VALIDATION AND DIAGNOSTIC UTILITY OF THE MINI-MENTAL STATE EXAMINATION AND MONTREAL COGNITIVE ASSESSMENT IN SCREENING FOR DEMENTIA WITHIN A MIXED CLINICAL SAMPLE

A Dissertation

Presented to

The Faculty of the Department

of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Katie McCulloch, M.A.

August 2014

VALIDATION AND DIAGNOSTIC UTILITY OF THE MINI-MENTAL STATE EXAMINATION AND MONTREAL COGNITIVE ASSESSMENT IN SCREENING FOR DEMENTIA WITHIN A MIXED CLINICAL SAMPLE

An Abstract of a Dissertation

Presented to

The Faculty of the Department

of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Katie McCulloch, M. A.

August 2014

Abstract

The Mini-Mental Status Examination (MMSE) and Montreal Cognitive Assessment (MoCA) are frequently utilized cognitive screening measures. The goal of the present study was to evaluate: (1) diagnostic utility values (e.g., sensitivity, specificity) of each measure, (2) cutoffs that maximize diagnostic accuracy within a mixed clinical sample, (3) the effect of base rates and severity of cognitive impairment on the efficacy of the screening measures, and (4) the relationship of the screening measure subscores to similar neuropsychological measures. The study included 218 veterans who completed the MMSE, MoCA, and neuropsychological testing. Empirically derived cutoffs across criterion variables – performance at least 1*SD* or 2*SD* below average on at least one neuropsychological domain, or dementia versus non-dementia diagnosis -- showed less than 24 and 25 as optimal for the MMSE with sensitivities ranging from 0.32 to 0.44 and specificities ranging from 0.78 to 0.87. Optimal cutoffs for the MoCA were 20, 21, and 25 with sensitivities ranging from 0.44 to 0.73 and specificities ranging from 0.57 to 0.83. Across criterion variables, the area under the receiver operating characteristic (ROC) curve (AUC) with the MMSE total score ranged between 0.59 and 0.70. The AUC of the MoCA ranged between 0.69 and 0.72, which was significantly greater than the MMSE when classifying patients based on the criterion of at least 1*SD* neuropsychological impairment. The MMSE and MoCA subtest scores showed poor convergent and discriminant validity relative to performance on neuropsychological domains, which indicates poor subscore interpretability. The study provides evidence that use of either the MMSE or MoCA increases classification accuracy beyond the base rate of dementia, although, of the two screening instruments, the MoCA has a relative advantage for classification accuracy at mild levels of neuropsychological impairment.

Acknowledgments

This material is the result of work supported with resources and the use of facilities at Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX.

Table of Contents

Title Page..... i

Signature Page…… ii

Abstract……... iii

Acknowledgements…… iv

List of Tables, Figures, and Appendices…..…….. v - vi

Introduction…… 1

Methods…….. 9

Participants……. 9

Materials and Procedures……... 10

Statistical Analyses……. 14

Results…… 16

Screening Test Total Score Comparison with Neuropsychological Testing….. 17

MMSE. ……..……..... 17

MoCA. ……..……..... 18

Screening Test Total Score Comparison with Diagnosis...…… 19

MMSE. ……..……..... 19

MoCA. ……..……..... 20

Classification accuracy among diagnostic groups. …… 21

Screening Test Subscores……..……. 21

MMSE. ……..... 21

MoCA. ……..... 22

Discussion……... 23

Tables……. 34-49

Figures……..... 50-51

Appendix A…… 52-53

Appendix B……. 54

Appendix C……. 55-56

References…….. 57-69

List of Tables, Figures, and Appendices

Table 1. ……. 34-36

Review of Existing Studies Investigating MMSE and MoCA Classification Accuracy

Table 2…….……..... 37

Participant Characteristics and Diagnoses

Table 3…… 38

Descriptives of Neuropsychological Measures in Dementia and Non-dementia Groups

Table 4……..……. 39-40

Frequency Distribution and Description of MMSE Subscores

Table 5…….…….……. 41

Frequency Distribution and Description of MoCA Subscores

Table 6…….…….. 42-43

Classification Accuracy Using MMSE Adjusted Total Scores and Three Criterion Variables

Table 7……..……. 44

Areas Under the Curve, Likelihood Ratios, Pre- and Post-test Probabilities, and Predictive Values for Selected Criteria and Base Rates

Table 8…….……..... 45-46

Classification Accuracy Using MoCA Adjusted Total Scores and Three Criterion Variables

Table 9……...…… 47

MMSE and MoCA Subscore Mean Ranks for Dementia Patients and Non-dementia Patients

Table 10..……..…….... 48

Spearman Correlations of Neuropsychological Composite Domains and MMSE Subscores Using all Participants

Table 11……..……... 49

Spearman Correlations of Neuropsychological Composite Domains and MoCA Subscores Using all Participants

Figure 1……. 50

Histogram of MMSE and MoCA adjusted total scores.

Figure 2……...…….. 51

Scatterplot of MMSE and MoCA scores among five diagnostic groups: individuals with dementia, MCI, a neurological disorder, a psychiatric disorder, and "normal" (i.e., no diagnosis) patients. Screening instrument cutoffs, derived with diagnosis as the criterion, are marked with lines to represent a MMSE cutoff at less than 24 and a MoCA cutoff at less than 20.

Appendix A……...…… 52-53

Copy of the Original MMSE

Appendix B…….…….. 54

Copy of the MoCA

Appendix C……..……. 55-56

Formulas to Calculate Classification Accuracy

Validation and Diagnostic Utility of the Mini-Mental State Examination and Montreal Cognitive Assessment in Screening for Dementia within a Mixed Clinical Sample

The baby boomer generation began turning 65-years old in 2011 and it is projected that by 2030 one in five residents of the United States will be over the age of 65 (Vincent & Velkoff, 2010). As the population ages, there are serious concerns about how to provide and fund health care, especially for individuals with long-term care needs such as dementia. The total cost of health care for individuals in the United States with dementia is projected to increase, unadjusted for inflation, from $200 billion in 2012 to $1.1 trillion in 2050 with 70% of these costs funded by Medicare and Medicaid programs (Alzheimer's Association, 2012).

In the coming years, as the proportion of older patients increases, accurate and early diagnosis of dementia will be increasingly important. Cognitive screening measures are widely utilized to quickly assess for symptoms of Alzheimer's disease and related dementias (ADRD). Neurology practice guidelines state that general cognitive screening instruments, as well as neuropsychological batteries, should be considered when evaluating for the presence of dementia (Petersen et al., 2001b). Screening measures can aid professionals in determining whether a referral for more extensive neuropsychological testing is necessary. Early detection can allow caregivers to formulate plans for future care and financial resources, alert physicians to start potentially beneficial medications, and motivate patients to improve their general physical health. Identifying true cognitive impairment is important for the reasons outlined above, but the ability to identify circumstances in which further testing is unnecessary is also important to minimize healthcare costs and reduce distress of patients and their loved ones.

Accurate diagnosis is complicated by varying symptomatic profiles among dementia subtypes. Causes of dementia can be associated with different cognitive, behavioral, neuropsychiatric, and motor symptoms. For example, cardinal signs of Lewy body dementia (i.e., fluctuating course, parkinsonian features, and hallucinations) are distinct from common characteristics of Alzheimer's dementia -- gradually declining course, prominent memory loss, and anomia (Lezak, Howieson, & Loring, 2004). In comparison, patients with vascular dementia can display impairment in a variety of cognitive domains, not necessarily involving memory, and the onset may be sudden with a course linked to the progression of cerebrovascular disease (Lezak et al., 2004). These symptom profiles may also evolve as the dementia progresses. For example, the most prevalent cause of dementia, Alzheimer's disease, is initially characterized by a gradual onset of memory impairment and, typically, later followed by impairments in language and visuospatial functioning (Mendez & Cummings, 2003). To further complicate diagnosis, an individual may have a mixed classification that reflects the presence of more than one cause of dementia. It is difficult to create a screening measure for dementia that would encompass all of these distinct cognitive profiles.

Another factor that complicates the effectiveness of a screening instrument for dementia is the broad spectrum of symptom severity that it must identify, which is particularly difficult when symptoms are of mild severity. Although there has been speculation for decades about a prodromal period preceding the development of Alzheimer's disease, interest increased following the formal adoption of mild cognitive impairment (MCI) criteria by the American Academy of Neurology (Petersen et al., 2001b). MCI is frequently suspected to represent early stages of ADRD (Petersen et al., 2001b; Shulman et al., 2006).

For example, Petersen and colleagues (1999) demonstrated that 12% of MCI patients converted to ADRD annually within a four-year follow-up interval compared to about 1-2% of controls, although many patients diagnosed with MCI may never progress to a future diagnosis of ADRD. There has also been speculation that patients classified with amnestic, single- or multi-domain, MCI subtypes are at relatively greater risk for converting to ADRD than individuals with non-amnestic MCI (Espinosa et al., 2013; Manly et al, 2008; Tabert et al., 2006), although this finding is not consistent and other factors associated with an increased risk of conversion have been documented (Boyle et al., 2005; Busse, Angermeyer, & Riedel-Heller, 2006; Conti et al., 2013; Xu et al., 2010). Using screening measures to detect mild cognitive impairment and early stages of dementia is complicated in practice and leaves much to be desired. Valcour, Masaki, Curb, and Blanchette (2000) estimated that as many as two-thirds of dementia cases seen in primary care settings are unrecognized, with the majority of missed cases being of mild or moderate severity. The National Institute on Aging and Alzheimer's Association workgroup (Albert et al., 2011) recommends that cognitive testing be used to assess the degree of cognitive impairment when evaluating for MCI and dementia, as informal memory tests are generally insensitive to early disease stages and frequently result in false negatives. Accurate detection of mild impairments, including MCI and early stages of dementia, is increasingly recognized as a goal when developing and utilizing cognitive screening measures.

As of yet, there is no consensus for a standard cognitive screening measure to detect MCI and ADRD. Cullen, O'Neill, Evans, Coen, and Lawlor (2006) describe an effective screen as one that first detects the presence of impairment and then is able to provide more information regarding the etiology. Therefore, an ideal cognitive screening test would be

both statistically robust as well as qualitatively rich. It is recommended that an effective screening test be validated in the intended population, be accurate across levels of impairment, be free of demographic biases, use profiles of impairment rather than cutoff scores, and represent cognitive domains across the spectrum of dementias (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). Logistically, it is also important for the measure to be easy and quick to administer. These are among many factors that each individual clinician must consider when choosing to use a screening measure. Clinicians must also consider the value of conventional measures, with large literature bases, versus newer measures that may reflect advanced understanding of MCI and ADRD. It is not likely that a single measure is the most appropriate choice in all situations, and it remains the clinician's role to choose what is most appropriate in each circumstance.

There are numerous cognitive screening measures available for professionals to use. To date, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975; Appendix A) is the most widely used cognitive screening measure (Shulman et al., 2006) and the 29th most popular neuropsychological measure (Rabin et al., 2005). Many professionals have familiarity and experience administering the instrument, which facilitates communication across settings and professionals. Since its development, the MMSE has been routinely utilized for clinical and research purposes and, subsequently, there is an extensive literature base associated with it. Since its inception, the MMSE has undergone only minor revisions to improve standardization; for example, it now includes fixed naming and memory objects. However, there continue to be drawbacks to its design. For example, the serial 7s and WORLD items continue to be treated as interchangeable in terms of scoring, despite evidence that serial 7s is more difficult than WORLD (Ganguli et al., 1990). Furthermore,

ceiling effects are consistently noted (Olson et al., 2011; Toglia, Fitzgerald, O'Dell, Mastrogiovanni, & Lin, 2011), such that the majority of patients have a high total MMSE score. There is a high proportion of language items on the MMSE that are rarely failed except in cases of profound impairment, and this contributes to the restricted range of scores (Tombaugh & McIntyre, 1992). The MMSE has been shown to be affected by age and education, which together account for an estimated 12% of the variance in total scores (Bravo & Herbert, 1997). Although the MMSE is frequently utilized, researchers have consistently demonstrated problems related to its design and psychometric characteristics.

The MMSE was originally developed to assess severity of cognitive impairment and document cognitive change over time but it is often used instead to screen for MCI and ADRD. The American Academy of Neurology guidelines for diagnosing MCI and dementia specifically recommend using the Mini-Mental State Examination, while acknowledging that age and education biases must be considered (Petersen et al., 2001a). A score of 23 or less has largely been accepted to represent cognitive impairment (Tombaugh & McIntyre, 1992), although this does not indicate the etiology of the failure. The MMSE can be administered to screen for a variety of causes of cognitive impairment other than dementia but there is some evidence that the MMSE is relatively insensitive for detecting milder forms of cognitive impairment, such as early stages of dementia (Ihl, Frolich, Dierks, Martin, & Maurer, 1992; Mitchell, 2009). It is important that clinicians base their predictions on norms that accurately reflect their entire population, including varying etiologies and severities of disease. Although the MMSE is a current standard for cognitive screening, there are several caveats regarding its use.

The Montreal Cognitive Assessment (MoCA) is another, more recent, cognitive screening instrument that has been gaining popularity (Nasreddine et al., 2005; Appendix B) and, to date, it has been cited in over 2,000 peer-reviewed journal articles. The MoCA was specifically designed to detect milder stages of dementia severity, including MCI and early stages of ADRD. In its original validation, Nasreddine and colleagues (2005) reported very high sensitivity (0.90-1.00) and specificity (0.87) for the MoCA in a sample of mixed Francophones and Anglophones, comparing performance of probable Alzheimer's disease or MCI against healthy controls. However, follow-up studies in English-speaking samples have generally yielded lower sensitivity and specificity values, with sensitivities ranging from 0.72 to 0.98 (Mickes et al., 2010; Whitney, Mossbarger, Herman, & Ibarra, 2012) and specificities ranging from 0.30 to 0.72 (Hoops et al., 2009; Mickes et al., 2010). Similar to the MMSE, the MoCA is suspected to be subject to age and education effects (Godefroy et al., 2011). Furthermore, many providers may be unfamiliar with the MoCA's design, scoring, and interpretation, which can hinder communication across professional settings.

In designing the MoCA to be sensitive to MCI and ADRD, the developers allocated a large proportion of the MoCA's total points to memory and executive functioning domains. The MoCA is divided into seven subscore domains: visuospatial/executive (letter-number sequencing, cube copy, and clock draw), naming (naming animals from a drawing), attention (repeating digits forward and backward, detecting the letter A in a sequence, serial subtraction by 7), language (sentence repetition, lexical fluency), abstraction (describing similarities of objects), delayed recall (recalling five words after five minutes), and orientation (four items for time and two items for place). Many of the MoCA subscores involve components of executive functioning. For example, the language domain assesses

lexical fluency, which requires basic language skills, but similar neuropsychological tests, such as FAS (Spreen & Benton, 1997; Spreen & Strauss, 1998) and Controlled Oral Word Association (COWA; Benton, Hamsher, & Sivan, 1994) have long been recognized as heavily influenced by executive functioning (Rabin, Barr, & Burton, 2005). Follow-up studies have typically focused on evaluating the MoCA by its total score. There has been very limited evaluation of the subscore domains, and this has been largely restricted to identifying the subscores failed most often in specific populations (Schweizer, Al-Khindi, & Macdonald, 2012; Sweet et al., 2011; Toglia et al., 2011; Whitney et al., 2012). Thus far, validation of MoCA subscore content and organization has not included comparison with neuropsychological testing in MCI and ADRD populations.

Although the MoCA was originally developed using patients with MCI and mild Alzheimer's disease, subsequent validation studies have been conducted in a variety of other patient populations, such as: patients with Huntington's disease (Videnovic et al., 2010), Parkinson's disease (Gill, Freshman, Blender, & Ravina, 2008; Hoops et al., 2009; Zadikoff et al., 2008), cerebrovascular disease (Dong et al., 2010; Pendlebury, Cuthbertson, Welch, Mehta, & Rothwell, 2010; Toglia et al., 2011), and brain metastases (Olson et al., 2011). Furthermore, the MoCA has been validated in a variety of cultural populations. The MoCA was originally developed in a sample of Canadian citizens and its validity has been established in a variety of other countries (e.g., Dong et al., 2010, 2013; Frietas et al., 2012a, 2012 b; Smith, Gildeh, & Holmes, 2007; Tsai et al., 2012). As of yet, its validity has been only minimally evaluated in dementia populations or within the United States (Hoops et al., 2006; Luis, Keegan, & Mullan, 2009; Mickes et al., 2010; Roalf et al., 2013; Whitney et al.,

2012). Therefore, nuances of interpreting the MoCA in American clinical practice may not be fully understood yet.

Thus far, there have been a small number of studies comparing the MMSE and MoCA in dementia assessment, which are summarized in Table 1. Study samples have been further restricted by the use of known-groups designs, which typically include only patients with a specific dementia subtype, MCI, or healthy controls. Such methodology is initially helpful for establishing construct validity of a measure but the external validity of classification accuracy results is limited by using distinctive groups and eliminating individuals who are likely to be misclassified. The majority of studies have used diagnosis as the criterion variable, although a few (Hoops et al., 2009; Whitney et al., 2012) have attempted to establish a criterion based on neuropsychological testing (i.e., ≥2 domains failed at ≤-1.5*SD*). Of the three studies that have attempted to make demographic adjustments to total scores, no study has applied a correction to the MMSE total score, despite adjusting MoCA scores (Dong et al., 2010; Freitas, Simões, Alves, Vincente, & Santana, 2012b; Nasreddine et al., 2005). The adjustments to MoCA total scores have been exclusively determined by educational attainment and no study has attempted to norm scores by the patient's age. Regarding MoCA educational adjustments, one study relied on its own corrections and two studies added one point if ≤12 years education, which is not the most appropriate adjustment (Johns et al., 2010). Despite the existence of several studies comparing the MMSE and MoCA, there are several methodological features that could be improved to better assess external validity.

The existing literature indicates that the MMSE and MoCA are both frequently utilized cognitive screening measures with unique limitations. In order to maximize effective

screening for dementia, it is imperative to assess which measure is more appropriate in clinical practice and how this decision is affected by unique factors of the setting. This study seeks to determine whether the MMSE or MoCA is more effective in screening patients with dementia in a mixed clinical sample. This will be determined by evaluating: (1) the diagnostic utility values (e.g., sensitivity, specificity, likelihood ratios, and pretest/posttest probabilities) of each screening measure, (2) cutoffs for each test that maximize diagnostic accuracy within a mixed clinical sample, (3) how the severity of cognitive impairment and base rates of impairment influence appropriate test selection, and (4) whether the subscore domains of each screening measure are consistent with comparable neuropsychological measures.

Methods

Participants

This study included 218 participants (204 males, 14 females) with a mean age of 66.5 years (*SD*=11.3). Participant demographic characteristics are presented in Table 2. Participants were referred to a cognitive disorders clinic at the Michael E. DeBakey Veteran Affairs Medical Center in Houston, Texas. Procedures were approved by the Institutional Review Boards at the Baylor College of Medicine and the University of Houston and further reviewed by the Research and Development Committee at the Michael E. DeBakey Veteran Affairs Medical Center. Participants were included if they were evaluated by the neuropsychology clinic between 2005 and 2012, received an MMSE during the neuropsychological assessment, and completed a MoCA during a cognitive disorders clinic evaluation (78% of participants had the MoCA administered first, 22% had the MMSE administered first, <1% completed both instruments on the same day). Participants were

excluded if they did not receive the MMSE and MoCA within a 6-month period (the mean interval was 60.5 days, *SD*=37.5 days, *n*=64), or if they demonstrated suboptimal performance validity during the neuropsychological evaluation (*n*=56). Suboptimal performance validity was defined as failure on the Test of Memory Malingering (Tombaugh, 1996), Word Memory Test (Green, 2003), or California Verbal Learning Test-II (CVLT-II) forced-choice recognition (Delis, Kramer, Kaplan, & Ober, 2000).

Materials and Procedure

Appendix A contains the original MMSE form and instructions (Folstein et al., 1975). The MMSE total score consists of 30 points spanning 11 subscore domains: orientation to time (5 points), orientation to place (5 points), registration (3 points), attention and calculation (5 points), recall (3 points), naming (2 points), repetition (1 point), verbal comprehension (3 points), written comprehension (1 point), writing (1 point), and construction (1 point). Possible scores range from 0 to 30 points with smaller values indicating cognitive impairment. Reported values of internal consistency, as measured by Cronbach's alpha, range between 0.60 and 0.89 (Larner, 2013; Mitchell, 2009; Mystakidou, Tsilika, Parpa, Galanos, & Vlahos, 2007; Toglia et al., 2011). Test-retest reliability within 24 hours is typically reported to be above 0.85 (Larner, 2013; Mitchell, 2009).

Appendix B contains the MoCA test form (Nasreddine et al., 2005). The MoCA test consists of 30 points across eight subscore domains: visuospatial/executive (5 points), naming (3 points), memory encoding (no points), attention (6 points), language (3 points), abstraction (2 points), delayed recall (5 points), and orientation (6 points). Similar to the MMSE, possible scores range from 0 to 30 points with smaller values indicating cognitive

impairment. During its initial validation, the internal consistency was reported to be 0.83 and test-retest reliability was 0.92 (Nasreddine et al., 2005).

Demographic information was collected from medical records and the clinical interview. MMSE, MoCA, and neuropsychological test scores were obtained from patient medical records. The MoCA administration closest in time to the neuropsychological evaluation was selected. Total and subscale scores from the MMSE and MoCA were recorded in the database. Total scores for both screening tests were adjusted for demographic variables, including age and education. The raw total score of the MMSE was corrected for age and education effects (Mungas, Marshall, Weldon, Haan, & Reed, 1996): MMSEadj = MMSEraw – [0.471*(Education-12)] + [-0.131*(Age-70)]. Then, the adjusted MMSE scores were rounded down to the nearest whole number. The MoCA raw total score was adjusted for education, as proposed by Nassreddine's most recent recommendations: add two points for four to nine years of education or add one point for 10 to 12 years of education (Johns et al., 2010). All patients had MMSE total score and subscores available. At the item level on the MMSE, the higher score of WORLD and serial 7's was utilized, as recommended by Tombaugh and McIntyre (1992). Twenty-one patients did not have subscores of their MoCA available.

Patients completed a battery of standardized neuropsychological instruments. Neuropsychological measures were administered and scored by a board-certified clinical neuropsychologist or trainees under his supervision. Composite scores for the following five cognitive domains (methodology described below) were calculated after reviewing correlations among measures in the neuropsychological battery.

Language: Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983; normed with Heaton, Miller, Taylor, & Grant, 2004), Wechsler Adult Intelligence Scales-III Similarities (WAIS-III; Wechsler, 1997b), and Wechsler Test of Adult Reading (WTAR; Holdnack, 2001).

Verbal memory: CVLT-II trials 1-5 total, long delay free recall, and d' scores (Delis et al., 2000); and Wechsler Memory Scales-III Logical Memory I and II total scores (WMS-III LM; Wechsler, 1997a).

Attention and processing speed: Trail Making Test A (Reitan, 1958; normed with Heaton et al., 2004), WAIS-III Digit Span, Symbol Search, and Digit-Symbol Coding (Wechsler, 1997b), and Wechsler Adult Intelligence Scales-IV Symbol Search and Coding (WAIS-IV; Wechsler, 2008).

Visuospatial processing: WAIS-III or WAIS-IV Block Design (Wechsler, 1997b; Wechsler, 2008) and Judgment of Line Orientation (JLO; Benton, Sivan, Hamsher, Varney, & Spreen, 1994; normed with Eslinger, Damasio, Benton, & Van Allen, 1985).

Executive functioning: Trail Making Test B (Reitan, 1958; normed with Heaton et al., 2004), WAIS-IV Digit Span and Visual Puzzles (Wechsler, 2008), FAS phonemic fluency (Spreen & Benton, 1997; Spreen & Strauss, 1998; normed with Heaton et al., 2004), and animal naming (Goodglass & Kaplan, 1972; normed with Heaton et al., 2004). The Digit Span subtest on the WAIS-III correlated with other measures of attention and processing speed, although on the WAIS-IV it is more strongly correlated with measures of executive functioning.

Available test scores within each composite neuropsychological domain were converted to *z*-scores and averaged. Not all patients had scores within each of the five

cognitive domains because a fixed-flexible neuropsychological battery was implemented, i.e., the specific tests administered differed somewhat across patients. For total score analyses of screening measures, 12 participants were excluded for having fewer than two composite neuropsychological scores available; however, these participants were utilized later during subscore analyses if they had undergone testing within that particular neuropsychological domain. Individuals missing at least one cognitive domain had lower total MMSE scores (M=23.83, SD=4.89) than those with scores from all domains (M=26.33, SD=3.05), $t(47)=3.13$, $p<.01$, when Levine's Test indicated unequal variances. Similarly, individuals with one or more missing domain scores had lower total MoCA scores (M=18.95, SD=4.74) than participants without missing domain scores (M=22.82, SD=3.98), $t(216)=5.40$, $p<.001$. Average performance on neuropsychological testing and composite domains is presented in Table 3 comparing patients with dementia and non-dementia diagnoses.

Clinical diagnosis for each patient was determined retrospectively by two board-certified neurologists. A randomized list of patients included in the study was provided to each neurologist. Diagnoses were usually assigned independently based on practice standards of the neurology department using information available in the patient's electronic medical records (i.e., progress notes, imaging, laboratory tests). As of yet, there is no definitive way to diagnose many forms of dementia or MCI while the patient is alive, although clinical presentation and medical history are commonly utilized to determine possible or probable diagnoses. The raters were blind to certain information to reduce bias and confounding effects in the following analyses. For example, performance on neuropsychological testing and screening measures was not considered in the determination of diagnosis. The raters also

did not consider records more recent than the later of each patient's MMSE or MoCA assessments in order to characterize the disorder or severity at the time of testing.

Statistical Analyses

Classification accuracy was evaluated by comparing MMSE and MoCA total scores against three criterion standards. The first and second criterion variables were based on neuropsychological test performance, with impairment defined at two levels to reflect varying levels of severity. The first level of severity, performance at least 1*SD* below average on at least one of the five composite neuropsychological domains, was intended to reflect impairment of mild severity. At the second level of severity, performance at least 2*SD* below average on at least one of the five composite neuropsychological domains, was selected to represent a greater degree of impairment. Although there is no consensus regarding the definition of impairment on neuropsychological testing, some consider between one and two standard deviations below the mean as borderline or mild impairment, whereas performance greater than two standard deviations below the mean would represent cognitive impairment in the moderate to severe range and convey a greater degree of confidence that performance is below normal limits (Iverson & Brooks, 2011). This distinction is also suggested by the DSM-5 to help clinicians differentiate severity of neurocognitive symptoms between Mild and Major Neurocognitive Disorders (American Psychiatric Association, 2013). The third criterion for classification accuracy was dementia diagnosis. The adjusted total scores for each screening measure were evaluated against the presence or absence of any dementia diagnosis in the sample.

Optimal classification was determined for each of these three criterion standards. A positive score on the predictor or criterion variable indicated the presence of cognitive

impairment, whereas a negative score suggested that the patient's cognitive ability was within normal limits. Test sensitivity was calculated as the number of patients who are classified as impaired by both the criterion and the screening test as a proportion of all patients classified as impaired on the criterion. Specificity was defined as the proportion of patients classified as not impaired by both the criterion and the screening test out of all patient classified as not impaired on the criterion. Indices of sensitivity and specificity were evaluated at each possible cutoff level. Receiver operating characteristic (ROC) curves were created and areas under the curve (AUC) were calculated, which is an overall measure of test performance that represents the relationship between sensitivity and specificity but is unaffected by cutoff selection. AUC values range from 0.5 to 1 with higher values indicating greater classification accuracy. MMSE and MoCA AUC were compared at each criterion using the Hanley and McNeil (1982, 1983) method, which assesses the statistical significance of differences in the areas under two ROC curves based on the same sample of participants. Youden's Index (Youden, 1950) was calculated to select optimal cutoffs for the screening tests. Once optimal cutoff values were selected, positive and negative predictive values (PPV and NPV) were calculated. PPV is the conditional probability that individuals who are classified as impaired on the screening test would also be classified as impaired on the criterion measure, and NPV is the conditional probability that individuals who were not classified as impaired on the screening test would also not be classified as impaired on the criterion measure. When calculating PPV and NPV in the present sample, the base rate of impairment was used as well as additional base rate values (e.g., 0.10, 0.35, 0.60) that were selected to simulate other samples. Positive and negative likelihood ratios (+LR, -LR) were calculated with optimal cutoffs for the screening tests to determine the change in probability

favoring the presence of dementia if the screening test score is positive (+LR) or the change in probability favoring the absence of dementia if the screening test score is negative (-LR). Additionally, pre-test probability and post-test probability were compared at the optimal cutoff values. Pre-test probability is the probability of dementia before the screening test is administered (i.e., the base rate of impairment), and post-test probability is the probability of dementia after interpretation of the score on the screening test (i.e., equivalent to PPV calculated with base rates of impairment). An increased post-test probability beyond the base rate of the condition within the sample provides an indication as to the effectiveness of the screening measure and may have bearing on the clinician's diagnostic confidence. Formulas discussed above are provided in Appendix C.

Efforts to describe MMSE and MoCA subscores were also undertaken. This was begun by comparing means on each subscore between those individuals with dementia and those with non-dementia diagnoses. Many subscore distributions were non-normal because of skewness or kurtosis. The mean, median, and mode values tended to be clustered at upper end of the score range (Tables 4 and 5). Given this violation of normality and the ordinal nature of scores, the Mann-Whitney U test (Mann & Whitney, 1947) was utilized to evaluate group differences in subscore performance. Spearman rank-order correlations, also non-parametric and appropriate for ordinal scores, were utilized to measure the association of scores on screening subtests with neuropsychological composite scores among all sample participants.

Results

The Spearman rank-order correlation between adjusted total MMSE and MoCA scores was 0.52. The distribution of all MMSE and MoCA adjusted total scores is displayed

in Figure 1. Both distributions were negatively skewed, but to different degrees (MMSE=-1.22, MoCA=-0.68). The MMSE's range appears to be restricted by a ceiling effect, as the majority of scores were distributed at the upper end, e.g., the mode of MMSE scores was 30. Furthermore, a Wilcoxon signed-rank test showed that the median rank-order of MMSE adjusted total scores (median=27.00, *M*=25.86, *SD*=3.59) was significantly greater than the median rank-order of MoCA scores (median= 22.50, *M*=22.09, *SD*=4.39), *Z*=-11.16, $p<.001$. Thus, the MMSE and MoCA distributions are different, and total scores cannot be interpreted as equivalent across measures.

Screening Test Total Score Comparison with Neuropsychological Testing

MMSE. First, adjusted MMSE total scores were evaluated against neuropsychological testing, such that a positive result was defined as scoring at least 1*SD* below the neuropsychological score in any composite cognitive domain. Using this criterion, the MMSE total score yielded an AUC of 0.59, which represents an overall measure of test performance. Sensitivity and specificity values at each possible cutoff are displayed in Table 6. The optimal total score cutoff on the MMSE was less than 25. This cutoff yielded a sensitivity, the proportion of patients who are classified as impaired by both the criterion and the screener out of all patients classified as impaired on the criterion, of 0.32 and a specificity, the proportion of patients classified as not impaired by both the criterion and the screener out of all patient classified as not impaired on the criterion, of 0.84. Using this sample's base rate of 0.73, the PPV, the conditional probability that individuals who are classified as impaired on the screening test were also classified as impaired on the criterion measure, was 0.84. The NPV, the conditional probability that individuals who were not classified as impaired on the screening test were also not classified as impaired on the

criterion measure, was 0.31. The positive likelihood ratio was 1.99, meaning that a patient with neuropsychological impairment is twice as likely to have a positive MMSE result than a cognitively intact person. The negative likelihood ratio was 1.23, meaning that a cognitively intact participant is 1.23 times as likely to have a negative MMSE result than someone with impairment. Relative to a pre-test probability of 73% (*n*=150 with one or more composite domains at least 1*SD* below the mean, *n*=56 with all composite scores above this cutoff), the post-test probability of impairment was 84%, meaning that use of the MMSE improves identification of impairment by 11% beyond the base rate alone (Table 7).

Then, MMSE total scores were compared with a greater degree of neuropsychological impairment, such that a positive result was defined as performance of at least 2*SD* below average on any cognitive domain. Using these criteria, the MMSE total score yielded an AUC of 0.64. The optimal total score cutoff on the adjusted MMSE was again 25, with a sensitivity of 0.44 and a specificity of 0.78. Using this sample's base rate of 0.29, the PPV was 0.45 and NPV was 0.77. The positive likelihood ratio was 2.02, whereas the negative likelihood ratio was 1.40. Compared with a pre-test probability of 29% (*n*=59 with one or more composite domains at least 2*SD* below the mean, *n*=147 with all composite scores above this cutoff), the post-test probability of impairment was 45%.

MoCA. MoCA total scores were evaluated against neuropsychological testing scores, such that a positive result was defined as performance 1 or more *SD* below the neuropsychological test performance in any composite neuropsychological domain. Using this criterion, the MoCA total score yielded an AUC of 0.69. Based on the Hanley and McNeil (1982, 1983) method, the AUC of the MoCA was significantly greater than that of the MMSE, z=2.23, p=.03. Sensitivity and specificity values at each possible cutoff are

displayed in Table 8. The optimal total score cutoff on the MoCA was less than 25, which yielded a sensitivity of 0.73 and a specificity of 0.57. For this sample's base rate of 0.73, the PPV was 0.82 and NPV was 0.44. The positive likelihood ratio was 1.71, whereas the negative likelihood ratio was 2.14. The pre-test probability of 73% and the post-test probability of impairment was 82% (Table 7).

Then, MoCA total scores were compared with a greater level of neuropsychological impairment, such that a positive result was defined as performance of at least 2*SD* below average on any cognitive domain. Using these criteria, the MoCA total score yielded an AUC of 0.72. The AUC of the MoCA was not significantly greater than that of the MMSE using this criterion, z=1.86, p=.06. The optimal total score cutoff on the MoCA was less than 21, which yielded a sensitivity of 0.56 and a specificity of 0.78. Using this sample's base rate of 0.29, the PPV was 0.51 and NPV was 0.81. The positive likelihood ratio was 2.56, whereas the negative likelihood ratio was 1.77. Compared with a pre-test probability of 29%, the post-test probability of impairment was 51%.

Screening Test Total Score Comparison with Diagnosis

MMSE. Among the three primary diagnostic categories (i.e., dementia, MCI, other diagnosis), one-way ANOVA yielded a significant between-group difference in MMSE adjusted total scores, $F(2, 217) = 17.52$, $p<.001$. Tukey post-hoc comparisons indicated that the dementia group had a lower MMSE score (M=23.86, SD=4.45) than the MCI (M=26.13, SD=2.73, $p<.001$) and other diagnosis groups (M=27.12, SD=2.81, $p<.001$). The MCI and other diagnosis groups did not differ significantly from each other.

MMSE adjusted total scores were evaluated against diagnosis, i.e., dementia versus non-dementia, in a classification accuracy analysis. The MMSE total score yielded an AUC of 0.70. Sensitivity and specificity values at each possible cutoff are displayed in Table 6. The optimal total score cutoff on the MMSE was less than 24, which yielded a sensitivity of 0.40 and a specificity of 0.87. For this sample's base rate of 0.28, the PPV was 0.46 and NPV was 0.84. The positive likelihood ratio was 2.99, whereas the negative likelihood ratio was 1.45. The pre-test probability of dementia diagnosis was 28%, and the post-test probability was 46% (Table 7).

MoCA. A one-way ANOVA yielded a significant difference among the three diagnostic groups (i.e., dementia, MCI, other), in MoCA adjusted total scores, $F(2, 215)=12.71$, $p<.001$. Similar to MMSE results, the Tukey post-hoc comparisons indicated that the dementia group had a significantly lower MoCA score ($M=19.86$, $SD=5.01$) than the MCI ($M=22.96$, $SD=3.20$, $p<.001$) and other diagnosis groups ($M=23.04$, $SD=4.20$, $p<.001$). The MCI and other diagnosis groups did not differ significantly.

MoCA total scores were also evaluated against diagnosis (i.e., dementia versus non-dementia) in a series of classification analyses. The MoCA total score yielded an AUC of 0.69, which was similar to the MMSE's AUC, $z=0.33$, $p=.74$. Sensitivity and specificity values at each possible cutoff are displayed in Table 8. The optimal total score cutoff on the MoCA was less than 20, with sensitivity of 0.44 and a specificity of 0.83. Using this sample's base rate of 0.28, the PPV was 0.49 and NPV was 0.79. The positive likelihood ratio was 2.51, whereas the negative likelihood ratio was 1.46. The pre-test probability of dementia diagnosis was 28%, and the post-test probability was 49% (Table 7).

Classification accuracy among diagnostic groups. In Figure 2, adjusted total scores of the MMSE and MoCA are displayed by diagnostic groups: dementia, MCI, neurological, psychiatric, and no diagnosis. The MoCA and MMSE demonstrated consistent levels of agreement among diagnostic groups, ranging from 75% (dementia) to 80% (MCI). However, both the MMSE and MoCA were least accurate in classifying the dementia group (40% correctly classified by the MMSE and 43% by the MoCA). The MMSE failed to identify nine individuals with dementia that were correctly classified by the MoCA, whereas the MoCA failed to identify seven individuals that were correctly classified by the MMSE. In contrast, the combination of optimal MMSE and MoCA adjusted scores correctly discriminated between 75% and 92% of patients within MCI, neurological, psychiatric, and no diagnosis groups from those patients with dementia. This discrepancy in accuracy for dementia vis-à-vis other diagnoses is likely a reflection of selecting cutoffs with high specificity for detecting dementia.

Screening Test Subscores

MMSE. Mean rank subscores on the MMSE were compared for individuals with and without dementia using the Mann-Whitney U test. The results are displayed in Table 9. There were significantly lower MMSE subscores for the dementia group on orientation to time, $p<.001$; attention/concentration, $p<.01$; delayed recall, $p=.04$; naming, $p=.04$; writing, $p=.02$; and drawing, $p<.01$. Although these results are statistically significant, the clinical significance is questionable, as none of the mean scores differed by as much as one point.

MMSE subscores then were correlated with neuropsychological composite domains among all participants. Results are displayed in Table 10. Four MMSE subscores -- naming,

comprehension, reading, writing -- were not significantly correlated with any neuropsychological domain, $p>.05$. Two of these subscores, naming and writing, differed between patients with dementia and non-dementia diagnoses (see above Mann-Whitney U results) but were not significantly correlated with any neuropsychological domain. It is unknown what these subscores are measuring to distinguish these groups. Of those subscores that were correlated with neuropsychological composite scores, three of the MMSE subscores (orientation to time and place, repetition) were most highly correlated with executive functioning, two (attention/concentration, delayed recall) with memory, one (registration) with visuospatial functioning, and one (drawing) with processing speed/attention. Furthermore, many were significantly correlated with more than one domain. For example, three subscores (orientation to time, repetition, and drawing) were significantly correlated with all five neuropsychological domains. One subscore (orientation to place) was correlated with four domains; two subscores (delayed recall, attention/concentration) were correlated with three domains; and registration was correlated with two domains. All in all, this pattern of correlations indicates poor convergent and divergent validity for MMSE subscores.

MoCA. Mean rank subscores on the MoCA were compared for individuals with and without dementia using the Mann-Whitney U test. Results are displayed in Table 9. Scores for the dementia group were significantly lower on visuospatial/executive, $p<.001$; naming, $p=.03$; language, $p<.01$; abstraction, $p=.03$; and orientation, $p<.001$, subtests. Despite statistical significance, the clinical significance is again limited, as the greatest difference among groups was a 1.15 point difference on the visuospatial/executive subscore.

Subsequently, MoCA subscores were correlated with neuropsychological composite domain scores for all participants. Results are displayed in Table 11. Each MoCA subscore was significantly correlated with at least one neuropsychological domain. Two of the MoCA subscores (language, attention) were most highly correlated with executive functioning; two subscores (naming, abstraction) with language; two (orientation, delayed recall) with memory; and one with visuospatial functioning (visuospatial/executive). Many subscores were significantly correlated with more than one domain. For example, three subscores (visuospatial/executive, attention, and language) were significantly correlated with all five neuropsychological domains; three subscores (orientation, abstraction, and naming) were correlated with three domains; and delayed recall was correlated with two domains. Given the non-normal nature of the majority of MMSE and MoCA subscore distributions, the lack of correlations between screening test subscores and neuropsychological scores, and the absence of consistent group differences in subscores, it was deemed inappropriate to assess the predictive value of MMSE and MoCA subscores.

Discussion

There is an increasing need for effective tools to identify dementia as the population of the United States ages. The MMSE and MoCA are two popular screening measures for detecting cognitive sequela associated with Alzheimer's disease and related dementia (ADRD). The current literature is limited by infrequent validation of the English version of the MoCA, use of predetermined cutoffs, lack of age and education corrections to total scores, and reliance on known-groups designs that limit external validity and artificially inflate classification accuracy values. The present study sought to address these issues by comparing screening tests after demographic (i.e., age, education) adjustments and by

deriving cutoffs to calculate classification values with multiple evaluation criteria (i.e., two levels of impairment measured by neuropsychological testing and patient diagnoses) within the context of a mixed clinical sample.

Both screening measures were shown to be capable of contributing valuable diagnostic information and, in fact, classification accuracy values were highly consistent between the MMSE and MoCA. For example, in this sample, when a positive result was obtained from the MMSE, the odds of making a correct diagnosis of cognitive impairment or dementia increased beyond the clinical base rate by 11-26%, depending on the criterion variable, whereas the MoCA increased the odds by 9-22%. However, the AUC is often used to compare various tests because other values of classification accuracy (e.g., sensitivity, specificity) are highly dependent upon the cutoff selected and whether sensitivity or specificity is favored at that cutoff. In this study, AUC values were shown to be greater for the MoCA when including a mild degree of neuropsychological impairment. Otherwise, when the criterion was severe neuropsychological impairment or dementia diagnosis the AUC was similar between the MMSE and MoCA, although there continued to be a general trend for MoCA AUC values to be larger. The development of the MoCA was aimed at creating a screening measure with improved classification accuracy, relative to the MMSE, to identify individuals with mild degrees of cognitive impairment (Nasreddine et al., 2005) and these results indicate that some success has been realized in achieving this goal.

Based on the three selected criterion, results yielded multiple cutoffs for each cognitive screening measure. Empirically-derived MMSE cutoffs (24, 25) in this study were generally consistent with fixed cutoffs (25, 26, 27) but lower than other empirically-derived cutoffs (28, 30) in the existing English-speaking literature comparing the MMSE and MoCA.

Similarly, empirically-derived MoCA cutoffs (20, 21, 25) in this study tended to be lower than the fixed (24, 26, 27) and derived (23, 26, 27) cutoffs proposed in this literature. Possible explanations for this discrepancy include the various severities of impairment among samples and the different methods to select an optimal cutoff. The present sample consists of patients who have proceeded through multiple levels of care to reach neuropsychological services (i.e., primary care and neurology services). These individuals likely represent more complicated, atypical, and impaired cases relative to those that can be successfully managed within primary care. Given this biasing factor, the MMSE and MoCA score distributions demonstrated in this study are likely lower than other studies including only early stages of dementia or utilizing known groups to generate optimal cut scores. Consequently MMSE and MoCA cutoffs would be lower. Additionally, procedures for optimal cutoff selection vary across studies, with many manuscripts not specifying what method was utilized. Several studies, similar to ours, utilized Youden's Index (Youden, 1950), which balances the relative importance of sensitivity and specificity to select the cutoff with the maximal combined value. However, depending on the characteristics of the sample, one may prefer to use a cutoff with greater sensitivity or specificity and to alter the cutoff selection accordingly (Grimes & Schulz, 2002; Sackett et al., 2000). For example, in primary care settings one would prefer to identify all possible cases of dementia and thus would select a higher cutoff (for greater sensitivity). In a specialty setting, where the emphasis is likely to be placed on differentiating dementia from other diagnoses, a lower cutoff would be favored (for greater specificity). For the purposes of this study, cutoffs favoring sensitivity and specificity equally were selected, although tables are provided to allow clinicians to select a cutoff that would be most appropriate to their setting. In addition

to these MMSE and MoCA cutoffs being lower than in other studies, MoCA cutoffs were lower than the MMSE cutoffs. This result was inconsistent with Nasreddine and colleagues' (2005) validation of the MoCA that proposed a cutoff of 26 as optimal for both the MMSE and MoCA. This mutual cutoff of 26 is what many researchers have utilized as a fixed cutoff when comparing the MMSE and MoCA. However, the consistently lower MoCA cutoffs of this study align with several follow-up studies that also derived cutoffs (Dong et al., 2010; Dong et al., 2013; Freitas, Simões, Alves, Duro, & Santana, 2012a; Hoops et al., 2009; Roalf et al., 2013; Tsai et al., 2012; Waldron-Perrine & Axelrod, 2012). It is imperative that clinicians select cutoffs derived from an appropriate sample in order to achieve comparable predictive ability and the most accurate classification.

Much of the existing literature comparing the MMSE and MoCA has failed to derive cutoffs and very few studies have attempted to adjust for age or education effects. There may be some discrepancy between total scores in this manuscript and those of other studies because age and education adjustments were utilized. Such adjustments result in a greater spread of scores within each distribution, and that may have contributed to a ceiling effect in MMSE scores. MoCA scores were also largely clustered in the top half of the score range, although to a lesser extent. The observed range restriction, in the context of a small possible total score range (i.e., 31-point range), limits the predictive ability of each measure, most notably the MMSE. As is currently recommended by Johns et al. (2010), the MoCA is not corrected for by age. Nasreddine and colleagues (2005) did not show age effects, although their sample was largely limited to older adults. There are methods, other than raw score addition, to account for age and education characteristics, e.g., normative tables that allow for a *z*-score conversion (Crum, Anthony, Bassett, & Folstein, 1993; Rossetti, Lacritz, Cullum,

& Weiner, 2011). Normative conversions may allow for easy interpretation within a neuropsychological evaluation in which the majority of tests are normatively scored. However, this comparison with norms is not standard practice for interpreting cognitive screening tests. It increases time and effort in interpretation, and it may be a technique that some providers find unfamiliar and difficult to use. There are also notable differences between screening tests secondary to literature base, as MMSE norms by Crum and colleagues (1993) have been cited and reevaluated more than the Rossetti et al. (2011) MoCA norms. There has been too little evaluation of raw score adjustments and normative score conversions to allow a determination as to what method best describes an individual's true cognitive state. Raw score adjustments for age and education may or may not be the most effective method for accounting for the influence of these variables, but were used in this study to reflect the way in which the majority of clinicians are likely to be using these instruments.

Cognitive screening instruments were not able to differentiate small degrees of impairment. Regarding the selection of multiple criterion variables, there was very little difference in cutoffs between levels of neuropsychological impairment (i.e., 25 for MMSE for both cutoffs versus 21 and 22 for MoCA cutoffs when specificity was favored). Further, when comparing total scores between diagnostic groups, screening tests were able to distinguish dementia patients from all other patients, although patients with MCI were indistinguishable from other non-dementia patients. Due to their ceiling effects and lack of score variability, the screening tests appear more appropriate for determining if there is some level of impairment that is outside normal limits. As the ceiling effect and restricted range seem to contribute to limited interpretability, several investigators have attempted to enhance

the MMSE by expanding the total point scale. For example, the MMX and 3MS have been developed, and they have been shown to improve classification ability over the original MMSE (Kaufer et al., 2008; Teng & Chui, 1987; Tombaugh, McDowell, Kristjansson, & Hubley, 1996), although these revised forms are not as frequently utilized. No corresponding studies have provided expanded scoring for the MoCA.

Using diagnosis as the criterion produced little change in cutoff score on the MMSE, although the optimal cutoff was higher on the MoCA when it was based on diagnosis instead of neuropsychological testing. This is likely related to the use of Youden's Index (Youden, 1950) as an objective method for selecting cutoff values, which does not favor sensitivity or specificity, rather than a consequence of the selection of criterion variable. The diagnosis-based MoCA cutoff favored sensitivity while all of the other MMSE and MoCA cutoffs more strongly favored specificity. If a cutoff favoring specificity was selected with the greatest subsequent Youden's Index value, the optimal cutoff would have been less than 22, rather than 25, which is much more consistent with the other empirically-derived cutoffs on the MoCA.

MMSE and MoCA subscores were evaluated for their construct validity and ability to distinguish individuals with and without dementia. Both tests contained subscores that did and did not distinguish these patient groups, although the MMSE tended to contain a greater proportion of poorly discriminating subscores (i.e., 4 of 10 MMSE subscores did not differentiate dementia and non-dementia patients versus 2 of 7 MoCA subscores). Subscores that do not differentiate patients are likely to contribute to the ceiling effects discussed previously. The implications of subscore differences between dementia and non-dementia patients are ambiguous, and these differences are of questionable practical utility. For

example, the group differences on subscores were statistically significant but were very small, typically less than one score point, making it difficult to differentiate groups reliably in clinical practice. Further, there was substantial intercorrelation among subscores and neuropsychological domains, and four subscores on the MMSE were not related to any domain. The domain-independent items are also likely to have contributed to the observed ceiling effect. Some subscores were not related to similar neuropsychological domains, indicating a lack of convergent validity, and many were also associated with unrelated neuropsychological domains, indicating a lack of discriminant validity. These findings are consistent with the results of Lam and colleagues (2013), who investigated the relationship of MoCA subscores to neuropsychological test performance and also found multiple intercorrelations. These cognitive screening tests do not provide valuable information about performance within specific cognitive domains and this emphasizes the important role of neuropsychological testing to describe specific areas of functioning, which can be essential for differential diagnosis. Overall, these findings militate against the practice of interpreting subscore performance; providers are encouraged to obtain a formal cognitive evaluation for descriptive information about specific neurocognitive functioning.

The cutoff values shown to be optimal in this study tended to favor specificity, which provides information about the ability of each test to correctly exclude a condition and suggests that follow-up testing is not necessary when a total score on the screening measure is above the cutoff. Although much attention is directed toward identifying impairment, being able to rule out cognitive difficulty and prevent unnecessary work-up is also important. As sensitivity was relatively deemphasized in cutoff selection and constituted small values, this reinforces the role of cognitive screening tests as indicators for more in-depth assessment

when a score below the cutoff is obtained, rather than tools for complex differential diagnosis. Although there is an increasing demand for all instruments to be sensitive to mild differences in the level of impairment, there are a number of psychometric limitations to cognitive screening tests (e.g., ceiling effects, limited point ranges) and some constraints (e.g., brevity of administration, appropriateness for a wide range of individuals) that can be avoided or minimized in more extensive neuropsychological testing. Furthermore, the attempt to discriminate between two groups implies that these groups are completely distinct (Grimes & Schulz, 2002), which is unlikely, as the definition of dementia is somewhat arbitrary and subject to continual revision. A cutoff at any value is likely to be imperfect, which implies the need for cautious interpretation, experienced clinical judgment, and utilization of other sources of information.

This study's design was intended to produce results that could be replicated in real-world settings. Known-groups designs, which compare groups that are expected to perform differently, are an efficacious initial method for establishing the concurrent validity of a measure. The existing research comparing the MMSE and MoCA has exclusively utilized predetermined groups. Known-group designs have inflated classification accuracy values secondary to the lack of atypical and difficult to classify individuals, as patients who are more likely to be misclassified are excluded (Sackett et al., 2000). Because the present study utilized a diverse clinical sample, classification sensitivity and specificity were smaller than in other studies. The lower classification values in this study probably approximate the values to be expected in actual clinical settings.

Limitations of the current study include characteristics of the sample. Although the availability of clinical data was an advantage for establishing external validity, it is

impossible for such results to be completely generalizable to all settings and samples. The sample in this study was selective, in that it consisted largely of male veterans and allowed inclusion of individuals with comorbid psychological and physical disorders, which may limit generalizability to other samples. However, the sample represented a broad range of ethnic diversity, included individuals of lower socioeconomic status, and encompassed a relatively broad age range of adults and older adults. Another limitation was that different patients had been administered a different set of neuropsychological tests. The test battery was altered as needed based on a variety of factors, such as the referral question and the patient's level of impairment. The ability to accurately assign clinical diagnoses may have been restricted by the neuropsychological and screening test performance that the raters disregarded. This was intended to guard against confounding our predictive results, although it may have hindered the raters' ability to accurately assign diagnoses as the sample largely consisted of atypical or complicated presentations. The diagnoses were also assigned to reflect the state of cognitive functioning at the time of testing, although there are not current diagnoses to illustrate individuals who progressed or remained stable regarding their cognitive functioning. Whereas the current sample was limited in some regards, the results continue to be relevant to other mixed clinical samples, particularly those with high base rates of dementia, because of the lack of predetermined groups.

Future work, outside the scope of this dissertation manuscript, will be necessary to describe other influential factors in how these screening tests can best be utilized and interpreted. For example, the present study removed individuals who performed below expectation on measures of performance validity. Providers utilizing the MMSE or MoCA may not be concurrently administering measures of performance validity, which could result

in ineffective utilization of resources when screening test results are positive due to invalidity. It may be possible to identify characteristics of screening test performance that differentiate individuals by performance validity, particularly in those subscore domains that are rarely failed. Furthermore, the present study evaluated classification accuracy based solely on cognitive screening performance, whereas a screening measure is intended to be one piece of data among several others (e.g., laboratory examinations, neuroimaging, physical examination, patient- and caregiver-report). Relative to a single test, multiple sources of information result in improved classification accuracy (Grimes & Schulz, 2002). It would be beneficial to understand how the use of a cognitive screening test increases predictive ability beyond that provided by other sources data that may be available. Furthermore, this study utilized a single cutoff to make a determination about cognitive impairment, although multiple cutoffs, representing clear versus questionable impairment, may be more descriptive for making clinically relevant decisions. If this specification is possible, it could improve appropriate referrals for follow-up neuropsychological testing when impairment is questionable and potentially save time and resources when impairment is well-established or performance is within normal limits on screening measures.

Overall, when evaluating the relative advantages and disadvantages of the MoCA and MMSE, the results showed more similarities than differences. For example, classification and predictive values were largely similar and both screening tests failed to demonstrate subscores that correlated well with neuropsychological testing domains. However, there are a few issues that differentiate the two measures. First, the MoCA possessed an advantage with its AUC when patients with mild neuropsychological impairment were included in the positive group and continued a trend for this with other criterion. Total scores on the MMSE

demonstrated a notable ceiling effect, which was not reflected to such a severe degree on the MoCA. Although the two screening measures are largely similar, these two clear advantages of the MoCA provide support for favoring its use over the MMSE. Many factors contribute to the practical utility of choosing one measure over another, and some of those factors are specific to each provider's setting and may be outside the scope of psychometric properties (e.g., cost, procedures of the institution, familiarity to referral sources). This study can provide information regarding the accuracy of the tests but, ultimately, it is each individual provider's responsibility to consider all the factors that may impact their use of these measures and to decide which test is preferable.

Several factors are presented in this study for clinicians to consider when selecting the MMSE versus the MoCA for dementia evaluation. This is the first study evaluating the MMSE and MoCA within a mixed clinical sample, rather than known-groups design, and to apply age- and education-adjustments to total scores on both screening measures. There was little difference in classification accuracy among the two cognitive screening measures, although the MoCA accounted for a greater AUC when utilizing mild impairment on neuropsychological testing as the criterion and demonstrated a less obvious ceiling effect in its total score distribution. Despite the demonstrated utility of these instruments, awareness of their limitations is also important. Neither test was shown to be sufficient to differentiate MCI from a mixed clinical sample or to diagnose dementia alone. Rather, each of these screening tests is one of several important tools that can be used to alert clinicians to the likelihood of impaired cognition, the possible need for specialized evaluation, and rule out gross cognitive impairment.

Table 1

Review of Existing Studies Investigating MMSE and MoCA Classification Accuracy

	MMSE			MoCA						
Authors	Sensitivity	Specificity	Cut (<)	Sensitivity	Specificity	Cut (<)	Derived or Fixed Cutoff?	Criterion	Sample	Age or Education Corrected?
Nasreddine et al. (2005)	0.78	1.00	26	1.00	0.87	26	MMSE fixed, MoCA derived,	Diagnosis (AD vs controls)	Canada - Probable AD (*n*=93), MCI (*n*=94), controls (*n*=90)	MMSE none, MoCA +1 if ≤12 years education
Smith, Gildeh, & Holmes (2007)	0.25	1.00	26	0.94	0.50	26	Fixed	Diagnosis (dementia vs controls)	United Kingdom – Dementia (*n*=32), MCI (*n*=23), psychiatric controls (*n*=12)	No
Hoops et al. (2009)	0.90	0.38	30	0.90	0.53	27	Derived	Diagnosis	USA - No cognitive disorder (*n*=92), MCI (*n*=23), PDD (*n*=17)	No
	0.92	0.42	30	0.86	0.72	26		Testing (≥2 of 4 domains ≤-1.5 *SD*)		
Luis et al. (2009)	0.36	0.96	25	0.97	0.35	27	Fixed	Diagnosis	USA - Probable AD (*n*=20), amnestic MCI (*n*=24), normal (*n*=74)	No
Dong et al. (2010)	0.86	0.82	25	0.90	0.77	22	Derived	Diagnosis	Singapore - Acute ischemic stroke or TIA (*n*=100)	MMSE none, MoCA +1 if ≤12 years education

Table 1 (Continued)

Authors	MMSE			MoCA			Derived or Fixed Cutoff?	Criterion	Sample	Age or Education Corrected?
	Sensitivity	Specificity	Cut (<)	Sensitivity	Specificity	Cut (<)				
Mickes et al. (2010)	0.85	0.32	26	0.98	0.30	26	Fixed	Diagnosis	USA - Mild to moderate Huntington's Disease (*n*=39), controls (*n*=73)	No
Larner (2011)	0.65	0.89	26	0.97	0.60	26	Fixed	Diagnosis	United Kingdom – Dementia (n=36), MCI (n=29), no dementia (n=85)	No
Freitas et al. (2012a)	0.58	0.88	26	0.78	0.98	17	Derived	Diagnosis	Portugal - Behavioral-variant FTD (*n*=50), AD (*n*=50), control (*n*=50)	No
Freitas et al. (2012b)	0.62	0.94	26	0.77	0.97	17	Derived	Diagnosis	Portugal – Probable AD (*n*=34), vascular dementia (*n*=34), control (*n*=34)	MMSE none, MoCA education (norms from Freitas Simões, Alves, & Santana, 2011)

Table 1 (Continued)

Authors	MMSE Sensitivity	MMSE Specificity	MMSE Cut (<)	MoCA Sensitivity	MoCA Specificity	MoCA Cut (<)	Derived or Fixed Cutoff?	Criterion	Sample	Age or Education Corrected?
Tsai et al. (2012)	0.95	0.98	26	0.98	0.95	22	Derived	Diagnosis	China-Taiwan – AD (*n*=98), MCI (*n*=71), controls (*n*=38)	No
Whitney et al. (2012)	0.52	0.77	27	0.86	0.57	26	Fixed	Testing (≥2 of 4 domains ≤-1.5 *SD*)	USA veterans - Cognitive impairment (*n*=29), controls (*n*=53)	No
				0.72	0.75	24				
Dong et al. (2013)	0.87	0.80	26	0.80	0.92	20	Derived	Diagnosis (MCI vs controls)	Singapore – No cognitive impairment (*n*=128), MCI (*n*=130)	No
Freitas, Simões, Alves, & Santana (2013)	0.85	0.93	26	0.88	0.98	17	Derived	Diagnosis (AD vs controls)	Portugal – MCI (*n*=90), AD (*n*=90), matched controls (*n*=180)	No
Roalf et al. (2013)	0.96	0.97	28	0.94	0.96	23	Derived	Diagnosis (AD vs controls)	USA - AD (*n*=321), MCI (*n*=126), controls (*n*=140)	No

Table 2

Participant Characteristics and Diagnoses

Participant Characteristics	
Age (years)	*M*=66.5 (*SD*=11.3)
Gender	
Male	*n*=204 (94%)
Female	*n*=14 (6%)
Education (years)	*M*=13.1 (*SD*=2.6)
Estimated premorbid intelligence	*M*=95.4 (*SD*=12.3)
Ethnicity	
Caucasian	*n*=138 (63%)
African-American	*n*=56 (26%)
Hispanic	*n*=22 (10%)
Other	*n*=2 (1%)
Diagnoses	
Dementia	*n*=63 (29%)
Vascular Dementia	*n*=25 (11%)
Alzheimer's Dementia	*n*=15 (7%)
Mixed Dementia	*n*=9 (4%)
Frontotemporal Dementia	*n*=5 (2%)
Lewy Body Dementia	*n*=4 (2%)
Corticobasal Degeneration	*n*=3 (1%)
Huntington's Disease	*n*=1 (<1%)
Parkinson's Disease Dementia	*n*=1 (<1%)
MCI	*n*=70 (32%)
Amnestic MCI	n=33 (15%)
Non-amnestic MCI	n=37 (17%)
Other	n=85 (39%)
Neurological Disorders	n=37 (17%)
Stroke	n=5 (2%)
Epilepsy	n=4 (2%)
HIV	n=4 (2%)
Multiple Sclerosis	n=4 (2%)
TBI	n=3 (1%)
Parkinson's Disease	n=2 (1%)
Other Neurological	n=15 (7%)
Psychiatric Disorder	n=34 (16%)
No Diagnosis	n=14 (6%)

Table 3

Descriptives of Neuropsychological Measures in Dementia and Non-dementia Groups

	Dementia (*n*=63)		Non-dementia (*n*=155)	
Neuropsychological Domains and Measures	Mean *z*-score (*SD*)	*n*	Mean *z*-score (*SD*)	*n*
Language	-0.72 (0.99)	55	-0.37 (0.89)	144
Boston Naming Test	-0.46 (1.06)	47	-0.10 (1.13)	138
WAIS-III Similarities	-0.53 (1.02)	12	-0.19 (0.84)	42
WTAR	-0.90 (1.08)	47	-0.60 (1.04)	131
Verbal Memory	-0.84 (0.85)	46	-0.59 (0.91)	140
CVLT-II Trials 1-5 Total	-0.90 (0.94)	45	-0.53 (1.07)	138
CVLT-II d'	-0.78 (1.00)	44	-0.67 (1.03)	138
WMS-III Logical Memory I	-0.75 (0.86)	32	-0.50 (1.08)	92
WMS-III Logical Memory II	-0.41 (1.07)	32	-0.38 (1.19)	92
Attention and Processing Speed	-1.25 (1.00)	54	-0.84 (0.82)	147
Trail Making Test A	-1.44 (1.15)	53	-0.86 (1.03)	141
WAIS-III Digit Span	-0.60 (0.79)	15	-0.51 (0.86)	44
WAIS-III Symbol Search	-2.17 (0.24)	2	-1.44 (0.19)	3
WAIS-III Digit-Symbol Coding	-1.28 (1.22)	13	-0.67 (0.94)	44
WAIS-IV Symbol Search	-1.04 (1.05)	29	-0.78 (0.96)	93
WAIS-IV Coding	-1.15 (1.05)	28	-0.84 (0.81)	92
Visuospatial	-1.12 (1.06)	46	-0.51 (1.14)	139
WAIS-III Block Design	-1.47 (0.77)	15	-0.32 (1.17)	44
WAIS-IV Block Design	-0.82 (0.94)	29	-0.32 (0.94)	91
Judgment of Line Orientation	-1.11 (1.58)	25	-1.04 (1.47)	78
Executive Functioning	-1.23 (0.86)	59	-0.74 (0.91)	152
Trail Making Test B	-1.43 (1.30)	49	-0.97 (1.19)	138
WAIS-IV Digit Span	-0.86 (0.91)	31	-0.79 (0.81)	97
WAIS-IV Visual Puzzles	-0.12 (1.10)	14	-0.02 (1.25)	41
FAS phonemic fluency	-1.24 (1.06)	57	-0.74 (1.20)	144
Animal Naming	-1.34 (1.12)	57	-0.63 (1.23)	143

Note. CVLT=California Verbal Learning Test; WAIS=Wechsler Adult Intelligence Scale; WMS=Wechsler Memory Scales; WTAR=Wechsler Test of Adult Reading.

Table 4

Frequency Distribution and Description of MMSE Subscores

MMSE Domains	Frequency			Mean	Median	Mode	Skewness	Kurtosis
	Value	*n*	%					
Orientation to Time	0	2	1	4.42	5.00	5.00	-2.03	4.19
	1	2	1					
	2	11	5					
	3	16	7					
	4	43	20					
	5	144	66					
Orientation to Place	0	0	0	4.88	5.00	5.00	-4.40	22.38
	1	0	0					
	2	2	1					
	3	2	1					
	4	16	7					
	5	198	91					
Registration	0	1	<1	2.96	3.00	3.00	-8.93	93.64
	1	0	0					
	2	5	2					
	3	212	97					
Serial 7's	0	13	6	3.33	4.00	5.00	-0.61	-0.80
	1	23	11					
	2	27	13					
	3	34	16					
	4	48	22					
	5	67	32					
WORLD	0	5	5	3.77	5.00	5.00	-0.96	-0.35
	1	6	6					
	2	13	13					
	3	13	13					
	4	8	8					
	5	54	55					
Delayed Recall	0	37	17	1.79	2.00	3.00	-0.35	-1.22
	1	48	22					
	2	57	26					
	3	76	35					
Naming	0	0	0	1.98	2.00	2.00	-7.23	50.70
	1	4	2					
	2	214	98					
Repetition	0	30	14	0.86	1.00	1.00	-2.12	2.51
	1	188	86					

Table 4 (Continued)

MMSE Domains	Frequency			Mean	Median	Mode	Skewness	Kurtosis
	Value	*n*	%					
Compre-hension	0	0	0	2.88	3.00	3.00	-2.30	3.32
	1	0	0					
	2	27	12					
	3	191	88					
Reading	0	3	1	0.99	1.00	1.00	-8.41	69.29
	1	215	99					
Writing	0	12	6	0.94	1.00	1.00	-3.93	13.56
	1	206	95					
Drawing	0	48	22	0.78	1.00	1.00	-1.35	-0.17
	1	169	78					

Table 5

Frequency Distribution and Description of MoCA Subscores

MoCA Domains	Frequency			Mean	Median	Mode	Skewness	Kurtosis
	Value	*n*	%					
Visuospatial/ Executive	0	6	3	3.43	4.00	4.00	-0.65	-0.47
	1	16	8					
	2	28	14					
	3	38	19					
	4	55	28					
	5	54	27					
Naming	0	2	1	2.73	3.00	3.00	-2.33	6.08
	1	5	3					
	2	37	19					
	3	152	78					
Attention	0	2	1	4.46	5.00	6.00	-0.79	-0.29
	1	8	4					
	2	17	9					
	3	24	12					
	4	35	18					
	5	42	21					
	6	69	35					
Language	0	16	8	2.04	2.00	3.00	-0.66	-0.53
	1	36	18					
	2	70	36					
	3	75	38					
Abstraction	0	35	16	1.45	2.00	2.00	-0.99	-0.64
	1	38	19					
	2	124	63					
Delayed Recall	0	56	28	1.91	2.00	0	0.34	-1.00
	1	28	14					
	2	43	22					
	3	33	17					
	4	22	11					
	5	15	8					
Orientation	0	1	<1	5.53	6.00	6.00	-2.79	9.02
	1	2	1					
	2	1	<1					
	3	6	3					
	4	14	7					
	5	26	13					
	6	147	75					

Table 6

Classification Accuracy Using MMSE Adjusted Total Scores and Three Criterion Variables

Criterion	Cutoff (<)	Sensitivity	Specificity
≥1 domain at < -1*SD*	14	0	1.00
	15	<.01	1.00
	16	.01	1.00
	17	.03	.96
	18	.04	.96
	19	.05	.96
	20	.06	.95
	21	.08	.93
	22	.13	.93
	23	.15	.91
	24	.23	.86
	25[a]	.32	.84
	26	.42	.66
	27	.49	.57
	28	.65	.46
	29	.78	.34
	30	.87	.25
	=30	1.00	0
≥1 domain at < -2*SD*	14	0	1.00
	15	.02	1.00
	16	.03	1.00
	17	.10	.99
	18	.10	.99
	19	.12	.97
	20	.14	.97
	21	.17	.95
	22	.20	.93
	23	.25	.91
	24	.36	.85
	25[a]	.44	.78
	26	.53	.65
	27	.59	.57

Table 6 (Continued)

Criterion	Cutoff (<)	Sensitivity	Specificity
≥1 domain at < -2*SD*	28	.75	.42
	29	.86	.29
	30	.90	.18
	=30	1.00	0
Dementia versus Other	12	0	1.00
	14	.02	1.00
	15	.03	1.00
	16	.05	1.00
	17	.13	1.00
	18	.13	.99
	19	.16	.99
	20	.18	.98
	21	.21	.97
	22	.27	.96
	23	.29	.92
	24[a]	.40	.87
	25	.47	.80
	26	.53	.66
	27	.65	.59
	28	.82	.46
	29	.92	.31
	30	.95	.19
	=30	1.00	0

[a] Represents the optimal cutoff score given the characteristics of the criterion variable.

Table 7

Areas Under the Curve, Likelihood Ratios, Pre- and Post-test Probabilities, and Predictive Values for Selected Criteria and Base Rates

					Sample Base Rate			Base Rate = 0.10			Base Rate = 0.35			Base Rate = 0.60		
Criterion	Cutoff (<)	AUC	+LR	-LR	Pre-TP	Post-TP/PPV	NPV	Pre-TP	Post-TP/PPV	NPV	Pre-TP	Post-TP/PPV	NPV	Pre-TP	Post-TP/PPV	NPV
MMSE																
≥1 domain at < -1*SD*	25	.59	1.99	1.23	0.73	0.84	0.31	0.10	0.18	0.92	0.35	0.52	0.70	0.60	0.75	0.45
≥1 domain at < -2*SD*	25	.64	2.02	1.40	0.29	0.45	0.77	0.10	0.18	0.93	0.35	0.52	0.72	0.60	0.75	0.48
Dementia vs Other	24	.70	2.99	1.45	0.29	0.55	0.78	0.10	0.25	0.93	0.35	0.62	0.73	0.60	0.82	0.49
MoCA																
≥1 domain at < -1*SD*	25	.69	1.71	2.14	0.73	0.82	0.44	0.10	0.16	0.95	0.35	0.48	0.80	0.60	0.72	0.59
≥1 domain at < -2*SD*	21	.72	2.56	1.77	0.29	0.51	0.81	0.10	0.22	0.94	0.35	0.58	0.77	0.60	0.79	0.54
Dementia vs Other	20	.69	2.51	1.46	0.29	0.51	0.78	0.10	0.22	0.93	0.35	0.58	0.73	0.60	0.79	0.49

Note. +LR=positive likelihood ratio, -LR=negative likelihood ratio, AUC=area under the curve, NPV=negative predictive value, Post-TP=post-test probability, PPV=positive predictive value, Pre-TP=pre-test probability.

Table 8

Classification Accuracy Using MoCA Adjusted Total Scores and Three Criterion Variables

Criterion	Cutoff (<)	Sensitivity	Specificity
≥1 domain at < -1*SD*	8	0	1.00
	10	<.01	1.00
	12	.01	1.00
	13	.03	.98
	14	.04	.98
	15	.07	.98
	16	.09	.96
	17	.11	.93
	18	.15	.89
	19	.20	.89
	20	.28	.86
	21	.37	.82
	22	.47	.77
	23	.54	.68
	24	.63	.63
	25[a]	.73	.57
	26	.84	.45
	27	.92	.38
	28	.97	.20
	29	.98	.13
	30	.99	.05
	=30	1.00	0
≥1 domain at < -2*SD*	8	0	1.00
	10	.02	1.00
	12	.03	1.00
	13	.07	.99
	14	.10	.99
	15	.12	.97
	16	.15	.95
	17	.17	.93
	18	.25	.90
	19	.32	.88

Table 8 (Continued)

Criterion	Cutoff (<)	Sensitivity	Specificity
≥1 domain at < -2*SD*	20	.48	.85
	21[a]	.56	.78
	22	.64	.69
	23	.71	.62
	24	.80	.54
	25	.85	.44
	26	.90	.30
	27	.93	.20
	28	.97	.10
	29	1.00	.07
	30	1.00	.03
	=30	1.00	0
Dementia versus Other	6	0	1.00
	8	.02	1.00
	10	.03	1.00
	12	.05	1.00
	13	.10	.99
	14	.13	.99
	15	.18	.99
	16	.21	.97
	17	.24	.96
	18	.31	.91
	19	.32	.87
	20[a]	.44	.83
	21	.47	.73
	22	.60	.65
	23	.66	.56
	24	.76	.49
	25	.84	.40
	26	.89	.28
	27	.94	.19
	28	.98	.10
	29	1.00	.06
	30	1.00	.03
	=30	1.00	0

[a] Represents the optimal cutoff score given the characteristics of the criterion variable.

Table 9

MMSE and MoCA Subscore Mean Ranks for Dementia Patients and Non-dementia Patients

Screening Test	Subscore Domain	Dementia Group Mean (*SD*)	Non-Dementia Group Mean (*SD*)	*p* [a]
MMSE	Orientation to Time	3.90 (1.40)	4.63 (0.69)	<0.001
	Orientation to Place	4.82 (0.53)	4.90 (0.37)	0.22
	Registration	2.98 (0.13)	2.96 (0.29)	0.52
	Attention/Concentration	3.53 (1.54)	4.26 (1.04)	<0.01
	Recall	1.55 (1.10)	1.88 (1.09)	0.04
	Naming	1.95 (0.22)	1.99 (0.08)	0.04
	Repetition	0.81 (0.40)	0.88 (0.32)	0.13
	Comprehension	2.92 (0.28)	2.86 (0.35)	0.22
	Reading	0.98 (0.13)	0.99 (0.11)	0.85
	Writing	0.89 (0.32)	0.97 (0.18)	0.02
	Drawing	0.62 (0.49)	0.84 (0.37)	<0.01
	Total Adjusted Score	23.77 (4.44)	26.69 (2.81)	<0.001
MoCA	Visuospatial/Executive	2.63 (1.52)	3.78 (1.18)	<0.001
	Naming	2.61 (0.64)	2.78 (0.51)	0.03
	Attention	4.22 (1.50)	4.56 (1.57)	0.08
	Language	1.75 (0.90)	2.16 (0.94)	<0.01
	Abstraction	1.29 (0.81)	1.52 (0.76)	0.03
	Delayed Recall	1.61 (1.63)	2.04 (1.59)	0.08
	Orientation	5.08 (1.43)	5.72 (0.66)	<0.001
	Total Adjusted Score	19.76 (4.98)	23.02 (3.77)	<0.001

[a] Probability levels from Mann-Whitney U test.

Table 10

Spearman Correlations of Neuropsychological Composite Domains and MMSE Subscores Using all Participants

MMSE Domains	Neuropsychological Domains				
	Processing Speed/ Attention	Executive Functioning	Memory	Language	Visuospatial Functioning
Orientation to Time	0.26 (p<.001)	0.36 (p<.001)	0.32 (p<.001)	0.29 (p<.001)	0.24 (p=.001)
Orientation to Place	0.16 (p=.03)	0.29 (p<.001)	0.15 (p=.04)	0.21 (p<.01)	0.11 (p=.14)
Registration	0.12 (p=.08)	0.16 (p=.02)	0.13 (p=.08)	0.12 (p=.11)	0.18 (p=.02)
Attention/ Concentration	0.11 (p=.12)	0.20 (p<.01)	0.21 (p=.01)	0.17 (p=.01)	0.11 (p=.13)
Recall	0.18 (p=.01)	0.23 (p=.001)	0.44 (p<.001)	0.12 (p=.10)	0.11 (p=.14)
Naming	-0.09 (p=.19)	0.00 (p=.99)	-0.08 (p=.25)	0.04 (p=.61)	-0.01 (p=.94)
Repetition	0.21 (p<.01)	0.27 (p<.001)	0.17 (p=.02)	0.16 (p=.02)	0.19 (p=.01)
Comprehension	0.13 (p=.07)	0.10 (p=.15)	0.09 (p=.23)	0.00 (p=.96)	0.03 (p=.67)
Reading	0.11 (p=.14)	0.06 (p=.38)	0.03 (p=.66)	0.14 (p=.05)	0.10 (p=.17)
Writing	0.03 (p=.70)	0.06 (p=.43)	0.09 (p=.20)	0.06 (p=.41)	0.02 (p=.76)
Drawing	0.33 (p<.001)	0.31 (p<.001)	0.27 (p<.001)	0.25 (p<.001)	0.30 (p<.001)

Table 11

Spearman Correlations of Neuropsychological Composite Domains and MoCA Subscores Using all Participants

	Neuropsychological Domains				
MoCA Domains	Processing Speed/ Attention	Executive Functioning	Memory	Language	Visuospatial Functioning
Visuospatial/ Executive	0.41 (p<.001)	0.37 (p<.001)	0.29 (p<.001)	0.28 (p<.001)	0.43 (p<.001)
Naming	0.16 (p=.04)	0.25 (p<.001)	0.08 (p=.29)	0.32 (p<.001)	0.10 (p=.20)
Attention	0.31 (p<.001)	0.36 (p<.001)	0.21 (p<.01)	0.32 (p<.001)	0.22 (p<.01)
Language	0.33 (p<.001)	0.40 (p<.001)	0.22 (p<.01)	0.17 (p=.03)	0.28 (p<.001)
Abstraction	0.20 (p<.01)	0.19 (p<.01)	0.15 (p=.06)	0.26 (p<.001)	0.13 (p=.09)
Delayed Recall	0.12 (p=.10)	0.21 (p<.01)	0.50 (p<.001)	0.06 (p=.42)	0.06 (p=.47)
Orientation	0.12 (p=.11)	0.24 (p=.001)	0.28 (p<.001)	0.23 (p<.01)	0.02 (p=.81)

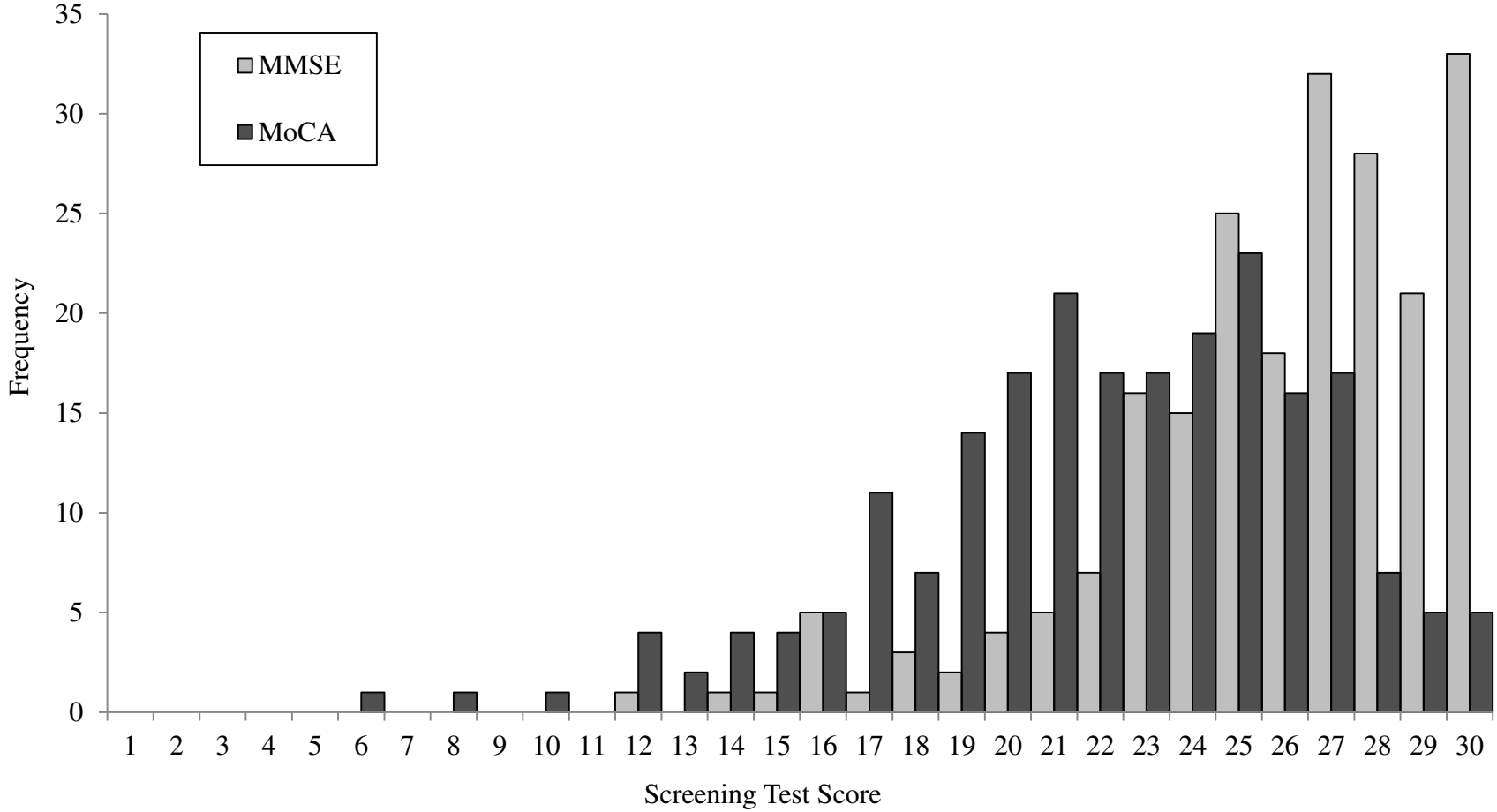


Figure 1. Histogram of MMSE and MoCA adjusted total scores.

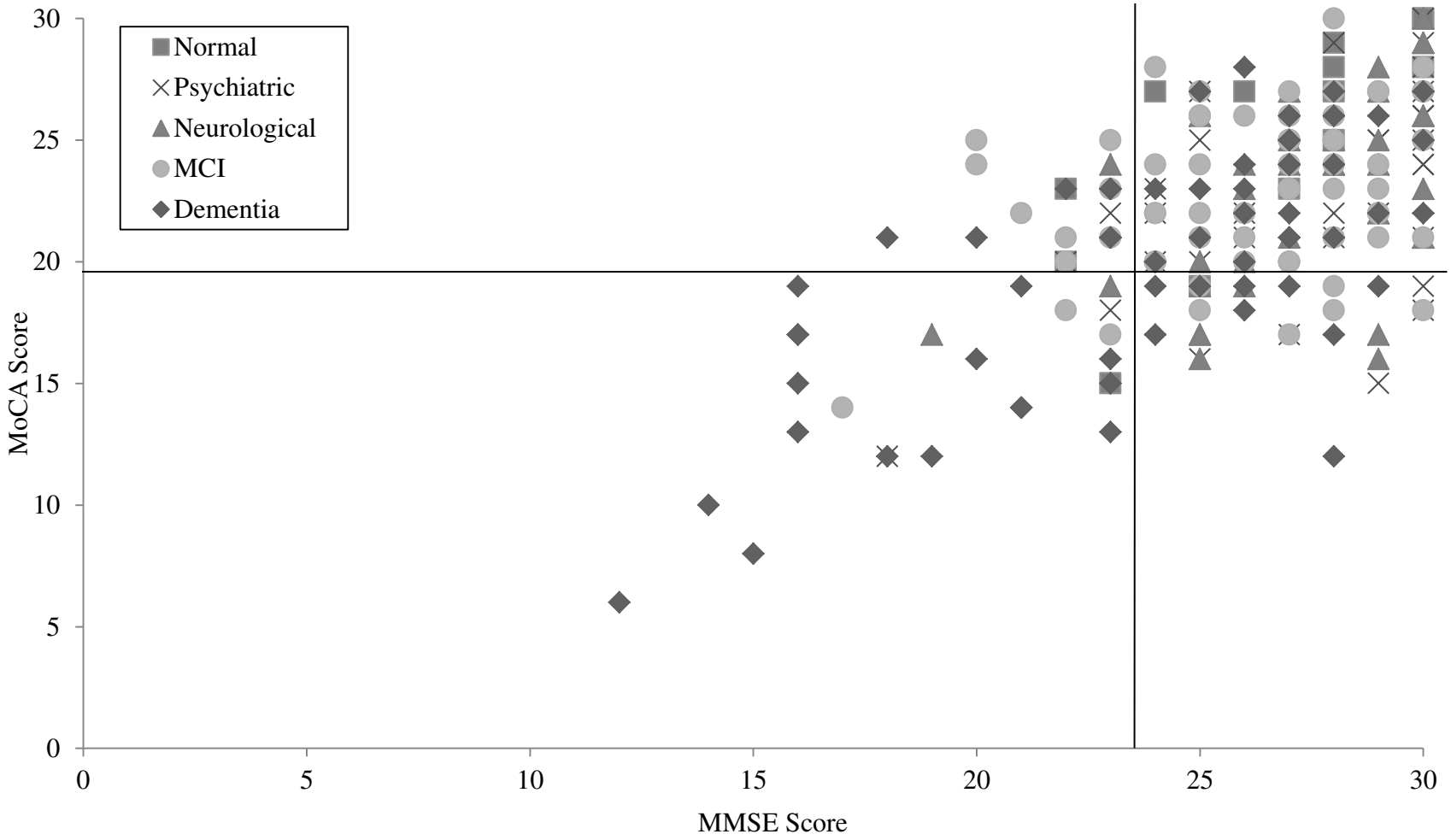


Figure 2. Scatterplot of MMSE and MoCA scores among five diagnostic groups: individuals with dementia, MCI, a neurological disorder, a psychiatric disorder, and "normal" (i.e., no diagnosis) patients. Screening instrument cutoffs, derived with diagnosis as the criterion, are marked with lines to represent a MMSE cutoff at less than 24 and a MoCA cutoff at less than 20.

Appendix A

Copy of the Original MMSE

Patient. . . . . .
Examiner . . . . . .
Date . . . . . , ...

"MINI-MENTAL STATE"

ORIENTATION

5 ( ) What is the (year) (season) (date) (day) (month)?
5 ( ) Where are we: (state) (county) (town) (hospital) (floor).

REGISTRATION

3 ( ) Name 3 objects: 1 second to say each. Then ask the patient all 3 after you have said them. Give 1 point for each correct answer. Then repeat them until he learns all 3. Count trials and record trials.

ATTENTION AND CALCULATION

5 ( ) Serial 7's. 1 point for each correct. Stop after 5 answers. Alternatively spell "world" backwards.

RECALL

3 ( ) Ask for the 3 objects repeated above. Give 1 point for each correct.

LANGUAGE

9 ( ) Name a pencil, and watch (2 points)
Repeat the following "No ifs, ands or buts." (1 point)
Follow a 3-stage command: "Take a paper in your right hand, fold it in half, and put it on the floor" (3 points)
Read and obey the following: CLOSE YOUR EYES (1 point)
Write a sentence (1 point)
Copy design (1 point)

_____Total score

ASSESS level of consciousness along a continuum ----------

Alert Drowsy Stupor Coma

INSTRUCTIONS FOR ADMINISTRATION OF
MINI-MENTAL STATE EXAMINATION

ORIENTATION

(1) Ask for the date. Then ask specifically for parts omitted, e.g., "Can you also tell me what season it is?" One point for each correct.

Appendix A (Continued)

(2) Ask in turn "Can you tell me the name of this hospital?" (town, county, etc.). One point for each correct.

REGISTRATION

Ask the patient if you may test his memory. Then say the names of 3 unrelated objects, clearly and slowly, about one second for each. After you have said all 3, ask him to repeat them. This first repetition determines his score (O-3) but keep saying them until he can repeat all 3, up to 6 trials. If he does not eventually learn all 3, recall cannot be meaningfully tested.

ATTENTION AND CALCULATION

Ask the patient to begin with 100 and count backwards by 7. Stop after 5 subtractions (93, 86,79,72,65). Score the total number of correct answers. If the patient cannot or will not perform this task, ask him to spell the word "world" backwards. The score is the number of letters in correct order. E.g. dlrow = 5, dlorw = 3.

RECALL

Ask the patient if he can recall the 3 words you previously asked him to remember. Score O-3.

LANGUAGE

Naming*:* Show the patient a wrist watch and ask him what it is. Repeat for pencil. Score O-2.
Repetition: Ask the patient to repeat the sentence after you. Allow only one trial. Score 0 or 1.
3-Stage command*:* Give the patient a piece of plain blank paper and repeat the command. Score 1 point for each part correctly executed.
Reading: On a blank piece of paper print the sentence "Close your eyes", in letters large enough for the patient to see clearly. Ask him to read it and do what it says. Score 1 point only if he actually closes his eyes.
Writing: Give the patient a blank piece of paper and ask him to write a sentence for you. Do not dictate a sentence, it is to be written spontaneously. It must contain a subject and verb and be sensible. Correct grammar and punctuation are not necessary.
Copying: On a clean piece of paper, draw intersecting pentagons, each side about 1 in., and ask him to copy it exactly as it is. All 10 angles must be present and 2 must intersect to score 1 point. Tremor and rotation are ignored. Estimate the patient's level of sensorium along a continuum, from alert on the left to coma on the right.

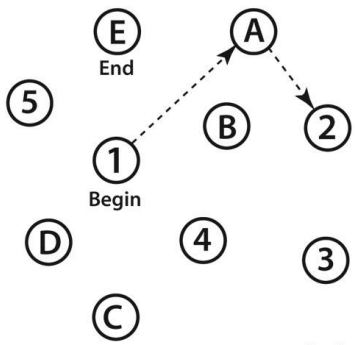
Appendix B

Copy of the MoCA

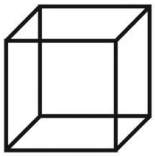
MONTREAL COGNITIVE ASSESSMENT (MOCA)
Version 7.1 Original Version

NAME :
Education : Date of birth :
Sex : DATE :

VISUOSPATIAL / EXECUTIVE — POINTS



Copy cube



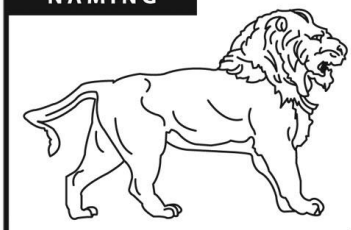
[ ] [ ]

Draw CLOCK (Ten past eleven) (3 points)

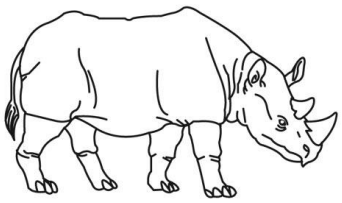
[ ] Contour [ ] Numbers [ ] Hands

__/5

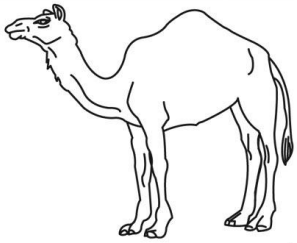
NAMING



[ ]



[ ]



[ ]

__/3

MEMORY Read list of words, subject must repeat them. Do 2 trials, even if 1st trial is successful. Do a recall after 5 minutes.

	FACE	VELVET	CHURCH	DAISY	RED
1st trial					
2nd trial					

No points

ATTENTION Read list of digits (1 digit/ sec.).

Subject has to repeat them in the forward order [ ] 2 1 8 5 4
Subject has to repeat them in the backward order [ ] 7 4 2

__/2

Read list of letters. The subject must tap with his hand at each letter A. No points if ≥ 2 errors

[ ] F B A C M N A A J K L B A F A K D E A A A J A M O F A A B

__/1

Serial 7 subtraction starting at 100 [ ] 93 [ ] 86 [ ] 79 [ ] 72 [ ] 65

4 or 5 correct subtractions: **3 pts**, 2 or 3 correct: **2 pts**, 1 correct: **1 pt**, 0 correct: **0 pt**

__/3

LANGUAGE Repeat : I only know that John is the one to help today. [ ]
The cat always hid under the couch when dogs were in the room. [ ]

__/2

Fluency / Name maximum number of words in one minute that begin with the letter F [ ] _____ (N ≥ 11 words)

__/1

ABSTRACTION Similarity between e.g. banana - orange = fruit [ ] train – bicycle [ ] watch - ruler

__/2

DELAYED RECALL

	Has to recall words **WITH NO CUE**	FACE [ ]	VELVET [ ]	CHURCH [ ]	DAISY [ ]	RED [ ]	Points for UNCUED recall only
Optional	Category cue						
	Multiple choice cue						

__/5

ORIENTATION [ ] Date [ ] Month [ ] Year [ ] Day [ ] Place [ ] City

__/6

© Z.Nasreddine MD www.mocatest.org Normal ≥ 26 / 30

TOTAL __/30

Add 1 point if ≤ 12 yr edu

Administered by: _____

Appendix C

Formulas to Calculate Classification Accuracy

		Criterion Measure	
		Positive (Impaired)	Negative (Not impaired)
Screening Measure	Positive (Impaired)	True Positive	False Positive
	Negative (Not impaired)	False Negative	True Negative

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}}$$

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{False Positives+True Negatives}}$$

$$\text{Positive Predictive Value} = \frac{\text{Base rate*Sensitivity}}{(\text{Base rate*Sensitivity})+\big((\text{1-Base rate})(\text{1-Specificity})\big)}$$

$$\text{Negative Predictive Value} = \frac{(\text{1-Base rate})*\text{Specificity}}{\big((\text{1-Base rate})*\text{Specificity}\big)+\big(\text{Base rate}*(\text{1-Sensitivity})\big)}$$

$$\text{Positive Likelihood Ratio} = \frac{\text{Sensitivity}}{\text{1-Specificity}}$$

$$\text{Negative Likelihood Ratio} = \frac{\text{Specificity}}{\text{1-Sensitivity}}$$

$$\text{Pretest Probability} = \frac{\text{Subjects positive on criterion}}{\text{Total subjects}}$$

$$\text{Posttest Probability} = \frac{\left[\left(\frac{\text{Base rate}}{\text{1-Base rate}}\right)*\text{Likelihood ratio}\right]}{\left[\left(\frac{\text{Base rate}}{\text{1-Base rate}}*\text{Likelihood ratio}\right)+1\right]}$$

Hanley and McNeil (1982, 1983): $z = \frac{AUC_1 - AUC_2}{\sqrt{SE_1^2 + SE_2^2 - 2rSE_1SE_2}}$

Appendix C (Continued)

r is the estimated correlation between the two areas under the curve, AUC1 and AUC2. It can be converted using Table 1 of Hanley and McNeil's (1983) manuscript with the average of the two AUCs and the average of the two correlation coefficients for ratings given to each patient group (i.e., positive and negative classification) by the two modalities.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., … Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging and Alzheimer's Association workgroup. *Alzheimer's and Dementia*, 1-10. doi: 10.1016/j.jalz.2011.03.008

Alzheimer's Association. (2012). 2012 Alzheimer's disease facts and figures. *Alzheimer's and Dementia: The Journal of the Alzheimer's Association*, *8*(2), 131-168. doi: 10.1016/j.jalz.2012.02.001

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Benton, A. L., Hamsher, K., & Sivan, A. B. (1994). Controlled Oral Word Association test. In *Multilingual Aphasia Examination* (3rd ed.). Iowa City, IA: AJA Associates.

Benton, A. L., Sivan, A. B., Hamsher, K., Varney, N. R., & Spreen, O. (1994). *Contributions to neuropsychological assessment* (2nd ed.). Orlando, FL: Psychological Assessment Resources.

Boyle, P. A., Wilson, R. S., Aggarwal, N. T., Arvanitakis, Z., Kelly, J., Bienias, J. L., & Bennett, D. A. (2005). Parkinsonian signs in subjects with mild cognitive impairment. *Neurology, 65*(12), 1901-1906. doi: 10.1212/01.wnl.0000190262.59672.4c

Bravo, G., & Hebert, R. (1997). Age- and education-specific reference values for the Mini-Mental and Modified Mini-Mental State Examinations derived from a non-demented elderly population. *International Journal of Geriatric Psychiatry, 12*, 1008-1018. doi: 10.1002/(SICI)1099-1166(199710)12:10<1008::AID-GPS676>3.0.CO;2-A

Busse, A., Angermeyer, M. C., & Riedel-Heller, S. G. (2006). Progression of mild cognitive impairment to dementia: A challenge to current thinking. *The British Journal of Psychiatry, 189*, 399-404. doi: 10.1192/bjp.bp.105.014779

Conti, M. Z., Vicini-Chilovi, B., Riva, M., Zanetti, M., Liberini, P., Padovani, A., & Rozzini, L. (2013). Odor identification deficit predicts clinical conversion from mild cognitive impairment to dementia due to Alzheimer's disease. *Archives of Clinical Neuropsychology, 28*(5), 391-399. doi: 10.1093/arclin/act032

Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2006). A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry, 78,* 790-799. doi: 10.1136/jnnp.2006095414

Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. *Journal of the American Medical Association, 269*(18), 2386-2391. doi: 10.1001/jama.1993.03500180078038

Delis, D. C., Kramer J. H., Kaplan E., & Ober, B. A. (2000). *California Verbal Learning Test (CVLT-II)* (2nd ed.). San Antonio, TX: The Psychology Corporation.

Dong, Y., Sharma, V. K., Chan, B. P., Venketasubramanian, N., Teoh, H. L., Seet, R. C. S., … Chen, C. (2010). The Montreal Cognitive Assessment (MoCA) is superior to the Mini-Mental State Examination (MMSE) for the detection of vascular cognitive impairment after acute stroke. *Journal of the Neurological Sciences, 299*, 15-18. doi: 10.1016/j.jns.2010.08.051

Dong, Y., Lee, W. Y., Hilal, S., Saini, M., Wong, T. Y., Chen, C. L., … Ikram, M. K. (2013). Comparison of the Montreal Cognitive Assessment and the Mini-Mental State

Examination in detecting multi-domain mild cognitive impairment in a Chinese sub-sample drawn from a population-based study. *International Psychogeriatrics, 25*(11), 1831-1838. doi: 10.1017/S1041610213001129

Eslinger, P. J., Damasio, A. R., Benton, A. L., & Van Allen, M. (1985). Neuropsychologic detection of abnormal mental decline in older persons. *Journal of the American Medical Association, 253*, 670-674.

Espinosa, A., Alegret, M., Valero, S., Vinyes-Junqué, G., Hernández, I., Mauleon, A., … Boada, M. (2013). A longitudinal follow-up of 550 mild cognitive impairment patients: Evidence for large conversion to dementia rates and detection of major risk factors involved. *Journal of Alzheimer's Disease, 34*(3), 769-780. doi: 10.3233/JAD-122002

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State – A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198. doi: 10.1016/0022-3956(75)90026-6

Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2011). Montreal Cognitive Assessment (MoCA): Normative study for the Portuguese population. *Journal of Clinical and Experimental Neuropsychology, 33*(9), 989-996. doi: 10.1080/13803395.2011.589374

Freitas, S., Simões, M., Alves, L., Duro, D., & Santana, I. (2012a). Montreal Cognitive Assessment (MoCA): Validation study for frontotemporal dementia. *Journal of Geriatric Psychiatry and Neurology, 25*(3), 146-154. doi: 10.1177/0891988712455235

Freitas, S., Simões, M., Alves, L., Vincente, M., & Santana, I. (2012b). Montreal Cognitive Assessment (MoCA): Validation study for vascular dementia. *Journal of the*

International Neuropsychological Society, 18, 1031-1040. doi: 10.1017/S135561771200077X

Freitas, S., Simões, M., Alves, L., & Santana, I. (2013). Montreal Cognitive Assessment: Validation study for mild cognitive impairment and Alzheimer disease. *Alzheimer Disease and Associated Disorders, 27*(1), 37-43. doi: 10.1097/WAD.0b013e3182420bfe

Ganguli, M., Ratcliff, G., Huff, F. J., Belle, S., Kancel, M. J., Fischer, L., & Kuller, L. H. (1990). Serial sevens versus world backwards: A comparison of the two measures of attention from the MMSE. *Journal of Geriatric Psychiatry and Neurology, 3*(4), 203-207. doi: 10.1177/089198879000300405

Gill, D. J., Freshman, A., Blender, J. A., & Ravina, B. (2008). The Montreal Cognitive Assessment as a screening tool for cognitive impairment in Parkinson's disease. *Movement Disorders, 23*, 1043-1046. doi: 1010.1002/mds.22017

Godefroy, O., Fickl, A., Roussel, M., Auribault, C., Bugnicourt, J. M., Lamy, C., … Petitnicolas, G. (2011). Is the Montreal Cognitive Assessment superior to the Mini-Mental State Examination to detect poststroke cognitive impairment?: A study with neuropsychological evaluation. *Stroke, 42,* 1712-1716. doi: 10.1161/STROKEAHA.110.606277

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders*. Philadelphia, PA: Lee and Febiger.

Green, P. (2003, revised 2005). *Word Memory Test for Windows: User's manual and program*. Edmonton, AB: Green's Publishing.

Grimes, D. A., & Schulz, K. F. (2002). Uses and abuses of screening tests. *The Lancet, 359*(9309), 881-884. doi: 10.1016/S0140-6736(02)07948-5

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating character (ROC) curve. *Radiology, 143*(1), 29-36. doi: 10.1148/radiology.143.1.7063747

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves defined from the same cases. *Radiology, 148*(3), 839-843. doi: 10.1148/radiology.148.3.6878708

Heaton, R. K., Miller, S. W., Taylor, M. J., Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead Reitan battery: demographically adjusted neuropsychological norms for African American and Caucasian adults*. Orlando, FL: Psychological Assessment Resources.

Holdnack, H.A. (2001). *Wechsler Test of Adult Reading manual*. San Antonio, TX: The Psychological Corporation.

Hoops, S., Nazem, S., Siderowf, A. D., Duda, J. E., Xie, S. X., Stern, M. B., & Weintraub, D. (2009). Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology, 73*, 1738-1745. doi: 10.1212/WNL.0b013e3181c34b47

Ihl, R., Frolich, L., Dierks, T., Martin, E., & Maurer, K. (1992). Differential validity of psychometric tests in dementia of the Alzheimer type. *Psychiatry Research, 44*, 93-106. doi: 10.1016/0165-1781(92)90044-4

Iverson, G. I., & Brooks, B. L. (2011). Chapter 32: Improving accuracy for identifying cognitive impairment. In M. R. Schoenberg & J. G. Scott (Eds.), *The little black book*

of neuropsychology: A syndrome-based approach (pp. 923-950). New York, NY: Springer.

Johns, E. K., Phillips, N. A., Nasreddine, Z., Chertkow, H., Bergman, L., Solomon, S., & Desormeau, J. (2010, April). *Level of education and performance on the Montreal Cognitive Assessment (MoCA©): new recommendations for education corrections*. Poster presented at the annual Cognitive Aging Conference, Atlanta, GA.

Kaplan E, Goodglass H, & Weintraub S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea and Febiger.

Kaufer, D. I., Williams, C. S., Braaten, A. J., Gill, K., Zimmerman, S., & Sloane, P. D. (2008). Cognitive screening for dementia and mild cognitive impairment in assisted living: Comparison of 3 tests. *Journal of the American Medical Directors Association, 9*, 586-593. doi: 10.1016/j.jamda.2008.05.006

Lam, B., Middleton, L. E., Masellis, M., Stuss, D. T., Harry, R. D., & Kiss, A. (2013). Criterion and convergent validity of the Montreal Cognitive Assessment with screening and standardized neuropsychological testing. *Journal of American Geriatrics Society, 61,* 2181-2185. doi: 10.1111/jgs.12541

Larner, A. J. (2011). Screening utility of the Montreal Cognitive Assessment (MoCA): In place of – or as well as – the MMSE? *International Psychogeriatrics, 24*, 391-396. doi: 10.1017/S1041610211001839

Larner, A. J. (Ed.). (2013). *Cognitive screening instruments: A practical approach*. London, UK: Springer-Verlag.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment (*4th ed.). New York, NY: Oxford Press.

Luis, C. A., Keegan, A. P., & Mullan, M. (2009). Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the southeastern US. *International Journal of Geriatric Psychiatry, 24*, 197-201. doi: 10.1002/gps.2101

Manly, J. J., Tang, M., Schupf, N., Stern, Y., Vonsattel, J. G., & Mayeux, R. (2008). Frequency and course of mild cognitive impairment in a multiethnic community. *Annuals of Neurology, 63*, 494-506. doi: 10.1002/ana.21326

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics, 18*(1), 50-60. Mendez, M. F., & Cummings, J. L. (2003). *Dementia: A clinical approach* (3rd ed.). Boston, MA: Butterworth-Heinemann.

Mickes, L., Jacobson, M., Peavy, G., Wixted, J. T., Lessig, S., Goldstein, J. L., & Corey-Bloom, J. (2010). A comparison of two brief screening measures of cognitive impairment in Huntington's disease. *Movement Disorders, 25*(13), 2229-2233. doi: 10.1002/mds.23181

Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research, 43*, 411-431. doi: 10.1016/j.jpsychires.2008.04.014

Mungas, D., Marshall, S. C., Weldon, H., Haan, M., & Reed B. R. (1996). Age and education correction of the Mini-Mental State Examination for English and Spanish-speaking elderly. *Neurology, 46*(3), 700-706. doi: 10.1212/WNL.46.3.700

Mystakidou, K., Tsilika, E., Parpa, E., Galanos, A., & Vlahos, L. (2007). Brief cognitive assessment of cancer patients: Evaluation of the Mini-Mental State Examination

(MMSE) psychometric properties. *Psycho-Oncology, 16*, 352-357. doi: 10.1002/pon.1090

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatric Society, 53*, 695-699. doi: 10.1111/j.1532-5415.2005.53221.x

Olson, R., Tyldesley, S., Carolan, H., Parkinson, M., Chhanabhai, T., & McKenzie, M. (2011). Prospective comparison of the prognostic utility of the Mini Mental State Examination and the Montreal Cognitive Assessment in patients with brain metastases. *Support Care Cancer, 19*, 1849-1855. doi: 10.1007/s00520-010-1028-1

Pendlebury, S. T., Cuthbertson, F. C., Welch, S. J. V., Mehta, Z., & Rothwell, P. M. (2010). Underestimation of cognitive impairment by Mini-Mental State Examination versus the Montreal Cognitive Assessment in patients with transient ischemic attack and stroke. *Stroke, 41*, 1290-1293. doi: 10.1161/STROKEAHA.110.579888

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 56*, 303-308. doi:10.1001/archneur.56.3.303

Petersen, R. C, Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., … Winblad, B. (2001a). Current concepts in mild cognitive impairment. *Archives of Neurology, 58*, 1985-1992. doi:10.1001/archneur.58.12.1985

Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cumming, J. L., & DeKosky, S. T. (2001b). Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review): Report of quality standards subcommittee of

the American Academy of Neurology. *Neurology, 56*, 1133-1142. doi: 10.1212/WNL.56.9.1133

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology, 20*, 33-65. doi: 10.1016/j.acn.2004.02.005

Reitan R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*, 271-276. doi: 10.2466/pms.1958.8.3.271

Roalf, D. R., Moberg, P. J., Xie, S. X., Wolk, D. A., Moelter, S. T., & Arnold, S. E. (2013). Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimer's and Dementia, 9,* 529-537. doi: 10.1016/j.jalz.2012.10.001

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the Montreal Cognitive Assessment (MoCA) in the population-based sample. *Neurology, 77,* 1272-1275. doi: 10.1212/WNL.0b013e318230208a

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-Based Medicine: How to Practice and Teach EBM* (2nd ed.). Philadelphia, PA: Churchill Livingstone.

Schweizer, T. A., Al-Khindi, T., & Macdonald, R. L. (2012). Mini-Mental State Examination versus Montreal Cognitive Assessment: Rapid assessment tools for cognitive and functional outcome after aneurysmal subarachnoid hemorrhage. *Journal of the Neurological Sciences, 316*, 137-140. doi: 10.1016/j.jns.2012.01.003

Shulman, K. I., Herrmann, N., Brodaty, H., Chiu, H., Lawlor, B., Ritchie, K., & Scanlan, J. M. (2006). IPA survey of brief cognitive screening instruments. *International Psychogeriatrics*, *18,* 281-294. doi: 10.1017/S1041610205002693

Smith, T., Gildeh, N., & Holmes, C. (2007). The Montreal Cognitive Assessment: Validity and utility in the memory clinic setting. *The Canadian Journal of Psychiatry, 52*(5), 329-332. Retrieved from http://www.stvincents.ie/dynamic/file/montreal_reference.pdf

Spreen, O., & Benton, A. L. (1977). *Manual of instructions for the Neurosensory Center Comprehensive Examination for Aphasia*. Victoria, BC: University of Victoria.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York, NY: Oxford University Press.

Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., … Devanand, D. P. (2006). Neuropsychological prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *Archives of General Psychiatry, 63*, 916-924. doi:10.1001/archpsyc.63.8.916

Teng, E. L., & Chui, H. C. (1987). The Modified Mini-Mental State (3MS) examination. *Journal of Clinical Psychiatry, 48*(8), 314-318. Retrieved from http://www.iowahealthcare.org/u serdocs/Documents/Teng_Chui-The_modified_mini-mental_state_3MS_examination.pdf

Toglia, J., Fitzgerald, K. A., O'Dell, M. W., Mastrogiovanni, A. R., & Lin, C. D. (2011). The Mini-Mental State Examination and Montreal Cognitive Assessment in persons with mild subacute stroke: Relationship to functional outcome. *Archives of Physical Medicine and Rehabilitation, 92*, 792-798. doi: 10.1016/j.apmr.2010.12.034

Tombaugh, T. N. (1992). The Mini-Mental State Examination: A comprehensive review. *Journal of the American Geriatric Society, 40*, 922-935. Retrieved from http://www.cfmal.com/Guardianship%20related%20references-1-11-10/Tombaugh%201992%20The%20MMSE%20-%20A%20comprehensive%20review.pdf

Tombaugh T. N. (1996). *Test of Memory Malingering (TOMM).* Boston, MA: Pearson Education Incorporated.

Tombaugh, T. N., McDowell, I., Kristjansson, B., & Hubley, A. M. (1996). Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data. *Psychological Assessment, 8*, 48-59. Retrieved from http://ww.feltoninstitute.org/articles/mmse%20normative%20data.pdf

Tsai, C. F., Lee, W. J., Wang, S. J., Shia, B. C., Nasreddine, Z., & Fuh, J. L. (2012). Psychometrics of the Montreal Cognitive Assessment (MoCA) and its subscales: Validation of the Taiwanese version of the MoCA and an item response theory analysis. *International Psychogeriatrics, 24*(4), 651-658. doi: 10.1017/S1041610211002298

Valcour, V. G., Masaki, K. H., Curb, D., & Blanchette, P. L. (2000). The detection of dementia in the primary care setting. *Archives of Internal Medicine, 160*, 2964-2968. doi:10.1001/archinte.160.19.2964

Videnovic, A., Bernard, B., Fan, W., Jaglin, J., Leurgans, S., & Shannon, K. M. (2010). The Montreal Cognitive Assessment as a screening tool for cognitive dysfunction in Huntington's disease. *Movement Disorders, 25*, 402-404. doi: 10.1002/mds.22748

Vincent, G. K., & Velkoff, V. A. (2010). The next four decades - The older population in the United States: 2010 to 2050: Population estimates and projections. Washington D.C:

US Census Bureau. Retrieved from http://www.census.gov/prod/2010pubs/p25-1138.pdf

Waldron-Perrine, B., & Axelrod, B. N. (2012). Determining an appropriate cutting score for indication of impairment on the Montreal Cognitive Assessment. *International Journal of Geriatric Psychiatry, 27*, 1189-1194. doi: 10.1002/gps.3768

Wechsler, D. (1997a). *Wechsler Memory Scale—third edition (WMS–III) administration and scoring manual* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Adult Intelligence Scale (WAIS–III)* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale (WAIS–IV)* (4th ed.). San Antonio, TX: The Psychological Corporation.

Whitney, K. A., Mossbarger, B., Herman, S. M., & Ibarra, S. L. (2012). Is the Montreal Cognitive Assessment superior to the Mini-Mental State Examination in detecting subtle cognitive impairment among middle-aged outpatient U.S. military veterans? *Archives of Clinical Neuropsychology, 27*(7), 742-748. doi: 10.1093/arclin/acs060

Xu, W., Caracciolo, B., Wang, H-X, Winblad, B., Bäckman, L., Qui, C., & Fratiglioni, L. (2010). Accelerated progression from mild cognitive impairment to dementia in people with diabetes. *Diabetes, 59*(11), 2928-2935. doi: 10.2337/db10-0539

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1): 32-25. doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Zadikoff, C., Fox, S. H., Tang-Wai, D. F., Thomsen, T., de Bie, R. M. A., Wadia, P., … Marras, C. (2008). A comparison of the Mini-Mental State Exam to the Montreal

Cognitive Assessment in identifying cognitive deficits in Parkinson's disease. *Movement Disorders, 23*, 297-299. doi: 10.1002/mds.21837