Dement Neurocogn Disord. 2024 Jul;23(3):117-126
https://doi.org/10.12779/dnd.2024.23.3.117
pISSN 1738-1495·eISSN 2384-0757

DND Dementia and Neurocognitive Disorder

Original Article

Check for updates

A Comparison of Item Characteristics and Test Information Between the K-MMSE~2:SV and K-MMSE

Jihyang Kim,[1] **Seungmin Jahng**,[2] **SangYun Kim**,[3] **Yeonwook Kang** [1]

[1]Department of Psychology, College of Social Sciences, Hallym University, Chuncheon, Korea
[2]Department of Psychology, College of Social Sciences, Sungkyunkwan University, Seoul, Korea
[3]Department of Neurology, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Korea

OPEN ACCESS

Received: Jul 4, 2024
Revised: Jul 17, 2024
Accepted: Jul 18, 2024
Published online: Jul 25, 2024

Correspondence to
Yeonwook Kang
Department of Psychology, College of Social Sciences, Hallym University, 1 Hallymdaehak-gil, Chuncheon 24252, Korea.
Email: ykang@hallym.ac.kr

ORCID iDs
Jihyang Kim
https://orcid.org/0009-0004-0658-7436
Seungmin Jahng
https://orcid.org/0000-0002-1392-2229
SangYun Kim
https://orcid.org/0000-0002-9101-5704
Yeonwook Kang
https://orcid.org/0000-0003-4426-9127

Funding
This study was funded by the Korean Dementia Association and the Inpsyt, Inc., a branch of Hakjisa publisher.

Conflict of Interest
The authors have no financial conflicts of interest.

ABSTRACT

Background and Purpose: The Korean-Mini Mental State Examination, 2nd edition (K-MMSE~2) was recently released. This study aimed to determine whether the K-MMSE~2: Standard Version (K-MMSE~2:SV) had the same test characteristics as the K-MMSE.
Methods: A total of 1,514 healthy community-based participants aged 19 to 90 years were administered the K-MMSE~2:SV Blue Form along with the language items from the K-MMSE. The item and test characteristics and test information for the K-MMSE~2:SV and K-MMSE were compared using Item Response Theory analysis.
Results: Item discriminations for the K-MMSE~2:SV and K-MMSE were above the moderate range for all items except Recall. Most of the items on the K-MMSE~2:SV and K-MMSE had item category difficulty in the very easy or easy range. The test information curve (TIC) showed that the K-MMSE~2:SV and K-MMSE provide almost the same amount of information (27.86 vs. 28.44), with both tests providing the most information at an ability level of –1.57. The generalizability (G) coefficient for the K-MMSE~2:SV and K-MMSE was 0.99.
Conclusions: These results indicate that the K-MMSE~2:SV and K-MMSE are equally optimal tests for screening for mild cognitive impairment and early dementia. Given that the amount of test information provided by the two tests was almost identical, the shapes of the TICs were very similar, and the G coefficient was close to 1, we can conclude that the K-MMSE and K-MMSE~2:SV are equivalent tests.

Keywords: Neuropsychological Tests; Mini Mental State Examination; Screening; Mild Cognitive Impairment; Dementia

INTRODUCTION

As the elderly population and the number of people with dementia increase, the need for cognitive screening tests that can screen for dementia and reliably measure its severity in a short time in community and clinical settings also increases.[1] The Mini-Mental State Examination (MMSE)[2] is one of the leading brief screening tests for assessing cognitive impairment. It has been widely used in a variety of areas related to dementia, including clinical settings,[3-6] epidemiological studies,[7,8] and clinical trials.[9,10]

Author Contributions

Conceptualization: Kang Y. Data curation: Kim S, Kang Y. Formal analysis: Kim J. Jahng S. Funding acquisition: Kim S, Kang Y. Methodology: Kim J, Jahng S, Kang Y. Project administration: Kang Y. Writing - original draft: Kim J. Writing - review & editing: Kim J, Jahng S, Kim S, Kang Y.

In 2010, 35 years after the MMSE was first published, a revised version, the Mini-Mental State Examination, 2nd edition (MMSE~2),[11] was released. The newly developed MMSE~2 consists of 3 versions: a brief version (MMSE~2:BV), a standard version (MMSE~2:SV), and an expanded version (MMSE~2:EV).[11] The authors of the MMSE~2 explained their reasons for developing 3 versions. First, the MMSE~2:BV was developed for use in rapid clinical assessment and for screening individuals in large population studies. Second, the MMSE~2:SV was developed to maintain equivalence to the original MMSE while changing some items to avoid standardized administration issues. Finally, a slightly longer version, the MMSE~2:EV, was developed. It is more sensitive to subcortical dementia and age-related changes. It is sufficiently difficult in that it does not have a ceiling effect.[11]

Recently, the Korean-Mini Mental State Examination, 2nd edition (K-MMSE~2) was standardized and released.[12] The K-MMSE~2 is also available in 3 standardized versions: a brief version (K-MMSE~2:BV), a standard version (K-MMSE~2:SV), and an expanded version (K-MMSE~2:EV). Of these, the K-MMSE~2:SV has received more attention than the other 2 newly developed versions because it will replace the Korean-MMSE (K-MMSE),[13] which has been used in Korea for more than 20 years. Although the K-MMSE~2:SV maintains the same structure and scoring system as the K-MMSE, some language items (naming and comprehension) and the order of items have been changed to match the MMSE~2:SV. The reliability and validity of each version of the K-MMSE~2 have been statistically verified by researchers.[12] However, psychometric similarities and differences between the K-MMSE and the K-MMSE~2:SV have not yet been reported. As the K-MMSE has been used in Korea for more than 20 years, it is necessary to clarify the relationship between the K-MMSE and the K-MMSE~2:SV to ensure the continuity of data.

Unlike the classical test theory (CTT), where item characteristics such as item difficulty and discriminant power vary depending on characteristics of the test taker, the item response theory (IRT) assumes that each item has immutable intrinsic characteristics. It analyzes items through an item characteristic curve that reflects these characteristics. As a result, the difficulty and discriminant power of items remain constant regardless of the characteristics of the test taker, which is the main strength of IRT.[14]

Several previous studies have used IRT to analyze the MMSE. Some of them have been designed to produce a single cut-off score for diagnosis rather than to characterize the items,[15-17] while others have used dichotomous IRT despite the fact that the MMSE contains polytomous items, making results unreliable.[18-20] In addition, to the best of the authors' knowledge, studies using IRT to examine test characteristics of the MMSE~2 or determine differences between the MMSE and the MMSE~2:SV have not been reported yet.

Therefore, the aim of this study was to analyze items and test characteristics of the K-MMSE and K-MMSE~2:SV using polytomous IRT. We also examined how well the K-MMSE and K-MMSE~2:SV worked as screening instruments using the test information curve (TIC). Finally, we checked the equivalency of the K-MMSE~2:SV and the K-MMSE to see if the K-MMSE~2:SV could be used as an equivalent to the K-MMSE despite changes in items.

METHODS

Participants

Participants were 1,514 subjects randomly sampled from all metropolitan cities and provinces across the country who participated in the K-MMSE~2 standardization study.[12] Based on Christensen's health screening criteria,[21] those with a history of neurological or psychiatric disorders or suspected brain injury, those with untreated chronic diseases such as hypertension, diabetes, and hyperlipidemia that might affect cognitive function, and those with visual or hearing impairments that could affect test performance were excluded. Ages of participants ranged from 19 to 90 years. Their education levels ranged from illiterate to post-graduate.

Measurements

Items of the K-MMSE and K-MMSE~2:SV were identical except for the naming and comprehension items. Thus, all participants were administered the K-MMSE~2:SV blue form along with the naming and comprehension items on the K-MMSE.

Statistical analysis

Paired t-test was used to compare K-MMSE and K-MMSE~2:SV scores. Item discriminations, item category difficulties, TICs, and item information curves (IICs) for the K-MMSE and K-MMSE~2:SV were analyzed using generalized partial credit model (GPCM) of the polytomous IRT.[22] In the IRT, item discrimination refers to how well an item differentiates between examinees' abilities. In tests with polytomous items, such as those of MMSE and MMSE~2, item difficulty is assessed by item category difficulty. This indicates the level of ability required to reach each response category. Item difficulty is expressed through the difficulty parameter for each category. Ability level (θ) represents the estimated ability parameter of the examinee. The TIC shows how accurately a test measures across the entire range of ability. IIC shows how well an item contributes to estimating ability along the ability continuum. The TIC is derived by summing individual IICs.[14,23] To determine the equivalence of the K-MMSE and the K-MMSE~2:SV, the generalizability (G) coefficient was obtained using the generalizability theory.[11,24-26]

Paired *t*-test was performed using the R Statistical Software (version 4.3.1; R Foundation for Statistical Computing, Vienna, Austria).[27] IRT analysis was conducted using the ltm R package (version 1.2.0; R Foundation for Statistical Computing).[28]

Ethics statement

The protocol of the present study was reviewed and approved by the Institutional Review Board (IRB) of Hallym University (HIRB-2019-44).

RESULTS

Demographic characteristics and K-MMSE and K-MMSE~2:SV scores

The mean age of participants was 63.89±16.75 years (range, 19–90 years). There were 641 men and 873 women. The mean level of education was 10.21±5.71 years (range, 0–25 years) (**Table 1**). K-MMSE and K-MMSE~2:SV scores are shown in **Table 2**. There was a significant difference (t=−9.02, p<0.001) between the Language score on the K-MMSE (7.31±1.04) and the Language score on the K-MMSE~2:SV (7.48±1.06). As a result, the mean score

https://doi.org/10.12779/dnd.2024.23.3.117

DND Dementia and Neurocognitive Disorder

Table 1. Distribution of participants by age and education

Age (yr)	Education (yr)					
	0	1–6	7–9	10–12	≥13	Total
19–29	0	0	0	24	78	102
30–39	0	0	0	7	25	32
40–49	0	0	5	52	81	138
50–59	1	24	31	79	100	235
60–69	15	77	60	73	87	312
70–79	87	128	60	73	88	436
80–90	70	82	22	41	44	259
Total	173	311	178	349	503	1,514

Table 2. Item and total scores of the K-MMSE and K-MMSE~2:SV

Item	K-MMSE	K-MMSE~2:SV
Registration	2.93±0.33	
Orientation to time	4.67±0.70	
Orientation to place	4.91±0.31	
Recall	1.99±1.00	
Attention & calculation	3.86±1.43	
Language	7.31±1.04	7.48±1.06
Drawing	0.86±0.35	
Total score	26.54±3.47	26.71±3.48

K-MMSE: Korean-Mini Mental State Examination, K-MMSE~2:SV: Korean-Mini Mental State Examination, 2nd edition: Standard Version.

of K-MMES~2:SV (26.71±3.48) was 0.17 points higher than K-MMSE (26.54±3.47). The difference between the two was statistically significant (t=−9.02, p<0.001).

Item characteristics

Item discrimination and item category difficulty for K-MMSE and K-MMSE~2:SV were analyzed with the GPCM. Results are shown in **Table 3**. Based on Seong's criteria,[29,30] of the 7 items on the K-MMSE and K-MMSE~2:SV, Registration, Orientation to time, Orientation to place, and Attention & calculation all had moderate discrimination, Language had high discrimination, and Drawing had very high discrimination. However, Recall had a low discrimination (0.59 with K-MMSE and 0.56 with K-MMSE~2:SV).

Table 3. GPCM item parameter estimates of the K-MMSE and K-MMSE~2:SV

Items	Item discrimination	Transition location (Item category difficulty)							
		0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8
K-MMSE									
Registration	1.07	−2.70	−2.61	−3.41					
Orientation to time	0.92	−3.36	−2.57	−2.80	−2.31	−1.74			
Orientation to place	1.04			−4.24	−3.43	−2.88			
Recall	0.59	−1.49	−1.30	−0.05					
Attention & calculation	1.03	−1.75	−1.42	−1.34	−0.58	−0.56			
Language	1.37			−3.89	−2.98	−2.29	−1.62	−1.29	−0.60
Drawing	2.40	−1.31							
K-MMSE~2:SV									
Registration	1.08	−2.68	−2.58	−3.37					
Orientation to time	0.93	−3.35	−2.53	−2.76	−2.29	−1.73			
Orientation to place	1.05			−4.22	−3.39	−2.85			
Recall	0.56	−1.50	−1.35	−0.08					
Attention & calculation	1.06	−1.73	−1.41	−1.34	−0.58	−0.53			
Language	1.48				−2.20	−1.72	−1.78	−1.34	−1.19
Drawing	2.52	−1.29							

GPCM: generalized partial credit model, K-MMSE: Korean-Mini Mental State Examination, K-MMSE~2:SV: Korean-Mini Mental State Examination, 2nd edition: Standard Version.

https://doi.org/10.12779/dnd.2024.23.3.117

DND Dementia and Neurocognitive Disorder

Item category difficulty is not represented by a single number when there are multiple item response categories, such as for items in MMSE. For the GPCM used in this analysis, the item difficulty parameter was expressed as the transition location where response probabilities intersect across one category to the next. Most items in K-MMSE and K-MMSE~2:SV had item category difficulty ranging from very easy (<−2.0) to easy (−2.0 to −0.5) level (**Table 3**) except for Recall. The item category difficulty for Recall in K-MMSE and K-MMSE~2:SV was easy for scores ranging from 0 to 1 and from 1 to 2, respectively. However, it was medium (−0.5 to 0.5) for scores ranging from 2 to 3.[29-31]

TICs

A TIC shows how accurately a test estimates ability over the entire range of ability levels (θ). That is, the higher the value of the information function, the more accurately the test estimates the participant's ability, which means a smaller standard error of the participant's ability estimate.[31] **Fig. 1** shows TICs for K-MMSE and K-MMSE~2:SV, both of which are positively skewed. Total test information values of K-MMSE and K-MMSE~2:SV were 28.44 and 27.86, respectively, showing a small difference of 0.58. Both tests were found to be the most informative and discriminating well between participants' abilities at an ability level of −1.57, although the K-MMSE~2:SV had a slightly higher information value than the K-MMSE at that ability level.

Table 4 shows values of the test information function of K-MMSE and K-MMSE~2:SV at each ability level. At ability levels of −2.0 to −1.0, the K-MMSE~2:SV had higher values of test information function than the K-MMSE. However, at other ability levels, the K-MMSE had similar or slightly higher values than the K-MMSE~2:SV. At ability levels 1.0 and above, both tests had very low values (<1.0) of test information function.

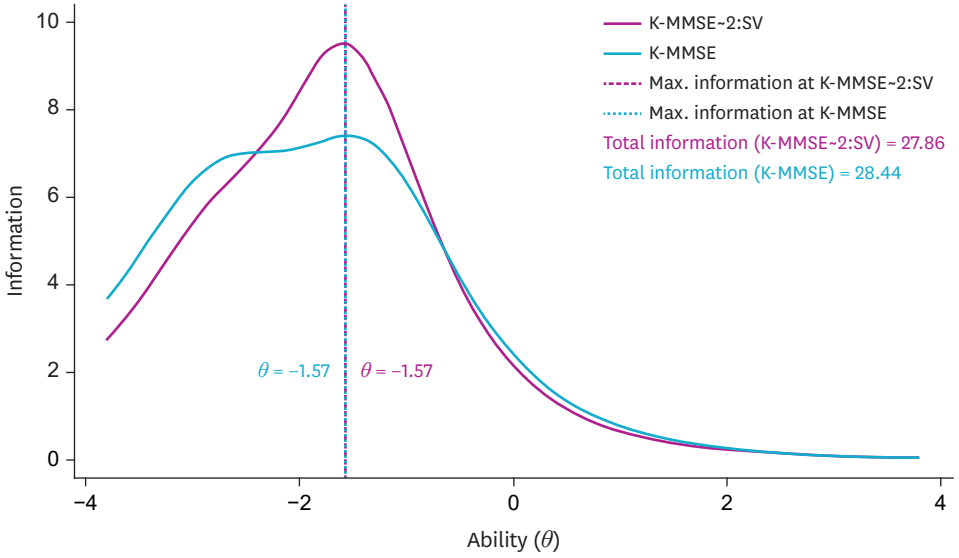

Fig. 1. Test information curves of the K-MMSE and K-MMSE~2:SV.
K-MMSE: Korean-Mini Mental State Examination, K-MMSE~2:SV: Korean-Mini Mental State Examination, 2nd edition: Standard Version.

https://doi.org/10.12779/dnd.2024.23.3.117

DND Dementia and Neurocognitive Disorder

Table 4. Values of the test information function at each ability level (θ)

Ability level (θ)	K-MMSE	K-MMSE~2:SV
−4.0	3.08	2.26
−3.5	4.68	3.65
−3.0	6.35	5.40
−2.5	7.02	6.76
−2.0	7.14	8.45
−1.5	7.40	9.42
−1.0	6.35	7.02
−0.5	4.17	4.06
0.0	2.42	2.17
0.5	1.37	1.17
1.0	0.78	0.66
1.5	0.46	0.39
2.0	0.28	0.24
2.5	0.18	0.16
3.0	0.11	0.10
3.5	0.08	0.07
4.0	0.05	0.05

K-MMSE: Korean-Mini Mental State Examination, K-MMSE~2:SV: Korean-Mini Mental State Examination, 2nd edition: Standard Version.

IICs

Fig. 2 shows IICs for 7 items in the K-MMSE~2:SV and the Language item in the K-MMSE. The IIC shows how well and precisely each item measures the cognitive function it is intended to measure at various levels of ability.

The IICs for all items except Recall were positively skewed. Registration, Orientation to time, and Orientation to place provided the most information with ability levels between −3 and slightly below −2, while Attention & calculation, Language, and Drawing provided the most information with ability levels between slightly above −2 and −1. However, the IIC for Recall had very low kurtosis, indicating that there was no difference in the amount of information provided across the ability range.

Peak information of the Language in the K-MMSE~2:SV was higher than that in the K-MMSE. The difference was even more pronounced between ability levels −2 and −1, which was why the test information of K-MMSE~2:SV was slightly higher than that of K-MMSE at that interval as shown in **Fig. 1**.

G coefficient

K-MMSE~2:SV has most contents of the K-MMSE, although a few items are modified. To assess the equivalence of these 2 tests using the generalizability theory, we calculated the G coefficient for the K-MMSE~2:SV and K-MMSE. The G coefficient was $\sigma^2(S)/[\sigma^2(S)+\sigma^2(E)]=11.81/[11.81+0.13]=0.99$, where $\sigma^2(S)$ was the variance component for participants and $\sigma^2(E)$ was the variance component for errors, showing a very high degree of equivalence between K-MMSE~2:SV and K-MMSE.

https://doi.org/10.12779/dnd.2024.23.3.117

DND Dementia and Neurocognitive Disorder

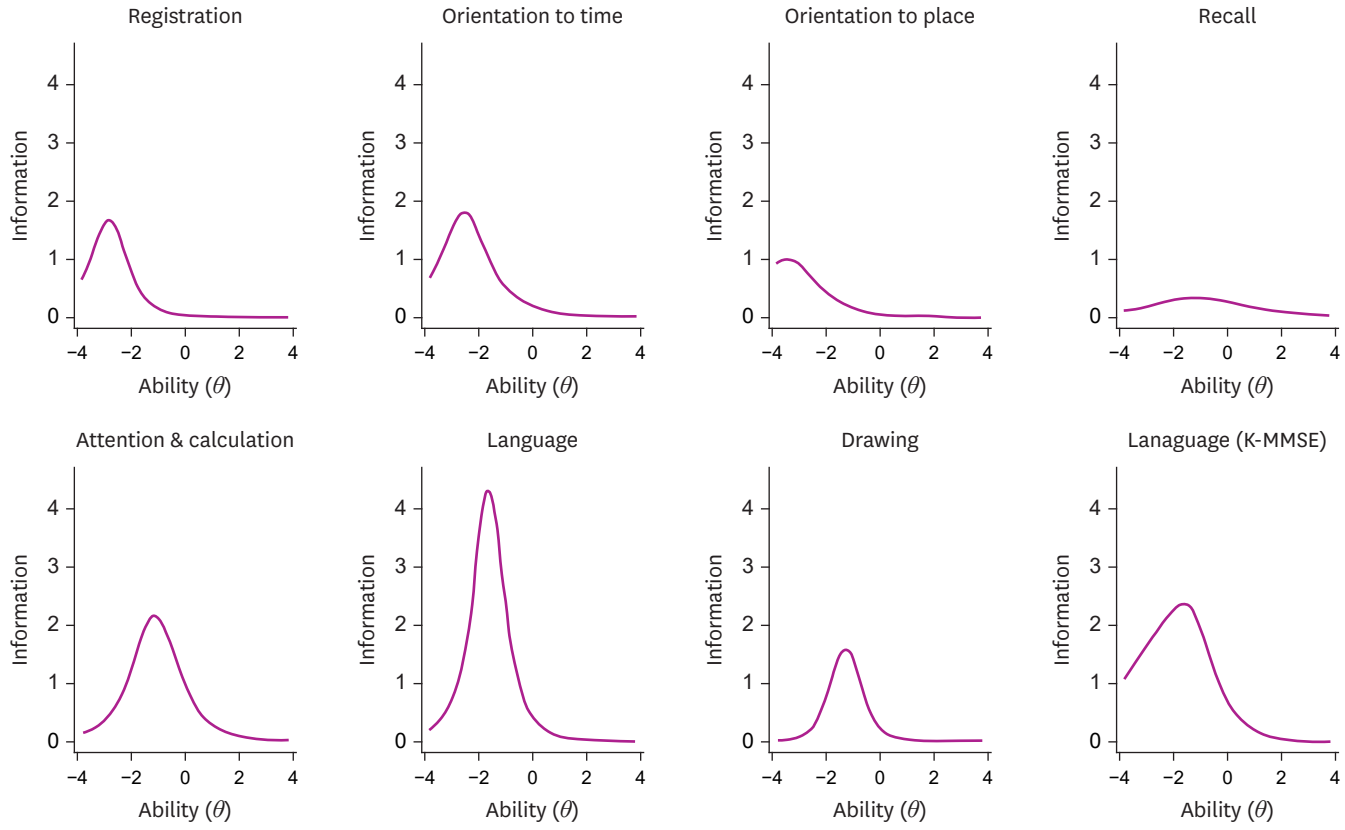

Fig. 2. IICs for the K-MMSE~2:SV and IIC for Language in the K-MMSE.
IIC: item information curve, K-MMSE: Korean-Mini Mental State Examination, K-MMSE~2:SV: Korean-Mini Mental State Examination, 2nd edition: Standard Version.

DISCUSSION

This study applied a generalized partial-score model of the multi-category item response theory to compare item and test characteristics of K-MMSE and K-MMSE~2:SV and to determine whether the K-MMSE~2:SV could be used as an equivalent test to the K-MMSE.

The mean score for the K-MMSE~2:SV was slightly higher than that for the K-MMSE. This was due to the fact that modified language items of the K-MMSE~2:SV were slightly easier than original language items of the K-MMSE.

Based on item parameter estimation results, both K-MMSE and K-MMSE~2:SV had very similar discrimination and item categorical difficulty (**Table 3**). All 6 items of the K-MMSE and K-MMSE~2:SV, with the exception of Recall, showed moderate to high item discrimination. Only Recall had a low item discrimination, meaning that the probability of a correct response to Recall did not vary significantly according to participant's ability level. This is an interesting and unexpected finding, because clinicians often consider Recall score to be the most important measure for cognitive screening among items in the MMSE. However, previous studies on normal people have shown that most errors on the MMSE occur in Recall, Attention & calculation, and Orientation to time, with recall showing the largest error.[32,33] This means that people with dementia or mild cognitive impairment as well as normal people have a harder time scoring high on Recall.

https://doi.org/10.12779/dnd.2024.23.3.117

In terms of item difficulty estimation, most items of the K-MMSE and K-MMSE~2:SV were easy or very easy items except Recall. In this study, especially when analyzing the K-MMSE and K-MMSE~2:SV using the GPCM, we were able to obtain richer information because we were able to estimate the difficulty of partial scores, which could not be estimated using classical test theory, i.e., we were able to compare the difficulty of scoring a 1 to the difficulty of scoring a higher score. As expected, for most items, the item category difficulty increased as the score increased. However, for Registration, the item category difficulty showed an opposite trend, with a score of 3 being easier to achieve than a score of 1 or 2. This result is likely due to the fact that Registration is an easy item. Most participants scored a 3 for Registration.

The test information function analysis showed that total amounts of information provided by the K-MMSE and K-MMSE~2:SV were almost the same (28.44 vs. 27.86) and that both tests provided maximum information at the same ability level (θ=–1.57), meaning that both tests provided the most information for discriminating the ability of a participant with an ability level of –1.57, i.e., mild cognitive impairment. A closer look at the test information function values presented in **Table 4** along with the TICs showed that for ability levels –2.0 to –1.0, the K-MMSE~2:SV was more informative than the K-MMSE. For ability levels –2.5 and below, the K-MMSE was slightly more informative than the K-MMSE~2:SV. For ability levels 1.0 and above, both tests provided very little information without showing a significant difference. These results suggest that neither the K-MMSE nor the K-MMSE~2:SV is sensitive for detecting cognitive decline or change in people with above-average cognitive ability. Many studies have already reported that the ceiling effect of the MMSE limits its use in people with higher cognitive abilities.[13,34-36] Based on the above results, we also expect that K-MMSE is slightly better than K-MMSE~2:SV in detecting people with very low cognitive ability (ability level –2.5 or lower). However, since the information gap between the 2 tests is not large enough to know if this is a significant difference, further studies in clinical settings with people with dementia at ability level –2.5 or lower are needed to confirm this difference in real-world practice.

The only differences between the K-MMSE and K-MMSE~2:SV were language items, namely naming and comprehension. In the K-MMSE, naming items were “pen” and “watch” which were asked after showing real objects to the examinee, whereas in the K-MMSE~2:SV blue form, they were replaced with “eye” and “ear,” with the examiner pointing to his or her own body parts and asking the examinee.[11,12] In the K-MMSE comprehension, three-step verbal commands were used: “Turn the paper over, fold it in half, and give it to me.” In the K-MMSE~2, it was replaced with an item that showed three shapes and asked the examinee to point to them in order as directed.[12] Item discrimination for Language in the K-MMSE~2:SV was slightly higher than that in the K-MMSE (1.48 vs. 1.37). Item category difficulty for Language was also slightly easier in the K-MMSE~2:SV than in the K-MMSE (**Table 3**), which likely contributed to the higher information value in the screening. Comparing the IICs for Language in the K-MMSE and K-MMSE~2:SV as shown in **Fig. 2**, we could see that both tests provided the most information at the ability range from –2 to –1. The amount of information was higher in the K-MMSE~2:SV than in the K-MMSE, which could also explain why the K-MMSE~2:SV had higher test information values at those ability ranges than the K-MMSE. Therefore, if the examiner’s goal is to screen out participants with mild cognitive impairment, more information can be obtained using the K-MMSE~2:SV.

Finally, the G coefficient for K-MMSE and K-MMSE~2:SV was very high at 0.99. This proves that the they are equivalent tests. This is the same as the G coefficient for the original MMSE

https://doi.org/10.12779/dnd.2024.23.3.117

and MMSE~2:SV.[11] These results mean that we can directly compare previously collected K-MMSE scores with newly collected K-MMSE~2:SV scores in longitudinal data.

In summary, the new K-MMSE~2:SV was found to be somewhat more informative than the K-MMSE for screening for mild cognitive impairment or early dementia due to changed language items. Nevertheless, given that the amount of test information provided by the 2 tests was almost identical, with shapes of their TICs being very similar and the G coefficient being very high, it can be concluded that the K-MMSE and K-MMSE~2:SV are equivalent tests.

REFERENCES

1. Kang Y. A normative study of the Korean-Mini Mental State Examination (K-MMSE) in the elderly. Korean J Psychol Gen 2006;25:1-12.
2. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-198. PUBMED | CROSSREF
3. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care 2002;40:771-781. PUBMED | CROSSREF
4. Mitchell AJ. *Cognitive Screening Instruments*. Cham: Springer, 2017;37-48.
5. Modrego PJ, Fayed N, Pina MA. Conversion from mild cognitive impairment to probable Alzheimer's disease predicted by brain magnetic resonance spectroscopy. Am J Psychiatry 2005;162:667-675. PUBMED | CROSSREF
6. Malloy PF, Cummings JL, Coffey CE, Duffy J, Fink M, Lauterbach EC, et al. Cognitive screening instruments in neuropsychiatry: a report of the Committee on Research of the American Neuropsychiatric Association. J Neuropsychiatry Clin Neurosci 1997;9:189-197. PUBMED | CROSSREF
7. Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. JAMA 1993;269:2386-2391. PUBMED | CROSSREF
8. Kase CS, Wolf PA, Kelly-Hayes M, Kannel WB, Beiser A, D'Agostino RB. Intellectual decline after stroke: the Framingham Study. Stroke 1998;29:805-812. PUBMED | CROSSREF
9. Raskind MA, Peskind ER, Wessel T, Yuan W. Galantamine in AD: a 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. Neurology 2000;54:2261-2268. PUBMED | CROSSREF
10. Rogers SL, Doody RS, Mohs RC, Friedhoff LT; Donepezil Study Group. Donepezil improves cognition and global function in Alzheimer disease: a 15-week, double-blind, placebo-controlled study. Arch Intern Med 1998;158:1021-1031. PUBMED | CROSSREF
11. Folstein MF, Folstein SE, White T, Messer MA. *MMSE-2 User's Manual*. Lutz: Psychological Assessment Resources, 2010.
12. Kang Y, Jahng S, Kim SY, Korean Dementia Association. *Korean-Mini Mental State Examination, 2nd Edition (K-MMSE~2)*. Seoul: Hakjisa, 2020.
13. Kang Y, Na DL, Hahn SH. A validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. J Korean Neurol Assoc 1997;15:300-308.
14. Seong TJ. Analysis of item and tests using classical test theory and item response theory. J Sci Educ Res Inst Korea Natl Univ Educ 1998:143-170.
15. Orlando Edelen MO, Thissen D, Teresi JA, Kleinman M, Ocepek-Welikson K. Identification of differential item functioning using item response theory and the likelihood-based model comparison approach. Application to the Mini-Mental State Examination. Med Care 2006;44:S134-S142. PUBMED | CROSSREF
16. Schultz-Larsen K, Kreiner S, Lomholt RK. Mini-Mental Status Examination: mixed Rasch model item analysis derived two different cognitive dimensions of the MMSE. J Clin Epidemiol 2007;60:268-279. PUBMED | CROSSREF
17. Teresi JA. Mini-Mental State Examination (MMSE): scaling the MMSE using item response theory (IRT). J Clin Epidemiol 2007;60:256-259. PUBMED | CROSSREF
18. Aiello EN, Esposito A, Pucci V, Mondini S, Bolognini N, Appollonio I. Italian telephone-based Mini-Mental State Examination (Itel-MMSE): item-level psychometric properties. Aging Clin Exp Res 2022;34:1259-1265. PUBMED | CROSSREF

19. Lou MF, Dai YT, Huang GS, Yu PJ. Identifying the most efficient items from the Mini-Mental State Examination for cognitive function assessment in older Taiwanese patients. J Clin Nurs 2007;16:502-508. PUBMED | CROSSREF
20. Medonça de Melo D, Gonçalves Barbosa AJ, Ribeiro de Castro N, Liberalesso Neri A. Mini-Mental State Examination in Brazil: an item response theory analysis. Paidéia 2020;30:e3014. CROSSREF
21. Christensen KJ, Multhaup KS, Nordstrom S, Voss K. A cognitive battery for dementia: development and measurement characteristics. Psychol Assess 1991;3:168-174. CROSSREF
22. Muraki E. A generalized partial credit model: application of an EM algorithm. Appl Psychol Meas 1992;16:159-176. CROSSREF
23. Hambleton RK, Swaminathan H, Rogers HJ. *Fundamentals of Item Response Theory*. Thousand Oaks: Sage Publications, 1991;91-108.
24. Brennan RL. Generalizability Theory. New York: Springer, 2001;4-20.
25. Cronbach LJ, Gleser GC, Nanda H, Rajaratnam N. *The Dependability of Behavioral Measurements: Theory of Generalizability for Score and Profiles*. New York: Wiley, 1972;263-344.
26. Shavelson RJ, Webb NM. *Generalizability Theory: A Primer*. Newbury Park: Sage Publications, 1991;83-98.
27. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2023.
28. Rizopoulos D. ltm: An R package for latent variable modelling and item response theory analyses. J Stat Softw 2006;17:1-25. CROSSREF
29. Seong TJ. *Understanding and Applying Item Response Theory*. 2nd ed. Paju: Kyoyookbook, 2016;46-49.
30. Seong TJ. *Modern Education Evaluation*. 3rd ed. Seoul: Hakjisa, 2010;274-275.
31. Baker PB. *The Basic of Item Response Theory*. 2nd ed. Washington, D.C.: ERIC Clearinghouse on Assessment and Evaluation, 2001;33-126.
32. Bleecker ML, Bolla-Wilson K, Kawas C, Agnew J. Age-specific norms for the Mini-Mental State Exam. Neurology 1988;38:1565-1568. PUBMED | CROSSREF
33. Galasko D, Klauber MR, Hofstetter CR, Salmon DP, Lasker B, Thal LJ. The Mini-Mental State Examination in the early diagnosis of Alzheimer's disease. Arch Neurol 1990;47:49-52. PUBMED | CROSSREF
34. Franco-Marina F, García-González JJ, Wagner-Echeagaray F, Gallo J, Ugalde O, Sánchez-García S, et al. The Mini-mental State Examination revisited: ceiling and floor effects after score adjustment for educational level in an aging Mexican population. Int Psychogeriatr 2010;22:72-81. PUBMED | CROSSREF
35. Pendlebury ST, Markwick A, de Jager CA, Zamboni G, Wilcock GK, Rothwell PM. Differences in cognitive profile between TIA, stroke and elderly memory research subjects: a comparison of the MMSE and MoCA. Cerebrovasc Dis 2012;34:48-54. PUBMED | CROSSREF
36. Piccinin AM, Muniz-Terrera G, Clouston S, Reynolds CA, Thorvaldsson V, Deary IJ, et al. Coordinated analysis of age, sex, and education effects on change in MMSE scores. J Gerontol B Psychol Sci Soc Sci 2013;68:374-390. PUBMED | CROSSREF