ORIGINAL RESEARCH

Comparison of Alternate and Original Items on the Montreal Cognitive Assessment

Elena Lebedeva, PhD[1], Mei Huang, BSc[1], Lisa Koski, PhD[1,2]

[1]*The Research Institute of the McGill University Health Centre, Montreal, QC;*
[2]*Department of Neurology & Neurosurgery, McGill University Health Centre, Montreal, QC*

DOI:http://dx.doi.org/10.5770/cgj.19.216

ABSTRACT

Background

The Montreal Cognitive Assessment (MoCA) is a screening tool for mild cognitive impairment (MCI) in elderly individuals. We hypothesized that measurement error when using the new alternate MoCA versions to monitor change over time could be related to the use of items that are not of comparable difficulty to their corresponding originals of similar content. The objective of this study was to compare the difficulty of the alternate MoCA items to the original ones.

Methods

Five selected items from alternate versions of the MoCA were included with items from the original MoCA administered adaptively to geriatric outpatients (N = 78). Rasch analysis was used to estimate the difficulty level of the items.

Results

None of the five items from the alternate versions matched the difficulty level of their corresponding original items.

Conclusions

This study demonstrates the potential benefits of a Rasch analysis-based approach for selecting items during the process of development of parallel forms. The results suggest that better match of the items from different MoCA forms by their difficulty would result in higher sensitivity to changes in cognitive function over time.

Key words: cognition, geriatrics, MoCA, cognitive impairment, alternate forms

INTRODUCTION

The MoCA test was developed to detect mild cognitive impairment (MCI) in older populations with suspected cognitive decline.[1,2] Practice effects in mixed groups of older persons with Alzheimer's dementia, mild cognitive impairment, or no cognitive impairment appear to be minimal, with estimates of around 1 point increase over a 1 month interval,[2,3] suggesting that the MoCA could be used to monitor cognitive status over time. However, it is increasingly used in settings and with populations far different from those for which it was developed, including younger populations who may be more likely to show practice effects with repeated administration.[4] Recently, two alternate English versions of the MoCA (i.e., Alternate Version 2 and Alternate Version 3), in which items from the original MoCA were replaced by alternate items with similar content, were developed with the goal of avoiding practice effects when the test is administered repeatedly within a short period of time (http://www.mocatest.org/). The reliability of these two alternate MoCA versions has not yet been demonstrated although German-language translations of the original version and two alternate versions were shown to yield similar total scores and to effectively discriminate MCI patients from healthy controls.[5]

In a previous study we presented the ranking of the original MoCA items by their difficulty level using the Rasch methodological approach.[6] The results of that analysis showed that items with equal scoring values on the original MoCA differ from each other significantly in their difficulty level and can be summed to yield a reliable estimate of cognitive ability across a wide range. To be interchangeable with the original MoCA, items from alternate versions of the MoCA must be of comparable difficulty to their content-matched originals. Any differences in difficulty level between the new and original items would lead to discrepancies in screening properties among the parallel forms, affecting the inter-form reliability that is a prerequisite for monitoring cognitive changes longitudinally. Although the problem could be solved by establishing separate cut-offs with comparable sensitivity/specificity for each of the MoCA test forms, high inter-form reliability is ensured by selecting alternate items that are matched to the original items by their difficulty. We hypothesized that inadequate matching of the difficulty of questions in the alternate MoCA tests could explain in part the low sensitivity to change in cognitive ability over time that was recently reported for alternate MoCA versions.[7]

The objective of this study was to begin to assess the difficulty level of items from the two alternate MoCA versions in relation to their corresponding content items from the original MoCA (Version 1). To accomplish this goal, selected items from two alternate English MoCA versions (i.e., Alternate MoCA Version 2 and Alternate MoCA Version 3) from the official MoCA test website (http://www.mocatest.org/) were administered, along with items from the original version of the MoCA, to people attending geriatric outpatient clinics. Subsequently, difficulty estimates of the items from the alternate MoCA versions were determined and compared to the original MoCA items using Rasch analysis.

METHODS

Study Sample

The study was approved by the McGill University Health Centre's Research Ethics Board. Seventy-eight individuals were recruited during a period from June to September 2012 at the Geriatric Outpatient Clinics of the McGill University Health Centre. Patients seen during their first or routine follow-up appointment who would normally undergo cognitive screening with the MoCA for clinical purposes were invited to participate in the study. In a previous study using Rasch analysis, we showed that the location of the Mini-Mental State Examination and MoCA items is overlapping.[8] Therefore, we did not seek to exclude individuals with a prior diagnosis of dementia who would be able to perform easier items from the MoCA. In keeping with normal clinical practice, a pre-existing diagnosis was not a prerequisite for cognitive screening, which increases the generalizability of our findings. Eligibility criteria included being able to understand the purpose of the research project, sign the informed consent form, and complete the test in English or French. To describe the characteristics of the sample, cognitive diagnoses were obtained from notes in the medical chart updated at the conclusion of the clinic visit. MCI and dementias were diagnosed following a clinical evaluation based on DSM-IV[9] and MCI criteria as described by Petersen,[10] respectively. On the basis of their clinical diagnosis, participants were classified into four groups: MCI (all types), Dementia (all types), Unspecified Cognitive Impairment (where no diagnostic conclusion had yet been reached), and Not Cognitively Impaired.

Items

We selected five items from the alternate MoCA Versions 2 and 3 that would be least likely to interfere with performance on the original test items (e.g., due to practice effects) or to lengthen test administration time. Two naming items from Alternate Version 2 (Bear, Hippo) were compared with original naming items Rhino and Camel. Two abstraction items (Version 2: How are Diamond/Ruby alike? Version 3: Trumpet/Piano?) were compared with the original's Train/Bicycle and Watch/Ruler. Copying a drawing of a Cylinder (Version 3) was compared with the original Cube item. The original MoCA was administered adaptively based on criteria described in our previous study in 322 patients.[11] Briefly, each participant draws a clock first and is subsequently classified into one of four cognitive ability ranges on the basis of performance on this item. They then complete the set of MoCA items from Version 1 that correspond to that difficulty level to obtain a more precise estimate of ability. For this study, the new items from the alternate MoCAs were administered at the same time as their corresponding originals (after translating into French, where required). Thus, each new item is completed by a subset of participants, and individuals with low cognitive ability are not required to attempt items beyond their ability (e.g., abstraction items).

Statistical Analysis

IBM SPSS Statistics version 20.0 was used for descriptive and correlation statistics. Using methods outlined previously,[8] Rasch analysis of the set of 33 items (original 28 plus 5 alternates) was performed to estimate the difficulties of the new items in relation to their corresponding original MoCA items using RUMM 2030 software (http://www.rummlab.com/). Briefly, Rasch analysis assesses whether a test or scale appropriately measures a single construct (e.g., cognitive ability) based on the participants' responses to test items when compared to the Rasch mathematical model (i.e., fit to Rasch model). It also assesses the psychometric properties of individual test items, including ordering test items by difficulty level, determining whether there are irrelevant or redundant test items, identifying test items that are influenced by other variables (e.g., sex, test language), again by comparing to what degree the data fit the ideal Rasch model. A cut-off of ± 2.5 was used to identify individual item misfit to the Rasch model. When a set of items (i.e., a test) fits the model, both the difficulty of an individual item and an individual's cognitive ability score is expressed in the same unit scale. Rasch analysis uses logits as the unit of measurement to describe the relative location of items and persons in relation to each other, where a value of 0 logits is given to an item in the middle of the spectrum of cognitive ability assessed by the MoCA, a value of +1 logits is given to items that lie 1 SD away from the middle in terms of their difficulty. For illustration, 0 logits converts to a score of 14 on the 0–30 point scoring system of the MoCA. The conventional cut-offs of 26 and 18 points for mild cognitive impairment and dementia[2] correspond to the logits scores +2.59 and +0.73, respectively.

RESULTS

Descriptive Statistics

Demographic and clinical characteristics of the sample are presented in Table 1. The mean age for the sample was 83.71

TABLE 1.
Demographic and clinical characteristics of the sample of 78

Age (years)	*Min – Max Mean±SD*	*69–95 83.71±5.66*
Sex	Female/Male (%)	41/59
Education (years)	Min – Max	0–25
	Mean±SD	12.71±4.90
	Missing	2
Language	English / French (%)	87.2/12.8
MMSE score	Median (Min – Max)	25 (14–30)
	N scoring < 26	44
Diagnosis	Not Cognitively Impaired	6 (7.7%)
	MCI	32 (41.0%)
	Dementia	39 (50.0%)
	Unspecified Cognitive Impairment	1 (1.3%)

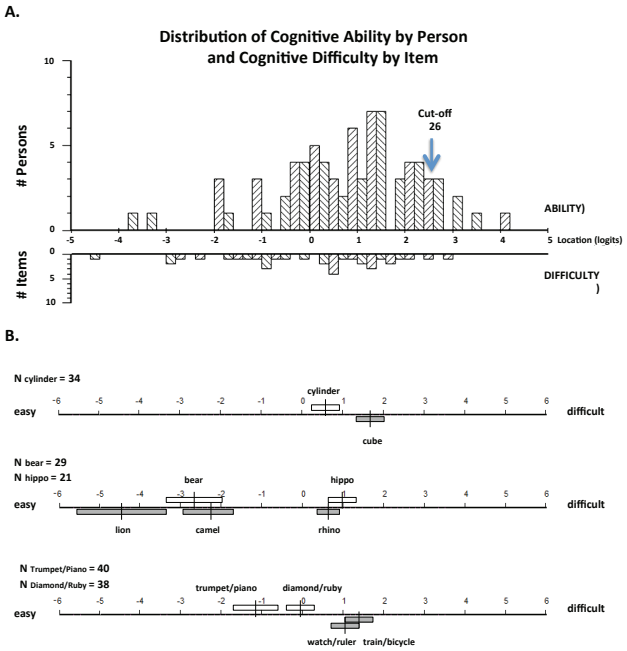

FIGURE 1. (A) Frequency distribution of the 78 individuals in top histogram and 33 MoCA items (28 originals + 5 alternates) in bottom histogram by location in logits on the ability/difficulty scale. The patient sample, and also the items, span a wide range of the construct of cognitive ability (± 4 SD). It can be seen that five people obtained a score between 0 and +0.2 logits, which corresponds approximately to a score of 16 on a 30-point scale. The mean ability of the whole sample is 0.82 logits, or approximately 18 on a 30-point scale. (B) Difficulties of the new items (white bars) and their corresponding original MoCA items (grey bars) as a function of the cognitive ability continuum represented by the full set of items. Error bars represent standard error of the estimate of item difficulty. N indicates the number of individuals who performed each alternate MoCA item.

± 5.66 years. On average, individuals comprising the sample had 12.71 ± 4.90 years of education. Fifty-nine percent were men, and 87.2% of participants took the test in English.

Rasch Analysis

Each item completed by a participant was scored as passed or failed except the Serial 7 subtraction task, which was scored using the 0–3 scoring system defined for the original MoCA. All item scores for the 28 original MoCA items and the five new items from the alternate MoCA versions were analyzed to evaluate the fit of the items to the Rasch model and to obtain estimates of the difficulty of each item. The overall test of item-trait interaction was non-significant (chi-square $p = .66$), and principal-components analysis revealed that the items yield a score that is predominantly a reflection of a single construct, cognitive ability, with little systematic correlation in the residuals (< 13% of the variance accounted for by a second factor). Each of the 33 items fit the Rasch model ($p > .05$, corrected), excepting a marginal fit for item copy Cube (fit residual 2.51, $p < .05$, corrected). The internal consistency of the set of items was 0.84.

Figure 1(A) shows the distribution of cognitive ability among the 78 participants (upper bars) and the distribution of difficulty among the 33 MoCA items (lower bars). Lower ability and easier items are on the left, and higher ability and harder items are on the right. Fit to the Rasch model means that an individual's ability is defined by the item(s) on which they have a 50% probability of responding correctly. The probability of their answering a question correctly increases beyond 50% for items progressively further to the left, and below 50% for items progressively further to the right on the graph. Figure 1(B) portrays the difficulty of the alternate MoCA items and their corresponding original items on a ruler-like scale in logit values, where increasingly negative numbers represent increasingly easier items, and increasingly positive numbers represent increasingly harder items. The Cylinder copying task was significantly easier than the Cube copying task by a difference of 1.08 logits, or 5.3 points on a 30-point scale. This suggests that while the Cylinder shows good fit to the Rasch model, it is too easy to be an appropriate alternate item for the Cube. In fact, 13 of 34 individuals (32.2%) who performed both tasks failed on copying a Cube but passed on copying a Cylinder.

Abstraction items such as Train/Bicycle and Watch/Ruler were 1.32 and 2.32 logits away towards the more difficult end from their respective alternate items, Diamond/Ruby and Trumpet/Piano. Fourteen out of 38 (36.8%) and 15 out of 40 (37.5%) individuals performed successfully Diamond/Ruby and Trumpet/Piano but failed Train/Bicycle and Watch/Ruler, respectively.

Even bigger discrepancies in location on the difficulty scale were observed for naming a Rhino and a Bear (3.33 logits, or 13.2 points difference), and naming a Camel and a

Hippo (3.30 logits difference). Seven out of 29 individuals (24.9%) who could not name the Rhino, named the Bear correctly, and 5 of 21 individuals (23.9%) who obtained a point for the Camel failed to name the Hippo. Thus the Bear is too easy to replace the Rhino, and the Hippo is too difficult to serve as an alternate item for the Camel.

DISCUSSION

In this study, we assessed the equivalency of original MoCA items to five alternate items selected from Alternate MoCA Version 2 and Alternate MoCA Version 3 by comparing their difficulty. All five of the tested alternate items were located more than 1 logit away from their corresponding original items. The percentage of individuals whose MoCA score was affected by the original to alternate item replacement ranged from 23.9%, if naming the Hippo was administered instead of naming a Camel, to 37.5%, if Trumpet/Piano was administered instead of Watch/Ruler.

These results suggest that a more systematic approach to selecting alternate MoCA items when replacing original MoCA items should be applied to ensure high parallel-forms reliability. Otherwise, measurement error introduced by administering alternate MoCA versions might be comparable to error introduced by practice effect. Offsetting the effects of an easier alternate item with the addition of a more difficult alternate item within the same alternate MoCA version may result in a reliable total score but will change the nature of the test in terms of how the components of the construct (i.e., cognitive ability) are covered. Alternate German MoCA versions show no increase in total scores over a 1-hour interval; however, significant differences between some domain scores are seen despite counterbalanced order of presentation (visuospatial/executive subtest version 3 vs. version 1, $t = 1.93$, $p = .03$).[12]

The main limitation of this study was its sample size, which resulted in high standard errors for the items' estimates. Adaptive administration means that different sets of people responded to different sets of questions. However, our previous work with larger samples that completed the full test showed that a person's performance on each MoCA test item relative to other test items is a function of cognitive ability rather than of potential confounders such as age or sex.[6] Studies including more individuals with different ability levels and testing other alternate MoCA items not evaluated in this study can give a better idea whether selection of appropriate items for the alternate versions based on estimates of their difficulty levels could be a beneficial approach to obtain high parallel-forms reliability.

The prevalence of uptake of the alternate versions of the MoCA in clinical practice settings is currently unknown. Our results have significant implications when considering increase in usage of alternate versions for longitudinal assessment in both research and clinical practice. We conclude that the use of alternate versions whose response patterns differ from the original MoCA may create more problems with reliability than it solves**.**

ACKNOWLEDGEMENTS

The authors would like to acknowledge funding from CIHR operating grant MOP-97810.

CONFLICT OF INTEREST DISCLOSURES

The authors declare that no conflicts of interest exist.

REFERENCES

1. Smith T, Gildeh N, Holmes C. The Montreal Cognitive Assessment: validity and utility in a memory clinic setting. *Can J Psychiatry*. 2007;52(5):329–32.
2. Nasreddine ZS, Phillips NA, Bedirian V, *et al*. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695–99.
3. Lee JY, Dong Woo L, Cho SJ, *et al*. Brief screening for mild cognitive impairment in elderly outpatient clinic: validation of the Korean version of the Montreal Cognitive Assessment. *J Geriatr Psychiatry Neurol*. 2008;21(2):104–10.
4. Debert CT, Benson BW, Dukelow S. Montreal cognitive assessment (MoCA): baseline evaluation of cognition in the athletic population. *Br J Sports Med*. 2013;47(5):e1.
5. Costa AS, Fimm B, Friesen P, *et al*. Alternate-form reliability of the Montreal cognitive assessment screening test in a clinical setting. *Dement Geriatr Cogn Disord*. 2012;33(6):379–84.
6. Koski L, Xie H, Finch L. Measuring cognition in a geriatric outpatient clinic: Rasch analysis of the Montreal Cognitive Assessment. *J Geriatr Psychiatry Neurol*. 2009;22(3):151–60.
7. Costa AS, Reich A, Fimm B, *et al*. Evidence of the sensitivity of the MoCA alternate forms in monitoring cognitive change in early Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2014;37(1-2):95–103.
8. Koski L, Xie H, Konsztowicz S. Improving precision in the quantification of cognition using the Montreal Cognitive Assessment and the Mini-Mental State Examination. *Int Psychogeriatr*. 2011;23(7):1107–15.
9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC: American Psychiatric Association, 2000.
10. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256(3):183–94.
11. Konsztowicz S, Xie H, Higgins J, *et al*. Development of a method for quantifying cognitive ability in the elderly through adaptive test administration. *Int Psychogeriatr*. 2011;23(7):1116–23.
12. Costa AS, Fimm B, Friesen P, *et al*. Alternate-form reliability of the Montreal Cognitive Assessment Screening Test in a clinical setting. *Dement Geriatr Cogn Disord*. 2012;33:379–84. Supplementary Material [cited 2015 July 15]. Available from: http://www.karger.com/ProdukteDB/miscArchiv/000/340/006/000340006_sm.html

Correspondence to: Lisa Koski, PhD, Psychology Department, P2.142, Allan Memorial Institute, 1025 Pine Avenue West, Montreal, QC H3A 1A1, Canada
E-mail: lisa.koski@mcgill.ca