Psicologia: Teoria e Pesquisa
2025, v.41, e41302

DOI: https://doi.org/10.1590/0102.3772e41302.en

School and Developmental Psychology

Instruments for Measuring Episodic Memory and its Psychometric Properties: a Systematic Review

Gabriela Campana Barbosa[1], Mateus Machado Martins[1] & Francis Ricardo dos Reis Justi[1]

[1]*Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil*

ABSTRACT – There are changes in episodic memory during healthy or pathological cognitive aging. This study aimed to verify which instruments were most used to assess this cognitive domain in older adults in the last ten years, evaluating their psychometric parameters. A systematic review was conducted according to the PRISMA recommendations, and 3,203 articles were identified in Pubmed, PsycINFO, Scopus, and Web of Science databases. Forty-eight of these met the eligibility criteria for analysis. The most frequently used instruments were RAVLT, FCSRT, and EBMT. The results suggested good psychometric properties for the episodic memory tests, although there is still room for adaptation, construction, and validation of new tests appropriated for the Brazilian context.

KEYWORDS: episodic memory, aging, psychometry, measurement

Instrumentos para Mensuração da Memória Episódica e suas Propriedades Psicométricas: Revisão Sistemática

RESUMO – A memória episódica apresenta mudanças durante o processo de envelhecimento cognitivo saudável ou patológico. Este estudo teve como objetivo verificar quais instrumentos foram mais utilizados para a avaliação desse domínio cognitivo em idosos nos últimos dez anos, avaliando seus parâmetros psicométricos. A partir de uma revisão sistemática, segundo as recomendações PRISMA, foram identificados 3203 artigos nas bases de dados Pubmed, PsycINFO, Scopus e Web of Science, dos quais 48 atenderam aos critérios de elegibilidade para análise. Os instrumentos utilizados com mais frequência foram o RAVLT, FCSRT e o EBMT. Os resultados indicam que os testes apresentam bons parâmetros psicométricos, porém verificou-se a necessidade de adaptação, construção e validação de novos instrumentos para o contexto brasileiro.

***PALAVRAS-CHAVES*:** memória episódica, envelhecimento, psicometria, medida

Aging is characterized as a natural process of development that encompasses expected physiological degenerations, such as a higher incidence of mental illnesses and behavioral disorders (Raj & Sapharina, 2021). Regarding cognitive functioning, old age is accompanied by age-related cognitive decline (ARCD). ARCD is a non-pathological condition, yet significant for decline in simple and advanced mental abilities, impairing various cognitive functions (Bensalem et al., 2019). Among these impairments, memory loss is a significant concern for older adults (Bensalem et al., 2019). Besides ARCD, other conditions more common in advanced ages, such as dementia, mainly Alzheimer's, also cause severe impairments in the memory abilities of these individuals.

Memory is one of human beings' most important cognitive functions (Júnior & Faria, 2015). According to Squire (1992), it can be differentiated into two subsystems: implicit memory, or non-declarative, and explicit memory, or declarative. In this regard, Tulving (1972) proposed dividing explicit memory into semantic and episodic memory, which is related to the objectives of the current study.

Episodic memory (EM) is the ability to mentally relive a subjective experience that occurred at a different historical point from the present (Tulving, 2002). Through episodic recall, it is possible to engage in a kind of mental time travel, allowing episodes experienced in the past to be consciously relived in the present through a method of searching and evoking a particular memory, referencing specific events

through the "what," "where," and "when" components of mnemonic representations (Baddeley, 2011; Rodrigues & Jaeger, 2018).

Regarding the development of EM, it is incontrovertible to state that there are changes and declines in this domain and its components throughout the aging process (Friedman et al., 2010; Nyberg et al., 2012). Beyond the expected changes during healthy aging, pathological processes can accelerate the cognitive decline that is already evident. Due to the complexity of EM functioning, a decline in this cognitive domain can be considered one of the earliest signs of cognitive impairments (Gifforda et al., 2015). Therefore, instruments capable of assessing such functions are of great importance for the overall health of the elderly population.

It is possible to consider that instruments assessing cognition are essential for identifying cognitive declines related to disorders (Tripathi et al., 2013). Thus, identifying their psychometric characteristics is essential to assess their quality and relevance. Through a systematic literature review, the present study aimed to ascertain which instruments have been most used, nationally and internationally, in assessing episodic memory over the last ten years. Additionally, it sought to verify the characteristics of the methodological designs of empirical studies that used instruments in assessing episodic memory, the frequency of publications in the field, and the psychometric parameters of the identified instruments, establishing comparison criteria among them.

METHOD

The systematic review was conducted by two researchers using the terms "elderly", "episodic memory", "measure*" and "psychom*" in English, Portuguese, and Spanish from PsycINFO, PubMed (Medline), Scopus, and Web of Science databases. The first two terms were combined with each subsequent one separately, using the boolean operator AND in each search. The initial search was conducted in August and September 2021 and included studies published between 2011 and 2021. Subsequently, recognizing the time gap between article submission and evaluation, a new search was conducted in October 2023 to update the results. This search covered studies published between 2021 and 2023. In both searches, the parameters of the PRISMA checklist were considered, as Galvão et al. (2015) outlined.

In order to be included in the research, an article had to contain the search terms "elderly", "episodic memory" and "measure*" or "elderly", "episodic memory" and "psychom*" in any combination of the following fields: title, abstract, and keywords. This condition was the first to be added to refine the results obtained in the searches. Additionally, articles meeting the following conditions were included: a) empirical articles and b) studies that evaluated episodic memory in older adults, considering subjects aged 60 and older.

After applying the inclusion criteria, articles were excluded if they: a) were duplicates, b) did not contain the descriptors mentioned above in the title, abstract, or keywords, c) were not freely accessible or could not be retrieved via Capes Periodicals Portal database, and d) were written in languages other than English, Portuguese and Spanish.

A total of 3,203 studies were found, of which 784 were automatically excluded by the bibliography management software for scientific article publication (EndNote) due to duplication. Thus, 2,419 studies were identified in the databases: 161 from PsycINFO, 1,137 from PubMed, 941 from Scopus, and 180 from Web of Science. Based on the reading of each study's title, keywords, and abstract, 76 articles were excluded based on the title and 27 based on the abstract, mainly because they did not specifically assess episodic memory. After reading the full articles, 58 were excluded for being in another language, not using a standardized neuropsychological instrument for episodic memory measurement, or not mentioning which instrument was used. Additionally, some articles were excluded at this stage for presenting aggregated results regarding participants' ages, which would hinder the specific analysis of the subjects of interest, and for not presenting quantitative episodic memory data. Thus, the final sample resulted in 48 articles, as seen in Figure 1, which illustrates and summarizes the process of obtaining the analyzed studies.

The analysis of the articles consisted of the following steps: 1) Verifying and analyzing the methodological design used in the studies (type of study, sample, methodology); 2) Identifying and discussing the psychometric parameters of the most commonly used instruments in EM assessment; 3) Examining the investigated populations; and 4) Conducting the analysis based on a discussion of the results in line with relevant literature.

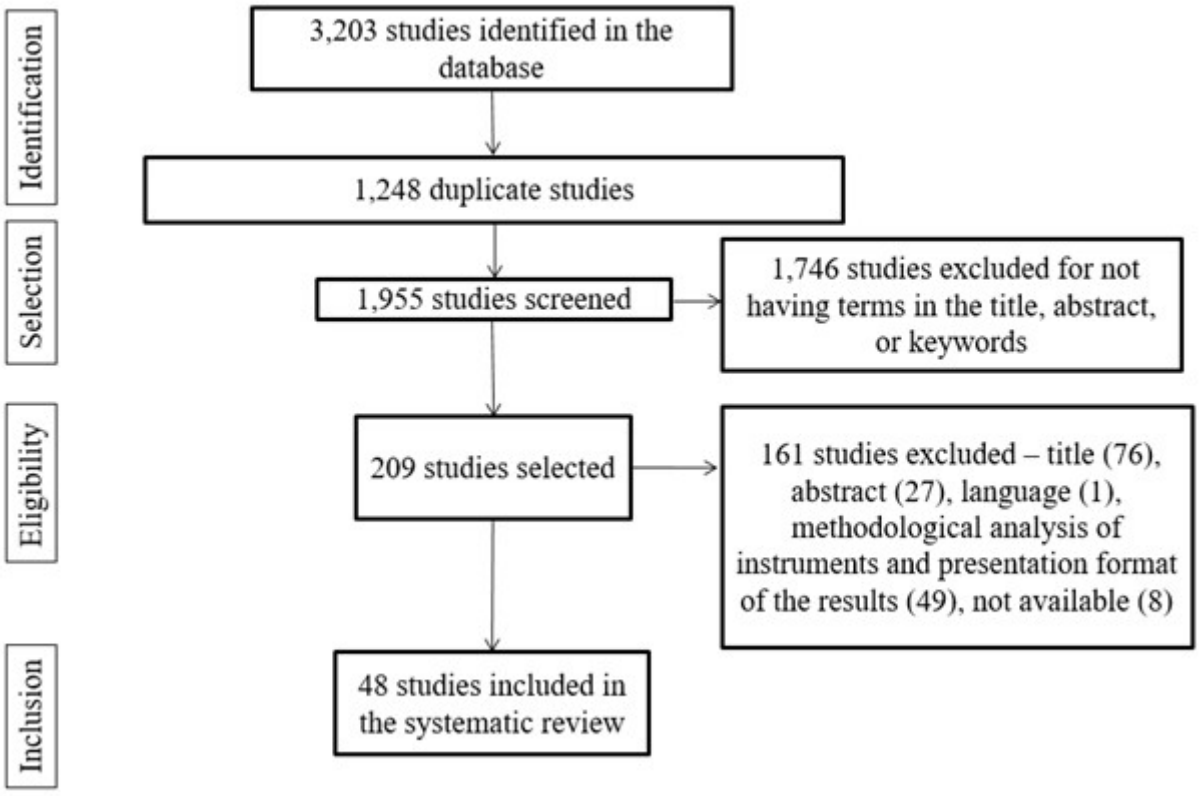

Figure 1
The Screening Process

RESULTS AND DISCUSSION

In total, 57,710 individuals participated in the 48 studies, with two articles not specifying the sample size. Of the 35 articles that reported gender differences in participation, it was possible to identify that most participants were women (56.6%). Healthy older adults and those with some form of cognitive impairment participated, with individuals with typical episodic memory development included in most articles (n=42). Only ten articles provided the minimum and maximum ages of participants, making it difficult to have a broader perception of the age range typically targeted by the tests.

Overall, from 2011 to 2023, there was a slight increase in the number of publications. Most of these are studies published after 2016 (n=25), which is the period with the highest peak of productions (n=8), but it still presents an incipient number of manuscripts considering the relevance of the topic in a society that is aging at an increasingly rapid pace (Cheng et al., 2020) and experiencing significant changes in episodic memory development both in typical aging and in pathological processes at this stage (Lee et al., 2016).

In every country where studies were identified, there were no significant changes in the number of publications over the years. In Brazil, articles were published in 2013, 2020, 2021, and 2023, with the years 2020 and 2021 having more publications, with two publications each. Six studies were found in the selected period, all using RAVLT for memory measurement. One factor that may explain this is the low number of standardized instruments for assessing episodic memory developed or adapted for the Brazilian context. No studies in this regard were identified in the search conducted. What was possible to verify was the work by De Paula et al. (2013), which aimed to investigate the validity and cutoff points for an unstructured neuropsychological assessment protocol for older adults. Considering the countries where the other studies were conducted, most of these studies(n = 23) were conducted in the United States, France, and China. Together, these three countries accounted for approximately 48% of the studies retrieved in this systematic review.

Despite many studies being identified in the present review, a more limited number of articles focused on the described topic. Episodic memory measurement was often analyzed secondarily, seeking to investigate its correlation with other relevant clinical and psychosocial conditions, such as diabetes (Guerrero-Berroa et al., 2015; Morris et al., 2016; Ravona-Springer et al., 2014), hereditary factors (Chen et al., 2020; Striepens et al., 2011), social support and stress (Ge et al., 2017), lifestyle (Klaming et al., 2017), engagement in cognitive and social activities (Grotz et al., 2017; Zhang et al., 2019), and education (Farina et al., 2020; Messinis et al., 2016; Meyer et al., 2020; Zhang et al., 2019; Volz-Sidiropoulou & Gauggel, 2012).

Due to this aspect, the analysis of the articles' methodological design indicated that most research was correlational, with descriptive objectives and cross-sectional designs (66.6%), as seen in Table 1. There was also a smaller number of longitudinal studies, which may have occurred because the search was focused on only one stage of development or because, as pointed out by some authors (Klaming et al., 2017; Chen et al., 2020), older adults are more prone to disorders and illnesses which makes it difficult to continue a research for long periods. Additionally, when conducted, such research may have a high dropout rate.

A few studies investigated predictive aspects of cognitive decline and dementia in older adults by considering physiological and genetic risk factors (Ernst et al., 2021; Riaz et al., 2021). Research of this type is vital as it allows for more accurate risk prediction and assists in developing more appropriate therapeutic or preventive strategies. Also presented in the selected articles there were comparative studies, validation/adaptation studies (n=3) (De Paula et al., 2013; Maillet et al., 2017; Messinis et al., 2016; Wang et al., 2019), construction of a cognitive assessment protocol (n=1) (Gevins et al., 2011), normative studies of instruments (n=1) (Ruan et al., 2020), as well as intervention studies (n=3) (Gross & Rebok, 2011; Requier et al., 2023; Rocha & Chariglione, 2021).

Table 1
Study Designs and Types of Articles Found (n=48)

Types of Study	The Design of Studies	
	Cross-sectional	Longitudinal
Correlational	19	12
Comparative	6	2
Validity Evidence	3	1
Standardization	1	0
Instrument Construction	1	0
Intervention	2	1
Total	32	16

Regarding the instruments, 19 standardized tests were found to be used for measuring episodic memory. It was possible to verify that, among those used by the authors, only one showed evidence of validity and adaptation to the Brazilian context. All the other tests with psychometric evidence of validity and/or reliability were used only in international studies. According to the criteria adopted, instruments specifically assessing episodic memory were prevalent. However, some test batteries that evaluate this cognitive domain, among others, were also used, such as the Cambridge Neuropsychological Test Automated Battery (CANTAB) and the Consortium to Establish a Registry for Alzheimer's Disease (CERAD).

The most frequently used instruments were, respectively, (1) the "Rey Auditory Verbal Learning Test" (RAVLT) by Rey (1964) (n=19), with some versions omitting André Rey's surname from the task title, using only AVLT; (2) the "Free and Cued Selective Reminding Test" (FCSRT) by Buschke (1984) (n=6); and (3) the "East Boston Memory Test" (EBMT) by Scherr et al. (1988) (n=5). Few studies among those analyzed used more than one instrument to assess episodic memory (n=11). Table 2 shows all the tests found, their frequencies, the target population evaluated by each of them in the articles, the ages for which they were used, and the total application time of these instruments.

Given the importance of measurement instruments with psychometric data for research development in the area of interest, it was chosen to highlight and describe the three most frequently used tests in the studies included in this systematic review. As mentioned, the first of these was RAVLT (Rey, 1964), used to assess episodic declarative memory, immediate memory span, verbal learning, susceptibility to proactive and retroactive interference, information retention, memory recall, and recognition (Lezak et al., 2004). As literature indicates (Cotta et al., 2012; De Paula & Malloy-Diniz, 2018), RAVLT was developed by Rey in 1958 and published in 1964 in his book, *L'examen clinique en psychologie*, with the initial objective of assessing learning and memory processes.

In Brazil, the instrument was translated, adapted, and standardized by Malloy-Diniz et al. (2000) for use with adolescents, adults, and older adults. Later, the same author conducted a new study with a sample of only typical older adults and developed another version of RAVLT, in which the original words were replaced with ones more representative of the Brazilian spoken language (Malloy-Diniz et al., 2007). In this new study, 262 community-dwelling older adults aged 60 to 89 participated, and the norms were stratified according to age ranges in 9-year intervals, with the presentation of norms based on percentiles (De Paula & Malloy-Diniz, 2018). According to Malloy-Diniz et al. (2007), Cronbach's alpha coefficients for different measures of learning across trials, proactive interference, retroactive interference, forgetting speed, and recognition memory were equivalent to 0.85.

De Paula et al. (2012) specifically evaluated the reliability (through Cronbach's alpha calculation) and the evidence of construct validity (obtained through its correlation with

Table 2
Instruments applied for measuring episodic memory, description of the investigated population, and duration of application for each of the tests

Instrument	Frequency	The investigated population	Age range (years)	Application time
CANTAB	1	Typical	64-93	30 min
CERAD	4	Typical and atypical (AD e MCI)	≥ 65	30 min
CVLT	5	Typical and atypical (AD e MCI)	≥ 60	15 a 20 min
DLM	1	Typical	≥ 60	30 min
EBMT	5	Typical	≥ 60	30 min
FCSRT	6	Typical and atypical (AD e MCI)	≥ 60	50 a 55 min
Five Word Test	1	Typical	≥ 60	5 min
HVLT	2	Typical and atypical (slightly impaired episodic memory)	> 65	10 min
HVLT-R	3	Typical	≥ 60	> 25 min
ISLT	1	Typical	65-88	25 min
KOLT	1	Typical	71-74	NE
MBT	2	Typical and atypical (AD, aMCI and subjective cognitive impairment)	65-85	26 min
Penn Face Memory Test	1	Typical and atypical (MCI)	≥ 60	> 15 min
RAVLT	19	Typical and atypical (AD e MCI)	≥ 60	40 min
Recognition Test	1	Typical	> 60	NF
TMA-93	1	Typical and atypical (AD)	> 60	10 min
VAT-E	3	Typical and atypical (AD e MCI)	≥ 65	> 15 min
VPAT	1	Typical	≥ 66	1 a 4 min
Wechsler Logical Memory Test	4	Typical and atypical (AD e MCI)	55-85	> 30 min

CANTAB: Cambridge Neuropsychological Test Automated Battery, CVLT: California Verbal Learning Test, HVLT: Hopkins Verbal Learning Test, HVLT-R: Hopkins Verbal Learning Test-Revised, RAVLT: Rey Auditory Verbal Learning Test, FCSRT: Free and Cued Selective Reminding Test, EBMT: East Boston Memory Test, CERAD: Consortium to Establish a Registry for Alzheimer's Disease, KOLT: Kendrick Object Learning Test, DLM: Delayed Logical Memory, VAT-E: Visual Association Test-Extended, VPAT: Verbal Paired Associates Test, ISLT: International Shopping List Test, TMA-93: The Memory Associate Test, MBT: Memory Binding Test. NF: not found. AD: Alzheimer's Disease, MCI: Mild Cognitive Impairment, aMCI: Amnestic Mild Cognitive Impairment.

the Mini-Mental State Examination – MMSE, and with the Clock Drawing Test – CDT; as well as exploratory factor analysis using the principal axis factoring method and oblique rotation) of RAVLT in a sample of 126 older Brazilian adults. Regarding the construct validity hypothesis, the correlations between the different components of the test with MMSE and CDT ranged from non-significant ($r = 0.08$, $p = 0.405$) to moderate ($r = 0.448$, $r^2 = 0.20$, $p < 0.001$). Additionally, the two screening instruments showed a strong correlation with each other ($r = 0.713$, $r2 = 0.50$, $p < 0.001$), highlighting the distinction between their constructs and those measured by RAVLT.

Still, in the study by De Paula et al. (2012) regarding the hypothesis testing on internal consistency, the procedure used indicated that the different items of RAVLT showed a consistent pattern of correlations (α=0.831), indicating a high level of internal consistency. According to the authors, the analysis of reliability through internal consistency, up to the time of the article's publication, was the only methodology used to assess the reliability of RAVLT in Brazil. This result emphasizes the importance of evaluating the stability and precision of RAVLT in Brazil.

In this line, De Sousa Magalhães et al. (2012) also provided evidence for the reliability of the RAVLT in

their study using, in addition to internal consistency, the test-retest methodology, with a mean interval of 35 days between applications. Additionally, evidence of convergent and divergent validity was also reported. For this purpose, performance on RAVLT was compared with the Benton Visual Retention Test (BVRT) and the Trail Making Test (TMT), respectively. All test-retest correlation coefficients were significant, ranging from 0.36 to 0.68. The value of the Cronbach's alpha coefficient was 0.80. RAVLT measures, A1-A5 and A7, did not significantly correlate with the TMT measures. Conversely, these measures exhibited significant and moderate correlations with BVRT measures (ranging from 0.37 to 0.44).

According to De Sousa Magalhães et al. (2012), RAVLT includes clear and specific instructions that minimize potential variations among researchers during the administration and scoring process. Additionally, Cotta et al. (2012) state that the test is considered a quick, easy-to-administer, and efficient instrument for measuring cognitive abilities and detecting episodic memory losses while also allowing for differentiation between normal aging, mild cognitive impairment, and Alzheimer's disease dementia with relatively high accuracy (De Paula et al., 2013). All of those point to the favorable cost-benefit ratio of the test, which seeks to maintain assessment efficiency by maintaining their practicality without compromising quality in clinical care. These aspects may have influenced its widespread use in the articles found in this review.

As presented in Meyer et al.'s study (2016), criticism regarding the use of RAVLT lies in the high level of cognitive demands and complexity of the activities performed in the test, which are often challenging for older adults with more pronounced cognitive impairments such as those with Alzheimer's disease (AD). In the case of Meyer et al.'s study (2016), a floor effect was found for RAVLT in the sample of individuals with some impairment. According to the authors, 90% of all patients in this group scored lower than what would be expected. The use of the Visual Association Test (VAT) was proposed as an alternative by the authors to minimize such biases, as it is a test of episodic memory that expands encoding specificity through the presentation of visual stimuli to the patient. As described by Meyer et al. (2016), it shows little to no floor effect in patients with mild cognitive impairment due to aging and mild AD. This property of VAT is an exciting finding and suggests the importance of conducting VAT adaptations in other countries such as Brazil.

The second most commonly used instrument was FCSRT (Buschke, 1984; Grober & Buschke, 1987), a list-learning test developed to assess verbal episodic memory. FCSRT is a memory test that controls attention and cognitive processing, requiring subjects to retrieve items in response to category cues presented during the learning process (Grober et al., 1992). Its performance has commonly been associated with preclinical and early-stage dementia in various longitudinal epidemiological studies (Grober et al., 2009).

Different versions of FCSRT employ different numbers of items to be learned (for example, 12 or 16). The 16-item version is called the Free and Cued Selective Reminding Test with Immediate Recall (FCSR-IR) (Grober & Buschke, 1987). Unlike RAVLT, the material to be remembered is not read aloud by the examiner. In this sense, the fact that older adults see the stimuli to be recalled instead of just hearing them can reduce biases regarding task understanding and correct reproduction of the presented stimulus. After all, it is important to consider that some older adults may have impairments in sensory functions due to expected aging-related changes, leading them to misinterpret or incorrectly reproduce the presented stimulus, consequently obtaining lower scores on the test not due to recall failures per se but to impairments in hearing during encoding.

As Grober et al. (2009) present, there are different forms of FCSRT-IR (A, B, and C), each containing 16 simple line-draw items belonging to different semantic categories. Like any neuropsychological test, particularly those with verbal content, FCSRT requires linguistic adaptation and standardization for different populations. This may explain why FCSRT is not widely used in other countries such as Brazil.

FCSRT does not include any assessment of recognition memory. The analyses here focus on free recall and cued recall tasks. Since these processes typically require greater self-initiated and active processing by subjects (Craik, 1986), measures of episodic memory that rely solely on recall tasks may underestimate the potential of the individuals assessed, making it difficult to achieve the clinical differentiation that the test aims to achieve. However, precisely because recognition memory remains more stable throughout the aging process (Shing et al., 2010), measuring recognition memory can contribute to the development of cognitive interventions aimed at enhancing and safeguarding cognitive functions that are still preserved in old age, thereby promoting well-being, quality of life, and longevity of older adults.

Normative data for the original version of FCSRT were part of Mayo's Older Americans Normative Studies Project (MOANS). The sample consisted of 734 individuals aged between 56 and 98 years. The results were presented by age intervals, as education and gender showed minimal influence on participants' performance on the test (Ivnik et al., 1997).

Grober et al. (2009) sought to analyze the psychometric properties of FCSR-IR by examining the internal consistency of the three forms of FCSRT-IR, the dimensionality of FCSRT-IR (using exploratory and confirmatory factor analysis), the information provided by the items (using item response theory), and criterion validity by deriving ROC curves separately for each FCSRT-IR form, and then for combined forms.

Overall, the results provided support for the psychometric robustness of FCSRT-IR. First, factor analyses indicated that the 16 pairs of unique category items in each of the three forms assessed a single construct or dimension presumed to be memory capacity. Furthermore, for all forms, the overall Cronbach's alpha reliability coefficient was good: 0.85, 0.86, and 0.88 (for Forms A, B, and C, respectively). Corrected item-total correlations were moderate for all items, ranging from 0.37 to 0.55 for Form A, 0.34 to 0.60 for Form B, and 0.35 to 0.65 for Form C. Additionally, FCSRT-IR demonstrated good concurrent criterion validity according to ROC curves derived separately for each FCSRT-IR form and then for the combined forms. The criterion was the presence or absence of dementia, which was determined by expert consensus using DSM-IV diagnostic criteria. The area under the curve (AUC) was similar for the three forms, ranging from 0.89 (95% CI: 0.84, 0.94) for Form A, 0.87 (95% CI: 0.80, 0.94) for Form B, and 0.91 (95% CI: 0.86, 0.96) for Form C. Using a cutoff score previously established by the authors to distinguish dementia cases from control cases, the sensitivity and specificity of the three forms were similar: for Form A, sensitivity = 0.75, specificity = 0.82; for Form B, sensitivity = 0.75, specificity = 0.86; and Form C, sensitivity = 0.77, specificity = 0.87.

Finally, another instrument used was EBMT (Scherr et al., 1988). The East Boston Memory Test is a brief measure of verbal episodic memory used to screen for memory impairment indicative of Alzheimer's disease (AD) and other forms of dementia. EBMT was first used by Scherr and colleagues in 1988 as part of a broader study evaluating changes in cognition in a sample of 3,682 older adults. Performance levels for the test were determined in five-year age bands for those aged 65-94 years and those over 95 years old. Scherr et al. (1988) also provided a cutoff score to predict memory impairment within their sample. They used scores of 5 and 6 points to predict performance without impairment, and scores of 4 points or less were considered suggestive of memory impairment (Gfeller & Horn, 1996).

No evidence of validity and reliability values for EBMT was found in the review conducted. Since the article in which the test was first used (Scherr et al., 1988) correlates the performance of older adults on various tests, only general results were provided. It was observed that the strong negative relationship of age with performance on EBMT persisted after adjusting for the effects of other variables. Additionally, multivariate analyses showed that fewer years of formal education, lower occupational prestige category, and physical functional disability on the Katz scale for assessment of activities of daily living were also associated with lower levels of performance on both EBMT and other instruments employed in the study by Scherr et al. (1988). A significant strength of this study, and consequently an advantage of the instrument developed through it, is its use in a large sample of individuals. In the present review, studies that used the EBMT tool had a significantly larger sample than others, averaging ,2891 participants per study, while the overall average number of participants per study was 678 individuals. As Scherr et al. (1988) pointed out, smaller sample sizes obtained in earlier studies may have limited the research's ability to examine the relationship between test performance and age in older adults, especially in those aged 65 or older, due to the reduced number of people in this specific age group. It is important to notice that EBMT has the shortest application time among the three most commonly found in the present systematic review. A summary of the advantages, disadvantages, and main characteristics of the three most used tests among the articles found in this systematic review can be seen in Table 3.

In addition to the aspects mentioned and discussed, it is essential to emphasize that some theories suggest that two distinct processes can control episodic memory: recall and a non-recall form termed familiarity (Mandler, 1980; Jacoby, 1991; Yonelinas, 1994). Both are essential parameters that act in stimulus recognition, belonging to the Dual Processing Theory (Gomes et al., 2014). However, as evidenced in this review, none of the current instruments encompasses the measurement of such parameters. The lack of information concerning these parameters is a significant gap because this information can indicate which memory processes are being used for information retrieval and which are more or less compromised. This information can contribute to explaining memory deficits characteristic of older individuals and aid in preventing future conditions; if a pattern different from expected for the individual's age is found from these assessments, more effective biopsychosocial interventions can be established. Thus, future research should consider ways to assess memory processes and not just memory performance.

Table 3
Comparison between RAVLT, FCSRT, and EBMT

Instruments	Application method	Assessed memory component	Advantages	Disadvantages
RAVLT	Auditory stimuli	Verbal episodic memory	Simple and practical application, distinguishes between healthy elderly groups and cases of cognitive impairment, adapted and validated for various countries, the wide age range of elderly included (up to 89 years), assesses recall and recognition memory.	It elicits high cognitive demands, criticism regarding the influence of education on scores, may reveal task-misunderstanding errors, and does not measure memory parameters.

Table 3
Cont.

Instruments	Application method	Assessed memory component	Advantages	Disadvantages
FCSRT	Visual and auditory stimuli	Verbal episodic memory	Effectively assesses free recall and cued recall processes, distinguishes between healthy elderly groups and cases of cognitive impairment, identifies dementia cases early, and reduces task-misunderstanding errors by also providing visual stimuli.	Does not provide measures of recognition memory, lengthy application, small sample, does not measure memory parameters.
EBMT	Auditory stimuli	Verbal episodic memory	Good screening tool for dementia cases, distinguishes between healthy elderly groups and cases of cognitive impairment, quick application, and large samples investigated.	May reveal task-misunderstanding errors, and does not measure memory parameters.

FINAL CONSIDERATIONS

The present study provided an overview of the psychometric properties of the instruments used to assess episodic memory. The results of this systematic review pointed to a slight increase in the number of empirical publications in the last decade, revealing researchers' interest in investigating episodic memory. It is one of the most complex and important cognitive functions, undergoing significant changes throughout the aging process. However, the number of studies in the area is still small.

Most of the identified studies analyzed typical older adults regarding episodic memory functioning, which may indicate the difficulty of some instruments in differentiating natural age-related decline from more complex cases of cognitive impairments. The prevalence of cognitively healthy individuals also highlights the relevance of studies on the topic. After all, despite the growing number of cases of dementia and mild cognitive impairment in the current aged population, a large proportion of individuals aged 60 and over are still not within this spectrum of impairment, allowing investments in health and interventions prioritizing the improvement of the quality of life of these individuals to be carried out.

Regarding the type of each publication, most studies were characterized as correlational, with fewer studies emphasizing changes in episodic memory over multiple years. The predominance of cross-sectional studies is a reality regarding research involving older adults, mainly due to the particularities that such population exhibit. As for the instruments, RAVLT was the most used in the selected articles. Its predominance may be related to the practicality and speed of its application, which do not affect the effectiveness of the measures presented. Furthermore, RAVLT involves both free recall tasks and recognition memory. However, there are criticisms regarding the complex cognitive demands that the instrument entails, especially for populations with Alzheimer's Disease, in addition to the remarks on the effects of formal education on RAVLT performance.

From the analyses conducted, it is highlighted that further research is necessary for more instruments to be adapted to the Brazilian context since all studies found in this review used the same test. Additionally, there is a need for measures that encompass the parameters of recall and familiarity, considering their relevance. Furthermore, more experimental studies are needed due to their importance in evaluating interventions that aim to delay decline and enhance preserved episodic memory processes. In this sense, only two of the retrieved studies were intervention studies. This result emphasizes the need for more intervention studies, mainly preventive interventions, with the aim of sophistication and safeguarding cognitive functions, aiding in older adults' well-being and quality of life.

One limitation of the present study concerns the inclusion criteria. Only published articles were admitted, excluding theses, dissertations, conference research reports, and other systematic reviews. Additionally, the exclusion of articles that presented aggregated results concerning the age of the participants may also have been a limitation, as the scope of longitudinal studies may have been affected.

REFERENCES

Baddeley, A. (2011). Memória episódica: organizando e lembrando. In A. Baddeley, M. Anderson, & M. W. Eysenck (Eds.), *Memória* (pp. 15–30). Artmed.

Bensalem, J., Dudonné, S., Etchamendy, N., Pellay, H., Amadieu, C., Gaud, S., et al. (2019). Polyphenols from grape and blueberry improve episodic memory in healthy elderly with lower level of memory performance: A bicentric double-blind, randomized, placebo-controlled clinical study. *Journals of Gerontology: Series A, Biological Sciences and Medical Sciences, 74*(7), 996-1007. https://doi.org/10.1093/gerona/gly166

Buschke, H. (1984). Cued recall in amnesia. *Journal of Clinical and Experimental Neuropsychology, 6*(4), 433–440. https://doi.org/10.1080/01688638408401233

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (Vol. 5, pp. 409–422). North-Holland.

*Chen, S. D., Li, H. Q., Shen, X. N., Li, J. Q., Xu, W., Huang, Y. Y., … & Alzheimer's Disease Neuroimaging Initiative. (2020). Genome-wide association study identifies *SLAMF1* affecting the rate of memory decline. *Journal of Alzheimer's Disease, 74*(1), 139–149. https://doi.org/10.3233/JAD-191214

* Cheng, X., Yang, Y., Schwebel, D. C., Liu, Z., Li, L., Cheng, P., … Hu, G. (2020). Population ageing and mortality during 1990–2017: A global decomposition analysis. *PLOS Medicine, 17*(6), e1003138. https://doi.org/10.1371/journal.pmed.1003138

Cotta, M. F., Malloy-Diniz, L. F., Nicolato, R., de Moraes, E. N., Rocha, F. L., & de Paula, J. J. (2012). O Teste de Aprendizagem Auditivo-Verbal de Rey (RAVLT) no diagnóstico diferencial do envelhecimento cognitivo normal e patológico. *Contextos Clínicos, 5*(1), 10–25. https://doi.org/10.4013/ctc.2012.51.02

*De Paula, J. J., Bertola, L., Avila, R. T., Moreira, L., Coutinho, G., de Moraes, E. N., … & Malloy-Diniz, L. F. (2013). Clinical applicability and cutoff values for an unstructured neuropsychological assessment protocol for older adults with low formal education. *PLOS ONE, 8*(9), e73167. https://doi.org/10.1371/annotation/2a374fe8-6d25-45cb-b97b-a5ca1529aa14

De Paula, J. J. D., Melo, L. P. C., Nicolato, R., Moraes, E. N. D., Bicalho, M. A., Hamdan, A. C., & Malloy-Diniz, L. F. (2012). Fidedignidade e validade de construto do Teste de Aprendizagem Auditivo-Verbal de Rey em idosos brasileiros. *Archives of Clinical Psychiatry (São Paulo), 39*, 19–23. https://doi.org/10.1590/S0101-60832012000100004

De Paula, J. J., & Malloy-Diniz, L. F. (2018). *Manual do Teste de Aprendizagem Auditivo-Verbal de Rey*. Vetor.

De Sousa Magalhães, S., Fernandes Malloy-Diniz, L., & Cavalheiro Hamdan, A. (2012). Validity convergent and reliability test-retest of the Rey Auditory Verbal Learning Test. *Clinical Neuropsychiatry, 9*(3), 129–137.

*Ernst, M. E., Ryan, J., Chowdhury, E. K., Margolis, K. L., Beilin, L. J., Reid, C. M., … & Murray, A. M. (2021). Long-term blood pressure variability and risk of cognitive decline and dementia among older adults. *Journal of the American Heart Association, 10*(13), e019613. https://doi.org/10.1161/JAHA.120.019613

*Farina, M., Breno Costa, D., Webber de Oliveira, J. A., Polidoro Lima, M., Machado, W. D. L., Moret-Tatay, C., … & Irigaray, T. Q. (2020). Cognitive function of Brazilian elderly persons: Longitudinal study with non-clinical community sample. *Aging & Mental Health, 24*(11), 1807–1814. https://doi.org/10.1080/13607863.2019.1636203

Friedman, D., de Chastelaine, M., Nessler, D., & Malcolm, B. (2010). Changes in familiarity and recollection across the lifespan: An ERP perspective. *Brain Research, 1310*, 124–141. https://doi.org/10.1016/j.brainres.2009.11.016

Galvão, T. F., Pansani, T. S. A., & Harrad, D. (2015). Principais itens para relatar revisões sistemáticas e meta-análises: A recomendação PRISMA. *Epidemiologia e Serviços de Saúde, 24*(2), 335–342. https://doi.org/10.5123/S1679-49742015000200017

*Ge, S., Wu, B., Bailey Jr, D. E., & Dong, X. (2017). Social support, social strain, and cognitive function among community-dwelling US Chinese older adults. *Journals of Gerontology: Series A, Biomedical Sciences and Medical Sciences, 72*(suppl_1), S16–S21. https://doi.org/10.1093/gerona/glw221

*Gevins, A., Ilan, A. B., Jiang, A., Chan, C. S., Gelinas, D., Smith, M. E., … & O'Hara, R. (2011). A method to combine cognitive and neurophysiological assessments of the elderly. *Dementia and Geriatric Cognitive Disorders, 31*(1), 7–19. https://doi.org/10.1159/000322108

Gfeller, J. D., & Horn, G. J. (1996). The East Boston Memory Test: A clinical screening measure for memory impairment in the elderly. *Journal of Clinical Psychology, 52*(2), 191–196. https://doi.org/10.1002/(SICI)1097-4679(199603)52:2<191::AID-JCLP10>3.0.CO;2-F

Gifforda, K., Dandan, L., Damona, S., Chapman, W., Romano, R., Samuel, L., & Jefferson, A. (2015). Subjective memory complaint only relates to verbal episodic memory performance in mild cognitive impairment. *Journal of Alzheimer's Disease, 44*(1), 309–318. https://doi.org/10.3233/JAD-140636

Gomes, C. F. A., Brainerd, C. J., Nakamura, K., & Reyna, V. F. (2014). Markovian interpretations of dual retrieval processes. *Journal of Mathematical Psychology, 59*, 50–64. https://doi.org/10.1016/j.jmp.2013.07.003

Grober, E., & Buschke, H. (1987). Genuine memory deficits in dementia. *Developmental Neuropsychology, 3*(1), 13–36. https://doi.org/10.1080/87565648709540361

Grober, E., Gitlin, H. L., Bang, S., & Buschke, H. (1992). Implicit and explicit memory in young, old, and demented adults. *Journal of Clinical and Experimental Neuropsychology, 14*(2), 298–316. https://doi.org/10.1080/01688639208402830

Grober, E., Ocepek-Welikson, K., & Teresi, J. A. (2009). The Free and Cued Selective Reminding Test: Evidence of psychometric adequacy. *Psychology Science Quarterly, 51*(3), 266–282.

*Gross, A. L., & Rebok, G. W. (2011). Memory training and strategy use in older adults: Results from the ACTIVE study. *Psychology and Aging, 26*(3), 503–517. https://doi.org/10.1037/a0022687

*Grotz, C., Matharan, F., Amieva, H., Pérès, K., Laberon, S., Vonthron, A. M., … & Letenneur, L. (2017). Psychological transition and adjustment processes related to retirement: Influence on cognitive functioning. *Aging & Mental Health, 21*(12), 1310–1316. https://doi.org/10.1080/13607863.2016.1220920

*Guerrero-Berroa, E., Ravona-Springer, R., Heymann, A., Schmeidler, J., Levy, A., Leroith, D., & Beeri, M. S. (2015). Haptoglobin genotype modulates the relationships of glycemic control with cognitive function in elderly individuals with type 2 diabetes. *Diabetologia, 58*(4), 736–744. https://doi.org/10.1007/s00125-014-3487-2

Ivnik, R. J., Smith, G. E., Lucas, J. A., Tangalos, E. G., Kokmen, E., & Petersen, R. C. (1997). Free and cued selective reminding test: MOANS norms. *Journal of Clinical and Experimental Neuropsychology, 19*(5), 676–691. https://doi.org/10.1080/01688639708403753

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

Júnior, C. A., & Faria, N. (2015). Memória. *Psicologia Reflexão Crítica, 28*(4). https://doi.org/10.1590/1678-7153.201528416

*Klaming, R., Annese, J., Veltman, D. J., & Comijs, H. C. (2017). Episodic memory function is affected by lifestyle factors: a 14-year follow-up study in an elderly population. Aging, Neuropsychology, and Cognition, 24(5), 528–542. https://doi.org/10.1080/13825585.2016.1226746

Lee, J. K., Wendelken, C., Bunge, S. A., & Ghetti, S. (2016). A time and place for everything: Developmental differences in the building blocks of episodic memory. *Child Development, 87*(1), 194–210. https://doi.org/10.1111/cdev.12447

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Maillet, D., Narme, P., Amieva, H., Matharan, F., Bailon, O., Le Clésiau, H., & Belin, C. (2017). The TMA-93: A new memory test for Alzheimer's disease in illiterate and less educated people. *American Journal of Alzheimer's Disease & Other Dementias®, 32*(8), 461–467. https://doi.org/10.1177/1533317517722630

Malloy-Diniz, L.F., da Cruz, M.F., Torres, V., & Cosenza, R. (2000). O teste de Aprendizagem Auditivo-Verbal de Rey: Normas para uma população brasileira. *Revista Brasileira de Neurologia*, 36(3),79-83.

Malloy-Diniz, L. F., Lasmar, V. A. P., Gazinelli, L. D. S. R., Fuentes, D., & Salgado, J. V. (2007). The Rey auditory-verbal learning test: Applicability for the Brazilian elderly population. *Brazilian Journal of Psychiatry, 29*, 324–329. https://doi.org/10.1590/S1516-44462006005000053

Mandler, G. (1980). Recognizing: the judgment of previous occurrence. *Psychological Review*, 87, 252–271. http://dx.doi.org/10.1037/0033-295X.87.3.252.

*Messinis, L., Nasios, G., Mougias, A., Politis, A., Zampakis, P., Tsiamaki, E., ... & Papathanasopoulos, P. (2016). Age and education adjusted normative data and discriminative validity for Rey's Auditory Verbal Learning Test in the elderly Greek population. *Journal of Clinical and Experimental Neuropsychology, 38*(1), 23–39. https://doi.org/10.1080/13803395.2015.1085496

*Meyer, S. R. A., Boelaarts, L., Lindeboom, J., De Jonghe, J. F., & Ponds, R. (2020). Episodic recognition memory based on incidental learning of visual associations is largely preserved compared to recall in amnestic mild cognitive impairment and mild Alzheimer's disease. *Applied Neuropsychology: Adult*, 1–9. https://doi.org/10.1080/23279095.2019.1703705

*Meyer, S. R., Spaan, P. E., Boelaarts, L., Ponds, R. W., Schmand, B., & de Jonghe, J. F. (2016). Visual associations cued recall: A paradigm for measuring episodic memory decline in Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 23*(5), 566–577. https://doi.org/10.1080/13825585.2015.1132667

*Morris, J. K., Vidoni, E. D., Mahnken, J. D., Montgomery, R. N., Johnson, D. K., Thyfault, J. P., & Burns, J. M. (2016). Cognitively impaired elderly exhibit insulin resistance and no memory improvement with infused insulin. *Neurobiology of Aging, 39*, 19–24. https://doi.org/10.1016/j.neurobiolaging.2015.11.005

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, 16(5), 292–305. https://doi.org/10.1016/j.tics.2012.04.005

Raj, D., Santhi, S., & Sapharina, G. J. S. (2021). Effectiveness of neurobic exercise program on memory and depression among elderly residing at old age home. *Journal of Complementary and Integrative Medicine, 17*(4). https://doi.org/10.1515/jcim-2019-0221

*Ravona-Springer, R., Heymann, A., Schmeidler, J., Sano, M., Preiss, R., Koifman, K., ... & Beeri, M. S. (2014). The ApoE4 genotype modifies the relationship of long-term glycemic control with cognitive functioning in elderly with type 2 diabetes. *European Neuropsychopharmacology, 24*(8), 1303–1308. https://doi.org/10.1016/j.euroneuro.2014.05.001

*Requier, F., Hendy, A. S., Schlosser, M., Demnitz-King, H., Whitfield, T., Chételat, G., ... & Collette, F. (2023). Association between meditative capacities and cognitive functions in healthy older adults naïve to meditation practice. *Mindfulness, 14*(3), 695–707. https://doi.org/10.1007/s12671-023-02077-9

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

*Riaz, M., Huq, A., Ryan, J., Orchard, S. G., Tiller, J., Lockery, J., ... & Lacaze, P. (2021). Effect of APOE and a polygenic risk score on incident dementia and cognitive decline in a healthy older population. *Aging Cell, 20*(6), e13384. https://doi.org/10.1111/acel.13384

*Rocha, F. D. S., & Chariglione, I. P. F. S. (2021). Episodic memory and elderly people: Main alterations from different cognitive interventions. *Psicologia: Teoria e Pesquisa, 36*, e3637. https://doi.org/10.1590/0102.3772e3637

Rodrigues, G. S., & Jaeger, A. (2018). O Uso de Tarefas Experimentais para o Estudo da Memória Episódica. *Ciências & Cognição*, *23*(1). Recuperado de http://revista.cienciasecognicao.org/index.php/cec/article/view/1371

*Ruan, Q., Xiao, F., Gong, K., Zhang, W., Zhang, M., Ruan, J., ... & Yu, Z. (2020). Demographically corrected normative Z scores on the neuropsychological test battery in cognitively normal older Chinese adults. *Dementia and Geriatric Cognitive Disorders, 49*(4), 375-383. https://doi.org/10.1159/000508171

Scherr, P. A., Albert, M. S., Funkenstein, H. H., Cook, N. R., Hennekens, C. H., Branch, L. G., ... & Evans, D. A. (1988). Correlates of cognitive function in an elderly community population. *American Journal of Epidemiology, 128*(5), 1084-1101. https://doi.org/10.1093/oxfordjournals.aje.a115279

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience & Biobehavioral Reviews, 34*(7), 1080-1091. https://doi.org/10.1016/j.neubiorev.2010.04.002

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4*(3), 232-243. https://doi.org/10.1162/jocn.1992.4.3.232

*Striepens, N., Scheef, L., Wind, A., Meiberth, D., Popp, J., Spottke, A., ... & Jessen, F. (2011). Interaction effects of subjective memory impairment and ApoE4 genotype on episodic memory and hippocampal volume. *Psychological Medicine, 41*(9), 1997-2006. https://doi.org/ 10.1017/S0033291711000067

Tripathi, R., Kumar, J. K., Bharath, S., Marimuthu, P., & Varghese, M. (2013). Clinical validity of NIMHANS neuropsychological battery for elderly: A preliminary report. *Indian Journal of Psychiatry, 55*(3), 279. https://doi.org/10.4103/0019-5545.117149

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *The Oxford handbook of memory* (pp. 245–261). Oxford University Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

*Volz-Sidiropoulou, E., & Gauggel, S. (2012). Do subjective measures of attention and memory predict actual performance? *Metacognition in older couples*. *Psychology and Aging, 27*(2), 440–448. https://doi.org/ 10.1037/a0025384.

*Wang, X., Li, T., Zhang, H., Sun, T., Xiong, L., Zhang, Y., ... & Wang, H. (2019). Validation of the Chinese version of the memory binding test for distinguishing amnestic mild cognitive impairment from cognitively normal elderly individuals. *International Psychogeriatrics, 31*(12), 1721-1730. https://doi.org/10.1017/S1041610219001649

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1341–1354. https://doi.org/10.1037/0278-7393.20.6.1341

*Zhang, W., Tang, F., Chen, Y., Silverstein, M., Liu, S., & Dong, X. (2019). Education, activity engagement, and cognitive function in US Chinese older adults. *Journal of the American Geriatrics Society, 67*(S3), S525–S531. https://doi.org/10.1111/jgs.15560

*Articles included in the systematic review

Conflict of interest
The authors have no conflicts of interest to declare.

Data availability statement
The data supporting the findings of this study can be requested from the corresponding author upon reasonable request

Editor-in-chief
Tiago Jessé Souza de Lima

Associate Editor
Isabelle Chariglione

Corresponding author
Gabriela Campana Barbosa
E-mail: gabriela_campanakta@hotmail.com

Submitted on
08/03/2023

Accepted on
31/10/2023