em relação ao QI total, caso a caso. Essa análise permitiu a visualização gráfica das diferenças entre as duas formas de medida e de um possível padrão de variação dos escores. Também possibilitou evidenciar o impacto que uma estimativa de QI errônea pode acarretar, o que não é identificado mediante análises da média dos escores de um grupo, visto que pode ocorrer a presença de mascaramento dos resultados em razão da existência de casos discrepantes.

Considerações finais

A literatura a respeito de formas curtas das escalas Wechsler apresenta controvérsias. Existem dúvidas em relação ao número de subtestes, quais seriam os mais adequados, a forma mais fidedigna de transformar os resultados em escores de QI, os

métodos de verificação de validade e fidedignidade, entre outras. Diante disso, percebe-se a relevância de estudos sobre métodos adequados para estimar rapidamente o QI.

Considerando as limitações do presente estudo, bem como das outras propostas de procedimento sugeridas, entende-se a complexidade de verificar a viabilidade de utilizar o QI estimado a partir de formas curtas. As estratégias de obtenção do QI estimado apresentadas, para serem consideradas viáveis e precisas, necessitariam ser desenvolvidas para a população brasileira da mesma forma que o foram para a população americana, bem como apresentar estudos sobre coeficientes de validade e fidedignidade em amostras maiores, comparando a viabilidade da aplicação de formas curtas. Como exemplo disso, em estudo recente, Coutinho (2009) conduziu análises psicométricas com escalas abreviadas, encontrando resultados de coeficientes de precisão altamente significativos para quatro formas abreviadas da WAIS-III. Essa investigação também indicou que apenas as composições com quatro e oito subtestes foram consideradas substitutas aceitáveis da versão completa do WAIS-III na avaliação do funcionamento intelectual.

Pelo presente estudo, pôde-se observar as limitações das estratégias utilizadas no Brasil para obtenção de estimativas de QI baseadas na composição dos subtestes Vocabulário e Cubos. Apesar da estratégia QI estimado 1 ter apresentado escores mais próximos aos escores da Escala Total, se comparada à estratégia QI estimado 2, ambas evidenciaram casos com discrepâncias significativas entre QI total e QI estimado. Portanto, na amostra estudada, as estratégias utilizadas para estimar o QI não foram adequadas em uma parcela importante de casos.

Este estudo consistiu em uma contribuição aos procedimentos iniciais de avaliação de validade e fidedignidade de formas curtas obtidas com base em dois subtestes. Destaca-se a importância de novas pesquisas sobre formas curtas da Escala Wechsler utilizando os métodos internacionalmente aplicados, com o objetivo de fornecer evidências de validade e fidedignidade de estimativas de QI para o contexto nacional.

Referências

- Albert, M. (1993). Neuropsychological and neurophysiological changes in healthy adult humans across the age range. *Neurobiology of Aging*, 14, 623–625.
- Alves, G. S., Alves, C. E. O., Lanna, M. E., Moreira,

- D. M., Engelhardt, E. & Laks, J. (2008). Subcortical ischemic vascular disease and cognition. A systematic review. *Dementia & Neuropsychologia*, 2(2), 82-90.
- Axelrod, B. N. (2002). Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms of estimating intellectual functioning. *Assessment*, 9(1), 17-23.
- Bland, J. M. & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307-310.
- Coutinho, A. C. A. M. (2009). Investigação psicométrica de quatro formas abreviadas do WAIS-III para avaliação da inteligência. *Resumos Mesas Redondas do IV Congresso Brasileiro de Avaliação Psicológica*, Campinas, 190-191.
- Cyr, J. J. & Brooker, B. H. (1984). Use of appropriate formulas for selecting WAIS-R short forms. *Journal of Consulting and Clinical Psychology*, 52(5), 903-905.
- Figueiredo, V. L. M. (2002). *WISC-III: Escala de Inteligência Wechsler para Crianças – adaptação brasileira da 3ª edição*. São Paulo: Casa do Psicólogo.
- Frankfort-Nachmias, C. & Nachmias, D. (1996). *Research Methods in the Social Sciences*. Londres: Arnold.
- Jeyakumar, S. L. E., Warriner, E. M., Raval, V., V. & Ahmad, S. A. (2004). Balancing the need for reliability and time efficiency: short forms of the Wechsler Adult Intelligence Scale-III. *Educational and Psychological Measurement*, 64, 71-86.
- Kaufman, J. C. & Kaufman, A. S. (2001) Time for changing of the guard: a farewell to short forms of intelligence tests. *Journal of Psychoeducational Assessment*, 19, 245-267.
- Kenny, J. T., Friedman, L., Findling, R. L., Swales, T. P., Strauss, M. E., Jesberger, J. A. & Schulz, S. C. (1997). Cognitive impairment in adolescents with schizophrenia. *American Journal of Psychiatry*, 154(11), 1613-1615.
- Mantoan, M. A. S., Silva, T. I., Alonso, N. B., Noffs, M. H. S., Marques, C. M., Rios, L. B., Azevedo, A. M., Westphal-Guitti, A. C., Sakamoto, A. C. & Yacubian, E. M. T. (2006). Neuropsychological assessment and quality of life in patients with refractory temporal lobe epilepsy related to hippocampal sclerosis. *Journal of Epilepsy and Clinical Neurophysiology*, 12(4), 201-206.
- Mejia, S., Pineda, D., Alvarez, L. M. & Ardila, A. (1998). Individual differences in memory and executive function abilities during normal aging. *International Journal of Neuroscience*, 95, 271-284.
- Nascimento, E. (2005). *WAIS-III: Escala de Inteligência Wechsler para Adultos – manual técnico*. São Paulo: Casa do Psicólogo.
- Nascimento, E. & Figueiredo, V. L. M. (2002a). A terceira edição das Escalas Wechsler de Inteligência. Em R. Primi (Org.). *Temas em Avaliação Psicológica* (pp. 61-79). Campinas: IBAP – Instituto Brasileiro de Avaliação Psicológica.
- Nascimento, E. & Figueiredo, V. L. M. (2002b). WISC-III e WAIS-III: alterações nas versões originais americanas decorrentes das adaptações para uso no Brasil. *Psicologia: Reflexão e Crítica*, 15(3), 603-612.
- Rabin, A. I. (1943). A short form of the Wechsler-Bellevue test. *Journal of Applied Psychology*, 27, 320-324.
- Ringe, W. K., Saine, K. C., Lacritz, L. H., Hynan, L. S. & Cullum, C. M. (2002). Dyadic short forms of the Wechsler Adult Intelligence Scale-III. *Assessment*, 9(3), 254-260.
- Sattler, J. M. (1988). *Assessment of children*. San Diego: Jerome M. Sattler.
- Silverstein, A. B. (1984). Estimating full scale IQs from short forms of Wechsler's Scales: linear scaling versus linear regression. *Journal of Consulting and Clinical Psychology*, 52(5), 919.
- Silverstein, A. B. (1990). Short forms of individual intelligence tests. *A Journal of Consulting and Clinical Psychology*, 2, 3-11.
- Tellegen, A. & Briggs, P. F. (1967). Old wine in new skins: grouping Wechsler subtests in new scales. *Journal of Consulting Psychology*, 31, 499-506.
- The Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence Manual*. San Antonio: Psychological Corporation.
- Thompson, A. P., Howard, D. & Anderson, J. (1986). Two- and four-subtest short forms of the WAIS-R: validity in a psychiatric sample. *Canadian Journal of Behavioral Science*, 18(3), 287-293.
- Thompson, S. J., Leigh, L., Christensen, R., Xiong, X., Kun, L. E., Heideman, R. L., Reddick, W. E., Gajjar, A., Merchant, T., Pui, C. H., Hudson, M. M. & Mulhern, R. K. (2001). Immediate neurocognitive effects of methylphenidate on

learning-impaired survivors of childhood cancer. *Journal of Clinical Oncology*, 19(6), 1802-1808.

Tramontina, S., Martins, S., Michalowski, M. B., Ketzer, C. R., Eizirik, M., Biederman, J. & Rohde, L. A. (2002). Estimated mental retardation and school dropout in a sample of students from state public schools in Porto Alegre, Brazil. *Revista Brasileira de Psiquiatria*, 24(4), 177-181.

Valerianni, M., Ranghi, F. & Giaquinto, S. (2003). The effects of aging on selective attention to touch: a reduced inhibitory control in elderly subjects? *International Journal of Psychophysiology*, 49, 75-87.

Watkins Jr., C. E. (1986). Validity and usefulness of WAIS-R, WISC-R, and WPPSI short forms: a

critical review. *Professional Psychology: Research and Practice*, 17(1), 36-43.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III*. San Antonio, TX: The Psychological Corporation.

Yates, D. B., Trentini, C. M., Tosi, S. D., Corrêa, S. K., Poggere, L. C. & Valli, F. (2006). Apresentação da Escala de Inteligência Wechsler abreviada (WASI). *Avaliação Psicológica*, 5(2), 227-233.

Recebido em junho de 2009

Reformulado em fevereiro de 2010

Aprovado em abril de 2010

Sobre as autoras:

Flávia Wagner é psicóloga, mestranda em Psicologia pelo Programa de Pós-Graduação em Psicologia da Universidade Federal do Rio Grande do Sul (UFRGS) e especialista em Psicologia Clínica – Ênfase em Avaliação Psicológica (UFRGS).

Josiane Pawlowski é psicóloga, doutoranda em Psicologia pelo Programa de Pós-Graduação em Psicologia da Universidade Federal do Rio Grande do Sul (UFRGS), mestre em Psicologia (UFRGS) e especialista em Psicologia Clínica – Ênfase em Avaliação Psicológica (UFRGS).

Denise Balem Yates é psicóloga, doutoranda em Psicologia pelo Programa de Pós-Graduação em Psicologia da Universidade Federal do Rio Grande do Sul (UFRGS), mestre em Psicologia (UFRGS) e especialista em Neuropsicologia (UFRGS).

Suzi Alves Camey é estatística, doutora em Estatística (USP) e mestre em Matemática (UFRGS). Atualmente é professora dos cursos de graduação em Estatística e pós-graduação em Epidemiologia da UFRGS,

Clarissa Marcell Trentini é psicóloga, doutora em Ciências Médicas: Psiquiatria pela Universidade Federal do Rio Grande do Sul (UFRGS) e professora do Instituto de Psicologia e do Programa de Pós-Graduação em Psicologia da UFRGS.