Internal Consistency and Test-retest Reliability of the Visual Memory Associative Test TMA-93

EMILIO FRANCO-MACÍAS (✉ efranco17@gmail.com)

Hospital Universitario Virgen del Rocío <https://orcid.org/0000-0003-4186-1037>

Silvia Rodrigo-Herrero

Hospital Juan Ramon Jimenez. Huelva

Andrea Luque-Tirado

Hospital Universitario Virgen del Rocío

Manuel Medina-Rodriguez

Hospital Universitario Virgen del Rocío

Eugenia Graciani-Cantisán

Hospital Universitario Virgen del Rocío

Didier Maillet

Hopital Saint Louis. Paris

María Bernal Sánchez-Arjona

Hospital Universitario Virgen del Rocío

Research

Keywords: TMA-93, internal consistency, test-retest reliability, binding, mild cognitive impairment

Posted Date: June 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31492/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Memory tests focused on binding may be more sensitive to diagnose Alzheimer's disease (AD) at an early phase. TMA-93 examines binding by pairs of drawings of semantically-related objects. Preliminary validation studies have confirmed good diagnostic accuracy for the test to discriminate between patients with cognitive impairment and healthy controls (HCs) in low-educated populations. This study aimed to evaluate the reliability (internal consistency and test-retest reliability) of the TMA-93 in a clinic setting with a high proportion of individuals with low-educational attainment.

Methods: The study was undertaken in a memory clinic of a university hospital in Southern Spain. The internal consistency of the 10 items that compose the TMA-93 was estimated on a cross-sectional, case-control study. 35 patients with amnesic Mild Cognitive Impairment and 40 healthy controls (HCs) matched by age, gender, and educational attainment were tested by the TMA-93. The participant's performance on each of the 10 items was scored from 0 to 3, according to the instructions for the test (maximum total score= 30 points). The internal consistency was estimated by Cronbach's alpha. In addition, "split-half reliability", by Spearman-Brown coefficient; "corrected item-total correlations", and item redundancy, by "Cronbach's alpha if item deleted", were analyzed. The test-retest reliability for the TMA-93 total score was studied in 51 HCs tested by the same examiner 2-4 months apart and quantified by the intraclass correlation coefficient (ICC).

Results: Internal consistency was "optimal" (Cronbach's alpha=0.936). Split-half reliability was also high (Spearman-Brown coefficient= 0.911). The item-total score correlations were always above 0.660. No item was redundant as Cronbach's alpha did not increase at deleting anyone. Test-retest reliability was "good" [ICC=0.802 (CI 95%=0.653-0.887)].

Conclusion: The internal consistency and test-retest reliability of the TMA-93 are "good". In addition to other psychometric properties, these results support the test for examining memory binding, particularly in contexts of low-educational attainment.

Background

Binding or associative learning is the recall of an image or a word facilitated by exposure to a second image or word, with which the first was paired and simultaneously encoded [1, 2]. This is a key mechanism in memory formation that deteriorates when there is damage in the hippocampus and other structures of the medial temporal lobe [3, 4]. Memory tests focused on binding may be more sensitive than episodic memory tests to diagnose Alzheimer's disease (AD) at an early phase [5]. The binding cost is the proportion between remembering isolated features (unbound task) and integrated one (bound task) and is significantly higher in AD patients and may be a useful clinical marker of the disease [6].

In Neuropsychology, binding ability can be examined by different tests. The Wechsler Memory Scale (WMS) assesses binding through the learning and recall of paired associated words (Wechsler, 1987) [7]. This subtest of the WMS discerns between easy (i.e. North/South) and difficult associations (i.e.

School/Cellar) [7]. The most recently published “Memory Binding Test” (MBT) examines associative memory through the recall of pairs of items that belong to the same semantic category (i.e. flea/ant = insects), but are presented in two different lists of words [8]. Both tests rely on verbal information and may be hard for low-educated patients. Testing binding by images rather than words could be more appropriate in individuals with low education level. The Memory Associative Test of the district of Seine-Saint-Denis (TMA-93) was recently developed in France for the early diagnosis of AD among low educated immigrants [9]. Briefly, during the encoding phase the patient is shown ten pairs of drawings of daily life that are semantically related (Fig. 1A). In the recall phase, only one of the two items is shown and the patient is asked to recall the missing item (Fig. 1B) (Maillet et al, 2017). In the original paper, the test demonstrated high diagnostic accuracy for discriminating AD patients from healthy controls in a sample of immigrant residents from a district in Paris (France) [9]. A posterior validation study in older educationally-diverse Spanish people demonstrated that the test is so sensitive as the picture version of the Free and Cued Selective Reminding Test (FCSRT) in discriminating between amnesic Mild Cognitive Impairment (aMCI) patients and healthy controls (HCs) [10]. In this study, the receiver operating characteristic (ROC) analysis determined an optimal area under the ROC curve (AUC) of 0.97 (95% CI, 0.89-1.00, $p < 0.001$) to distinguish between the two groups [10].

Another asset of the TMA-93 is its short administration time. The average time needed for its completion was 6 min for aMCI patients and only 3 min for HCs [10]. In this sense, the test could be feasible for screening memory complaints in busy primary care and general neurology outpatient settings with limited face-to-face time per patient.

There are normative studies for the TMA-93 from French [11] and Spanish population [12]. From these studies, we know that the test ran with a ceiling effect which is moderated by age and education [12]. This ceiling effect has been considered an advantage for the test in diagnosing aMCI since a small number of errors can be a worrisome result for a patient [10].

These potential uses and advantages of the TMA-93 encourages the completion of the development of the test, including the study of other psychometric properties as reliability. The aim here was to study internal consistency and test-retest reliability of the TMA-93.

Methods

Design

Internal consistency was evaluated in an extension of a cross-sectional, case-control study with convenience sampling originally designed to compare the discriminative validity of the TMA-93 to distinguish aMCI patients from HCs [10].

The test-retest reliability was evaluated in healthy controls who were randomly selected from an ongoing normative study [12] and invited to repeat the TMA-93 conducted by the same examiner between 2 and 4 months after the initial examination.

Study population

The sample of the case-control study consisted of 75 participants from an urban area of Spain. They comprised 2 groups: 35 patients with a MCI and 40 HCs matched for age, gender, and educational level (incomplete primary studies, only primary studies completed, and higher than primary studies). All participants were older than 50 years and spoke Spanish as their native language. All patients were selected by convenience sampling of consecutive cases who were diagnosed of aMCI at the Memory Unit of the Hospital Virgen del Rocio (Seville, Spain). The procedures had consisted of general, neurological, neuropsychological, laboratory, and neuroimaging examinations. Neuropsychological evaluation included “Spanish Version of the Informant-Based AD8 Questionnaire” (AD8) [13], “Phototest” [14], “Delayed Matching-to-sample Task 48” (DMS-48) [15], “Free and Cued Selective Reminding Test” (FCSRT) [16], “Geriatric Depression Scale 15 items” (GDS-15) [17], and “Interview for Deterioration in Daily Living Activities in Dementia” (IDDD) [18]. The diagnosis of aMCI had been made according to the National Institute on Aging and Alzheimer’s Association (NIA-AA) recommendations [19] and operationally put into practice as follows: (a) memory complaints corroborated by a reliable informant (score equal or above 3 on AD8), (b) objective memory impairment measured by a score equal or below the 10-percentile on set 2 of DMS-48 (this score being lower than that on set 1), and (c) no significant functional decline for activities of daily living (score up to 39 on IDDD was allowed).

HCs for the case-control study were recruited among the caregivers and relatives of patients attending the center. They met the following inclusion criteria: (a) absence of memory complaints, (b) absence of objective memory impairment (DMS-48 set 2 score equal or above the 25-percentile), and (c) intact level of independence in activities of daily living (score between 33 and 36 on IDDD).

HCs for the test-retest reliability were recruited among the participants in the Spanish normative study for the TMA-93 [12]. Inclusion criteria for this study were: 1) age equal or above 50; 2) no cognitive complaints; 3) score equal or above percentile 10 according to normative data for the Phototest in Spain (Carnero Pardo et al, 2018); and 4) independent level of functioning [12]. 51 randomly selected HCs were invited to repeat the TMA-93 conducted by the same examiner between 2 and 4 months after the initial examination for studying the test-retest reliability.

Instrument: TMA-93

In both studies, TMA-93 was administered following the instructions given by its authors [9]. During the encoding phase, subjects were shown and asked to name 10 pairs of real life semantically-related objects presented as drawings in cards (tree/ bird, bed/ bedside lamp, boat/ fish, dog/ sheep, foot/ trousers, knife/ apple, glasses/ book, hand/ watch, car/ car keys, flower/ sun). The examiner specifically asked the participants to memorize the pairs of drawings (Fig. 1A). Next, the first associative memory trial was administered: examinees were shown only one of the two objects of each pair and asked to recall the missing one (Fig. 1B). After each subject’s response (regardless accuracy) or a period of up to 5 seconds, we displayed the pair again. This protocol was repeated for the 9 remaining pairs.

The maximum score of 30 points was granted only when the participant produced 10 correct responses in this first trial, in which case, the second and the third trials were omitted. Otherwise, the participants were scored from 0 to 9 based on their number of correct answers in this first trial and were administered a second similar trial with the same 10 pairs of drawings. If a subject correctly recalled the 10 missing objects in this second trial, s/he was given 20 points: 10 points corresponding to the second trial, and 10 more corresponding to the third trial, which was cancelled.

The score of each of the 10 items of the TMA-93 ranged from 0 to 3 and these scores were used for estimating the internal consistency of the test.

Statistical analyses

Descriptive results are shown as frequency (percent) for dichotomous and categorical variables, mean ($\pm$ SD, range) for normally-distributed continuous variables, and median [interquartile range (IQR), range] for non-normally distributed continuous variables. Between group comparisons of continuous variables were performed with Student's t test or its non-parametric alternative Mann-Whitney U test. Between group comparisons of categorical variables were performed with Chi square test. The diagnostic accuracy of the TMA-93 was estimated by the area under curve (AUC) using receiver operating characteristic (ROC) curve analysis and classified as excellent (> 0.9), good ($0.8-0.9$), and poor (< 0.8).

Internal consistency was estimated by Cronbach's alpha. Values of Cronbach's alpha above 0.70 were considered acceptable, between 0.90 and 0.95 were considered "optimal", and above 0.95 were interpreted as indicative of "item redundancy" [20, 21]. In addition, "split-half reliability" was analyzed considering the first five pairs of drawings of the TMA-93 as a half and the last five ones as the other half and considering the correlation between each other by Spearman-Brown coefficient. "Corrected item-total correlations" were calculated and a value below 0.40 was considered as indicative of item redundancy [21]. Item redundancy was also evaluated by "Cronbach's alpha if item deleted" [22], considering any item as redundant if Cronbach's alpha increased at deleting it.

Test-retest reliability was estimated by the intra-class correlation coefficient (ICC) and categorized as: excellent ($ICC > 0.90$), good ($ICC 0.71-0.90$), moderate ($ICC 0.51-0.70$), mediocre ($ICC 0.31-0.50$) or bad/null ($ICC < 0.31$) [21]. Finally, we created the variable "total score time 2 minus total score time 1" and analyzed its distribution. Statistical significance was set at a $p < 0.05$ and all estimates were obtained with a 95% confidence interval (CI 95%). All statistical analyses were run in SPSS version 25 (IBM, USA).

Results

Socio-demographics characteristics and neuropsychological background for the cross-sectional study are shown in Table 1. For the total sample ($n = 75$), 46 participants were females. Their average was 74.6 (SD = 5.9, range = 51–84). With regards to educational attainment, 31 individuals (41.3%) had not completed primary studies (Table 1).

Table 1
SOCIO-DEMOGRAPHICS CHARACTERISTICS AND NEUROPSYCHOLOGICAL BACKGROUND FOR THE
CROSS-SECTIONAL STUDY

	Total, n = 75	HCS, n = 40	aMCI, n = 35	p
Age	74.6 ± 5.9 (51–84)	74.7 ± 6.3 (51–83)	74.6 ± 5.4 (65–84)	0.706
Gender	46 (61.3%)	21 (52.5%)	25 (71,4%)	0.093
Female	29 (38.7%)	19 (47.5%)	10 (19.6%)	
Male				
Educational attainment	31 (41.3%)	12 (30%)	19 (54.3%)	0.052
<first grade	19 (25.3%)	14 (35%)	5 (14.3%)	
First grade	25 (33.3%)	14 (35%)	11 (31.4%)	
>first grade				
GDS-15 items (total score)	1, 0–3, (0–10)	1, 0–3, (0–6)	2, 1–4, (0–10)	0.245
Phototest (total score)	31,8 ± 7,5 (13–52)	36,3 ± 5,7 (26–52)	27,1 ± 6,3 (13–41)	< 0.001
DMS 48	45, 41–47, (31–48)	47, 46–47,5, (41–48)	41, 35–44, (31–47)	< 0.001
Set 1 score			36, 30–39, (26–45)	
Set 2 score	43, 36–47, (26–48)	47, 45–48, (40–48)		
TMA-93 (total score)	24, 14–29, (0–30)	29, 25,2–30, (14–30)	13, 6–20, (0–28)	< 0.001
FCSRT	23, 11–28, (1–39)	27, 24,2–32, (18–39)	10, 6–17, (1–35)	< 0.001
TFR	45, 36–48, (12–48)	47, 46–48, (42–48)	36, 24–43, (12–48)	< 0.001
TR	0,90, 0,66 – 1, (0,18 – 1)	0,95, 0,90 – 1, (0,76 – 1)	0,63, 0,42 – 0,81, (0,18 – 1)	< 0.001
ISC				< 0.001
Legend: Results are shown as median, interquartile range, and (range) for non-normal distributed variables and mean ± SD, and (range) for normal distributed variables.				
Abbreviations: TFR (Total Free Recall of the FCSRT); TR (Total Recall of the FCSRT); ISC (Cueing Sensitivity Index of the FCSRT)				

There were not significant differences in age, gender, educational attainment, or total score on GDS-15 between aMCI and HCs groups (Table 1). aMCI patients scored significantly lower than HCs on Phototest,

DMS-48, FCSRT, and TMA-93 (Table 1). The ROC curve analysis determined an AUC of 0.94 (95% CI, 0.89–0.99) for discriminating between groups.

Internal consistency was “good” (Cronbach’s alpha = 0.936). Split-half reliability considering the first five pairs of drawings of the TMA-93 as a half and the last five ones as the other half was also high (Spearman-Brown coefficient = 0.911). Corrected item-total score correlations ranges from 0.661 for the pair “hand-watch” to 0.837 for the pair “flower-sun” (Table 2). There was no redundancy of any item as Cronbach’s alpha did not increase at deleting anyone (Table 2).

Table 2
“CORRECTED ITEM-TOTAL CORRELATIONS AND CRONBACH’S ALPHA IF ITEM DELETED”

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
tree / bird	18,6667	67,793	,773	,927
bed / bedsidelamp	18,4933	69,199	,773	,928
boat / fish	18,5067	68,929	,744	,929
dog / sheep	18,4800	71,253	,677	,932
foot / trousers	18,7200	68,366	,722	,930
knife / apple	18,5867	68,921	,770	,928
glasses / book	18,9333	68,441	,719	,930
hand / watch	18,6267	70,940	,661	,933
car / car keys	18,6400	70,098	,754	,929
flower / sun	19,1867	65,262	,837	,924
Legend: Corrected Item-Total Correlation was never lower than 0.400. Cronbach’s alpha was not above 0.936 at deleting any item. Both results demonstrated no redundancy of any item.				

Table 3
SOCIO-DEMOGRAPHIC CHARACTERISTICS AND NEUROPSYCHOLOGICAL BACKGROUND OF THE TEST-RETEST STUDY

Age	64.8 ± 8.9 (50-86)
Gender	Female, n=30 (58.8%) Male, n=21 (41.2%)
Educational attainment	<first grade, n=13 (25.5%) First grade, n=16 (31.4%) >first grade, n=22 (43.1%)
Phototest	37.8 ± 4.8 (27-47)
TMA-93	29, 28-30, (23-30)
Legend: Results are shown as median, interquartile range, and (range) for non-normal distributed variables and mean ± SD, and (range) for normal distributed variables.	

Socio-demographics characteristics and neuropsychological background for the test-retest reliability study are shown in Table 3. Their average was 64.8 (SD = 8.9, range = 50–86). 13 subjects (25.5%) had not completed primary studies, and 30 (58.8%) were females (Table 3). Test-retest reliability was “good” [ICC = 0.802 (CI 95%=0.653–0.887)]. The “total score time 2 minus total score time 1” variable showed a non-normal, right asymmetric, and leptokurtic distribution (median = 0, IQR = 0–1, Range=-3-3) (Fig. 2). There were four atypical observations: two of them scored three points more at the retest and the remaining two scored two and three point less, respectively (Fig. 3).

Discussion

To our knowledge, this is the first study focused on internal consistency and test-retest reliability of the TMA-93, the French visual binding test [9]. Internal consistency was optimal and test-retest reliability was good.

Internal consistency among the 10 pairs of semantically-related drawings of real life objects that compose the test was optimal (Cronbach’s alpha = 0.936). This result means the 10 items of the test are highly correlated each other and measure the construct of interest, visual binding, in a homogeneous way [23].

“Corrected Item-Total Correlation” is the correlation of the item designed with the summated score for all other items. A rule-of-thumb states that this value should be at least 0.40 to rule out item redundancy [21]. Every item of the TMA-93 fulfilled the rule. In the same way, Cronbach alpha did not increase at deleting any of the ten pairs, so, again, no item redundancy could be demonstrated.

Split-half testing is another measure of internal consistency. This method measures the extent to which all parts of the test contribute equally to what is being measured. Here, the TMA-93 was virtually split into two halves and the scores from both parts of the test were correlated. We found a strong correlation between the two halves that indicates that participants of the cross-sectional study performed equally well (or as poorly) on both halves of the test.

The TMA-93 showed a “good” test-retest reliability by the intra-class correlation coefficient [ICC = 0.802 (CI 95%=0.653–0.887)]. Having good test-retest reliability signifies the internal validity of the TMA-93 and ensures that the measurements obtained in one setting are both representative and stable over time [24].

Analyzing the distribution of the “total score time 2 minus total score time 1” variable, there were four atypical observations. These four outliers probably precluded that the test-retest reliability of the TMA-93 could be upgraded to “excellent”. Two of them scored 3 points more at the retest. On the opposite side, two outliers scored 2 and 3 points less, respectively, at the retest. The former could be explained by practice effect and the latter by cognitive decline, but a more global explanation could be that binding is somewhat changeable and dynamic, making it difficult to achieve better test-retest reliability.

Both samples were composed of a relatively high percentage of low-educated participants. Lack of education remains as a limitation in many elderly Spanish people since they had limited access to primary school in the aftermath of the Spanish Civil War (1936–1939). Although the situation has significantly improved in recent years, 59% of the population over 65 years of age in Spain did not complete primary studies [25]. Neuropsychological examination must comply with this handicap. Accordingly, the design of this study included tests based on images as Phototest [14], DMS-48 [15], and the picture version of the FCSRT [16]. Being a specific memory test, its short administration time is an advantage for the TMA-93 in comparison to DMS-48 and FCSRT. Here, the test was again demonstrated accurate and feasible to be administered to low-educated individuals. Despite of this feasibility, the TMA-93 total score should be expected lower in low-educated individuals according to normative studies [11, 12].

In addition to optimal diagnostic accuracy previously reported for the TMA-93, the good reliability here demonstrated encourages the completion of the development of the test. Next steps will be phase II and phase III validation studies including AD biomarkers and comparing the diagnostic accuracy of the test with that of the standard memory instruments on samples organized by educational attainment.

Conclusion

This study focused on internal consistency and test-retest reliability of the TMA-93 shows a good reliability for the instrument. In addition to diagnostic accuracy and short administration time, this good reliability is another asset to choose the TMA-93 for testing binding, particularly in individuals with low educational attainment.

Declarations

Ethics approval and consent to participate

Both studies were approved by the ethics committee of the Hospital Virgen del Rocio (Seville, Spain) and conducted according to the World Medical Association Declaration of Helsinki. All participants accepted the study procedures by signing an informed consent.

Consent for publication

It was included in the informed consent

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

Didier Maillet is the author of the TMA-93.

Funding

The test-retest study was based on a normative study supported by Hoffmann-La Roche.

Authors contribution

SRH and EFM designed the study and wrote the paper. EGC, ALT, and SRH collected the data. EFM and MBS were responsible for carrying out the statistical analysis. D. Maillet assisted with writing the article.

Acknowledgements

We are thankful to patients and families involved in research in our memory unit.

References

1. Van Geldorp B, Parra MA, Kessels RP. Cognitive and neuropsychological underpinnings of relational and conjunctive working memory binding across age. *Memory*. 2015;23(8):1112–22.

2. Gramunt N, Sánchez-Benavides G, Buschke H, Lipton RB, Masramon X, Gispert JD, Peña-Casanova J, Fauria K, Molinuevo JL. Psychometric Properties of the Memory Binding Test: Test-Retest Reliability and Convergent Validity. *Journal of Alzheimer's Disease*. 2016;50:999–1010.
3. Paller KA: Binding memory fragments together to form declarative memories depends on cross-cortical storage. In: *Perspectives from Cognitive Neuroscience. Handbook of Binding and Memory*. Edited by Zimmer HD, Mecklinger A, Lindenberger U. New York; Oxford University Press; 2006: 527–544.
4. Mayes A, Montaldi D, Migo E. Associative memory and the medial temporal lobes. *Trends Cogn Sci*. 2007;11:126–35.
5. Parra MA, Abrahams S, Fabi K, Logie R, Luzzi S, Della Sala S. Short-term memory binding deficits in Alzheimer's disease. *Brain*. 2009;132:1057–66.
6. Cecchini MA, Yassuda MS, Bahia VS, de Souza LC, Guimaraes HC, Caramelli P, Carthery-Goulart MT, Patrocínio F, Foss MP, Tumas V, Lima-Silva TB, Brucki SMD, Nitrini R, Della Sala S, Parra MA. Recalling feature bindings differentiates Alzheimer's disease from frontotemporal dementia. *J Neurol*. 2017;264:2162–9.
7. Wechsler D. Wechsler Memory Scale. Revised Edition. New York: The Psychological Corporation; 1987.
8. Buschke H. The rationale of the Memory Binding Test. In: *Dementia and Memory*. Edited by Nilsson LG, Ohta N. New York: Psychology Press; 2014: 55–71.
9. Maillet D, Narme P, Amieva H, Matharan F, Bailon O, Le Clésiau H, Belin C. The TMA-93: A New Memory Test for Alzheimer's Disease in Illiterate and Less Educated People. *Am J Alzheimers Dis Other Dement*. 2017;32:461–7.
10. Rodrigo-Herrero S, Carnero-Pardo C, Méndez-Barrio C, De Miguel-Tristancho M, Graciani-Cantisán E, Sánchez-Arjona MB, Maillet D, Jiménez-Hernández MD, Franco-Macías E. TMA-93 for Diagnosing Amnesic Mild Cognitive Impairment: A Comparison with the Free and Cued Selective Reminding Test. *Am J Alzheimers Dis Other Dement*. 2019;34:322–8.
11. Dessi F, Maillet D, Metivet E, Michault A, Le Clésiau H, Ergis AM, Belin C. (2009). Assessment of episodic memory in illiterate elderly. *Psychol NeuroPsychiatr Vieil*. 2009; 7: 287–296.
12. Rodrigo-Herrero S, Sánchez-Benavides G, Ainz-Gómez L, Luque-Tirado A, Graciani-Cantisán E, Bernal Sánchez-Arjona M, Maillet D, Jiménez-Hernández MD, Franco-Macías E. Norms for Testing Visual Binding Using the Memory Associative Test (TMA-93) in Older Educationally-Diverse Adults. *J Alzheimers Dis*. 2020 Apr 24. Doi:10.3233/JAD-191235. [Epub ahead of print].
13. Carnero Pardo C, De la Vega Cotarelo R, López Alcalde S, Martos Aparicio C, Vélchez Carrillo R, Mora-Gavilán E, Galvin JE. Assessing the diagnostic accuracy (DA) of the Spanish version of the informant-based AD8 questionnaire. *Neurologia*. 2013;28(2):88–94.
14. Carnero-Pardo C, Lopez-Alcalde S, Allegri RF, Russo MJ, Sáez-Zea C, Montiel-Navarro L, Del Saz P, Fera-Vilar I, Pérez Navarro MJ, Ruiz-Giménez J, et al. Diagnostic accuracy of the Phototest for cognitive impairment and dementia. *Neurologia*. 2007;22(10):860–9.

15. Barbeau E, Didic M, Tramonì E, Felician O, Joubert S, Sontheimer A, Ceccaldi M, Poncet M. Evaluation of visual recognition memory in MCI patients. *Neurology*. 2004;62:1317–22.
16. Grober E, Lipton R, Hall C, Crystal H. Memory impairment on free and cued selective reminding predicts dementia. *Neurology*. 2000;54:827–32.
17. Yesavage JA, Brink TL, Rose TL, Lum O. Development and validation of a geriatric depression scale: a preliminary report. *J Psychiat Res*. 1983;17:37–49.
18. Teunisse S, Derix MM. Measurement of activities of daily living in patients with dementia living at home: development of a questionnaire. *Tijdschr Gerontol Geriatr*. 1991;22(2):53–9.
19. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging and Alzheimer's Association workgroup. *Alzheimers Dement*. 2011;7(3):270–9.
20. Nunnally JC, Bernstein IH. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.
21. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 3rd edition. Philadelphia: F.A. Davis Company; 2015.
22. Gliem JA, Gliem RR. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability for Likert-type scales. *Midwest Research to Practice Conference in Adult, Continuing, and Community Education*, Columbus, 82–88.
23. Welch S, Comer J. (1988). *Quantitative Methods for Public Administration: Techniques and Applications*. Editorial Books/Cole Publishing Co. U.S.A.
24. Guttman L. A basis for analyzing test-retest reliability. *Psychometrika*. 1945;10:255–82.
25. De la Fuente A, Doménech R. El nivel educativo de la población en España y sus regiones: 1960–2011. *Investigaciones Regionales*. 2016;34:73–94.

Figures

Figure 1

PAIRS OF SEMANTICALLY-RELATED DRAWINGS OF THE TMA-93 LEGEND: In the codification phase, the semantically-related drawings are presented in pairs (Figure 1A). In the recall phase, the subject has to recall the missing object (Figure 1B).

Figure 2

Histogram representing the distribution of the “total score time 2 minus total score time 1” variable.

Figure 3

Boxplot chart showing the distribution of the “total score time 2 minus total score time 1” variable. There are four outliers.