

Kelly Cristina Atalaia-Silva<sup>1,II</sup>

Roberto Alves Lourenço<sup>II,III</sup>

# Tradução, adaptação e validação de construto do Teste do Relógio aplicado entre idosos no Brasil

## Translation, adaptation and construct validation of the Clock Test among elderly in Brazil

---

### RESUMO

**OBJETIVO:** Poucos estudos no Brasil investigaram a validação de testes cognitivos em amostras populacionais de idosos e nenhum deles analisou as propriedades psicométricas do Teste do Relógio de Tuokko. O objetivo do estudo foi traduzir e adaptar o modelo desse teste ao contexto brasileiro e avaliar sua validade de construto.

**MÉTODOS:** Estudo transversal com uma amostra populacional constituída por 353 sujeitos de Juiz de Fora (MG), 2004-2005. Para avaliar as validades convergente e discriminante, utilizou-se a correlação de Pearson. Os subtestes do Teste do Relógio foram correlacionados com os instrumentos de referência: cubos e dígitos, e mini-exame do estado mental para avaliar a validade convergente. A validade discriminante foi investigada por meio da correlação dos subtestes com a *Center for Epidemiologic Studies Depression Scale*.

**RESULTADOS:** Da amostra, 74,1% eram do sexo feminino, com idade entre 63 e 107 anos (73,8+8,5) e média de escolaridade foi 7,4 anos (DP=4,7). Obteve-se validade convergente, com correlações estatisticamente significativas entre todos os subtestes ( $p \leq 0,01$ ). Quanto à validade discriminante, o único subteste que teve correlação significativa com a escala de referência foi o “indicação das horas” ( $p \leq 0,05$ ).

**CONCLUSÕES:** O Teste de Relógio traduzido e adaptado em uma amostra comunitária de idosos mostrou ser um instrumento de rastreio cognitivo breve, com boa validade de construto quando analisado com outros dados da literatura. Futuros trabalhos devem investigar outras propriedades psicométricas, como as validades de critério e de conteúdo.

**DESCRITORES:** Idoso. Demência, diagnóstico. Testes Neuropsicológicos. Tradução (Produto). Estudos de Validação.

<sup>1</sup> Faculdade de Psicologia. Instituto de Ciências Humanas. Universidade Federal de Juiz de Fora. Juiz de Fora, MG, Brasil

<sup>II</sup> Programa de Pós-graduação em Ciências Médicas. Faculdade de Ciências Médicas (FCM). Universidade do Estado do Rio de Janeiro (UERJ). Rio de Janeiro, RJ, Brasil

<sup>III</sup> Departamento de Medicina Interna. FCM. UERJ. Rio de Janeiro, RJ, Brasil

#### Correspondência | Correspondence:

Kelly Cristina Atalaia-Silva  
Rua Santo Antônio 640/602 - Centro  
36015-001 Juiz de Fora, MG, Brasil  
E-mail: kcasilva@yahoo.com.br

---

## ABSTRACT

**OBJECTIVE:** In Brazil, not many studies have investigated the validation of cognitive tests in the ageing population and none of them has analyzed the psychometric properties of Tuokko's Clock Test. The objective of the study was to translate and adapt the test to the Brazilian context, and to assess its construct validation.

**METHODS:** This is a cross-sectional population-based study, involving 353 elderly patients from Juiz de Fora (Southeastern Brazil), from 2004-2005. To assess convergent and divergent validities Pearson's correlation statistics was used. The Clock Test subtests were correlated with these reference instruments: mini-mental state examination, digits, block design to evaluate the convergent validity. the divergent validity was assessed by comparing the subtests to the Center for Epidemiologic Studies Depression Scale.

**RESULTS:** In the sample, 74.1% were women, aged between 63 and 107 years (73.8±8.5), average schooling was 7.4 years (SD=4.7). In regard to convergent validity, significant correlations were found between all CT subtests and MMSE, Digits, and Block Design ( $p \leq 0.01$ ). As for the divergent validity, the only subtest that had significant association with the reference scale was the "clock setting" ( $p \leq 0.05$ ).

**CONCLUSIONS:** The translated and validated Clock Test into a community sample of aged showed to be a brief screening instrument, with good construct validity when compared to other studies. Future research must investigate other psychometric properties, such as content and criterion validities.

**DESCRIPTORS:** Aged. Dementia, diagnosis. Neuropsychological Tests. Translations. Validation Studies.

---

## INTRODUÇÃO

Inicialmente, o envelhecimento populacional foi observado principalmente em países desenvolvidos, mas atualmente, nos países em desenvolvimento é que a população idosa tem aumentado de modo acelerado.<sup>24</sup> A transição demográfica é acompanhada pela transição no perfil de morbi-mortalidade, que pode ser retratada pelo aumento na prevalência de doenças crônico-degenerativas, como a doença de Alzheimer.

Os processos demenciais estão entre os principais problemas de saúde da população idosa. A identificação precoce das demências é crucial para melhores resultados no tratamento. Algumas medidas permitem retardar o período de perda funcional, aliviar o sofrimento desnecessário para o paciente e sua família, além de aumentar a qualidade de vida das pessoas envolvidas. Entre essas medidas, citam-se as farmacológicas (galantamina, rivastigmina e donepezil)<sup>15</sup> e as não-farmacológicas, como a reabilitação cognitiva.<sup>3</sup>

O diagnóstico da síndrome demencial é um problema freqüentemente observado, sobretudo na doença de Alzheimer, na qual o diagnóstico de certeza só pode ser obtido em estudos *post mortem*. Sendo assim, adota-se

como padrão o diagnóstico clínico de demência realizado com base em critérios diagnósticos produzidos por agências internacionais (DSM-IV,<sup>2</sup> CID-10<sup>13</sup> e *National Institute of Neurological and Communicative Disorders and Stroke e Alzheimer's Disease and Related Disorders Association*).

Os testes neuropsicológicos são instrumentos que auxiliam o diagnóstico precoce de processos demenciais. Aqueles utilizados para rastreio cognitivo são caracterizados por serem de fácil aplicação e por necessitarem de pouco tempo de execução. São importantes, uma vez que a maior parte dos casos de demência pode ser identificada em locais de atendimento onde há escassez de tempo e de recursos humanos especializados, tais como os ambientes de assistência primária à saúde.

Apesar da falta de um modo único padronizado de administração e correção, o Teste do Desenho do Relógio (TDR) é amplamente aceito como um instrumento de rastreio cognitivo. Os métodos mais empregados na clínica, de acordo com Braunberger,<sup>8</sup> são os propostos por: Shulman et al,<sup>17</sup> Sunderland al,<sup>21</sup> Wolf-Klein et

---

<sup>8</sup> Braunberger P. Clock drawing test. [S.l.: s.ed.]; 2001 [citado 2004 nov 14]. Disponível em: [http://199.243.225.113/clinicalassistant/scales/clock\\_drawing\\_test.htm](http://199.243.225.113/clinicalassistant/scales/clock_drawing_test.htm)

al,<sup>26</sup> Mendez et al,<sup>12</sup> Tuokko et al,<sup>23</sup> Manos & Wu,<sup>11</sup> e Shua-Haim et al.<sup>16</sup> As diferenças variam em função da instrução dada para realizar a tarefa, da hora a ser marcada, bem como do sistema de escore a ser utilizado. Do ponto de vista do domínio cognitivo aferido, supõe-se que o TDR avalie habilidades visuo-espaciais, habilidades construtivas e funções executivas.<sup>20</sup>

Em um estudo de revisão, Shulman<sup>18</sup> observou que a maioria dos estudos do TDR apresentou medidas de sensibilidade e especificidade em torno de 85% e grande aceitabilidade de execução pelos pacientes. Todavia, apesar de sua importância, , escassos estudos no Brasil têm investigado seus métodos de aplicação, suas escalas de correção e suas características psicométricas, sobretudo em relação à validade do instrumento.<sup>5,10,a</sup>

A validade de construto, por verificar a hipótese de legitimidade da representação comportamental dos traços latentes,<sup>14</sup> talvez seja a forma mais fundamental de estudo em instrumentos neuropsicológicos e tem sido empregada para avaliação do TDR por vários autores.

Hadjistavropoulos et al,<sup>6</sup> para avaliarem a validade de construto do TDR, investigaram as validades convergente e discriminante. O escore no TDR correlacionou-se significativamente com os subtestes Dígitos e Cubos da escala WAIS-III, e não se correlacionou com a escala de humor de Kozma & Stones.<sup>8</sup>

Uma forma ampliada do uso do relógio para testagem cognitiva foi desenvolvida por Tuokko et al.<sup>23</sup> Os autores levaram em consideração o fato de que o conceito de tempo é abstrato e as habilidades para desenhar, indicar e ler determinados horários parecem ser particularmente sensíveis na identificação de comprometimento cognitivo.

O novo teste, chamado de Teste do Relógio de Tuokko (TR), é um instrumento clínico e de pesquisa desenvolvido para detectar prejuízo cognitivo, com ênfase na investigação das habilidades visuo-espaciais, construcionais, perceptivo-visuais e conceitual-abstratas. Os autores acrescentaram mais duas tarefas – a indicação e a leitura das horas –, acreditando que essas contribuiriam para discernir melhor o comprometimento cognitivo em estágios iniciais, quando comparadas ao desenho do relógio isoladamente.

Segundo Tuokko et al,<sup>23</sup> o TR é um instrumento econômico, confiável, válido e de grande utilidade no diagnóstico de demências, sobretudo quando utilizado conjuntamente com outras técnicas de avaliação cognitiva e de entrevista com o paciente e seus familiares.

Por ser um instrumento mais abrangente e que utiliza tarefas que envolvem desenho, indicação e leitura das horas, o TR pode ser útil na identificação precoce de suspeição de demência também em indivíduos pouco

escolarizados. Isso porque o desempenho nos instrumentos de avaliação cognitiva é influenciado pela escolaridade, uma variável muito importante quando se trata deste tipo de estudo em populações, como a avaliada no presente estudo. Deste modo, torna-se importante o desenvolvimento de pesquisas brasileiras que investiguem o TR.

O objetivo do presente estudo foi traduzir e adaptar operacionalmente o modelo de teste do relógio de Tuokko ao contexto brasileiro e avaliar sua validade de construto.

## MÉTODOS

O TR foi aplicado em amostra de população de idosos que vivem na comunidade, entre junho de 2004 e fevereiro de 2005. O estudo faz parte do projeto Estudo dos Processos de Envelhecimento Saudável (PENSA), conduzido na cidade de Juiz de Fora (MG).

A tradução, adaptação e validação do TR foram autorizadas pela autora do instrumento original.<sup>8</sup> Um dos autores (KCAS) traduziu o instrumento que foi avaliado por um segundo tradutor (brasileiro com domínio do idioma inglês).

A forma de aplicação e correção do TR apresenta elevado custo para a realidade brasileira, pois cada kit é constituído por sete folhas picotadas de papel cartão com carbono. Na versão original o sujeito não poderia ver o que havia acabado de fazer. Esse procedimento era possível por causa das folhas picotadas que eram destacadas à medida que o sujeito terminava um desenho.

Na adaptação operacional do TR, reduziu-se o número de folhas e de material utilizado na confecção do teste, visando à adequação do instrumento às limitações operacionais da pesquisa.

O pré-teste incluiu uma amostra de 20 idosos de uma unidade ambulatorial de hospital universitário e resultou em mudanças no anteparo do subteste indicação das horas.

Esse anteparo é formado por abas. Para manipular o anteparo, o examinador deve levantar suas abas em ordem crescente e abaixá-las assim que o examinando acabar de desenhar, a fim de que o mesmo não veja o que acabou de fazer e nem qual será o próximo desenho a ser executado (Figura 1). Com o instrumento ajustado, o TR foi aplicado em outros 20 idosos de um centro de convivência para idosos vinculado às atividades de extensão de uma universidade. Após o término de cada testagem, os sujeitos foram indagados sobre o entendimento das tarefas propostas pelo TR. Nenhuma outra alteração operacional foi necessária pois nenhum idoso relatou qualquer dificuldade de entendimento ou execução.

<sup>a</sup> Okamoto IH. Aspectos cognitivos da doença de Alzheimer no teste do relógio: avaliação de amostra da população brasileira [tese de doutorado]. São Paulo: Universidade Federal de São Paulo; 2001.

Entre os meses de junho e julho de 2004, os pesquisadores de campo – quatro entrevistadores e um recrutador – foram treinados nos procedimentos de aplicação do instrumento. Na última etapa o TR foi corrigido por dois aferidores. A padronização nos procedimentos de correção entre os aferidores resultou em valor kappa inter-aferidor de 0,95 de concordância inter e intra-aferidor em amostra de 20 sujeitos. A concordância teste-reteste com os dois aferidores e 20 testes, no intervalo de dois meses, teve valores de kappa 1,0 e 0,94 para cada aferidor.

A amostra da pesquisa maior (PENSA) foi constituída por idosos da comunidade, residentes em bairros de Juiz de Fora que representavam 15% ou mais de sua população. Utilizando o procedimento porta-a-porta do Instituto Brasileiro de Geografia e Estatística, foram incluídos na amostra 956 idosos.<sup>4</sup>

Desse total de idosos uma amostra aleatória simples de 40% dos casos foi selecionada para participar do presente estudo. Os critérios de inclusão foram: ser capaz de ouvir e entender o suficiente para participar da entrevista e assinar o termo de consentimento livre e esclarecido.

O total de sujeitos selecionados foi 382, dos quais 353 aceitaram participar do estudo (92,4%). Os motivos de não participação foram: 27,7% de falecimento; 24,1% de incapacidade; 34,5% de recusa com perda de seguimento; e 13,7% de mudança de residência.

O TR envolve três tarefas empíricas: o desenho do relógio (*clock drawing*), a indicação das horas (*clock setting*) e a leitura das horas (*clock reading*). O subteste desenho do relógio consiste em um círculo pré-desenhado no qual o examinando deve colocar os números e os ponteiros indicando 11:10h (Figura 2). O modo de correção desse subteste utiliza um detalhado sistema qualitativo e quantitativo dos erros cometidos pelo examinando. Esse sistema de escore classifica os erros em sete categorias amplas: omissão, ausência, perseveração, distorção, substituição, adição e rotação. Para fazer a classificação e a pontuação, são analisados erros em relação ao desenho da face do relógio, dos números, dos ponteiros e dos espaços entre os números. Atribui-se um ponto para cada erro e o escore máximo é ilimitado, embora raramente sejam vistos valores superiores a 31 pontos.

O subteste indicação das horas é formado por cinco círculos nos quais é solicitado ao examinando que faça os ponteiros indicando 3:00h, 9:15h, 1:00h, 11:10h e 7:30h (Figura 1). A correção deste subteste consiste na atribuição de um ponto para cada ponteiro colocado corretamente, e um ponto adicional quando o tamanho do ponteiro das horas for menor que o ponteiro dos minutos (total=15 pontos).

O subteste leitura das horas (Figura 3) é formado por um caderno onde estão desenhados relógios indicando os mesmos horários do subteste indicação das horas, no qual o sujeito deve dizer qual o horário está sendo representado em cada relógio. A correção é feita atribuindo-se um ponto para cada ponteiro lido corretamente e um ponto adicional quando os dois ponteiros são lidos de forma correta (total=15 pontos).

Para avaliar a validade de construto do TR, utilizaram-se os instrumentos: mini-exame do estado mental (MEEM), o Dígitos, o Cubos e a *Center for Epidemiologic Studies Depression Scale* (CES-D). O MEEM é composto de 30 itens, com subtestes que avaliam orientação espaço-temporal, memória imediata, evocação, memória de procedimento e linguagem. O escore varia de zero a 30 pontos, sendo adotados os pontos de corte de Almeida:<sup>1</sup> 19/20 para indivíduos sem escolaridade e 23/24 para indivíduos com alguma escolaridade formal.

Os subtestes Dígitos e Cubos fazem parte da *Wechsler Adult Intelligence Scale – Third Edition* (WAIS-III). O primeiro tem como objetivo avaliar recordação, repetição imediata, atenção e memória de trabalho. Sua pontuação é feita a partir da soma entre a ordem direta e a inversa dos dígitos, e seu valor máximo é 30. O subteste Cubos avalia organização perceptual e visual, capacidade de abstração, formação de conceito não-verbal e visualização espacial. O valor máximo desse subteste é 68.

A CES-D compreende itens que avaliam sintomas depressivos presentes na semana precedente à aplicação do instrumento. A pontuação pode variar entre zero e 60. Foi validada para o Brasil por Silveira & Jorge.<sup>19</sup> O ponto de corte adotado foi 16.

O escore do subteste do desenho do relógio foi convertido em acertos para ficar na mesma escala que os outros subtestes. Para comparar as características da população de estudo com a população original do PENSA, utilizou-se o teste t pareado para a variável contínua “escolaridade” e McNemar para a variável dicotômica “sexo”.

As validades convergente e discriminante foram analisadas utilizando-se a correlação de Pearson. Os subtestes do TR foram correlacionados aos subtestes Cubos e Dígitos da escala WAIS-III e o MEEM para avaliar a validade convergente. A validade discriminante foi investigada por meio da correlação dos subtestes do TR com a CES-D.

Os dados foram analisados com auxílio do *software* estatístico SPSS versão 10.0.

O estudo foi aprovado pelo Comitê de Ética em Pesquisa do Hospital Universitário de Juiz de Fora (nº 463.148.2004).

## RESULTADOS

Todos os idosos completaram o teste e tiveram seus dados analisados, sendo 74,1% do sexo feminino, com idade variando entre 63 e 107 anos (média de 73,8 anos; desvio-padrão de 8,5; mediana 73; e moda 63). Houve predomínio de idades avançadas (63,1% dos idosos tinham mais de 70 anos de idade). Cerca de 7,9% dos idosos declararam ter concluído curso superior (Tabela 1).

Quando comparada com a população que compôs o estudo original PENSA, não houve diferenças nas variáveis sexo e escolaridade. A média e o desvio-padrão da variável escolaridade na amostra do PENSA foi 7,1±4,4 e na amostra do presente estudo foi 7,4±4,6 ( $t=0,907$ ;  $p<0,36$ ). O sexo feminino representou 71,8% da amostra do PENSA, enquanto que, no presente estudo, 74,1% eram mulheres (taxa de concordância=84%;  $p<0,38$ ).

No subteste indicação das horas, houve concentração nos valores altos da escala (média de 10,3 e DP=3,9).

**Tabela 1.** Características sociodemográficas e de saúde auto-referida e funcional da amostra. Juiz de Fora, MG, 2005. N=353

Variável	n	%
Idade		
60-69	130	36,8
70-79	139	39,3
≥80	84	23,9
Estado Civil		
Casado	171	48,5
Viúvo	134	37,8
Solteiro	48	13,7
Escolaridade		
0-4 anos	148	41,8
5-8 anos	81	23,1
9-11 anos	49	13,8
≥12 anos	75	21,3
Saúde percebida		
Muito boa	95	26,8
Boa	193	54,8
Razoável	68	19,0
Péssima	13	3,7
Metamemória		
Muito boa	92	26,2
Boa	193	54,8
Razoável	68	19,0
Independente ABVD	332	94,0
Independente AIVD	265	75,2

ABVD: atividades básicas de vida diária; AIVD: atividades instrumentais de vida diária

O mesmo foi observado no subteste desenho do relógio (média de 3,9 e DP=5,7). O subteste leitura das horas foi o que apresentou desempenho mais homogêneo entre os idosos (média de 12,5 e DP=3,4). Em relação ao tempo de aplicação, o TR teve um tempo médio de 4,9 minutos.

O desempenho cognitivo da amostra quando avaliado pelo MEEM teve média de 25,3 (DP=3,3), mediana de 26 e moda de 27. No ponto de corte 23/24, 21,2% da amostra estariam com comprometimento cognitivo.

No subteste Cubos, a média foi de 16,9 (DP=8,4) com mediana 16 e moda 12. No subteste Dígitos média foi de 11,4 (DP=3,5), mediana 11 e moda 10.

O desempenho na CES-D apresentou média de 15,8 (DP=8,5), mediana 12 e moda 10. No ponto de corte 16, sugerido pelo estudo brasileiro de validação da escala, 31,5% dos idosos seriam considerados suspeitos de depressão.

As correlações entre os subtestes do TR, e esse último e os testes MEEM, Cubos, Dígitos e CES-D podem ser vistos na Tabela 2.

A consistência interna entre os subtestes do TR, quando avaliada pelo alfa de Cronbach, foi 0,773.

## DISCUSSÃO

Este é o primeiro estudo brasileiro que utilizou o TR em uma amostra populacional de idosos, realizando sua tradução, adaptação operacional e validação de construto.

Herdman et al<sup>7</sup> sugerem que numa abordagem universalista de adaptação transcultural é necessário investigar as equivalências conceitual, de item, semântica, operacional, de mensuração, e a equivalência funcional do instrumento.

**Tabela 2.** Resultados da correlação entre o Teste do Relógio e outros testes de avaliação cognitiva e de depressão. Juiz de Fora, MG, 2005. N=353

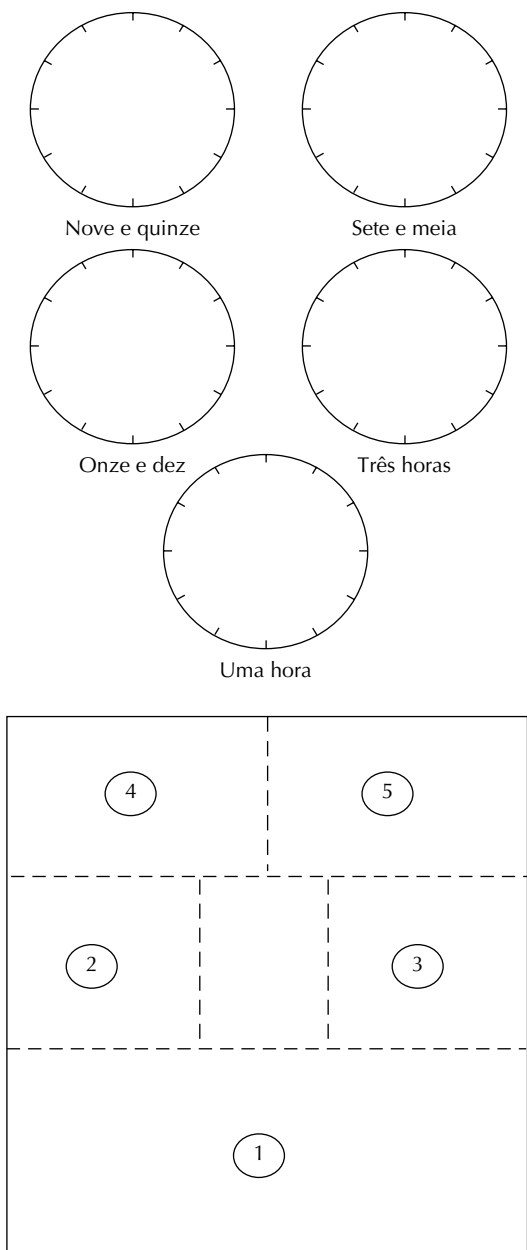
Teste	Teste do Relógio		
	Indicação das horas	Leitura das horas	Desenho do relógio
Indicação das horas	-	0,508*	0,376*
Leitura das horas	-	-	0,354*
MEEM	0,464*	0,389*	0,227*
Cubos	0,476*	0,361*	0,202*
Dígitos	0,399*	0,373*	0,118
CES-D	-0,140**	-0,064	-0,103

\*  $p<0,01$

\*\*  $p<0,05$

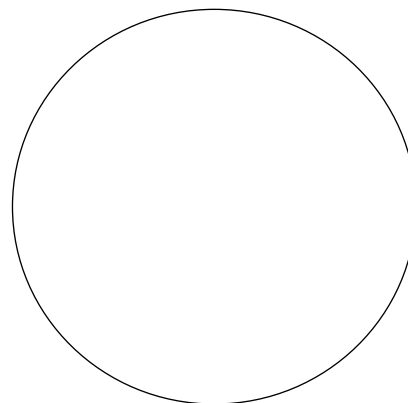
MEEM = Mini-exame do Estado Mental

CES-D = Center for Epidemiologic Studies Depression Scale

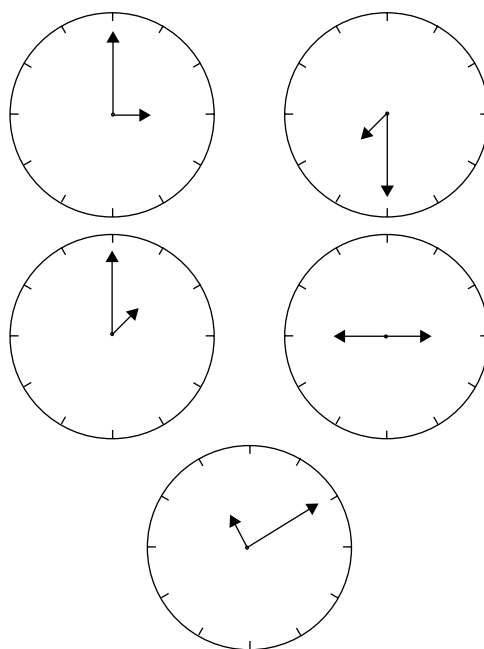


**Figura 1.** Modelo de subteste indicação das horas e do aparato desenvolvido.

No presente estudo, não consideramos necessária a execução de todas as etapas propostas por esses autores, pelas seguintes razões: o TR é simples em sua aplicação, não tem diversidade semântica e não é uma escala com itens dissertativos. Isso porque ele é composto por comandos objetivos para realização de tarefas, que não dependem da interpretação subjetiva do avaliador ou do examinando. Por outro lado, assumimos a equivalência conceitual do TR nas culturas canadense e brasileira, onde o conceito de horário como medida do tempo e o uso de relógios analógicos como instrumentação na contagem das horas são pressupostos básicos.



**Figura 2.** Modelo do subteste desenho do relógio.



**Figura 3.** Modelo do subteste leitura das horas.

Pela mesma razão, a impressão inicial foi de que não seria necessário realizar a tradução e a retrotradução por tradutores independentes. Durante o pré-teste, os examinandos declararam que as instruções do teste eram de fácil compreensão, não havendo problemas na execução das tarefas após um primeiro comando. Parece-nos que tal facilidade confirma a nossa hipótese inicial quanto ao tipo de tradução empregada. Porém, possivelmente tal facilidade tenha sido influenciada pelo fato dos indivíduos desta amostra apresentarem escolaridade superior à observada em outros estudos populacionais brasileiros.<sup>10</sup> A realização de grupos focais com idosos de diferentes níveis educacionais provavelmente fornecesse elementos mais objetivos, que dessem suporte a esse nosso pressuposto.

No entanto, a possibilidade de utilizar formas similares do instrumento, no que se refere ao formato, às instruções, ao modo de aplicação e de correção – equivalência operacional – era uma tarefa fundamental. Já que a versão original do TR exigiria custos elevados, a adaptação operacional foi essencial, devido a recursos limitados para a realização do estudo. De toda forma, tal adaptação operacional também é necessária para que tais testes cognitivos não tenham custos muito altos, e possam ser adquiridos e utilizados em serviços públicos de saúde no Brasil. A versão adaptada do presente estudo parece não ter interferido no resultado, pois a validade de construto encontrada foi semelhante à canadense.

O tempo médio de aplicação do TR (4,9 minutos), coloca-o como um instrumento de rastreio “curto”.<sup>9</sup> Outro ponto positivo é que o TR, por suas características de simplicidade e rapidez, pode auxiliar no diagnóstico de comprometimento cognitivo em locais onde existe restrição de tempo e carência de especialistas na área de avaliação cognitiva. Todavia não existem outros estudos sobre a mensuração do tempo de aplicação do TR, sendo o presente estudo o primeiro a investigar tal variável.

Como uma primeira abordagem da equivalência de mensuração do TR, foi analisada sua validade de construto. Por se tratar da primeira pesquisa brasileira com esse instrumento, acreditávamos que seu estudo em amostra populacional seria mais interessante por permitir descrever sua validade de construto e seu perfil de desempenho em uma amostra com cognição supostamente normal. A partir dos dados já coletados, os dados normativos serão tema de artigo futuro. Em outro momento, quando dados clínicos sobre o desempenho cognitivo dos indivíduos que compõem a amostra estiverem disponíveis, a validade de critério deverá ser analisada.

A correlação com escalas de avaliação cognitiva já validadas no Brasil (MEEM e subtestes da escala WAIS-III) sugere que o TR seja capaz de medir o que ele se propõe a mensurar. Identificamos correlações positivas e significativas, porém com força de associação entre fraca e moderada. Isso leva a pensar que talvez os construtos avaliados pelos testes TR, MEEM, Dígitos e Cubos sejam semelhantes, mas não os mesmos. O

subteste Dígitos é mais específico para avaliar memória de trabalho e o Cubos para planejamento visuo-espacial. Este é um problema comumente constatado na área de psicometria, visto que não há testes específicos para avaliar construtos iguais. Entretanto, outros estudos<sup>22,23</sup> que avaliaram a correlação entre esses subtestes de avaliação cognitiva também encontraram associações fracas e moderadas. Isso pode indicar que talvez a associação entre estas variáveis realmente se comporte da maneira observada no presente trabalho, uma vez que realmente existem diferenças entre os construtos avaliados separadamente por instrumento.

Por outro lado, Kurzman<sup>a</sup> preconiza que, em geral, as correlações da validade convergente são altas quando elas são calculadas em amostras clínicas, ao contrário do que é observado em amostras populacionais. Isso pode ocorrer porque a variação dos escores em sujeitos normais é restrita. O estudo canadense do TR também encontrou associações moderadas em suas correlações.<sup>23</sup>

Na validade discriminante, os subtestes desenho do relógio e leitura das horas não correlacionaram com a medida de sintomas depressivos. No estudo canadense do TR, nenhum dos subtestes correlacionou com a escala de humor.<sup>23</sup> Todavia, no presente estudo, o subteste indicação das horas apresentou correlação negativa e significativa com a CES-D ( $r=-0,118$ ;  $p \leq 0,05$ ), o que significa que quanto mais sintomas depressivos os sujeitos apresentavam, pior eram os seus desempenhos neste subteste. Isso pode ser explicado pelo fato de que este subteste exija maior capacidade atenta, capacidade esta frequentemente alterada em indivíduos deprimidos. Entretanto, a força da correlação entre a CES-D e o subteste indicação das horas foi muito fraca, representando uma associação espúria.

Em conclusão, o TR traduzido e adaptado em uma amostra comunitária de idosos brasileiros é um instrumento de rastreio cognitivo breve que apresentou boa validade de construto quando analisado com outros dados da literatura. Suas características psicométricas referentes à validade de conteúdo e de critério merecem ser analisadas em futuros trabalhos, assim como seus valores de sensibilidade e de especificidade.

<sup>a</sup> Kurzman D. The construct validity of the Clock Test in normal and demented adults. [Dissertação de mestrado]. Victoria: University of Victoria; 1992.

## REFERÊNCIAS

1. Almeida OP. Mini-exame do estado mental e o diagnóstico de demências no Brasil. *Arq Neuropsiquiatr*. 1998;56(3-B):605-12.
2. American Psychiatric Association. DSM-IV – TR. Manual de diagnóstico e estatística das perturbações mentais. Lisboa: Climepsi; 2002.
3. Ávila R, Miotto EC. Funções executivas no envelhecimento normal e na doença de Alzheimer. *J Bras Psiquiatr*. 2003;52(1):53-61.
4. Batistoni SST, Neri AL, Cupertino APFB. Validade da escala de depressão do *Center for Epidemiological Studies* entre idosos brasileiros. *Rev Saude Publica*. 2007;41(4):598-605. DOI: 10.1590/S0034-89102007000400014
5. Fuzikawa C, Lima-Costa MF, Uchoa E, Barreto SM, Shulman K. A population based study on the intra and inter-rater reliability of the clock drawing test in Brazil: the Bambuí health and ageing study. *Int J Geriatr Psychiatry*. 2003;18(9):450-6. DOI: 10.1002/gps.863
6. Hadjstavropoulos T, Tuokko H, Beattie BL. The clock test: construct validity for a multidimensional assesment method [Abstract]. *Can Psychol*. 1991;32(2a):219.
7. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQOL instruments: the universalist approach. *Qual Life Res*. 1998;7(4):323-35. DOI: 10.1023/A:1008846618880
8. Kozma A, Stones MJ. The measurement of happiness; development of the Memorial University of Newfoundland Happiness Scale (MUNSH). *J Gerontol*. 1980;35(6):906-12.
9. Lorentz WJ, Scanlan JM., Borson S. Brief screening tests for dementia. *Can J Psychiatry*. 2002;47(8):723-33.
10. Lourenço RA, Ribeiro-Filho ST, Moreira IFH, Paradela EMPS, Miranda AS. The clock drawing test: utility in elderly with low education level. In: Annual Meeting Abstract Book, Las Vegas, USA. *J Am Geriatr Soc*. 2004 Apr;52(4):S58-S58.
11. Manos PJ, Wu R. The ten-point clock test: a quick screen and grading method for cognitive impairment in medical and surgical patients. *Int J Psychiatry Med*. 1994;24(3):229-44.
12. Mendez MF, Ala T, Underwood KL. Development of scoring criteria for the clock drawing task in Alzheimer's disease. *J Am Geriatr Soc*. 1992;40:1095-9.
13. Organização Mundial da Saúde. Classificação estatística internacional de doenças. 10. ed. São Paulo: EDUSP; 2007.
14. Pasquali L. Psicometria: teoria dos testes na psicologia e na educação. Petrópolis: Vozes; 2003.
15. Rockwood K, Fay S, Song X, MacKnight C, Gorman M. Video-Imaging Synthesis of Treating Alzheimer's Disease (VISTA) Investigators. Attainment of treatment goals bay people with Alzheimer's disease receiving galantamine: a randomized controlled trial. *CMAJ*. 2006;174(8):1099-105. DOI: 10.1503/cmaj.051432
16. Shua-Haim J, Koppuzha G, Shua-Haim V, Gross J.A simple scoring system for clock drawing in patients with Alzheimer's disease. *Am J Alzheimer's Dis*. 1997;12(5): 212-5. DOI: 10.1177/153331759701200505
17. Shulman KI, Shedletsky R, Silver IL. The challenge of time: clock-drawing and cognitive function in the elderly. *Int J Geriatr Psychiatry*. 1986;1(2):135-40. DOI: 10.1002/gps.930010209
18. Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry*. 2000;15(6):548-61. DOI: 10.1002/1099-1166(200006)15:6<548:AID-GPS242>3.0.CO;2-U
19. Silveira D, Jorge M. Propriedades psicométricas da Escala de Rastreamento para Depressão, CES-D, em populações clínica e não clínicas de adolescentes e adultos jovens. *Rev Psiquiatr Clin*. 1997;25(5):251-61.
20. Spreen O, Strauss E. A compendium of neuropsychological tests. Administration, norms, and commentary. New York: Oxford University Press; 1998.
21. Sunderland T, Hill JL, Mellow AM, Lawlor B, Gundersheime J, Newhouse P, et al. Clock drawing in Alzheimer's disease: a novel measure of dementia severity. *J Am Geriatr Soc*. 1989;37:725-9.
22. Taylor JL, Miller TP, Tinklenberg JR. Correlates of memory decline: a 4-year longitudinal study of older with memory complaints. *Psychol Aging*. 1992;7(2):185-93. DOI: 10.1037/0882-7974.7.2.185
23. Tuokko H, Hadjstavropoulos T, Miller, JA, Beattie BL. The clock test: a sensitive measure to differentiate normal elderly from those with Alzheimer's disease. *J Am Geriatr Soc*. 1992;40:579-84.
24. Veras RP. Desafios e conquistas advindas da longevidade da população: o setor saúde e suas necessárias transformações. In: Veras RP, organizador. Velhice numa perspectiva de futuro saudável. Rio de Janeiro: UERJ; 2001. p.11-32.
25. Wolf-Klein GP, Silverstone FA, Levy AP, Brod MS, Breuer J. Screening for Alzheimer's disease by clock drawing. *J Am Geriatr Soc*. 1989;37:730-4.