

Iara Souza Santos^{1,2}, João Vitor da Silva Viana^{1,2}, Aline Siqueira de Souza², Giovanna Correia Pereira Moro^{1,2}, Caio Peixoto Tavares², Isabelle de Aguiar Maia³, Guilherme Vieira Rodrigues da Costa³, Alberto Fonseca Lopes Rodrigues³, Joice Coutinho de Alvarenga², Gabriela Tomé Oliveira Engemann², Marco Aurélio Fagundes Angelo⁴, Débora Marques de Miranda⁵, Marco Aurélio Romano-silva⁵, Jonas Jardim de Paula⁶, Luiz Armando Cunha de Marco⁵, Maria Aparecida Camargos Bicalho^{6,7}, Bernardo de Mattos Viana⁶.

¹Undergraduate Medicine, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ²Cog-Aging Research Group, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brazil, ³Older Adult's Psychiatry and Psychology Extension Program (PROEPSI), School of Medicine, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brazil, ⁴Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ⁵Neurotec R National Institute of Science and Technology (INCT-Neurotec R), Faculty of Medicine, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brazil, ⁶Molecular Medicine Postgraduate Program, School of Medicine, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brazil, ⁷Department of Internal Medicine, School of Medicine, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

#10305

BACKGROUND

Informant-based instruments assess cognitive and functional decline in neurocognitive disorders. The CDFAQ-IV evaluates six cognitive domains and functional abilities through ordinal informant reports; however, its inter-rater and test-retest reliability have not been extensively examined. Our objective was to evaluate the inter-rater agreement and reliability, and test-retest reliability of the CDFAQ-IV.

METHODS

This cross-sectional study from the Cog-Aging cohort comprised two analyses. Inter-rater reliability was assessed in 16 informants of participants (cognitively unimpaired, mild cognitive impairment, or mild dementia) by evaluators independently recording responses and alternating interview conduction. Test-retest reliability involved 36 independent informants reassessed by the same evaluator after 2-12 weeks. The CDFAQ-IV covers six domains - Complex Attention (CA), Executive Function (EF), Learning and Memory (LM), Language (L), Perceptual-motor (PM), and Social Cognition (SC) - each with five items, yielding section and total scores. Analyses used weighted Cohen’s Kappa (item-level) and Intraclass Correlation Coefficients (ICC[2,1] for inter-rater; ICC[3,1] for test-retest), with absolute agreement and single measures, for section and total scores.

The CDFAQ-IV demonstrated excellent inter-rater reliability and moderate test–retest stability, supporting its use as a reliable informant-based tool for assessing cognitive and functional decline.

RESULTS

In the inter-rater sample, participants (50% female) averaged 71.9±6.6 years of age and 10.1±6.4 years of education. Informants (68.8% female; mainly children or spouses) averaged 55.6±17.6 years of age and 13.1±6.2 years of education. Item-level agreement was mostly almost perfect ($\kappa \geq 0.81$). Items L2 and L3 showed substantial agreement, whereas PM3-PM5 were non-estimable due to absent variability. Section-level ICCs ranged from 0.960 to 1.000, and total score reliability was excellent (ICC=0.995; 95% CI: 0.986-0.998). In the test-retest sample, participants (63.9% female) averaged 75.8±8.3 years of age and 8.5±6.4 years of education. Informants (69.4% female; mainly children or spouses) averaged 53.9 ±15.2 years of age and 13.4±5.2 years of education. Item-level stability ranged from poor to substantial, with predominantly fair-to-moderate agreement, especially in EF, L, and SC. Section-level reliability was mostly moderate (0.552-0.719), with lower stability in PM (ICC=0.488). The total score demonstrated moderate test-retest reliability (ICC=0.692; 95% CI:0.474-0.830).

Section	Item	Inter-rater		Test-retest	
		Kappa	Interpretation	Kappa	Interpretation
Complex Attention	1	1.000	Almost perfect	0.649	Substantial
	2	0.887	Almost perfect	0.541	Moderate
	3	0.887	Almost perfect	0.219	Fair
	4	1.000	Almost perfect	0.416	Moderate
	5	1.000	Almost perfect	0.372	Fair
Executive Function	1	1.000	Almost perfect	0.688	Substantial
	2	1.000	Almost perfect	0.211	Fair
	3	1.000	Almost perfect	0.632	Substantial
	4	1.000	Almost perfect	0.450	Moderate
	5	1.000	Almost perfect	0.293	Fair
Learning and Memory	1	1.000	Almost perfect	0.586	Moderate
	2	0.900	Almost perfect	0.582	Moderate
	3	1.000	Almost perfect	0.376	Fair
	4	0.881	Almost perfect	0.029	Slight
	5	1.000	Almost perfect	0.553	Moderate
Language	1	1.000	Almost perfect	0.429	Moderate
	2	0.709	Substantial	0.272	Fair
	3	0.784	Substantial	0.706	Substantial
	4	0.848	Almost perfect	0.118	Slight
	5	1.000	Almost perfect	0.634	Substantial
Perceptual-motor	1	1.000	Almost perfect	0.464	Moderate
	2	1.000	Almost perfect	0.272	Fair
	3	NA	Not estimable	0.571	Moderate
	4	NA	Not estimable	-0.038	Poor
	5	NA	Not estimable	0.000	Slight
Social Cognition	1	1.000	Almost perfect	0.384	Fair
	2	1.000	Almost perfect	0.545	Moderate
	3	0.879	Almost perfect	0.371	Fair
	4	1.000	Almost perfect	0.745	Substantial
	5	1.000	Almost perfect	0.726	Substantial

Section	Inter-rater ICC (95% CI)	Interpretation	Test-retest ICC (95% CI)	Interpretation
Complex Attention	0.993 (0.980-0.998)	Excellent	0.559 (0.286-0.748)	Moderate
Executive Function	1.000	Perfect	0.719 (0.515-0.846)	Moderate
Learning and Memory	0.991 (0.974-0.997)	Excellent	0.577 (0.311-0.760)	Moderate
Language	0.960 (0.890-0.986)	Excellent	0.552 (0.277-0.743)	Moderate
Perceptual-motor	0.966 (0.908-0.988)	Excellent	0.488 (0.194-0.701)	Poor
Social Cognition	0.995 (0.985-0.998)	Excellent	0.713 (0.505-0.843)	Moderate
Total	0.995 (0.986-0.998)	Excellent	0.692 (0.474-0.830)	Moderate

Characteristic	Inter-rater sample (n=16)	Test–retest sample (n=36)
Participants		
Age, years, mean ± SD	71.9 ± 6.57	75.8 ± 8.29
Education, years, mean ± SD	10.1 ± 6.42	8.5 ± 6.39
Female sex, n (%)	8 (50.0%)	23 (63.9%)
Male sex, n (%)	8 (50.0%)	13 (36.1%)
Diagnostic category, n		
Cognitively unimpaired	5 (31.3%)	10 (27.8%)
Mild cognitive impairment	5 (31.3%)	12 (33.3%)
Mild dementia	6 (37.5%)	14 (38.9%)
Informants		
Age, years, mean ± SD	55.6 ± 17.57	53.9 ± 15.21
Education, years, mean ± SD	13.1 ± 6.20	13.4 ± 5.24
Female sex, n (%)	11 (68.8%)	25 (69.4%)
Male sex, n (%)	5 (31.2%)	11 (30.6%)
Relationship to participant, n (%)		
Daughters	3 (18.8%)	16 (44.4%)
Sons	2 (12.5%)	6 (16.7%)
Wives	5 (31.3%)	7 (19.4%)
Husbands	3 (18.8%)	5 (13.9%)
Daughter-in-law	1 (6.3%)	—
Granddaughter	1 (6.3%)	1 (2.8%)
Friend	1 (6.3%)	1 (2.8%)

CONCLUSIONS

Preliminary results indicate that CDFAQ-IV had excellent inter-rater reliability and predominantly almost perfect item-level agreement, with fair-to-moderate temporal stability.

REFERENCES

- Diniz, M.F. (2021). Construção de um questionário de avaliação de domínios cognitivos e funcionalidade baseados nos critérios diagnósticos do DSM-5 para transtornos neurocognitivos. Universidade Federal de Minas Gerais, Master's Dissertation.
- American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 5th ed. American Psychiatric Association.
- Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163.



CONTACT

Iara Souza Santos
Email: iarass2307@gmail.com

Bernardo de Mattos Viana, MD, Ph.D
Email: bernardomviana@yahoo.com