Adaptation of the Montreal Cognitive Assessment for Elderly Filipino Patients

以菲律宾老年患者为对象的蒙特利尔认知评估量表诊断改编版

JC Dominguez, MGS Orquiza, JR Soriano, CD Magpantay, RC Esteban, ML Corrales, ER Ampil

Abstract

Objective: The Montreal Cognitive Assessment (MoCA) is an instrument that aids clinicians in detecting mild cognitive impairment and early Alzheimer’s disease. The study aimed to adapt the MoCA for use in the Philippines, and determine its psychometric validity when used in the Filipino setting.

Methods: The MoCA was adapted by a multidisciplinary team working at the Memory Center of St. Luke’s Medical Center, Manila, the Philippines. Contextual adaptation, rather than direct translation, was done. Pilot testing of the Filipino version of the MoCA (MoCA-P) was done on 12 grade 6 pupils and subsequently on 14 cognitively intact elderly people. Reliability testing of the MoCA-P was done on 25 elderly people by trained psychologists. Internal consistency, inter-rater and intra-rater reliability, as well as convergent and divergent validity of the MoCA-P were determined.

Results: The MoCA-P yielded a high level of internal consistency (Cronbach’s α = 0.938). Inter-rater and intra-rater reliability were 0.887 (p ≤ 0.05) and 0.969 (p ≤ 0.05), respectively. The MoCA-P correlated negatively with the Epworth Sleepiness Scale (r = -0.233) and had a positive but low correlation with the Mini-Mental State Examination (r = 0.555).

Conclusion: Contextual translation and pilot testing yielded several modifications of the MoCA. The MoCA-P is a reliable instrument for use in elderly Filipino patients. Further diagnostic validation of the MoCA-P to establish cutoff scores that would discriminate elderly individuals with normal cognition from those with dementia is needed to establish the clinical utility of the test.

Key words: Aged; Alzheimer disease; Neuropsychological tests

摘要

目的：蒙特利爾認知評估（MoCA）量表能協助臨床醫生檢測輕度認知障礙和早期阿茲海默症。本研究旨在把MoCA量表改編菲律賓語版本（MoCA-P），並檢視於當地醫療服務體系使用的心理效度。

方法：由稱蒙特利爾和聖路加醫療中心的記憶中心跨部門團隊，採用語境順應而非直接翻譯的方法改編MoCA量表。先於12名6年級學生和14名認知功能正常的長者進行試點測試，然後由受過訓練的心理學家向25名長者進行信度測試，以檢視量表的內部一致性、施測者間信度、施測者內信度，以及聚合效度和分歧效度。

結果：MoCA-P量表產生高內部一致性水平（克隆巴赫系數 = 0.938）。施測者間信度和施測者內信度分別為0.887（p ≤ 0.05）和0.969（p ≤ 0.05）。MoCA-P量表與愛普沃斯嗜睡量表呈負相關（r = -0.233），與簡易精神狀態量表則呈正面但低相關性（r = 0.555）。

結論：MoCA-P量表因應語境順應和試點測試作若干修改，它是診斷菲律賓老年患者的可靠工具。為實踐研究的臨床應用，須進一步驗證MoCA-P的診斷效度，從而建立可區別認知功能正常的長者與腦退化症患者的切點。

關鍵詞：長者、阿茲海默病、神經心理學測試
Introduction

Dementia is a progressive brain dysfunction involving memory, language, gnosia, praxis, and executive function. The cognitive impairment is sufficient to interfere with a patient’s ability to function independently. Dementia also causes disturbing behaviours that impair relationships with others and lead to distress to the family and the affected individual. The most common form of dementia is Alzheimer’s disease (AD), and age is clearly associated with an increased risk for AD.

Since AD develops insidiously and progresses gradually, it is not uncommon that it goes unnoticed for many years before disturbing symptoms that merit medical attention occur. At the onset, when symptoms are mildest and commonly related to memory decline, such change is typically dismissed as part of normal ageing. However, early diagnosis is important as it provides an opportunity for early drug intervention, which has been shown to improve cognition and function, and may delay deterioration or institutionalization in countries where elderly people are eventually cared for in nursing homes. Aside from early treatment, early diagnosis promotes acceptance and better understanding of the disease, its symptoms, and its course. Early diagnosis enables families to better care for their relative and plan ahead for social, financial, and legal issues that may arise from a patient’s disabilities.

The cognitive domains that are affected in AD are orientation, memory, attention, executive function, language, visuospatial function, visuomotor speed, and visuomotor construction. Involvement of these domains is defined and documented through the use of neuropsychological tests that could help to identify both the impaired and preserved skills. Given that the core feature of AD is cognitive impairment, neuropsychological tests are ancillary to the diagnosis of AD. Most tests are developed in western English-speaking populations, where they have been developed from clinical experience and validated in the local population. Test performance is known to be influenced by age, education, and language as observed for the Mini-Mental State Examination (MMSE). Values and beliefs can also affect neuropsychological testing. The issue of how culture and language affect test performance has been a topic of concern in past decades. Floor effects have been apparent when English tests were administered to non-native populations in the US, particularly because of the high cultural bias of some of the items found in the tests. Thus, particularly for screening instruments not only for AD, but also for other types of dementia such as vascular cognitive impairment by some professional groups and researchers.

Methods

Study Design
This was a cross-sectional study conducted at the Memory Center of St. Luke’s Medical Center in Manila, the Philippines. This centre is a specialized referral unit providing assessment and treatment for patients with dementia and other related disorders. Ethics approval for the study was granted by the St. Luke’s Medical Center Institutional Ethics Review Committee. Permission to adapt the MoCA was obtained from Prof. Ziad Nasreddine, the primary author of the MoCA.

Study Population
The study participants for the pilot test were recruited from grade 6 pupils in a public elementary school and cognitively intact elderly individuals. The pre-testing participants were elderly individuals who were different from those in the pilot tests. The elderly people were recruited from the Senior Citizen Registry of Marikina, which is an outreach community of the St. Luke’s International Institute for Neurosciences. All study participants provided informed consent before any testing was done.

Procedure
The original version of the MoCA was translated and adapted by a multidisciplinary team comprising a neurologist, a geriatrician, 2 psychologists, a geriatric psychiatrist, an anthropologist, a speech and language therapist, and a nurse. The members were specifically chosen because of their knowledge of the concepts and objectives underlying the test items. All team members are bilingual (English and Filipino) but their mother tongue is Filipino. Since the intended users, i.e. the testers who are medical / allied medical professionals and are English-literate, the guideline for the test administrators was retained in English, and only the test instructions that would be read to the participants were translated into Filipino. To ensure the applicability...
of test items to the Filipino elderly, contextual adaptation rather than simple direct translation was done, with careful consideration of the Filipino culture. The translation was compared with the original15 to verify cross-cultural equivalence. This translation was the first version of the MoCA adaptation. This version was then critiqued independently by a second multidisciplinary team with the same qualifications as the translating team. The translation was also back-translated by a nurse and physician who were not working at the memory centre and were unacquainted with the intent of the test.

The critique and comments from the back-translation were discussed by all members of the 2 teams and amendments were introduced. The result was the second version of the MoCA adaptation. This version was then pilot-tested by grade 6 pupils for comprehensibility of instructions and ease of delivery. The second version was also pilot-tested by cognitively intact elderly people. The elderly participants were assessed by a neurologist to ascertain absence of dementia based on the National Institute of Non-communicative Disorders and Stroke — Alzheimer’s Disease and Related Disorders Association clinical criteria.1 Observations from the pilot tests — particularly semantic, experiential, and conceptual equivalence with the original test — were incorporated in the second version. The result was the third version of the adaptation, which was referred to as the Filipino version of the MoCA (MoCA-P). All disagreements throughout the process were resolved by consensus. The guidelines to preserve equivalence in cross-cultural adaptation of health-related measures were followed.35

The cognitively intact elderly people who pre-tested the MoCA-P returned for re-testing after 2 weeks to assess retest reliability. The pre-testing was videotaped and the tapes were viewed and rated independently by other testers in order to assess inter-rater reliability. The pre-test participants were also administered the MMSE-P14 and the Epworth Sleepiness Scale to assess for convergent and divergent validity, respectively. The testers were psychologists with experience of MoCA test administration.

**Statistical Analysis**

Internal consistency of the MoCA-P was computed by Cronbach’s α. Test-retest reliability was assessed using intraclass correlation coefficients for scores at baseline and at 2 weeks. Pearson’s correlation scores were calculated to assess the relationship between the MoCA-P and MMSE-P, and between the MoCA-P and Epworth Sleepiness Scale. All statistical analyses were done using the Statistical Package for the Social Sciences Windows version 17 (SPSS Inc, Chicago [IL], US). A p value of less than 0.05 was considered statistically significant for all the analyses.

**Results**

Twelve grade 6 pupils and 14 cognitively intact elderly individuals were included in the pilot test and 25 cognitively intact elderly individuals were included in the pre-test. The elderly participants were community-dwelling and were representative of the diversity of Filipino communities in terms of education and ethnicity (Table). All of the elderly participants were fluent in the Filipino language.

The MoCA translation was fully comprehensible to the grade 6 pilot participants. However, it was apparent from the responses of the elderly pilot participants that linguistic and cultural modifications had to be made. None among the elderly participants named the rhinoceros correctly because it seemed that the animal was not within the realm of their reality and vocabulary. Inability to name the rhinoceros by the elderly participants appeared to be a generational issue as all the grade 6 pupils in the pilot test were able to name the animal with ease. Three animals were recommended as a replacement for rhinoceros — carabao (water buffalo), crocodile, and owl. Carabao was too common, hence easy to identify, while a crocodile’s picture could be mistakenly named as lizard or alligator. The owl, which is an indigenous Philippine fauna used with medium frequency and is distinctly identifiable by its prominently big round eyes, was chosen by consensus.

In the language subtest, letter fluency with ‘F’ was not possible as there are no Filipino words that begin with this letter other than the word Filipino. Thus, the letter B was chosen. It was also found that word-for-word translation of the 2 English sentences for repetition unusually lengthened the text in Filipino, and the long sentences became challenging tongue twisters. Therefore, the 2 sentences were carefully abridged while retaining their meaning and the number of words in the original version.

In the abstraction subtest, the consensus was to change clock to weighing scale (‘timbangan’). This was done because the elderly pilot participants did not consider a clock to be a measuring instrument like a ruler. For these participants, a clock tells the time, but does not measure time. The elderly participants also categorised the paired items based on concrete and physical attributes. For example, a common response regarding the similarity of a bicycle and train was that both are made of metal (vs. the correct original response that both are modes of transportation). The participants also categorised items from the viewpoint of what is immediately practical, as in orange and banana are food (vs. the correct original response that both are fruits). The responses were not the best answers, but were nevertheless correct. By consensus agreement, these responses were considered acceptable and participants would not be penalised for not thinking in concepts familiar to western populations. These responses were therefore to be rated correct in the MoCA-P.

Some modifications of the original test instructions were also introduced. The test sheet was folded crosswise to hide the pictures of the animals to be named because the participants were distracted by them. Another modification was made to the abstraction subtest, for which participants needed repetition to understand the instruction. Therefore, repetition of the instructions up to 2 times was allowed and
was incorporated in the instructions to testers. The resultant MoCA-P version finalised after the pilot testing is shown in the Figure.

The MoCA-P yielded a high level of internal consistency (Cronbach’s $\alpha = 0.938$). Re-test reliability was done at a mean of 15.31 days (standard deviation [SD], 6.79 days) apart, although 4 of the 25 participants failed to return for re-testing. The mean MoCA scores were 16.14 and 16.81 for baseline and follow-up, respectively, with a mean change of -0.76 (SD, 2.26) points. The inter-rater and intra-rater reliability were 0.887 ($p \leq 0.05$) and 0.969 ($p \leq 0.05$), respectively. The MoCA-P correlated negatively with the Epworth Sleepiness Scale ($r = -0.233$) and had a positive but low correlation with the MMSE-P ($r = 0.555$). This low convergent validity could be due to the fact that, although both tests are measures of cognition and should be highly correlated, the MMSE-P is limited in measures of executive function, which are found in the MoCA-P.

**Discussion**

At least some degree of familiarity of participants with test items is essential for a test to be valid. In the Korean version of the MoCA, ‘velvet’ and ‘daisy’ were changed to ‘silk’ and ‘azalea’, respectively, because the original items were unfamiliar to Korean elderly people. The elderly participants employed a different cognitive approach to western populations when determining the relationship of 2 seemingly obtuse objects as evidenced by the answers to the similarity of bicycle and train or orange and banana. Certain responses are therefore rated as correct in the MoCA-P, even if different from the original version. This issue is of particular importance in this instance when a response can only be rated as right or wrong, unlike in protocols where responses are graded in a 3-point scale from best to inappropriate answers. The same is true of the rating for the cube performance, which can only be scored as either right or wrong, whereas in other protocols such as the Alzheimer’s Disease Assessment Scale–Cognitive Subscale, the answers are rated with respect to lines, face orientation, parallelism, and dimensionality. Subjectivity was therefore observed in the testers when rating these tests, including the simple clock drawing where the tester could be lenient or strict.

The need for modification of the test items and instructions in the MoCA-P supports the recommendation of the International Working Group for the Harmonization of Dementia Drug Guidelines that test instruments should be validated in the population where the tests will be used, particularly in populations for whom language and cultural experiences differ from those of the West. Valid assessment of elderly Filipino patients can now be done using the MoCA-P, and since the MoCA has also been validated in

---

**Table. Demographic profile of pilot and pre-test groups.**

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Pilot testing of grade 6 pupils (n = 12)</th>
<th>Pilot testing of elderly people† (n = 14)</th>
<th>Pre-testing of elderly people† (n = 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5 (42%)</td>
<td>7 (50%)</td>
<td>19 (76%)</td>
</tr>
<tr>
<td>Male</td>
<td>7 (58%)</td>
<td>7 (50%)</td>
<td>6 (24%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>11.5 ± 0.5</td>
<td>69.6 ± 4.4</td>
<td>72.6 ± 7.2</td>
</tr>
<tr>
<td>10-12</td>
<td>12 (100%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>60-69</td>
<td>-</td>
<td>5 (36%)</td>
<td>10 (40%)</td>
</tr>
<tr>
<td>70-79</td>
<td>-</td>
<td>9 (64%)</td>
<td>12 (48%)</td>
</tr>
<tr>
<td>80-89</td>
<td>-</td>
<td>0</td>
<td>3 (12%)</td>
</tr>
<tr>
<td>Education level (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary (1-6)</td>
<td>6 ± 0</td>
<td>9.8 ± 3.5</td>
<td>8.2 ± 3.7</td>
</tr>
<tr>
<td>High school (7-10)</td>
<td>12 (100%)</td>
<td>4 (28%)</td>
<td>9 (36%)</td>
</tr>
<tr>
<td>College (11-14)</td>
<td>-</td>
<td>5 (36%)</td>
<td>12 (48%)</td>
</tr>
<tr>
<td>Postgraduate (≥ 15)</td>
<td>-</td>
<td>0</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tagalog</td>
<td>12 (100%)</td>
<td>9 (64%)</td>
<td>9 (36%)</td>
</tr>
<tr>
<td>Bicolano</td>
<td>0</td>
<td>4 (29%)</td>
<td>4 (16%)</td>
</tr>
<tr>
<td>Bisaya</td>
<td>0</td>
<td>1 (7%)</td>
<td>5 (20%)</td>
</tr>
<tr>
<td>Kapampangan</td>
<td>0</td>
<td>0</td>
<td>3 (12%)</td>
</tr>
<tr>
<td>Ilocano</td>
<td>0</td>
<td>0</td>
<td>4 (16%)</td>
</tr>
</tbody>
</table>

* Data are shown as No. (%) of subjects or mean ± standard deviation.
† All patients were cognitively intact.
other Asian countries, it can be used in cross-cultural studies of dementia in Asia.

This study has some limitations. Firstly, in adapting an existing instrument, conclusions based on concepts in the original instrument’s source culture may be heightened despite differences or absence of these concepts in the Filipino culture, as was evident in the abstraction subtest. Thus, an implicit limitation in this adaptation is the restriction in discovering or identifying patterns unique to the Filipino elderly in exchange for content validity. Secondly, the subjectivity in rating responses justifies the development of an instruction manual to standardise test administration, rating, and interpretation.

Conclusion

Contextual translation and pilot testing yielded several modifications in the adaptation of the MoCA for use in the
Philippines. Pre-testing study showed that the MoCA-P is a reliable instrument for use in the Filipino elderly. Due to the variability of responses that can be considered as correct by testers and subjectivity in rating responses, a manual of instructions must be developed to standardise test administration, rating, and interpretation. Further diagnostic validation of the MoCA-P to establish cutoff scores that would discriminate elderly individuals with normal cognition from those with dementia is needed to establish the clinical utility of the test.

Acknowledgement

This study would not have been possible without the support of the St. Luke’s Medical Center Research and Biotechnology Division, the Philippine Council for Health Related Diseases of the Philippine Department of Science and Technology, and Prof. Ziad Nasreddine who granted permission to adapt the MoCA.

Declaration

This study was supported by St. Luke’s Medical Center and the Philippine Council for Health and Research Development, Philippine Department of Science and Technology. The authors declared that they have no disclosures to make.

References