

Estudo de fidedignidade do instrumento neuropsicológico *Iowa Gambling Task*

Reliability study of the neuropsychological test Iowa Gambling Task

Caroline de Oliveira Cardoso¹, Janaína Castro Núñez Carvalho², Charles Cotrena¹, Daniela di Giorgio Schneider Bakos³, Christian Haag Kristensen¹, Rochele Paz Fonseca¹

RESUMO

Objetivo: O objetivo do presente estudo foi verificar evidências de fidedignidade do instrumento neuropsicológico *Iowa Gambling Task* (IGT) a partir do método teste-reteste. **Método:** Participaram 50 indivíduos saudáveis, de 19 a 75 anos de idade, com no mínimo cinco anos de educação formal. A aplicação foi realizada de forma individual, em dois encontros, com intervalo de um a seis meses entre teste e reteste. **Resultados:** Os resultados evidenciaram uma correlação positiva moderada significativa entre teste-reteste no cálculo global. Na análise por segmentos, os blocos 4 e 5 apresentaram uma correlação positiva moderada, mas não foram observadas correlações significativas nos blocos 1, 2 e 3. **Conclusão:** Esses dados corroboram estudos atuais que encontraram correlações moderadas entre teste-reteste em medidas de funções executivas e sugerem que o IGT pode ser empregado para avaliar o processo de tomada de decisão de forma confiável ao longo do tempo, desde que sejam considerados estudos de fidedignidade com populações saudáveis mais amplas e com populações clínicas.

Palavras-chave

Funções executivas, tomada de decisão, *Iowa Gambling Task*, fidedignidade.

ABSTRACT

Objective: This study aimed at evaluating reliability evidence of neuropsychological instrument *Iowa Gambling Task* (IGT) by the means of the test-retest method. **Method:** The sample was comprised of 50 healthy subjects, 19-75 years of age, with at least five years of formal education. The assessment was done individually, in two meetings with an interval from 1 to 6 months between test and retest. **Results:** The findings showed a significant moderate positive correlation between test-retest in the overall calculation. In the analysis by segments, a significant moderate positive correlation was found for blocks 4 and 5, whereas there were no significant correlations for blocks 1, 2 and 3. **Conclusion:** These data corroborate recent studies that found moderate correlations between test-retest measures of executive functions, and suggest that the IGT can be used to assess decision making over time if specific analyses are taken into account.

Keywords

Executive functions, decision making, *Iowa Gambling Task*, reliability.

Recebido em
26/7/2010
Aprovado em
5/10/2010

1 Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS), Departamento de Psicologia.

2 Instituto de Educação Superior de Brasília, Departamento de Psicologia.

3 Universidade Luterana do Brasil, Departamento de Psicologia.

Endereço para correspondência: Caroline de Oliveira Cardoso
Programa de Pós-Graduação em Psicologia – Cognição Humana, Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS)
Av. Ipiranga, 6681, prédio 11, sala 932 – 90619-900 – Porto Alegre, RS
Telefone: (51) 3320-3500 (ext. 7742)/Telefax: (51) 3320-3633
E-mail: carolinecardoso@yahoo.com.br

INTRODUÇÃO

A avaliação neuropsicológica corresponde ao método de investigação do funcionamento cerebral com procedimentos comportamentais¹ e está ganhando cada vez mais espaço como procedimento adjunto à prática médica para a análise e o diagnóstico de disfunções cerebrais. No contexto clínico, é vastamente reconhecido que muitos distúrbios psiquiátricos apresentam algumas alterações cognitivas e a exploração dessas funções mentais torna-se extremamente útil para a compreensão da etiologia e da fisiopatologia dos transtornos^{2,3}.

Dos processos cognitivos investigados pela neuropsicologia, as funções executivas (FE) são as que mais exigem exploração em nível conceitual e empírico, uma vez que elas desempenham um papel-chave na caracterização de diferentes doenças neurológicas e neuropsiquiátricas⁴. A denominação FE é utilizada para designar uma ampla variedade de funções cognitivas, incluindo planejamento, flexibilidade, raciocínio verbal, tomada de decisão, entre outras. Em sentido geral, esses processos têm por objetivo facilitar a adaptação do indivíduo a situações novas⁵.

Dentre os diversos componentes que fazem parte das FE, destaca-se o processo de tomada de decisão. A avaliação neuropsicológica desse componente é de extrema importância no âmbito clínico, já que a tomada de decisão parece estar alterada em diversos transtornos neuropsiquiátricos². Trata-se de um processo dinâmico que favorece a eleição da alternativa mais adaptada entre inúmeras opções de respostas, valorizando sua influência em ações futuras⁶. Durante muito tempo, a tomada de decisão foi estudada na psicologia cognitiva e na economia como sendo o resultado de uma escolha racional das alternativas disponíveis. Na década de 1990, ao propor a Hipótese do Marcador Somático, Damásio⁷ destaca a emoção e seus correlatos biológicos como aspectos fundamentais na tomada de decisão.

Damásio⁷ salienta que, quando uma pessoa necessita tomar uma decisão, antes de fazer uma análise de custo e benefício da situação, o indivíduo se depara com um estado afetivo, ou melhor, com uma sensação corporal automática, conhecida como um conjunto de “marcadores somáticos”. Essas respostas somáticas são sinais emocionais que antecipam as possíveis consequências de distintas alternativas, induzindo à decisão⁸. A eleição do que é vantajoso ou desvantajoso é baseada nas experiências passadas de cada indivíduo, naquilo que foi sendo registrado ao longo de sua vida como prazeroso e desprazeroso. Segundo o autor, uma das principais estruturas cerebrais envolvidas nesse mecanismo é o córtex pré-frontal ventromedial.

Na tentativa de testar empiricamente sua hipótese teórica, Bechara *et al.* desenvolveram o paradigma clínico do IGT⁹. Esse instrumento neuropsicológico vem sendo um grande referencial na avaliação da tomada de decisão sob incerteza, um dos componentes das FE^{9,10}.

O IGT examina o comportamento do indivíduo em um jogo de cartas composto por quatro baralhos. Ao longo da tarefa, o examinando deve escolher uma carta dentre quatro baralhos ao longo de 100 jogadas, com o objetivo de acumular o máximo de dinheiro possível. São quatro baralhos de cartas, sendo dois vantajosos (C e D), resultando em ganhos monetários a longo prazo e baixa perda de dinheiro, e os outros dois baralhos, desvantajosos (A e B), pois trazem ganhos de muito dinheiro a curto prazo, porém com perda monetária mais frequente e intensa⁹. Por envolver escolhas monetárias em curto e longo prazo, permite classificar o comportamento decisório em adaptativo ou prejudicado⁸.

Esse paradigma clínico foi inicialmente desenvolvido para avaliar a tomada de decisão em pacientes com lesão no córtex pré-frontal ventromedial¹⁰. Os autores perceberam que esses pacientes, apesar de apresentarem funcionamento intelectual preservado, não tinham os marcadores somáticos e por isso apresentavam “miopia para o futuro”, ou seja, denotavam uma incapacidade de antecipar as consequências de suas ações.

Outras pesquisas comprovam que o IGT pode ser aplicado na avaliação de diversas patologias que supostamente estão associadas a prejuízos no processo de tomada de decisão³, como é o caso de dependentes químicos¹¹, pacientes com transtorno obsessivo-compulsivo¹², indivíduos com transtorno de déficit de atenção com hiperatividade (TDAH)¹³, entre outros. Estudos demonstram que, diferentemente dos indivíduos normais que no decorrer do teste tendem a escolher os baralhos mais vantajosos, os pacientes com algum transtorno neuropsiquiátrico mantêm um padrão de escolha imediatista, ou seja, tendem a optar por ganhos maiores em curto prazo, apresentando um comportamento considerado prejudicado¹⁴.

No Brasil, o IGT vem sendo objeto de estudo em dois grupos de neuropsicologia. No Sudeste, estudos realizados por Malloy-Diniz *et al.*¹³ buscaram desenvolver e adaptar a versão original em inglês do IGT para o português brasileiro. E no sul do país, Schneider⁸ também se interessaram em adaptar e normatizar o instrumento para a população local. O processo de normatização do IGT ainda se encontra em andamento no país. Pesquisas atuais foram realizadas objetivando avaliar o papel das variáveis idade e escolaridade no desempenho do IGT, bem como buscar evidências de validade e correlacionar o IGT com outros instrumentos que mensuram diferentes componentes das FE¹⁵.

Apesar de ser um instrumento internacionalmente utilizado para avaliar a tomada de decisão, o IGT ainda carece de evidências de fidedignidade e de dados normativos tanto com relação a populações clínicas quanto a participantes saudáveis. No processo de adaptação e validação de um instrumento neuropsicológico estrangeiro para a realidade local, é necessária a busca de evidências de parâmetros psicométricos para que o instrumento seja considerado fidedigno

e válido^{16,17}. Em geral, o primeiro procedimento empírico a ser conduzido na adaptação de um instrumento é o estudo de fidedignidade. Para a busca de evidências de fidedignidade temporal (estabilidade), um dos métodos mais conhecidos é o do teste-reteste¹⁶.

Na avaliação da fidedignidade de instrumentos psicológicos, as medições são muito propensas a erros em virtude da natureza dos construtos avaliados e de inúmeros fatores externos que se tornam difíceis de controlar¹⁷. Na confiabilidade de instrumentos destinados a mensurar as FE, alguns autores apontam que há uma baixa correlação entre teste-reteste¹⁸, por causa das peculiaridades do construto avaliado. Estudos atuais, porém, vêm desafiando essa ideia ao encontrarem uma correlação moderada positiva entre teste-reteste de medidas de FE^{19,20}. Poucas investigações apresentaram um posicionamento consensual sobre qual índice entre teste e reteste pode ser considerado suficiente para atestar sobre a confiabilidade de alguma tarefa de FE.

No caso do IGT, no manual americano do instrumento não consta uma análise da sua confiabilidade nem se comenta sobre a aplicabilidade com a finalidade de reavaliação¹⁴. Dessa forma, até onde se sabe, não parece haver nenhum estudo que tenha se proposto a examinar diretamente a fidedignidade desse instrumento. Apenas há na literatura pesquisas que de forma indireta aplicaram o IGT em dois momentos distintos a uma mesma população, buscando avaliar o efeito de aprendizagem^{21,22}.

Diante do provável desconhecimento da estabilidade temporal da avaliação com o IGT, torna-se bastante relevante investigar esse parâmetro em busca de um melhor entendimento quanto a efeitos de aprendizagem e de exposição anterior ao teste – sobretudo ao se considerar a natureza executiva dessa tarefa. Tais resultados podem contribuir para a reflexão sobre os processos de avaliação e reavaliação clínica de pacientes neurológicos e/ou psiquiátricos

MÉTODO

Participantes

Participaram deste estudo 50 indivíduos entre 19 e 75 anos de idade, com no mínimo cinco anos de educação formal, sexo feminino (70% – n = 35) e masculino (30% – n = 15). Foram selecionados em ambientes universitários, de trabalho e em centros de convivência, sendo a amostragem por conveniência. Ingressaram na amostra somente os participantes que, além dos critérios de idade e escolaridade, não apresentassem: (a) quaisquer distúrbios sensoriais (auditivos e/ou visuais) não corrigidos; (b) sinais sugestivos de quadros neurológicos e/ou psiquiátricos, examinados pelo questionário sociocultural e de aspectos de saúde²³; (c) sinais de depressão mensurados pelo Inventário Beck de Depressão (BDI-II) (Beck; Steer; Brown, 1996, adaptado ao português brasileiro

por Finger e Argimon²⁴, com escore > 19); (d) demência, em triagem feita pelo Miniexame do Estado Mental (MEEM – adaptado para a população local por Chaves e Izquierdo²⁵, com escore ≥ 24); (e) histórico de alcoolismo, triados com a escala CAGE (versão utilizada no estudo de Amaral e Malbergier²⁶, com escore ≥ 2); (f) relato de uso abusivo atual ou prévio de drogas ilícitas ou de benzodiazepínicos (avaliado por autorrelato em um questionário sociocultural); e (g) escores ponderados inferiores a sete nos subtestes Vocabulário e Cubos do WAIS-III²⁷. A amostra foi composta apenas por participantes brasileiros natos e falantes do português brasileiro. A tabela 1 apresenta a caracterização sociodemográfica e clínica da amostra.

Tabela 1. Caracterização da amostra

	Variáveis de caracterização da amostra	M (dp)
Variáveis sociodemográficas	Idade (anos)	36,50 (19,55)
	Escolaridade (anos)	14,34 (3,92)
	Escore socioeconômico (escore médio)	27,86 (7,46)
	Hábitos leitura e escrita (frequência)	17,10 (5,43)
Variáveis clínicas	Escore MEEM	28,48 (1,69)
	Escore BDI-II	6,58 (4,353)
	Vocabulário WAIS-III (escore ponderado)	10,78 (1,71)
	Cubos WAIS-III (escore ponderado)	13,02 (2,51)

M: média; dp: desvio-padrão.

Os dados expostos na tabela 1 permitem observar que os participantes são predominantemente adultos jovens e de idade intermediária, de alta escolaridade, de classe socioeconômica média B1 no Brasil (escore socioeconômico médio entre 29 e 34 pontos) e apresentam alta frequência de hábitos de leitura e escrita. Clinicamente, quanto aos instrumentos utilizados para a verificação de critérios de inclusão e de caracterização da amostra, os participantes não apresentaram sinais sugestivos de demência nem de depressão, com um escore nos subtestes de inteligência acima da média. Além disso, não houve mudanças relatadas pelos participantes entre teste e retestes quanto a saúde geral e eventos de vida, conforme questionário estruturado efetuado.

Instrumentos e procedimentos

De acordo com os aspectos éticos envolvidos em investigações com seres humanos, a participação dos indivíduos foi voluntária, e os instrumentos foram administrados individualmente em um ambiente apropriado. Os participantes que aceitaram participar deste estudo assinaram um Termo de Consentimento Livre e Esclarecido após aprovação pelo Comitê de Ética em Pesquisa da Pontifícia Universidade Católica do Rio Grande do Sul (protocolo número 042/2009).

As avaliações com o IGT foram realizadas em dois encontros: o primeiro (teste) com duração de aproximadamente

uma hora e meia, em que foram aplicados, além do IGT, os instrumentos de caracterização da amostra e de critérios de inclusão; o segundo momento (reteste) durou aproximadamente 20 minutos, com apenas a aplicação do IGT, além de uma breve entrevista sobre mudanças na saúde geral e de eventos de vida entre teste e reteste. A sessão de reteste ocorreu entre um e seis meses após a primeira sessão ($M = 3,84$, $DP = 2,01$).

Aplicou-se uma versão computadorizada do IGT, a partir da versão atualmente utilizada pelo autor do instrumento, que foi adaptada para o português brasileiro por Schneider⁸. A tarefa envolve um jogo de cartas e avalia o processo de tomada de decisão do indivíduo em termos de busca ou aver-são ao risco, durando aproximadamente 20 minutos.

Critérios de análise de dados

Os dados foram analisados descritiva e inferencialmente. Foi utilizado o pacote estatístico SPSS na versão 17.0, com um nível de significância de 5%. Os resultados do IGT foram analisados segundo duas pontuações. Uma das interpretações mais utilizadas na literatura diz respeito ao cálculo global do examinando, que é derivado do cálculo da subtração do número de cartas retiradas dos baralhos desvantajosos do número de cartas retiradas dos baralhos vantajosos: $[(C+D)-(A+B)]$. Essa pontuação é denominada de score com base no cálculo global. Conforme o ponto de corte a partir de uma distribuição binomial adotada por Denburg *et al.*²⁸, os participantes têm seu desempenho classificado em prejudicado, adaptado ou limítrofe. No caso de escores acima de +18, são considerados adaptados, entre -18 e +18, o desempenho é classificado como limítrofe e abaixo de -18, prejudicado⁸.

O segundo critério de análise, comumente mencionado na literatura^{9,10}, refere-se ao score por bloco que é obtido em cada um dos cinco blocos de 20 cartas, ao longo das 100 jogadas. O cálculo empregado para a obtenção desse resultado é o mesmo utilizado para o cálculo global $[(C+D)-(A+B)]$ realizado a cada bloco de 20 cartas. A partir desse score, é possível analisar a curva de aprendizagem durante a execução da tarefa.

Para a análise teste-reteste, foi utilizado um estudo de correlação entre os dois momentos como sugerido na literatura¹⁷. Utilizou-se a técnica de correlação linear de Pearson para verificar a fidedignidade do IGT (estabilidade temporal), já que a distribuição da amostra foi paramétrica ($p = 0,200$). Os escores foram correlacionados considerando-se o cálculo global e o cálculo de cada um dos blocos de 20 jogadas que compõem o instrumento, com um nível de significância de $p \leq 0,05$. Além disso, considerando-se a variabilidade no intervalo de tempo entre teste-reteste, optou-se por realizar uma análise de correlação parcial, controlando para tempo de intervalo. A interpretação da intensidade das correlações adotada neste estudo é aquela sugerida por Levin e Fox²⁹. Baseado em outros artigos, iremos utilizar coeficiente de correlação de Kendall.

RESULTADOS

Com base na correlação entre teste-reteste para o cálculo total do instrumento, o coeficiente linear de Pearson foi $r = 0,43$, $p = 0,002$. Esse achado corresponde a uma correlação positiva moderada significativa entre a primeira e a segunda aplicação. Considerando-se a variabilidade entre os participantes quanto ao intervalo de tempo teste-reteste, optou-se por verificar a correlação parcial para o cálculo total, controlando para o tempo. Essa análise reteve o mesmo coeficiente de correlação ($r = 0,43$; $p = 0,002$), sugerindo que a relação linear anteriormente verificada se mantém mesmo enquanto se controla para os efeitos da variável tempo.

Logo após, analisou-se qualitativamente o desempenho da amostra. Na tabela 2, é possível observar a frequência absoluta e o percentual de participantes por classificação de desempenho de acordo com o cálculo global (prejudicado, limítrofe e adaptado), tanto no teste como no reteste.

Tabela 2. Classificação de desempenho baseado no cálculo global do IGT

Desempenho	Teste		Reteste	
	F	%	f	%
Prejudicado	8	16%	11	22%
Limítrofe	26	52%	19	38%
Adaptado	16	32%	20	40%

Classificação de desempenho baseada em Denburg *et al.* (2001).

Conforme apresenta a tabela 2, a maioria da amostra distribuiu-se nas classificações limítrofe e adaptada, tanto no teste como no reteste. Por meio do teste de correlação de Kendall, verificou-se uma associação positiva entre as frequências de classificação de desempenho no teste e reteste, $\tau = 0,49$, $p < 0,001$.

Em complementaridade, na tabela 3 são apresentados os coeficientes de correlação do teste e do reteste de cada um dos 5 blocos de 20 cartas que compõem a tarefa. São, ainda, visualizados os valores de p de cada correlação.

Tabela 3. Correlação teste-reteste para os segmentos do IGT

Segmentos IGT	r	p
Segmento 1	0,08	0,606
Segmento 2	0,09	0,530
Segmento 3	0,16	0,260
Segmento 4	0,43	0,002
Segmento 5	0,47	0,001

R: coeficiente de correlação de Pearson.

No que diz respeito à análise de correlação de cada um dos cinco segmentos que compõem a tarefa, evidencia-se

que tanto o bloco 4 quanto o bloco 5 apresentaram uma correlação positiva, moderada e significativa entre teste e reteste. Já os blocos 1, 2 e 3 não se correlacionaram significativamente.

DISCUSSÃO

A proposta deste estudo foi verificar evidências de fidedignidade do IGT pelo método teste-reteste. Em relação à variável cálculo global, que avalia o processamento da tomada de decisão no instrumento, os resultados evidenciaram que o IGT apresentou uma correlação positiva, moderada e significativa. Na análise da frequência de classificação de desempenho, baseada no cálculo global, também se encontrou uma correlação moderada entre o teste e o reteste. Com esses achados, pode-se argumentar que o IGT avalia o processo de tomada de decisão de forma confiável e precisa ao longo do tempo.

Se por um lado não há registro na literatura de pesquisas que verificaram diretamente a confiabilidade do IGT a fim de comparar com os achados encontrados, por outro lado, Ernst *et al.*²¹ e Verdejo-Garcia *et al.*²² em seus estudos fizeram aplicações repetidas do IGT. Verdejo-Garcia *et al.*²², ao aplicarem o IGT duas vezes em um intervalo de 25 dias a pacientes com abstinência de cocaína, usuários de maconha e grupo controle, evidenciaram que todos os grupos apresentaram um efeito de aprendizagem sobre a segunda administração do IGT. Resultado semelhante foi encontrado no estudo de Ernst *et al.*²¹. Com intervalo de uma semana, os autores reaplicaram o IGT em uma amostra de adolescentes com distúrbio de comportamento, adultos com transtorno por uso de substância, adultos e adolescentes saudáveis e constataram que todos os grupos melhoraram sua *performance* em um segundo momento, exceto os adolescentes com distúrbio de comportamento.

É possível questionar se o padrão de correlações identificado no presente estudo representa, de fato, um argumento favorável à confiabilidade do IGT – visto que os coeficientes encontrados foram moderados. No entanto, alguns estudos sugerem que as evidências de fidedignidade de testes que medem as FE podem ser menos satisfatórias do que aquelas encontradas em estudos com outros instrumentos neuropsicológicos, por causa das características inerentes ao próprio construto examinado¹⁸, que é multidimensional e bastante complexo. Para avaliar as FE com sensibilidade e especificidade, faz-se necessária uma tarefa nova que exija certo esforço cognitivo demandando um processamento do tipo controlado. Desse modo, quando há uma situação de reteste, a tarefa em questão já não é mais uma novidade, envolvendo ativação de estratégias já aprendidas (mais automáticas) e diminuindo a necessidade de utilizar processos executivos perante situações inéditas e inovadoras que demandariam adaptação às mudanças do ambiente¹⁸.

De acordo com essa proposta, no estudo desenvolvido por Paolo *et al.*³⁰, ao verificarem a fidedignidade do *Wisconsin*

Card Sorting Test (WCST), com o intervalo de um ano entre teste-reteste, referiram uma baixa estabilidade temporal. Os autores encontraram correlações variando entre 0,12 (na escala aprendendo a aprender) e 0,66 (número total de erros), com uma correlação média de 0,46. Esse achado pode sugerir que o reteste não mede a capacidade de resolução de problemas da mesma maneira que aquela empregada no teste.

De forma similar, outro instrumento que apresentou também baixa correlação entre teste-reteste foi o *Behavioral Assessment of the Dysexecutive Syndrome* (BADS). A estabilidade temporal do BADS foi estudada por Wilson *et al.*³¹, que administraram o instrumento em 29 participantes saudáveis e com intervalo entre teste-reteste de 6 a 12 meses. A correlação teste-reteste variou entre 0,08 (*Rule Shift Cards*) e 0,71 (*Key Search*), evidenciando baixa à moderada correlação¹⁹.

Em contrapartida, outros estudos vêm desafiando essa visão de que os instrumentos que avaliam o construto FE não são confiáveis e válidos ao serem aplicados mais de uma vez^{19,20}. Com o intuito de verificar essa ideia, Ettenhofer *et al.*¹⁹ investigaram a fidedignidade de três instrumentos neuropsicológicos amplamente utilizados para medir as FE – WCST, teste de Stroop e *Trail Making Test* (TMT). O intervalo de tempo entre teste e reteste foi de quatro a oito semanas. Os resultados mostram que no WCST o número médio de erros perseverativos foi significativamente menor, indicando um melhor desempenho no reteste. As correlações tanto do WCST ($r = 0,50, p < 0,01$) como do Stroop ($r = 0,68, p < 0,01$) foram moderadas, diferentemente do TMT, que apresentou baixa estabilidade temporal ($r = 0,23, p < 0,05$).

No entanto, outro estudo recente envolvendo o TMT encontrou alta correlação tanto no TMT parte A como no TMT parte B. Dikmen *et al.*³², em um intervalo de 11 meses, evidenciaram uma alta correlação na parte A ($r = 0,79$) e na parte B ($r = 0,89$). Esse achado demonstrou que o TMT pode ser um instrumento consistente entre as aplicações.

Langenecker *et al.*²⁰, também objetivando investigar a fidedignidade de testes de exame das FE, desenvolveram um estudo com *Parametric Go/No-Go* (PGNG). O instrumento foi criado para avaliar atenção e alguns componentes das FE, em três níveis diferentes de dificuldade. Com intervalo estabelecido entre teste-reteste de três semanas, demonstraram uma variação na correlação de $r = 0,57 (p < 0,01)$ para $r = 0,83 (p < 0,001)$, evidenciando alta confiabilidade e desempenho bastante estável ao longo das três semanas, exceto no nível 1 de dificuldade, em virtude do efeito de teto ($r = -0,11, p = 0,59$).

Ao longo do que foi exposto, percebe-se que algumas pesquisas sugerem que as FE podem ser medidas de forma confiável em pelo menos dois momentos diferentes, o que vai ao encontro dos resultados achados no estudo em questão, com o IGT. Apesar de todas as pesquisas analisadas proporem a investigação da fidedignidade de algum instrumento neuropsicológico de FE, cabe ressaltar que os instrumentos referidos, por exemplo, WCST e TMT, buscam

mencionar componentes das FE frias, uma vez que estão relacionadas à circuitaria dorsolateral-estriatal. Já o IGT mede uma FE quente, subsidiada pelo circuito ventromedial-estriatal. Além disso, evidencia-se a heterogeneidade de métodos e o quanto isso pode influenciar a interpretação dos dados. Entre as diferenças mais notáveis está o intervalo que cada autor estabelece entre o teste-reteste. De acordo com Urbina¹⁷, a correlação entre teste-reteste diminui progressivamente à medida que o intervalo aumenta. Sabe-se que, se o intervalo for longo, pode ocorrer, por exemplo, a mudança no estado cognitivo do examinando; em contrapartida, se o intervalo entre teste-reteste for muito curto, o resultado pode ser influenciado pelo efeito de aprendizagem, ou seja, o indivíduo pode ter seu desempenho beneficiado no reteste por sua aprendizagem no teste ser muito recente. Em estudos de fidedignidade com instrumentos que avaliam FE ainda não há um consenso na literatura sobre a faixa de intervalo entre teste e reteste¹⁷. Optou-se, assim, neste estudo pelo intervalo de tempo utilizado em um importante instrumento nacional que avalia, entre outras funções cognitivas, as FE (Instrumento de Avaliação Neuropsicológica Breve NEUPSILIN³³), de um a seis meses. No presente estudo, no entanto, a correlação teste-reteste do IGT não foi afetada pela variabilidade do tempo de intervalo entre as aplicações, conforme análise da correlação controlando-se o fator tempo.

Além de verificar a correlação do cálculo global, outro critério de análise utilizado neste estudo foi a correlação por segmentos, que proporciona maior compreensão do comportamento e aprendizado dos participantes ao longo da tarefa. Enquanto os dois últimos blocos, 4 e 5, apresentaram uma estabilidade temporal moderada entre o teste e o reteste, não houve correlação significativa nos três primeiros blocos do IGT. Na medida em que os primeiros blocos são guiados predominantemente por processos emocionais implícitos, em que os examinandos ainda não conseguem obter padrão de reforços e punições, a baixa relação entre teste-reteste pode se dever à característica mais exploratória desse primeiro momento. Na medida em que a tarefa vai se tornando mais explícita, com a aprendizagem, o examinando vai aumentando a capacidade de discriminar os baralhos vantajosos dos desvantajosos e, dessa forma, conseguiriam replicar em outro momento o comportamento semelhante ao da primeira vez.

No entanto, a análise por segmentos necessita ainda ser mais bem investigada, já que estudos atuais indicam que o primeiro bloco, principalmente, não apresenta uma correlação significativa com o cálculo total da tarefa, sugerindo um funcionamento diferenciado no processo de tomada de decisão avaliado pelo instrumento¹⁵.

Embora outros estudos prévios de fidedignidade de testes que examinam as FE também apresentam uma variação grande entre teste e reteste³¹, salienta-se que uma possível limitação do presente estudo seja a amplitude de intervalo

do tempo entre teste e reteste, de um a seis meses. Além disso, a variação da faixa etária da amostra, 19 e 75 anos, também pode ser considerada outra limitação da pesquisa. No entanto, como a análise teste-reteste pressupõe correlação do indivíduo com ele mesmo, não há restrição quanto à homogeneidade da amostra.

CONCLUSÃO

Os dados encontrados nesta pesquisa corroboram os resultados de estudos atuais que encontraram correlações moderadas entre teste-reteste de medidas de FE. Embora a relação não tenha sido forte entre os dois momentos, o IGT parece avaliar o processo de tomada de decisão de forma confiável ao longo do tempo. Novos estudos se fazem necessários para a confirmação dos dados aqui apresentados e em busca de identificação de escores mais representativos de subprocessos da tomada de decisão na execução do IGT, principalmente nos primeiros blocos da tarefa. Junto a isso, fazem-se relevantes buscas por evidências comportamentais e de neuroimagem funcional, com populações saudáveis e clínicas, sobre a utilização do IGT em mais de um momento. Investigações com populações psiquiátricas, por suas alterações funcionais frontais cada vez mais mencionadas na literatura, mostram-se cruciais.

De modo geral, no âmbito clínico, a psiquiatria pode se beneficiar dos estudos com instrumentos neuropsicológicos e das hipóteses que descrevem as disfunções cognitivas com base no conhecimento da relação normal do cérebro, uma vez que é a partir desse conhecimento que se permite decifrar as possíveis etiologias neuroanatômicas e a fisiopatologia dos transtornos psiquiátricos. Diante de novos estudos que confirmam a consistência do IGT com o passar do tempo, esse instrumento poderia ser utilizado em processos de reavaliação do paciente, o que é muito relevante para a clínica neuropsiquiátrica, na medida em que os pacientes podem ser examinados e terem seu desempenho de tomada de decisão comparado em dois momentos de acompanhamento ou de tratamento medicamentoso e/ou psicoterápico. Nesse contexto, pesquisas com o IGT em amostras com quadros psiquiátricos de eixo I e de eixo II são desafiadoras e, concomitantemente, muito promissoras para a consolidação da ainda recente interface entre neuropsicologia e psiquiatria.

REFERÊNCIAS

1. Lezak MD, Howieson DB, Loring DW. *Neuropsychological assessment*. New York: Oxford University Press; 2004.
2. Jollant F, Guillaume S, Jaussent I, Bellivier F, Leboyer M, Castelnau D, et al. Psychiatric diagnoses and personality traits associated with disadvantageous decision-making. *Eur Psychiatry*. 2007;22(7):455-61.

3. Kapczinski N, Peuker AC, Narvaez J. Aplicações do exame neuropsicológico à psiquiatria. In: Malloy-Diniz L, Fuentes D, Mattos P, Abreu N (Org.). Avaliação neuropsicológica. Porto Alegre: Artmed; 2010.
4. Natale L, Teodoro M, Barreto V, Haase V. Propriedades psicométricas de tarefas para avaliar funções executivas em pré-escolares. *Psicol Pesq.* 2008;2(02):23-5.
5. Van Der Linden M, Ceschi G, Zermatten A, Dunker D, Perroud A. Investigation of response inhibition in obsessive-compulsive disorder using the Hayling test. *J Int Neuropsychol Soc.* 2005;11:776-83.
6. Clark L, Manes F, Antoun N, Sahakian BJ, Robbins TW. The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia.* 2003;41:147-83.
7. Damásio A. O erro de Descartes: emoção, razão e o cérebro humano. São Paulo: Companhia das Letras; 1996.
8. Schneider D. Iowa Gambling Task: considerações desenvolvimentais e implicações neuropsicológicas e psicométricas [tese]. Porto Alegre (RS): Universidade Federal do Rio Grande do Sul; 2008.
9. Bechara A, Damásio A, Damásio H, Anderson S. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition.* 1994;50:7-15.
10. Bechara A. Iowa Gambling Task. Professional Manual. Psychological Assessment Resources, Inc.; 2007.
11. Bechara A, Damásio H. Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia.* 2002;40:1675-89.
12. Cavedini P, Riboldi G, Annucci A, Belotti P, Cisima M, Bellodi L. Decision-making heterogeneity in obsessive-compulsive disorder: ventromedial prefrontal cortex function predicts different treatment outcomes. *Neuropsychologia.* 2002;40(2):205-11.
13. Malloy-Diniz L, Leite W, Moraes P, Correa H, Bechara A, Fuentes D. Brazilian Portuguese version of the Iowa Gambling Task: transcultural adaptation and discriminant validity. *Rev Bras Psiq.* 2008;30(2):144-8.
14. Buelow M, Suhr J. Construct validity of the Iowa Gambling Task. *Neuropsychol Rev.* 2009;19:1-13.
15. Carvalho JN. Tomada de decisão no Iowa Gambling Task: estudos comparativos quanto às variáveis idade e escolaridade e estudo correlacional [dissertação]. Porto Alegre (RS): Pontifícia Universidade Católica do Rio Grande do Sul; 2010.
16. Pasquali L. Psicometria: teoria dos testes na psicologia e na educação. Petrópolis: Vozes; 2003.
17. Urbina S. Essentials of psychological testing. New Jersey: John Wiley & Sons; 2004.
18. Jelicic M, Henquet C, Derix M, Jolles J. Test-Retest stability of the behavioural assessment of the dysexecutive syndrome in a sample of psychiatric patients. *Intern J Neurosci.* 2001;110:73-8.
19. Ettenhofer M, Hambrick D, Abeles N. Reliability and stability of executive functioning in older adults. *Neuropsychology.* 2006;20(5):607-13.
20. Langenecker S, Zubieta J, Young E, Akil H, Nielson K. A task to manipulate attentional load, set-shifting, and inhibitory control: convergent validity and test-retest reliability of the Parametric Go/No-Go Test. *J Clin Exp Neuropsychol.* 2007;29(8):842-53.
21. Ernst M, Grant SJ, London ED, Contoreggi CS, Kimes AS, Spurgeon L. Decision making in adolescents with behavior disorders and adults with substance abuse. *Am J Psychiatry.* 2003;160:33-40.
22. Verdejo-García A, Benbrook A, Funderburk F, David P, Cadet J-L, Bolla KI. The differential relationship between cocaine use and marijuana use on decision-making performance over repeat testing with the Iowa Gambling Task. *Drug Alcohol Depend.* 2007;90:2-11.
23. Pawlowski J. Questionário sociocultural e aspectos da saúde; 2007.
24. Finger I, Argimon I. Validação de constructo do Inventário de Depressão de Beck – II (BDI-II) em uma população universitária. In: III Mostra de Pesquisa da Pós-Graduação PUCRS. Porto Alegre, Brasil; 2008.
25. Chaves ML, Izquierdo Y. Differential diagnosis between dementia and depression: a study of efficiency increment. *Acta Neurol Scand.* 1992;85:378-82.
26. Amaral R, Malbergier A. Avaliação de instrumento de detecção de problemas relacionados ao uso do álcool (CAGE) entre trabalhadores da Prefeitura do Campus da Universidade de São Paulo. *Rev Bras Psiq.* 2004;26(3):156-63.
27. Nascimento E. WAIS-III: Escala de Inteligência Wechsler para Adultos. São Paulo: Casa do Psicólogo; 2004.
28. Denburg NL, Bechara A, Cole C, Tranel D. Iowa Gambling Task performance in elderly persons predicts susceptibility to the influence of misleading advertising. In: Annual meeting of the Society for Neuroscience. San Diego (USA): SN; 2001.
29. Levin J, Fox J. Estatística para ciências humanas. 9. ed. São Paulo: Prentice Hall; 2004.
30. Paolo AM, Axelrod BN, Troster AI. Test-retest stability of the Wisconsin Card Sorting Test. *Assessment.* 1996;3:137-43.
31. Wilson BA, Evans JJ, Emslie H, Alderman N, Burgess P. The development of an ecologically valid test for assessing patients with a dysexecutive syndrome. *Neuropsychol Rehabil.* 1998;8:213-28.
32. Dikmen SS, Heaton R, Grant I, Temkin N. Test-retest reliability and practice effects of expanded Halstead-Raitan Neuropsychological Test Battery. *J Int Neuropsychol Soc.* 1999;9:346-56.
33. Fonseca R, Salles J, Parente M. Instrumento de Avaliação Neuropsicológica Breve NEUPSI-LIN. São Paulo: Vetor Editora; 2009.