



Italian telephone-based Mini-Mental State Examination (Itel-MMSE): item-level psychometric properties

Edoardo Nicolò Aiello¹ · Antonella Esposito² · Veronica Pucci^{3,4} · Sara Mondini^{3,4} · Nadia Bolognini^{2,5} · Ildebrando Appollonio⁶

Received: 20 September 2021 / Accepted: 20 November 2021 / Published online: 8 January 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Background The Italian telephone-based Mini-Mental State Examination (Itel-MMSE), despite being psychometrically sound, has shown relevant ceiling effects, which may negatively impact the interpretation of its scores. In address to overcome such an issue, this study aimed at providing item-level insights on the Itel-MMSE through Item Response Theory (IRT) analyses.

Methods Five-hundred and sixty-seven healthy Italian adults (227 males, 340 females; mean age: 51 ± 17 years, range 18–96; mean education: 13.31 ± 4.3 years). A two-parameter logistic IRT model was implemented to assess item discrimination and difficulty of the Itel-MMSE. Construct unidimensionality, statistical independence of items, and model and item fit were tested. Informativity levels were also assessed graphically.

Results With respect to the Itel-MMSE total score, ceiling effects were found in 92.7% of participants. Unidimensionality was violated; both model and item fit were poor; a few items showed statistical dependence. Both the whole test and its items proved to be scarcely informative, especially for medium-to-high levels of ability, except for attention and spatial orientation subtests, which consistently yielded the highest discriminative capability.

Discussion The Itel-MMSE appears to be most informative in low-performing healthy individuals. However, the present findings should not lead practitioners to aprioristically equate ceiling effects/low informativity to clinical uselessness. Items assessing attention and, to a lesser extent, spatial orientation appear to be the most informative.

Keywords Telephone-based · Mini-Mental State Examination · Cognitive screening · Item response theory

Introduction

In the context of remote neurological and geriatric health-care, telephone-based cognitive screening tests allow to the cognitive evaluation assessment of populations having poor access to in-person services due to either logistical issues (e.g., motor disabilities) or unequal geographical coverage of clinics [1]. Moreover, telephone-based cognitive screening are useful for the implementation of epidemiological studies on cognitive impairment [2] and dementia prevention campaigns [3].

Within the international literature on telephone-based cognitive screening tools, versions of the renowned Mini-Mental State Examination (MMSE [4]) administrable via the telephone are highly represented [5, 6]. In Italy, the telephone-based MMSE (Itel-MMSE) has shown good validity, reliability, and diagnostic properties [7, 8]. However, it also showed relevant ceiling effects [8], which may negatively

✉ Edoardo Nicolò Aiello
e.aiello5@campus.unimib.it

¹ PhD Program in Neuroscience, School of Medicine and Surgery, University of Milano-Bicocca, Via Cadore 48, 20900 Monza, Italy

² Department of Psychology, University of Milano-Bicocca, Milan, Italy

³ Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata (FISPPA), University of Padova, Padua, Italy

⁴ Human Inspired Technology Research Centre (HIT), University of Padova, Padua, Italy

⁵ Neuropsychological Laboratory, IRCCS Istituto Auxologico Italiano, Milan, Italy

⁶ Neurology Section, School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

impact the interpretation of its scores, as it has been highlighted for the in-person version of the MMSE [9].

A possible approach to overcome such issues is to deliver item-level information in the framework of Item Response Theory (IRT). IRT allows assessing item measurement properties in relation to the underlying trait that they are meant to target [10]. When approaching cognitive screening through IRT, two item features are of major interest: difficulty, that is the level of latent ability required for an individual to “pass” an item, and discrimination, namely the capability of an item to discriminate between individuals having different levels of ability [10].

In this view, IRT would undoubtedly help practitioners gain as much information as possible from each item when an overall score is not sufficient to draw clinical judgments on individual cognitive profiles [11], as in the case of ceiling effects.

With respect to the Italian paper-and-pencil MMSE, IRT has proved to represent a successful approach to reveal item-level features based on demographic confounders [12]. Consistently, this study was intended to provide IRT-based, item-level psychometric insight into the ITEL-MMSE in an Italian population sample with the aim of helping practitioners interpret its outcome.

Methods

Participants

Five-hundred and sixty-seven individuals (227 males, 340 females; mean age: 50.99 ± 17.04 years, range = 18–96 years; mean education: 13.31 ± 4.25 years, range = 1–26 years) with no history of neurological, psychiatric, severe general medical conditions (i.e., organ/system failures, non-compensated metabolic disorders), active psychopharmacological therapies, or hearing deficits were recruited from different regions of Italy (445 participants recruited in Northern, 122 in Center, and 286 in South Italy; see Supplementary Table 1 for sample stratification for age, education, and sex). Participants were recruited between 2020 and 2021 *via* word-to-mouth advertising through personal acquaintances of researchers from the University of Milano-Bicocca and the University of Padua. Medical history was investigated by means of a semi-structured interview exploring the areas of neurological, psychiatric, general medical and psychopharmacological history.

This study was conducted in accordance with the Declaration of Helsinki and approved on behalf of the Ethical Committees of the University of Milano-Bicocca and the University of Padua. Participants provided informed consent to participation.

Materials

The ITEL-MMSE by Metitieri et al. [7] depletes items on reading, writing, oral comprehension, and constructional praxis; moreover, the “place” item for spatial orientation is dropped, and only one naming-to-description item is maintained. As for the in-person format, if at least one error is committed on the serial subtraction task, an alternative serial spelling task is administered to test sustained attention. The total score of the test ranges from 0 to 22.

Procedures

An in-depth sound-check was preliminarily carried out to ensure a good quality of the call. Participants were instructed on those actions that were required to execute the tasks. The presence of an informant was required to make sure that no facilitations occurred during the test administration, as well as to confirm address information (which allowed to test spatial orientation).

Statistics

R 4.0.1 [13] and SPSS 27 [14] were adopted for the statistical analyses.

Each of the 22 items was dichotomized and cognitive efficiency was addressed as the latent ability. In respect to sustained attention tasks, only serial subtraction was taken into account (as being administered first and thus yielding no missing values).

Item difficulty and discrimination were examined by means of a two-parameter logistic IRT model through the R package *mirt* [15]. “Canonical” difficulty was addressed as ranging from -4 to $+4$ [16]. Cut-offs for “high” and “very high” discrimination were set at ≥ 1.5 and ≥ 1.7 , respectively [16].

According to the guidelines proposed by Şahin and Anıl [17], the minimum sample size for the accurate estimation of a two-parameter logistic IRT model (i.e., including item difficulty and discrimination) was set at $N \approx 500$ on the basis of the number of items of the ITEL-MMSE ($n = 22$).

mirt [15] was also adopted to check for IRT assumptions.

Unidimensionality (i.e., whether the test actually measures a unique latent ability/cognitive efficiency) was tested through an exploratory factor analysis [10]. Local independence (i.e., whether items are statistically independent of each other, net of their association with the latent trait) [10] was checked *via* Yen’s Q3 statistic (i.e., the correlation between item residuals, which quantify the discrepancy between expected and observed values) [10, 18]. Each item yielding $Q3 \geq .36$ toward the others was judged as locally dependent [19]. Item fit (i.e., the degree of consistency between items and the test as a

whole in measuring the latent trait) and model fit (i.e., whether the estimated two-parameter logistic model is consistent with observed data) were judged by assessing the root mean square error of approximation (RMSEA); higher RMSEA values are suggestive of a deviation from the ideal fit [20]. RMSEA $\geq .06$ were addressed as indexing poor item/model fit.

Results

Mean aggregated Itel-MMSE scores were the following: total: $21.54 \pm .88$ (range = 14–22); orientation: $8.84 \pm .42$ (range = 6–9); attention: $4.93 \pm .38$ (range = 1–5); memory: $5.84 \pm .47$ (range = 2–6); and language: $1.93 \pm .28$ (range = 0–2).

Ceiling effects for the total score were cumulatively detected in 90.2% of participants (69.7% participants scoring the maximum, 20.5% scoring 21 out of 22).

Cronbach's α was acceptable (0.65). However, the exploratory factor analysis suggested a modest violation of the unidimensionality assumption (Supplementary Table 2), with several items loading < 0.3 . Accordingly, most items showed poor fit (Table 1), as well as the model as a whole (RMSEA = 0.062). Partial local dependence was detected for the first 3 items assessing spatial orientation and the last 2 serial subtractions ($0.48 \leq Q3 \leq 0.75$). Other items proved to be locally independent ($Q3 \leq 0.28$).

A summarization of item difficulty and discrimination is shown in Table 2.

As for difficulty, estimates overall proved to be unreliable (i.e., extremely high difficulty in spite of clear ceiling effects), whereas serial calculation items and, to a lesser extent, spatial orientation ones proved to be the most discriminative, as visually suggested by item characteristic functions (i.e., a graphical representation of difficulty and discrimination parameters; Fig. 1) and item information curves (representing the extent of informativity of each item in respect to individuals' level of ability; Fig. 2).

The visual representation of the expected total score is suggestive of a strong ceiling effect and a general scarce discriminative capability, except for individuals with low ability levels (Supplementary Fig. 1). Consistently, the test information curve (visually representing the extent to which the test as a whole is informative in respect to the individual's ability) showed informativity peaks and low standard errors for low levels of ability, whereas poor informativity and unreliable estimates for ability levels were equal to/greater than the mean (Fig. 3).

Discussion

The present work provides Italian practitioners and researchers with item-level psychometric insights on the Itel-MMSE from a large Italian population sample, in order to ease the

Table 1 Item fit values for the Itel-MMSE

	Items	RMSEA
1	Temporal orientation—1	0
2	Temporal orientation—2	0.03
3	Temporal orientation—3	–
4	Temporal orientation—4	0
5	Temporal orientation—5	0
6	Spatial orientation—1	–
7	Spatial orientation—2	–
8	Spatial orientation—3	–
9	Spatial orientation—4	0.02
10	Immediate recall—1	–
11	Immediate recall—2	0.01
12	Immediate recall—3	0.01
13	Serial subtraction—1	0.02
14	Serial subtraction—2	0.02
15	Serial subtraction—3	0
16	Serial subtraction—4	0.09
17	Serial subtraction—5	0.02
18	Delayed recall—1	0.03
19	Delayed recall—2	0.05
20	Delayed recall—3	0
21	Naming	0
22	Repetition	0.04

RMSEA Root Mean Square of Error Approximation of the Itel-MMSE

RMSEA values are referred to $S\text{-}\chi^2$ statistics. Adequately fitting items ($\text{RMSEA} \leq 0.06$) are in bold

interpretation of its outcomes in both clinical and experimental settings.

The Itel-MMSE proved to be scarcely informative of individuals' cognitive profiles as far as the most of its items are concerned, except for those assessing sustained attention (i.e., serial subtraction), as well as for certain ones exploring spatial orientation. Moreover, these items tended to yield the highest level of informativity for low levels of ability, this being especially true for spatial orientation ones. Consistently, as a whole, the Itel-MMSE was most informative in low-performing individuals. In this regard, previous evidence on the clinical usability of the Itel-MMSE showed that this test may offer a valid estimate of the cognitive *status* in patients with mild, moderate, and even severe dementia [7], while healthy individuals score the maximum on the Itel-MMSE [8]. Further studies are needed in order to determine to which degree of cognitive impairment the Itel-MMSE is sensitive, since the previous findings suggest it might be able to detect also mild deficits [7], at variance with the present results.

The poor features of the Itel-MMSE in the healthy population are attributable to the low across-individual variability

Table 2 Item difficulty and discrimination for the IteI-MMSE

Subtest	Items (N)	Difficulty	Discrimination
Orientation	Temporal (5)	4.01 ± 0.95 (2.42–4.95)	0.38 ± .23 (0–0.58)
	Spatial (4)	9.03 ± 5.94 (4.76–17.81)	2.05 ± 2.58 (0.28–5.79)**
Memory	Immediate recall (3)	4.7 ± 0.84 (3.75–5.38)	0.22 ± 0.29 [(- 0.04)–(0.16)]
	Delayed recall (3)	3.26 ± .36 (2.88–3.58)	0.21 ± 0.44 [(- 0.3)–(0.53)]
Attention	Serial calculation (5)	6.12 ± 3.26 (2.39–11.36)	5.24 ± 3.82 (2.06–11.51)**
Language	Naming (1)	3.42 [†]	0.34
	Repetition (1)	3.51 [†]	0.56

“Canonical” difficulty was addressed as ranging from - 4 to + 4[†] [16]. Cut-offs for “high” and “very high” discrimination were set at ≥ 1.5* and ≥ 1.7**, respectively [16]. Item difficulty refers to the level of latent ability required for an individual to “pass” an item. Item discrimination refers to the capability of an item to discriminate between individuals having different levels of ability

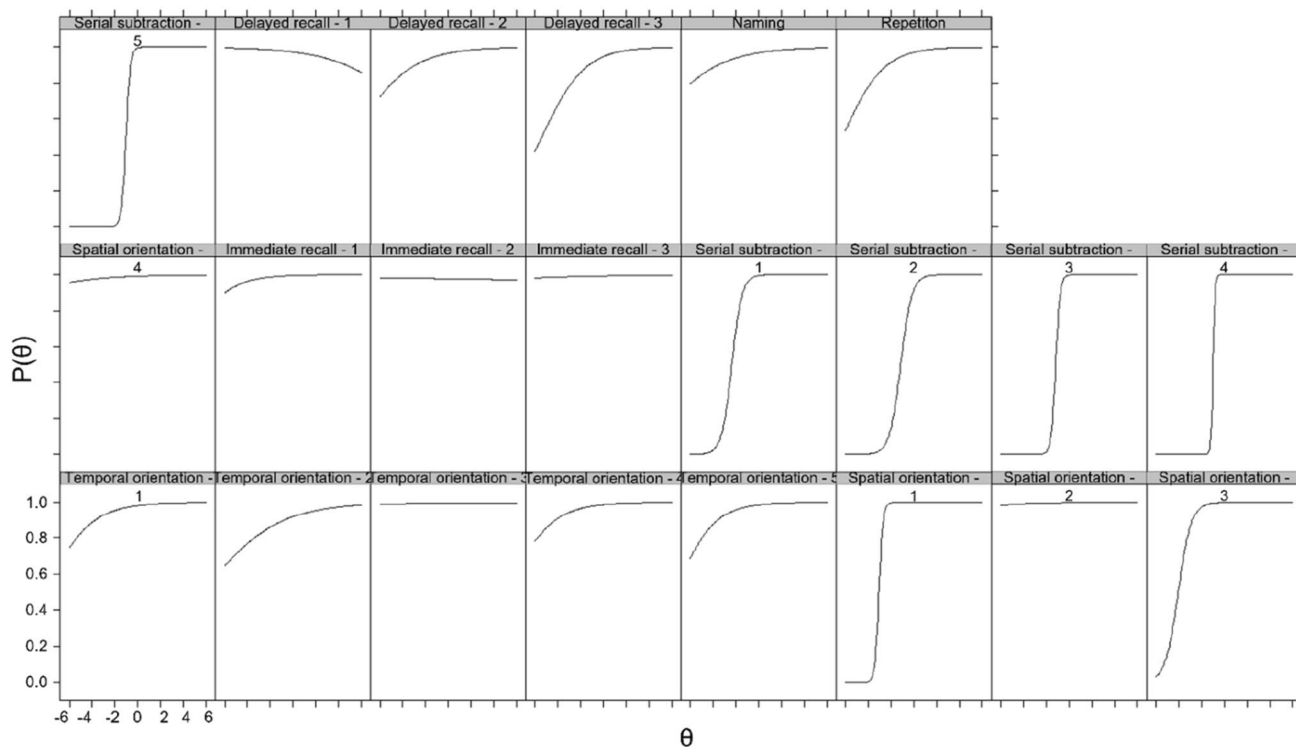


Fig. 1 Item characteristic functions for IteI-MMSE items. On x-axis, levels of ability θ (theoretically ranging from $-\infty$ to $+\infty$, whereas conventionally ranging from - 6 to + 6) are expressed in the logarithm of the odds ratio: above-zero values indicate levels of ability above the mean; below-zero values indicate levels of ability below the mean. On y-axis, the probability of a correct response $P(\theta)$ (ranging

from 0 to 1). Flatter curves are suggestive of low difficulty (ceiling effect) and scarce discriminative capability. By contrast, the steeper portions of the curve in relation to a given level of ability represent the ability level(s) at which the test is most discriminative. The items that prove to be the most difficult and discriminative are those of *Serial subtraction* and *Spatial orientation -1* and *-3*

in scores, which may be in turn related to strong ceiling effects [9]. However, these findings should not lead practitioners to aprioristically equate ceiling effects/low informativity to clinical uselessness. Indeed, cognitive measures showing ceiling effects in healthy individuals might not necessarily behave the same when applied to patients [21]. For instance, a given patient failing 2 out of 3 immediate recall items will be undoubtedly judged as more impaired than another scoring the maximum, even though this subset

of items shows near-zero variability and markedly goes to ceiling in unimpaired individuals.

Further investigations on the IteI-MMSE are thereupon needed to determine whether such inferences actually apply to target clinical populations [22]. In this last respect, it might be of interest for future studies to explore the potential of the IteI-MMSE in detecting cognitive complaints in healthy adults and in elderly related to the COVID-19 pandemic and its social restrictions [23, 24].

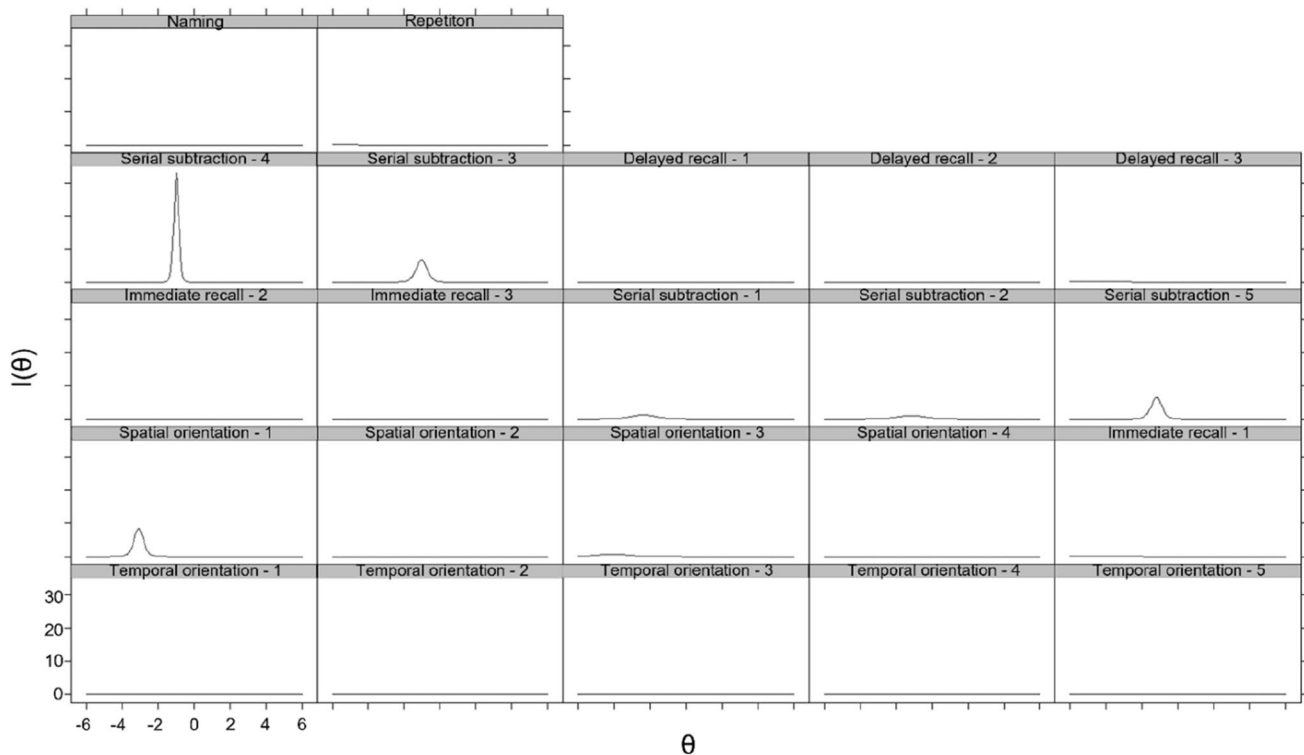


Fig. 2 Item information curves for ITEL-MMSE items. On x -axis, levels of ability θ (conventional range = -6 to $+6$) are expressed in the logarithm of the odds ratio: above-zero values indicate levels of ability above the mean; below-zero values indicate levels of ability below the mean. On y -axis, the level of informativity of a given item $I(\theta)$ (ranging from 0 to $+\infty$). Curves peaking index items that are

informative to an extent. The items that prove to be the most informative with respect to the individual level of cognitive functioning are *Spatial orientation—1* and *Serial subtraction -3, -4, and -5*. These items are more informative in low performers (i.e., low ability levels, $-4 < \theta < 0$)

Furthermore, the present item-level investigation shed further light on the factorial structure underlying the MMSE, regardless of its version (in presence *vs.* telephone-based). Indeed, findings herewith reported underscore the multidomain nature of the ITEL-MMSE, with attention items somehow “carrying” the whole test as far as informativity is concerned, in agreement with previous contributions within the field of clinical usability of the in-person MMSE [25, 26].

A major limitation of the present study should be borne in mind, namely, the fact that two-parameter logistic model estimates might have been distorted due to the poor fit of both the model and the items. However, at an explorative and descriptive level, it is believed that information here reported may come in handy both for practitioners to orient themselves in adopting the ITEL-MMSE as a telephone-based cognitive screening tool, and to promote further studies on the statistical properties and usability of this instrument.

A more appropriate standardization study of the ITEL-MMSE in line with current methodological standards on remote cognitive testing [27, 28] is desirable. Indeed, it remains to be derived norms specific to the remote modality of delivery [27], to assess its construct and criterion validity,

also exploring the association with other telephone-based measures of cognitive status, as well as with in-person cognitive screening tests [29, 30], including the assessment of the equivalence between the ITEL-MMSE and its in-person version by adopting ad-hoc statistical methods [31]. Such a standardization study should also include the following aspects: the exploration of both the inter-rater and the test–retest reliability of the ITEL-MMSE, considering the potential subjectivity in scoring approaches with remote assessments [30]; the assessment of the diagnostic properties of the ITEL-MMSE mostly relevant for screening tests (i.e., sensitivity, specificity, positive and negative predictive values and likelihood ratios) [22, 32], as well as of its ability to discriminate cases from controls and its sensitivity to detect changes over time [22, 32]. With respect to psychometrics, diagnostics and norms for telephone-based cognitive screening tests, two recent works by Aiello et al. [33, 34] can be taken as righteous approaches of standardization.

In conclusion, this study suggests that, when interpreting ITEL-MMSE scores, clinicians should focus on those items targeting attention and, to a lesser extent, spatial orientation;

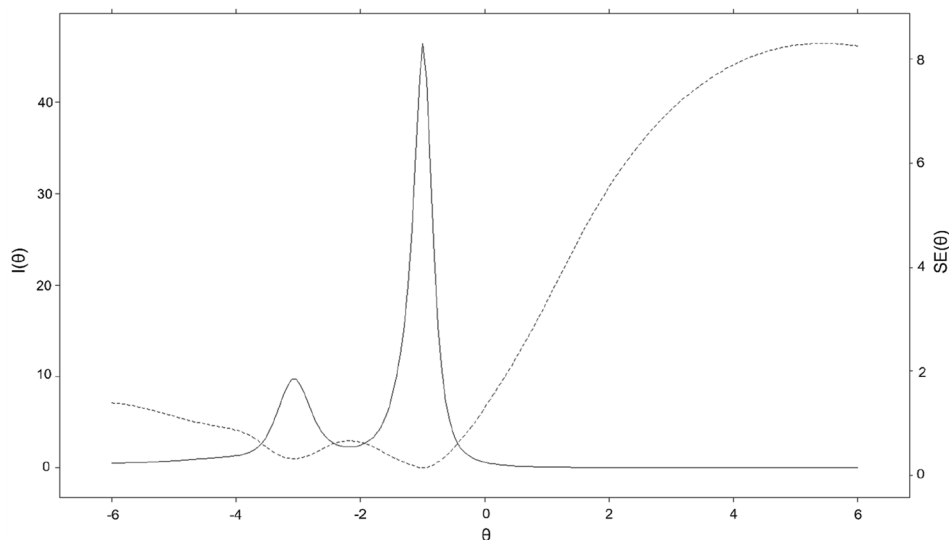


Fig. 3 Test information curve and standard error for the Itel-MMSE. On x -axis, levels of ability θ (conventional range = -6 to $+6$) are expressed in the logarithm of the odds ratio: above-zero values index levels of ability above the mean; below-zero values index levels of ability below the mean. On y -axis, the level of informativity of the test as a whole $I(\theta)$ (ranging from 0 to $+\infty$) is reported on the left, whereas the standard errors of the informativity estimated $SE(\theta)$

(ranging from 0 to $+\infty$) is reported on the right. The solid line represents the test information curve, whereas the dotted one the standard error of estimates. Peaks for the solid line represent the levels of ability at which the test is the most informative. By contrast, peaks for the dotted line shows at which levels of ability informativity estimates are unreliable. The total Itel-MMSE score is highly informative in low-performer individuals (i.e., low levels of ability, $-4 < \theta < 0$)

overall, the test appears to be more useful for screening cognition in low-performing individuals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40520-021-02041-4>.

Acknowledgements The authors thank Dr. Tiziana Metitieri and Dr. Paolo Caimi for providing us with the permission to use the Italian telephone-based Mini-Mental State Examination.

Author contributions ENA contributed to conceptualization, statistical analyses, and drafting, revision; AE contributed to conceptualization and statistical analyses; VP and SM were involved in data collection, drafting, and revision; NB and IA were involved in conceptualization, data collection, drafting, and revision.

Funding This work did not receive any specific funding.

Data availability Data collected and analyzed within the present study are openly accessible at <https://osf.io/jswh7/>

Declarations

Conflict of interest E.N.A., N.B., and I.A. are editing the Italian standardization of the Telephone Interview for Cognitive Status on behalf of Giunti Psychometrics. A.E., V.P., and S.M. have nothing to declare.

Statement of human and animal rights All procedures performed in this study that involved human participants were in accordance with the ethical standards of the Ethics Committee of the Universities of Milano-Bicocca (Ethics Committee I.D.: RM-2121-382, 19 February, 2021) and Padova (Ethics Committee I.D.: 4017, 24 March,

2021) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This study received ethical approval.

Informed consent Participants provided informed consent.

References

- Gatz M, Reynolds CA, John R, Johansson B, Mortimer JA, Pedersen NL (2002) Telephone screening to identify potential dementia cases in a population-based sample of older adults. *Int Psychogeriatr* 14:273–289
- Herr M, Ankri J (2013) A critical review of the use of telephone tests to identify cognitive impairment in epidemiology and clinical research. *J Telemed Telecare* 19:45–54
- Crooks VC, Clark L, Petitti DB, Chui H, Chiu V (2005) Validation of multi-stage telephone-based identification of cognitive impairment and dementia. *BMC Neurol* 5:1–8
- Folstein MF, Folstein SE, McHugh PR (1975) “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189–198
- Castanho TC, Amorim L, Zihl J, Palha JA, Sousa N, Santos NC (2014) Telephone-based screening tools for mild cognitive impairment and dementia in aging studies: a review of validated instruments. *Front Aging Neurosci* 6:16
- Elliott E, Green C, Llewellyn DJ, Quinn TJ (2020) Accuracy of telephone-based cognitive screening tests: systematic review and meta-analysis. *Curr Alzheimer Res* 17:460–471
- Metitieri T, Geroldi C, Pezzini A, Frisoni GB, Bianchetti A, Trabucchi M (2001) The Itel-MMSE: an Italian telephone version of the Mini-Mental State Examination. *Int J Geriatr Psychiatry* 16:166–167

8. Vanacore N, De Carolis A, Sepe-Monti M, Bomboi G, Stazi A, Bianchetti A, Giubilei F (2006) Validity of the Italian telephone version of the Mini-Mental State Examination in the elderly healthy population. *Acta Neurol Belg* 106:132–136
9. Franco-Marina F, García-González JJ, Wagner-Echeagaray F, Gallo J, Ugalde O, Sánchez-García S, Espinel-Bermúdez C, Juárez-Cedillo T, Rodríguez MÁ, García-Peña C (2010) The Mini-mental State Examination revisited: ceiling and floor effects after score adjustment for educational level in an aging Mexican population. *Int Psychogeriatr* 22:72–81
10. Baylor C, Hula W, Donovan NJ, Doyle PJ, Kendall D, Yorkston K (2011) An introduction to item response theory and Rasch models for speech-language pathologists. *Am J Speech Lang Pathol* 20:243–259
11. Aiello EN, Gramegna C, Esposito A, Gazzaniga V, Zago S, Difonzo T, Maddaluno O, Appollonio I, Bolognini N (2021a) The Montreal Cognitive Assessment (MoCA): updated norms and psychometric insights into adaptive testing from healthy individuals in Northern Italy. *Aging Clin Exp Res*. <https://doi.org/10.1007/s40520-021-01943-7>
12. Crane PK, Gibbons LE, Jolley L, Van Belle G, Selleri R, Dalmondo E, De Ronchi D (2006) Differential item functioning related to education and age in the Italian version of the Mini-mental State Examination. *Int Psychogeriatr* 18:505–515
13. R Core Team (2021). R: A Language and environment for statistical computing. (Version 4.1.0) [Computer software]. Retrieved from <https://cran.r-project.org>
14. IBM Corp (2021) IBM SPSS Statistics for Windows, Version 27.0. IBM Corp, Armonk
15. Chalmers RP (2012) mirt: a multidimensional item response theory package for the R environment. *J Stat Softw* 486:1–29
16. Arifin WN, Yusoff MSB (2017) Item response theory for medical educationists. *Edu Med J* 9:69–81
17. Şahin A, Anıl D (2017) The effects of test length and sample size on item parameters in item response theory. *Edu Sci Theory Pract* 17:321–335
18. Yen WM (1993) Scaling performance assessments: strategies for managing local item dependence. *J Educ Meas* 30:187–213
19. Tan Q, Cai Y, Li Q, Zhang Y, Tu D (2018) Development and validation of an item bank for depression screening in the Chinese population using computer adaptive testing: a simulation study. *Front Psychol* 9:1225
20. Xia Y, Yang Y (2019) RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: the story they tell depends on the estimation methods. *Behav Res Methods* 51:409–428
21. Ilardi CR, Chieffi S, Scuotto C, Gamboz N, Galeone F, Sannino M, Garofalo E, La Marra M, Ronga B, Iavarone A (2021) The frontal assessment battery 20 years later: normative data for a shortened version (FAB15). *Neurol Sci*. <https://doi.org/10.1007/s10072-021-05544-0>
22. Aiello EN, Rimoldi S, Bolognini N, Bolognini N, Appollonio I, Arcara G (2021b) Psychometrics and diagnostics of Italian cognitive screening test: a systematic review. *Neurol Sci* 1–25. <https://doi.org/10.1007/s10072-021-05683-4>
23. Maggi G, Baldassarre I, Barbaro A, Cavallo ND, Cropano M, Nappo R, Santangelo G (2021) Mental health status of Italian elderly subjects during and after quarantine for the COVID-19 pandemic: a cross-sectional and longitudinal study. *Psychogeriatrics* 21:540–551
24. Santangelo G, Baldassarre I, Barbaro A, Cavallo ND, Cropano M, Maggi G, Nappo R, Trojano L, Raimo S (2021) Subjective cognitive failures and their psychological correlates in a large Italian sample during quarantine/self-isolation for COVID-19. *Neurol Sci*. <https://doi.org/10.1007/s10072-021-05268-1>
25. Brugnolo A, Nobili F, Barbieri MP, Dessi B, Ferro A, Girtler N, Palummeri E, Partinico D, Raiteri U, Regesta G, Servetto G (2009) The factorial structure of the mini mental state examination (MMSE) in Alzheimer's disease. *Arch Gerontol Geriatr* 49:180–185
26. Shigemori K, Ohgi S, Okuyama E, Shimura T, Schneider E (2010) The factorial structure of the mini mental state examination (MMSE) in Japanese dementia patients. *BMC Geriatr* 10:1–7
27. Postal KS, Bilder RM, Lanca M, Aase DM, Barisa M, Holland AA, Lacritz L, Lechuga DM, McPherson S, Morgan J, Salinas C (2021) Inter organizational practice committee guidance/recommendation for models of care during the novel coronavirus pandemic. *Arch Clin Neuropsychol* 36:17–28
28. Zanin E, Aiello EN, Diana L, Fusi G, Bonato M, Niang A, Ognibene F, Corvaglia A, De Caro C, Cintoli S, Marchetti G, Bonato M, Vestri A (2021) Tele-neuropsychological assessment tools in Italy: a systematic review on psychometric properties and usability. *Neurol Sci*: 1–14. <https://doi.org/10.1007/s10072-021-05719-9>
29. Binng D, Splonskowski M, Jacova C (2020) Distance assessment for detecting cognitive impairment in older adults: a systematic review of psychometric evidence. *Dement Geriatr Cogn Dis*. <https://doi.org/10.1159/000511945>
30. Hunter MB, Jenkins N, Dolan C, Pullen H, Ritchie C, Muniz-Terrera G (2021) Reliability of telephone and videoconference methods of cognitive assessment in older adults with and without dementia. *J Alzheimer's Dis*. <https://doi.org/10.3233/JAD-210088>
31. Lakens D (2017) Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Person Sci* 8:355–362
32. Booth T, Murray A, Muniz-Terrera G (2021) Are we measuring the same thing? Psychometric and research considerations when adopting new testing modes in the time of COVID-19. *Alzheimers Dement* 17:251–254
33. Aiello EN, Esposito A, Giannone I, Diana L, Appollonio I, Bolognini N (2021c) Telephone Interview for Cognitive Status (TICS): Italian adaptation, psychometrics and diagnostics. *Neurol Sci*. <https://doi.org/10.1007/s10072-021-05729-7>
34. Aiello EN, Esposito A, Giannone I, Diana L, Woolley S, Murphy J, Christodoulou G, Tremolizzo L, Bolognini N, Appollonio I (2021d) ALS Cognitive Behavioral Screen-Phone Version (ALS-CBS™-PhV): norms, psychometrics, and diagnostics in an Italian population sample. *Neurol Sci*. <https://doi.org/10.1007/s10072-021-05636-x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.