THE COLORADO COGNITIVE ASSESSMENT (CoCA): DEVELOPMENT OF A SCREENING TOOL THAT PROVIDES PROCESS INFORMATION AND IS SCORED USING MODERN PSYCHOMETRICS

by

ASHITA SHYAM GURNANI

A dissertation submitted to the Graduate Faculty of the University of Colorado Colorado Springs in partial fulfillment of the requirement for the degree of Doctor of Philosophy

Department of Psychology

2018
This dissertation for the Doctor of Philosophy degree by

Ashita Shyam Gurnani

has been approved for the

Department of Psychology

by

Brandon Gavett, Chair

Molly Maxfield

Dan Segal

Lori James

Julie Horwitz

Date 03/27/2018
Gurnani, Ashita Shyam (Ph.D., Psychology)

The Colorado Cognitive Assessment (CoCA): Development of a Screening Tool that Provides Process Information and is Scored Using Modern Psychometrics

Dissertation directed by Assistant Professor Brandon Gavett

**ABSTRACT**

Early and accurate diagnosis of dementia is imperative for patients’ overall well-being and improved quality of life. Current cognitive screening tools lack sufficient diagnostic accuracy to consistently diagnose mild and uncommon presentations of dementia. Additionally, most current screening instruments are not designed to assist with differential diagnosis and consist of tasks that are influenced by sociodemographic factors. Thus, the purpose of the present study was to develop and preliminarily validate a cognitive screening instrument that can facilitate early and accurate diagnosis of mild and atypical presentations of dementia. The Colorado Cognitive Assessment (CoCA) differs from other cognitive screens in that it assesses a broad range of cognitive domains; consists of tasks that were designed to be minimally influenced by education, sex, and culture; contains tasks that maximize attainment of qualitative features of performance to guide differential diagnosis; and is validated using modern psychometrics. The confirmatory factor analysis (CFA) model of the CoCA revealed excellent fit in a sample of 151 community dwelling older adults. Measurement equivalence analyses revealed that the items on the CoCA were invariant to sex, age, education, and mood. In comparison, a CFA model of the Montreal Cognitive Assessment (MoCA) had worse fit and was biased by age, education, and depressive symptomatology. The CoCA demonstrated adequate convergent validity with the MoCA and NAB Judgment subtest,
and divergent validity with the Geriatric Depression Scale-15 (GDS-15) and Geriatric Anxiety Scale-10 (GAS-10). Results provide preliminary evidence for the CoCA as a reliable and comprehensive cognitive screen. Future research is needed to validate the CoCA cross-sectionally and longitudinally in diverse non-clinical and clinical samples.

*Keywords:* Cognitive screening, dementia, mild cognitive impairment, differential diagnosis, CFA, measurement invariance
ACKNOWLEDGEMENTS

I would like to thank everyone who contributed to the completion of this project. To my committee members, thank you for your insightful comments and suggestions that helped shape this project. This project has been very exciting as it blended all aspects of clinical work and research that interest me. This is what also made the project a daunting task. The incredible support of my mentor, Brandon Gavett, has made the completion of this project possible. Brandon, thank you for helping me every step of the way; for challenging me to think about every missed detail and think some more about any detail that I thought I had figured out; for making me use statistical analyses I had never heard of; and most of all, for being so excited about the project all the time. I moved to the USA with the goal of becoming a neuropsychologist to not only be involved in the early and accurate diagnosis of dementia in a clinical capacity, but also with respect to research. And while I do not know if this project achieves that goal, I am so glad that my dissertation focuses on that very topic and I cannot thank you all enough for helping me through the process.
# TABLE OF CONTENTS

## CHAPTER

### I. INTRODUCTION .................................................................1

- Cognitive Continuum .........................................................3
- Importance of the Screening of Cognitive Impairment ..........4
- Cognitive Screening Instruments .........................................5
- Differential Diagnosis .......................................................20
- Development of the Colorado Cognitive Assessment (CoCA) ....28
- Psychometric Properties of the CoCA .................................39
- The Present Study ............................................................44

### II. METHOD ..................................................................46

- Participants .................................................................46
- Measures ..................................................................46
- Procedure ...............................................................59
- Statistical Analyses .....................................................60

### III. RESULTS .................................................................67

- Participants .................................................................67
- Modeling .................................................................67
- CoCA Descriptive Statistics ........................................74
- Measurement Invariance (MI) .........................................75
- Reliability and Validity ...............................................78
Comparison to the MoCA ................................................................. 81

IV. DISCUSSION .................................................................................. 85

Validation of the CoCA ................................................................. 85

Measurement Invariance of the CoCA ................................. 90

Reliability and Validity of the CoCA ......................................... 94

Advantages of the CoCA .............................................................. 98

Limitations and Future Directions .......................................... 101

Conclusion ....................................................................................... 103

REFERENCES ...................................................................................... 105

APPENDIX A: CoCA INSTRUMENT ......................................................... 132

APPENDIX B: CoCA PROCESS VARIABLES RECORD FORM .............. 143

APPENDIX C: COLORADO COGNITIVE ASSESSMENT (CoCA)
INSTRUCTION AND SCORING MANUAL ............................................... 146

APPENDIX D: MONTRAL COGNITIVE ASSESSMENT (MoCA) .......................... 165

APPENDIX E: GERIATRIC DEPRESSION SCALE (SHORT FORM) ............ 166

APPENDIX F: GERIATRIC ANXIETY SCALE – 10 ........................................ 167

APPENDIX G: DEMOGRAPHICS QUESTIONNAIRE .................................... 168

APPENDIX H: IRB APPROVAL .................................................................. 169
LIST OF TABLES

TABLE

1. Sample Demographic Characteristics ............................................................68
2. Model Fit Indices for Each Model of the CoCA ............................................69
3. Indicators Used in Each CFA Model of the CoCA ........................................70
4. Unstandardized and Standardized Parameter Estimates for the CFA Model of the CoCA .................................................................74
5. Descriptive Data for Each Subtest of the Colorado Cognitive Assessment (CoCA) .................................................................75
LIST OF FIGURES

FIGURE

1. Model to be validated in the future using confirmatory factory analysis (CFA) .................................................................41

2. Proposed model of the CoCA that was planned for the current study and examined using confirmatory factory analysis (CFA) .........................42

3. CFA model of The Montreal Cognitive Assessment .................................66

4. Best fitting model of the CoCA using confirmatory factor analysis (CFA) .72

5. Model of the CoCA using confirmatory factory analysis (CFA) ..............73

6. Distribution of global factor scores of the CoCA ................................76

7. Step 1 of multiple-indicators multiple-causes (MIMIC) modeling ............77

8. Step 3 of multiple-indicators multiple-causes (MIMIC) modeling ............79

9. Scatterplot showing the relationships between the global factor score on the CoCA and total scores on the Montreal Cognitive Assessment (MoCA; A), NAB Judgement Subtest (B), Geriatric Depression Scale-15 (GDS-15; C), and Geriatric Anxiety Scale-10 (GS-10; D) ..................80

10. Correlogram showing correlations between global factor score on the CoCA, age, education, total Montreal Cognitive Assessment (MoCA) score, total Geriatric Depression Scale-15 (GDS-15) score, total Geriatric Anxiety Scale-10 (GAS-10) score, and total NAB Judgement score ..........81
CHAPTER I
INTRODUCTION

Dementia, a clinical syndrome characterized by concomitant decline in cognitive capacity and functional ability, has a global prevalence of 5%-7% in individuals above the age of 60 (Prince et al., 2013). Advancing age is the biggest risk factor for dementia with an estimated 22% of the world’s population to be accounted for by individuals above the age of 60 (Alzheimer’s Association [AA], 2016). Early diagnosis of dementia is imperative in order to reduce caregiver burden and costs, improve planning, initiate treatment options that are most effective in the initial stages of the disease, and optimize overall well-being (Galvin et al., 2008). Detection of dementia is associated with improvement in patients’ knowledge and awareness of the disease, which likely results in greater professional help-seeking behaviors from patients and family members and provides the individual with the opportunity to be involved in determining their own care (Department of Health, 2009). Given the rise in the frequency of dementia and increase in the collateral damage as a result of this diagnosis, several non-specialist avenues of medicine that are frequented by older adults, including primary care physicians and geriatric specialist services, play a pivotal role in screening of cognitive impairment (Milne, Culverwell, Guss, Tuppen, & Whelton, 2008; Sheehan, 2012).

However, primary care medical records of 50% to 80% of individuals with dementia lack any mention of this diagnosis, potentially due to providers having reduced confidence in the results of cognitive screening instruments that tend to have high false
positive error rates. False positives could lead to distress and stigma often associated with a dementia diagnosis along with unnecessary treatments and increased costs associated with diagnostic work-up (Boustani, Peterson, Hanson, Harris, & Lohr, 2003; Ismail, Rajji, & Shulman, 2009; Milne et al., 2008; Philips, Walters, Biju, & Kuruvilla, 2016). The high percentage of undetected dementia cases also reflects difficulty in identifying high-risk individuals given the variability in the symptoms of dementia along with the absence of validated instruments for dementia screening (Alzheimer’s Society, 2007; Milne et al., 2008). Varying etiologies of dementia, other than Alzheimer’s disease (AD), are often not taken into consideration due to lack of relevant training about dementia and its causative pathologies amongst primary care providers (Bradford, Kunik, Schulz, Williams, & Singh, 2009; Cullen, O’Neill, Evans, Coen, & Lawlor, 2007). Identification of accurate changes in cognitive functioning requires psychometrically sound serial assessments (Gavett, Ashendorf, & Gurnani, 2015). Most cognitive assessments are scored using classical test theory (CTT), which often violates the assumption that the test possesses linear scaling properties (Mungas & Reed, 2000). Linear scaling properties, which are highly desirable for serial assessments, can be achieved by scoring tests using modern psychometric methods, such as confirmatory factor analysis (CFA; Brown, 2015). Additionally, CFA helps in determining how a test should be scored while demonstrating superior psychometric properties over traditional approaches to test development such as CTT (Brown, 2015). Thus, there is a need for a psychometrically robust cognitive screening tool that examines a broad range of cognitive domains and is minimally influenced by culture and education while maximizing the amount of information it provides about an individual’s cognition and functioning (Shulman, 2000).
Cognitive Continuum

The spectrum of cognitive decline consists of age-related cognitive changes on one end and dementia on the other end of the continuum, with mild cognitive impairment (MCI) in the middle (Hachinski, 2008). Typically, dementia is the clinical syndrome used to describe the progressive deterioration in cognition and functional abilities, more than what is expected for normal aging, with a wide range of symptoms that could be caused by various irreversible neurodegenerative pathologies, such as AD and reversible conditions, such as thyroid problems and vitamin deficiencies (AA, 2016; McKhann et al., 2011). Common irreversible causes of dementia include Lewy body disease, cerebrovascular disease, and fronto-temporal lobar degeneration, with AD being the most well-known and most frequent cause of dementia (AA, 2016). Though MCI has the same etiological underpinnings as dementia, the degree of cognitive decline is not substantial enough to constitute a diagnosis of dementia (Peterson et al., 2009). There is considerable variability amongst providers’ and clinicians’ diagnosis of MCI and its subsequent conversion to dementia, potentially due to the diagnosis resting on clinical judgment regarding the severity of functional impairment that constitutes a diagnosis of dementia (Nestor, Scheltens, & Hodges, 2004; Peterson et al., 2009). Given considerable heterogeneity in the cognitive functioning of older adults without dementia, boundaries between cognitive decline associated with normal aging and with dementia are often blurred (Galvin et al., 2005). The conception of the cognitive continuum suggests that cognitive ability is continuous and allows for more variability in cognitive functioning; dementia, however, is better characterized as a categorical construct since it represents a decline in cognition that reduces one’s ability to function independently (Gavett & Stern,
As a result, screening tools designed to diagnose dementia have to be adapted to detect the earliest changes in both cognition and daily functioning, with increased focus on the development, validation, and refinement of screening instruments to identify MCI before it progresses to dementia (Kalbe et al., 2004; Winblad et al., 2004).

**Importance of the Screening of Cognitive Impairment**

The increased economic, social, and psychological burden that occurs with dementia has resulted in policy-related guidelines consistently emphasizing the importance of early detection and intervention of people with dementia, recommending the use of standardized screening instruments to facilitate this process (National Institute for Health and Clinical Excellence and Social Care Institute for Excellence, 2006). Identification and characterization of cognitive deficits is primarily dependent on cognitive assessments given the lack of a reliable biomarker to distinguish between normal aging and dementia (Salmon & Bondi, 2009). Clinical assessment of dementia is also important in order to rule out reversible causes of the disease that are responsible for approximately 9% - 20% of dementia cases (Tripathi & Vibha, 2009). Early detection of dementia is beneficial for several reasons: (1) it provides information about the individual’s cognitive, behavioral, and affective features and associated implications, facilitating better understanding of the patient’s abilities and allowing the patient and his or her family to prepare emotionally, mentally, legally, and financially; (2) it allows for the identification of physical and functional risk factors such as falls and driving; and (3) it increases specificity when enrolling dementia patients and healthy controls in clinical trials evaluating the efficacy of novel disease-modifying drugs (Shulman et al., 2006).
Cognitive screening is often considered the first, and sometimes the final step in the detection of dementia. Given the inadequate number of neurologists and dementia care specialists, primary care providers are often the first to observe an individual with signs of dementia and frequently the ones to make the diagnosis (Galvin & Sadowsky, 2012; Lorentz, Scanlan, & Borson, 2002). In addition to dementia detection, routine screening of dementia in primary care settings can alert providers about the possibility of cognitive changes in older adults. Psychologists and neuropsychologists, particularly those working with older adults who may not be able to tolerate lengthy test batteries, may also benefit from an efficient screening tool that is diagnostic in nature and provides the most amount of qualitative information about performance in a relatively short amount of time. On a societal level, routine practice of dementia screening can help facilitate and expedite the translation of research initiatives into clinical practice, along with the development and implementation of standards for dementia care across various settings (Borson, Scanlan, Brush, Vitaliano, & Dokmal, 2000; Lorentz et al., 2002).

Though several national associations specializing in geriatrics and neurology encourage physicians to screen for dementia as early as possible, dementia still remains underdiagnosed, potentially due to the limited ability of screening instruments to detect subtle signs of dementia (Boise, Camicioli, Morgan, Rose, & Congleton, 1999; Wilkinson, Sganga, Stave, & O’Connell, 2005; Woods et al., 2003).

**Cognitive Screening Instruments**

The primary goal of cognitive screening tools is to provide an accurate estimate of the probability and severity of cognitive impairment based on a person’s test score in reference to the normative population for that test. An individual’s test score determines
the course of action taken by the provider administering the cognitive screen; for example, a provider may choose to make a diagnosis, refer the patient to another level of screening, or order a neuropsychological evaluation (Cullen et al., 1997). Given the important and evolving role of screening instruments, a large number of screening instruments have been developed over the past several years.

**Review of current cognitive screening instruments.** Over 30-40 cognitive screening tests have been developed over the past two decades to improve the efficacy of cognitive assessments. Generally, two types of screening instruments are used for the cognitive assessment of dementia - very short screening tests that are domain-specific and consist of a couple of items such as the Memory Impairment Screen (MIS; Buschke et al., 1999); and tests of global cognitive ability such as the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) and the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005). Sensitivity and specificity are two values that help gauge the diagnostic accuracy of cognitive screening instruments. Sensitivity is the probability that an individual tests positive or meets criteria for a disease when the disease is present. Specificity is the probability that an individual tests negative or does not meet criteria for a disease when the disease is absent (Parikh, Mathai, Parikh, Sekhar, & Thomas, 2008). In general, sensitivity and specificity values below .80 are considered inadequate; however, the higher the sensitivity and specificity values, the better the test is at diagnostic classification.

**Domain specific screening tests.** Shorter screening tests that assess one cognitive domain are used alone or in conjunction with tests of global cognitive ability, and can be used to make decisions surrounding further assessment. One such test that is primarily
used to assess memory function is the MIS, which assesses free and cued recall of four words (Buschke et al., 1999). The MIS has good reliability and construct validity and has been shown to detect dementia with 80% sensitivity and 96% specificity, with minimal impact of education, age, and sex (Buschke et al., 1999). In addition to the MIS having low sensitivity for detection of dementia, it also only assesses memory processes, making it uninformative about performance in other cognitive domains and increasing the likelihood of missing atypical presentations of AD and dementias due to frontal lobe degeneration (Velayudhan et al., 2014). The Executive Interview (EXIT-25) is a 25-item screening instrument that assesses a variety of executive functions including motor and verbal sequencing, fluency, inhibition, utilization behavior, primitive reflexes, and persistence (Royall, Mayurin, & Gray, 1992). The EXIT-25 is used to assist with early and accurate diagnosis of varying dementia etiologies and has demonstrated strong correlations with other neuropsychological tests of executive functioning (Marin, Butters, Mulsant, Pollock, & Reynolds, 2003; Royall et al., 1992). Though it has been found to reduce diagnostic misclassification (Royall, Cabello, & Polk, 1998; Royall, Chiodo, & Polk, 2000; Royall, Mahurin, Cornell, & Gray, 1993), it only provides information about executive dysfunction, making its utility in the diagnosis of dementia dependent on other tests of global cognition or memory (Stokholm, Vogel, Gade, & Waldemar, 2006).

Another very brief screening test that appears to primarily measure visuoconstructional and visuospatial abilities is the Clock Drawing Test (CDT; Mainland & Shulman, 2013; Sunderland et al., 1989). An experienced and trained clinician may obtain a general idea of additional aspects of cognitive functioning upon further examination of the CDT as it also taps into other cognitive skills such as verbal
understanding, long term memory, planning, abstract thinking, and several frontal/executive skills. The advantage of the CDT over other instruments is its potential to be information-rich in a very brief administration time (Aprahamian, Martinelli, Neri, & Yassuda, 2010; Storey, Rowland, Basic, & Conforti, 2002). Several studies and systematic reviews have consistently shown that the CDT demonstrates adequate sensitivity and specificity in detecting moderate to severe levels of dementia (Kato et al., 2013; Pinto & Peters, 2009; Shulman, 2000). Several scoring systems with varying degrees of complexity that provide qualitative and quantitative information exist for the CDT; though the scoring systems tend to be well correlated, there is no real consensus on which criteria are the most valid (Hubbard et al., 2008). Given its advantages, the CDT is consistently recommended for use with a test of global cognitive ability such as the MMSE (Cacho et al., 2010; Kato et al., 2013; Zhou & Jia, 2008). Despite its advantages of brevity and acceptability to patients, the CDT as a stand-alone measure does not possess adequate accuracy to diagnose dementia etiologies, though it has been shown to distinguish between healthy controls and AD with adequate sensitivity and specificity (Nair et al., 2010; Rubinova et al., 2014; Shulman, 2006). The CDT alone is less robust in distinguishing between MCI and AD, with a tendency to result in an increased number of false positive and false negative errors when screening for MCI (Nair et al., 2010). When very brief administration time is of primary importance, multiple systematic reviews have identified the MIS and CDT, separately or together, as having great utility in primary care settings due to their practicality, feasibility, range of applicability, and psychometric properties (Brodaty et al., 2006; Holsinger, Deveau, Boustani, & Williams, 2007; Lorentz et al., 2002; Milne et al., 2008).
Tests of global cognitive ability. Screens that tend to be longer, consisting of tasks that assess a wide variety of cognitive domains and even different aspects of one domain to evaluate core components of dementia syndromes are often used alone to detect dementia and identify its level of severity. Commonly used tests include the MMSE, Cognitive Abilities Screening Instrument (CASI; Teng et al., 1994), MoCA, Mini-Cog (Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000), St. Louis University Mental Status Examination (SLUMS; Tariq, Tumosa, Chibnall, Perry, & Morley, 2006), and Addenbrooke’s Cognitive Examination-Revised (ACE-R; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006). Cognitive screens can be administered in-person or on the phone, with the latter helping with the management of some of the practical challenges that one may face with in-person evaluations (Knopman et al., 2010). One telephone-administered cognitive screen demonstrating strong psychometric properties that provides a global composite score of cognitive functioning and has been validated in a large sample with a wide age range is the Brief Test of Adult Cognition by Telephone (BTACT; Brim, Ryff, & Kessler, 2004; Ryff & Lachman, 2007; Tun & Lachman, 2006). The BTACT assesses various aspects of cognitive functioning, including verbal memory, working memory, verbal fluency, inductive reasoning, processing speed, and task switching ability (Tun & Lachman, 2006). While common cognitive screens such as the MMSE are similar to the BTACT in that they assess global cognitive functioning, the BTACT can be scored using modern psychometrics, facilitating interpretation of performance (Gavett et al., 2013). Further, scoring of the BTACT also allows the examiner to correct for the effect of age, education, gender, and occupation; the ability to account for the effect of demographic variables on performance makes the BTACT
superior to other screening measures, such as the MMSE, at capturing one’s true cognitive ability (Gurnani, John, & Gavett, 2015).

Despite several studies showing that performance on the MMSE is impacted by education, literacy, age, culture, social class, and language (Freidl et al., 1996; Mungas, Marshall, Weldon, Hann, & Reed, 1996; Reed, Kligman, & Abyad, 1995), the MMSE has been consistently shown as the most widely used cognitive instrument, with one study reporting that 51% of primary care practices administer the MMSE alone and 80% employ the MMSE along with another dementia screening tool (Cullen et al., 2007; Milne et al., 2008). Briefly, the MMSE assesses orientation; word recall; attention; language comprehension, production, and repetition; and visuospatial ability. The MMSE has been shown to have poor reliability but good construct validity and sensitivity for detecting moderate to severe cognitive impairment, with sensitivity improving with increasing levels of impairment (Gavett et al., 2015; Tombaugh & McIntyre, 1992). A recent meta-analysis revealed that the sensitivity and specificity of the MMSE in detecting dementia was 81% and 89%, respectively (Tsoi, Chan, Hirai, Wong, & Kwok, 2015), indicating that the MMSE demonstrates relatively equal accuracy in identifying dementia and cognitively normal individuals.

The MMSE has been criticized for having reduced sensitivity in detecting MCI and for exhibiting ceiling effects in mild stages of AD and floor effects in advanced stages of dementia, resulting in a disproportionate number of false negative and false positive errors (Nelson, Fogel, & Faust, 1986; Simard & van Reekum, 1999). The MMSE is frequently used for repeated assessments in research and clinical practice despite its practice effects, which affect accurate interpretation of cognitive change over time.
Considering its limitations, the reliance on the MMSE to detect and gauge the severity of cognitive impairment is disproportionate, emphasizing the importance of developing and utilizing other measures with greater diagnostic accuracy and reduced sociodemographic biases.

The Modified Mini-Mental State Examination (3MS; Teng & Chui, 1987) was developed in response to the criticisms of the MMSE by adding four tests to the MMSE, including date and place of birth, word fluency, similarities, and delayed recall of words (Teng & Chui, 1987). The 3MS exhibits comparable reliability, sensitivity, and specificity with the MMSE, though low education reduces specificity and improves sensitivity, and increasing age reduces specificity (Tombaugh, McDowell, Kristjansson, & Hubley, 1996). Like the MMSE, the 3MS is minimally useful in detecting mild levels of cognitive impairment, though the additional tests slightly improve its utility as a screening instrument (McDowell, Kristjansson, Hill, & Herbert, 1997). The General Practitioner Assessment of Cognition (GPCOG, Brodaty, Low, Gibson, & Burns, 2002) and The Cognitive Abilities Screening Instrument (CASI, Teng et al., 1994) have classification accuracies similar to the MMSE though they allow for a wider range of scores, exhibiting greater utility in detecting milder forms of dementia (Lorentz et al., 2002). Briefly, the GPCOG includes a clock drawing task, an orientation based task, a test for the recall of a previously read address, and a short questionnaire asking the informant to compare the patient’s current functioning to his or her functioning five to ten years ago (Brodaty et al., 2002). The CASI assesses orientation; word recall; basic attention and concentration; language naming, production, and comprehension; and visuospatial ability (Teng et al., 1994). In comparison to the MMSE, the GPCOG has
been found to be more suitable as a dementia screening measure in general practice based on its simplicity, administration time of less than five minutes, and diagnostic accuracies similar to the MMSE (Brodaty et al., 2006; Holsinger et al., 2007; Lorentz et al., 2002).

In comparison to the MMSE, the Mini-Cog has better sensitivity and specificity in detecting dementia even in mildly impaired individuals and has been consistently identified as a useful screening tool for dementia in primary care settings (Brodaty et al., 2006; Holsinger et al., 2007; Lorentz et al., 2002; Milne et al., 2008). The Mini-Cog consists of two tasks: a three-word recall task and a clock drawing task (Borson et al., 2000). The Mini-Cog is preferred in primary care settings due to its evaluation of all aspects of memory and other frontal-executive and visuospatial abilities along with requiring minimal training and shorter administration time in comparison to the ACE-R and MoCA (Shulman et al., 2006). However, it has been shown to have lower diagnostic accuracy in early stages of dementia in comparison with the CASI, although with minimal influence of education and language on performance (Borson et al., 2000).

Given the reduced accuracy exhibited by multiple screening measures in detecting dementia in its earliest stages, the MoCA, SLUMS, and ACE-R were each developed to identify MCI. The SLUMS consists of tasks of orientation, object recall, animal fluency, basic working memory, clock drawing, simple language comprehension, and story memory (Tariq et al., 2006). The MoCA consists of tasks of executive functioning, visuospatial ability, naming, immediate and delayed word recall, attention and vigilance, language repetition, lexical fluency, abstract reasoning, and orientation (Nasreddine et al., 2005). The ACE-R consists of tasks assessing orientation; word recall; basic attention and concentration; immediate and delayed recall of an address; phonemic and animal
fluency; language production, comprehension and repetition; and visuospatial, visuoconstructional, and visuoperceptual abilities (Mioshi et al., 2006). In comparison to the MMSE, these tests assess a wider range of cognitive abilities and exhibit superior diagnostic accuracies in identifying dementia and MCI (Mioshi et al., 2006; 2007; Smith, Gildeh, & Holmes, 2007; Tariq et al., 2006).

The advantage of the SLUMS over other comparative cognitive screens is the use of cutoffs based on years of education completed; in comparison, a major drawback of the SLUMS is the effect of racial and ethnic differences on performance and lack of validity and adequate psychometric data (Stewart, O’Riley, Edelstein, & Gould, 2012). In contrast, the MoCA and ACE-R have been validated in various cultural backgrounds and also in individuals with vascular dementia (VaD), frontotemporal dementia (FTD), AD, MCI, and Parkinson’s disease dementia (PDD; Alexopoulos et al., 2010; Freitas, Simoes, Alves, Duro, & Santana, 2012; Freitas, Simoes, Alves, & Santana, 2013; Freitas, Simoes, Alves, Vicente, & Santana, 2012; Hoops et al., 2009; Reyes et al., 2009; Stewart et al., 2012). In addition, the ACE-R was created to cover several domains and has been shown to differentiate between various dementia etiologies while possessing strong psychometric properties (Mioshi et al., 2006; Velayudhan et al., 2014), though it demonstrates reduced correlations with basic and instrumental activities of daily living in AD and FTD (Mioshi et al., 2007). Of all the cognitive screening tests reviewed, it appears that the ACE-R has the most evidence supporting its ability to diagnose dementia and differentiate between the various etiologies, resulting in it being utilized in primary care and specialty settings (Crawford, Whitnall, Robertson, & Evans, 2012; Hsieh, Schubert, Hoon, Mioshi, & Hodges, 2013; Larner & Mitchell, 2014).
Though recently developed screening measures such as the MoCA and the ACE-R possess strong psychometric properties and cover a wide variety of cognitive domains to potentially identify early and atypical forms of dementia and assist with differential diagnosis, research does not indicate the presence of a direct relationship between the number of key cognitive abilities assessed and dementia detection. Rather, the utility of assessing a broad range of domains lies in the resulting process variables such as approach to the test including strategies used and errors made, that add to the total quantitative score (Cullen et al., 2007). Overall, screens assessing a wide range of cognitive abilities are likely to be of more utility in detecting mild and unusual cases of dementia and for differential diagnosis as they provide process information about the patient’s presenting symptoms that take precedence over a cut-off score (Cullen et al., 2007). Generally, most cognitive instruments are validated in dementia populations with screens sensitive to identifying AD and not necessarily having utility in picking up non-AD etiologies. Of all the cognitive instruments covered in this section, the ACE-R is the only screening tool that has been specifically created and validated for differential diagnosis as it provides cognitive profiles that can assist with estimating the likelihood of a particular dementia etiology (Cullen et al., 2007). Though the importance of differential diagnosis and detection of MCI is increasing, the MMSE currently appears to be the most utilized cognitive screening instrument despite its insensitivity to change in scores over time, heavy influence of education, age, culture, and language, and reduced emphasis on frontal and executive systems (Shulman et al., 2006).

**Identified essential characteristics for a cognitive screening instrument.**

Cognitive screening typically dictates the likelihood of the patient obtaining appropriate
investigation and treatment, thereby underscoring the importance of statistical robustness of the instrument (Cullen et al., 2007). In order for a test to distinguish cognitive impairment from normal age-related cognitive decline with increased objectivity, it should be able to achieve high sensitivity and specificity with a high positive predictive value in the minimum amount of time possible (Cullen et al., 2007; Sheehan, 2012). The initiative to maximize detection with the use of psychometrically sound instruments in a time-limited manner has led to the development of extremely brief screening tasks that have high predictive value but also possess several drawbacks (Cullen et al., 2007).

Specifically, in an attempt to maximize detection of dementia with limited time, many tests overemphasize detection of episodic memory impairment, which is considered to be the earliest cognitive feature of AD, the most common cause of dementia. As a result, tests neglect assessment of other domains such as executive functioning and language, which may be impaired early in other etiologies of dementia, such as cerebrovascular disease or fronto-temporal lobar degeneration (De Koning, Van Kooten, & Koudstaal, 1998). The disproportionate emphasis on memory dysfunction to obtain a diagnosis of dementia has led to what is known as the “Alzheimerization” of dementia, thereby influencing the construction of screening tests such that important domains affected in other forms of cognitive impairment are not examined thoroughly (Royall, 2003). A screen with several cognitive domains that provides rich qualitative information is especially useful to medical providers that are primarily interested in differential diagnosis and in further investigating individuals with mild or atypical presentations of dementia. Providers examining individuals suspected of having difficulty with functional and decision making capacity are likely to rely on clinical judgment more than cutoffs
and cognitive index scores, thus, a screening tool that assesses breadth in executive functioning ability quantitatively and qualitatively is likely to be useful (Cullen et al., 2007). Screens assessing a broad range of cognitive domains also allow referring clinicians to provide more information about a person’s cognitive profile, providing neuropsychologists with guidelines for further assessment (Teng et al., 1994).

Another problem with screening tools lies in their scoring; determination of cognitive impairment is based on comparing the person’s total score to a standard cut off. This method is flawed in that it ignores the iterative process of arriving at a diagnosis by carefully evaluating the cognitive profile of an individual (Cullen et al., 2007). Frequently, qualitative information about a person’s cognitive functioning is missed as a result of increased emphasis on the total score in comparison with the standard cutoff (Shulman et al., 2006). Screening instruments providing qualitative information about relevant cognitive domains are of much utility in smaller clinics that lack necessary resources to administer a full neuropsychological battery (Teng et al., 1994). Individuals with severe cognitive impairment or those with personality and behavioral characteristics that refuse to cooperate with several hours of testing are often only able to tolerate a cognitive screening instrument. Given the above-mentioned cases, it is imperative that cognitive screening tools make available the possibility of obtaining indices or structured information about key cognitive domains (Cullen et al., 2007).

Though the role of a cognitive screen is not to serve as a substitute for neuropsychological assessment, it is relatively common for neuropsychologists to administer a screening instrument at the beginning of their evaluation (Teng et al., 1994). The goal of the screening tool in this case is to identify the presence of any impairment
and indicate a likely cause for the presenting deficits that can then guide further assessment (Cullen et al., 2007). As established, current screening measures struggle with early detection of dementia. Another cause of reduced identification and misdiagnosis of dementia is the presence of several factors that introduce bias, including differences in level and type of formal education, language, and culture-specific variables (Parker & Philp, 2004). With the increase in cultural diversity, adaptation and development of tests such that they are minimally influenced by sociodemographic variables to improve diagnostic accuracy is increasingly gaining relevance in clinical and research settings.

**Biases in cognitive testing.** It has been argued that culture, including one’s values, beliefs, and styles of behavior, influences brain organization and impacts cognitive abilities and strategies (Ardila, 2005; Eviatar, 2000). There is widespread consensus that neuropsychological tests lack adequate diagnostic accuracy when used to assess individuals who are not Caucasian, well-educated, and middle to upper class (Manly, 2008). For example, Manly, Byrd, Touradji, and Stern (2010) have shown that the quality of education along with the process of acculturation affected African Americans’ neuropsychological test performance. There is a well-established association between level of education and cognitive performance, with education accounting for the effects of culture in some cases (Parker & Philp, 2004). For example, individuals with lower levels of education perform worse on the MoCA; as a result, a point is added to the total MoCA score for individuals with 12 or fewer years of education (Nasreddine et al., 2005). A substantial portion of older adults migrated to the USA when they were adults, thereby acquiring lower fluency and literacy in English than individuals born and raised in the USA, in addition to reduced familiarity with the predominant culture of the USA.
Language and literacy are important aspects of culture that play into performance on several test items, as is familiarity with the predominant American culture, thereby proving to serve as a disadvantage to older adults that may have migrated from another culture (Parker & Philp, 2004). For example, the MMSE is strongly influenced by culture even after translations and adjustments to the cut-off point; it has a high false positive rate for minimally educated individuals and increased specificity for non-whites (Richards et al., 2000; Stewart, Johnson, Richards, Brayne, & Mann, 2002).

Cognitive testing in an unfamiliar language or culture can also lead to exacerbation of test anxiety, resulting in under-performance (Parker & Philp, 2004). For example, anxiety due to being tested in another culture or language is one of the reasons that cognitively intact Spanish-speaking individuals are more likely than English-speaking participants to miss items on cognitive tests and be classified as impaired (Jacobs et al., 1997). Translation of test items and adaptation of a screening tool to account for cultural background effects does not necessarily eliminate the bias (Rait et al., 1999; 2000). Several tests have been presented as culture-free or culture-fair primarily because they consist of a larger percentage of non-verbal items, though research has consistently shown that non-verbal tasks may require abilities and skills that are influenced by culture (Crampton & Jerabek, 2000; Miller, 1973). Map- and figure-copying tests and use of time limits are examples of non-verbal tests and test components influenced by culture, because of the abilities and strategies valued in each culture (Miller, 1973; Rosselli & Ardila, 2003). Thus, it is evident that education and a cultural background that deviates from the predominant Western culture in the USA are likely to impact cognitive performance, resulting in a screening test not capturing one’s true
ability. Additionally, cultural variables may strongly influence performance on particular tasks, potentially having the ability to highlight deficits in one cognitive domain versus another primarily due to the task-culture interaction, potentially skewing the understanding of one’s dementing process towards a specific etiology.

Another demographic variable that has been shown to consistently affect cognitive performance is sex, with literature suggesting that females perform better on verbal tasks and males perform better on visuospatial tasks (Kramer, Delis, & Daniel, 1988; Herlitz, Airaksinen, & Nordstrom, 1999). Studies have shown that females are better at immediate and delayed word and story recall tasks (Herlitz et al., 1999; Kramer et al., 1988;), contributing to them being better than males at tasks of verbal episodic memory (Herlitz et al., 1997; 1999). In contrast, males tend to outperform females on tasks of visuospatial ability (Halpern, 1997; Linn & Petersen, 1985). Thus, females tend to perform better on tasks of verbal episodic memory whereas males excel on tasks of visuospatial episodic memory (Lewin, Wolgers, & Herlitz, 2001). With respect to a cognitive screen, items on the MMSE, including the serial subtraction and language items, have been shown to be biased by sex (Jones & Gallo, 2002).

Along with easily identifiable sources of possible bias such as sex and education, mood can also bias cognitive functioning. It has been established that depression impacts verbal and visual episodic memory performance (Goodwin, 1997) as well as executive functioning abilities such as psychomotor speed and attention (Austin, Mitchell, & Goodwin, 2001). The number of cognitive abilities and the degree to which they are affected is dependent on the severity of depression and how it is being managed. Similarly, anxiety can also affect aspects of cognitive performance such as attention,
inhibition, and processing speed (Eysenck, Derakshan, Santos, & Calvo, 2007). Though anxiety and depression can impact cognitive performance, it is important to identify whether items on cognitive tests are in fact biased by mood symptoms.

**Differential Diagnosis**

Most neurodegenerative etiologies of dementia, including AD, FTD, VaD, and dementia with Lewy bodies (DLB), have distinct neuropathologies, but are all generally identified by progressive impairment in cognitive, behavioral and motor functions (Alloul et al., 1998; Evans et al., 1989).

**Cognitive profiles of neurodegenerative disorders.** AD is characterized by prominent deficits in language, memory, visuospatial abilities, semantic knowledge, executive functioning, and attention (Salmon & Bondi, 2009). In comparison to AD, DLB is characterized by greater impairment on tasks of attention, executive functioning, and visuospatial functioning (Gurnani & Gavett, 2016; Tiraboschi & Guerra, 2010). However, these deficits are apparent for typical presentations of AD and DLB, and the core clinical features of DLB that do not typically overlap with AD such as cognitive fluctuations, visual hallucination, and parkinsonism are present at a low rate (McKeith et al., 2005). A typical clinical manifestation of FTD in its early stages includes personality changes with apathy, disinhibition, impaired judgment, stereotypy, and progressive memory problems that tend to be associated with impulsivity and an inability to monitor responses (Elfgren et al., 1994; Hutchinson & Mathais, 2007). Neuropsychologically, compared to AD, individuals with FTD benefit from cues and show greater impairments in executive functioning and language with more intact visuospatial ability and verbal memory (Hou, Carlin, & Miller, 2004, Elfgren et al., 1994).
Similar to FTD, the clinical presentation of VaD overlaps with AD to some extent, though the onset of VaD in many cases is acute with stepwise progression of symptoms and relatively preserved episodic memory along with a history of cerebrovascular disease (Micieli, 2006). When comparing the cognitive functioning of AD and VaD, some studies suggest that individuals with AD show more memory problems with decreased delayed recall and increased word finding difficulty, while VaD is characterized by greater executive dysfunction (Looi & Sachdev, 1999; Mathias & Burke, 2009). However, other studies have indicated that the two diseases do not differ with respect to their cognitive functioning (Fahlander, Wahlin, Almkvist, & Backman, 2002). Though the clinical course of AD is better characterized than other forms of dementia due to its disproportionate prevalence compared to other dementing processes, the insidious onset and gradual cognitive decline of most neurodegenerative diseases makes accurate diagnosis challenging (Karantzoulis & Galvin, 2011; Salmon & Bondi, 2009).

**Importance of differential diagnosis.** AD can have diverse clinical presentations, especially in the early stages of the disease, with the possibility of other dementia etiologies mimicking AD at the time of initial evaluation and further complicating differential diagnosis (Karantzoulis & Galvin, 2011). Though neurodegenerative disorders have some distinct pathologies, diagnosis in a majority of cases is made on the basis of clinical symptomatology alone (Johns et al., 2009). Core clinical features that can help differentiate between various dementia etiologies may not be apparent until further along in the disease process. For example, despite considerable clinical and research efforts in operationalizing cognitive profiles of dementing diseases
such as FTD and VaD, neuropsychological criteria cannot consistently accurately
distinguish between AD, VaD, and FTD (Kertesz, Nadkarni, Davidson, & Thomas,
2000). With the exception of brain biopsy, there are no definitive in vivo biological
markers for the diagnosis of AD, making it challenging to accurately distinguish AD
from other neurodegenerative diseases and from normal aging (Kraybill et al., 2005;
Mollenhauer et al., 2010; Salmon & Bondi, 2009; Sloane et al., 2002).

Differential diagnosis is important since treatment for neurodegenerative
conditions offers modest improvement of symptoms, making it important to acquire
better knowledge of factors that may regulate the process of aging. The accurate
diagnosis of the dementia etiology is imperative for prognosis, symptom management,
and application of disease-specific treatments.

For example, individuals with DLB respond better than AD to cholinesterase
inhibitor treatment with early aggressive treatment offering modest control over some of
the cognitive, psychotic, and delirium-type features seen in DLB (Galvin et al., 2008;
McKeith et al., 2005). Accurate diagnosis is also imperative to help connect individuals
with appropriate resources and support services, thereby improving and maintaining
quality of life (AA, 2016). In addition to neurodegenerative causes of dementia that are
irreversible, 20 percent of dementia causes are reversible and treatable such as delirium,
depression, and side-effects of medications (Gershon & Herman, 1982; Personal Social
Services Research Unit, 2005). For example, depression is a common risk factor for AD
and approximately 40% of individuals with AD are also significantly depressed (AA,
2017). To complicate the picture further, depression is often difficult to distinguish from
AD because individuals with depression are often vulnerable to tasks of episodic memory
that are also typically impaired in AD (Zakzanis, Leach, & Kaplan, 1998). Since depression is treatable and reversible, it is important that individuals are not falsely identified as having a neurodegenerative disease (Swainson et al., 2001). Overall, dementing illnesses present with similar clinical presentations early in the disease process when therapeutic targets have maximum efficacy, highlighting the importance of neuropsychological tests in detecting subtle changes in cognitive ability to detect dementia, followed by using neuropsychological test results to characterize various causes of dementia.

**Importance of process scores in differential diagnosis.** The diagnosis of dementia, or even MCI, and its underlying etiology, is primarily dependent on traditional quantitative neuropsychological test scores that are sometimes unable to capture subtle cognitive changes occurring in the earliest stages of the pre-clinical phase of a disease (Au & Devine, 2013). Underlying cognitive processes can be better understood by analyzing process variables such as the nature of generated errors to better understand underlying cognitive processes (Ashendorf, Swenson, & Libon, 2013; Kaplan, 1988). One such technique of examining process variables is the Boston Process Approach (BPA), which has demonstrated its utility in clinical and research settings in differential diagnosis and conversion of MCI to dementia (Au & Devine, 2013).

Process scores have been derived from frequently used neuropsychological tests and have been shown to detect changes in several cognitive domains before the presentation of clinical dementia. In the domain of memory, process scores associated with early detection include those observed on memory recognition trials (Bielak, Hultsch, Kadlec, & Strauss, 2007), errors and patterns of saving and forgetting on a serial
list learning test (Libon et al., 2011), and proactive interference on a verbal memory test (Loweinstein et al., 2004). Non-demented older adults with relatively intact memory performance but with presence of intrusion errors as evident on cued recall trials are more likely to progress to AD (Bondi, Salmon, Galasko, Thomas, & Thal, 1999). Process score analysis of executive functioning measures including errors of commission on tests of inhibition and cognitive flexibility (Wetter et al., 2005), non-perseverative errors (Nagahama et al., 2003), and cognitive asymmetry as depicted on dual-task executive function tests (Houston et al., 2005), have all shown increased utility in detecting pre-clinical AD.

Though the majority of research has focused on the conversion of MCI to AD, progression to a non-AD dementia is also a likely possibility that is often ignored or minimally investigated. Libon et al. (2010) were able to extract three MCI profiles – amnestic, dysexecutive, and multi-domain – from performance on tasks of executive control, confrontation naming, verbal fluency, and serial list learning. Process score analysis of cognitive performance on a few tasks can be used to further understand neuropsychological profiles; for example, process variables on a serial list-learning test can guide identification of MCI subtypes, though examination of various constructs provides a more reliable estimate of underlying pathology (Libon et al., 2010; 2011). For example, examination of performance on serial list learning, category fluency, and confrontation naming, has been used to identify distinct pathologies of MCI. Reduced savings, increased forgetting, high number of intrusion errors, and excessive responses to recognition foils on serial list learning, as well as reduced output on category fluency, with intact performance on lexical fluency and naming, is characteristic of an amnestic
type of MCI that is typically linked with medial temporal lobe pathology (Libon et al., 2010; 2011). In comparison, mild impairment on category fluency along with reductions in confrontation naming and mental control, and relatively intact performance on serial list-learning with endorsement of interference foils on delayed recall is indicative of dysexecutive MCI that is linked with frontal lobe pathology (Baldo, Delis, Kramer, & Shimamura, 2002; Libon et al., 2010; 2011). Though mixed MCI is typically related to subcortical pathology, it presents like dysexecutive MCI with mildly impaired performance on naming and category fluency. Process scores on serial list learning, such as relatively intact savings and fewer intrusion errors that are more subordinate in nature with endorsement of recognition foils, contribute to a more distinct dysexecutive MCI profile (Libon et al., 2010; 2011; Salmon & Bondi, 2009).

Error analysis guides differential diagnosis of neurological conditions when total achievement scores are unable to detect executive dysfunction (Possin & Kramer, 2013). Normative or quantitative neuropsychological data is important in providing overall information about cognitive domains, but as described above, process variables provide information over and above that offered by a statistical score (Ashendorf et al., 2013). For example, while the quantitative score on a test of concept formation may suggest difficulties with abstract reasoning, an analysis of errors helps differentiate between VaD and AD; in-set errors (for example, stating the similarity between a dog and a lion is that they eat, thereby providing a superordinate categorical response) are more common in AD whereas out-of-set errors (for example, stating that the similarity between eyes and ears is that we see with eyes and hear with ears, thereby describing how the items in the presented word pair are different) are indicative of VaD (Giovanetti et al., 2001).
Similarly, the quantitative score on a test of verbal fluency may indicate cognitive impairment, but the discrepancy in performance on category and lexical fluency suggests the presence of AD whereas reductions in both lexical and category fluency are indicative of VaD (Herbert, Brookes, Markus, & Morris, 2014). In addition, individuals with certain etiologies are likely to make specific types of errors irrespective of the test. For example, individuals with VaD produce more perseverative errors than those with AD. Additionally, perseverative responses in VaD resemble persistent motor responses whereas perseverations in AD are semantically based (Lamar, Price, Giovannetti, Swenson, & Libon, 2010). Thus, the amount of information obtained from process score analysis of test performance is also dependent on the ability of the examiner and the cognitive test to capture process variables.

On tests of learning and memory, individuals with VaD exhibit the same level of impairment on free recall trials as in AD. Similar to Huntington’s disease (HD) and Parkinson’s disease (PD), individuals with VaD have higher saving scores and recognition discrimination scores than AD (Libon et al., 1998; 2001). In comparison, AD results in the production of initial intrusion errors followed by perseveration of these intrusion errors across learning trials (Davis, Price, Kaplan, & Libon, 2002). List learning tests such as the California Verbal Learning Test (CVLT) have been imperative in facilitating the inclusion and analysis of process variables, resulting in examination of differential performance in various dementia etiologies (Delis, Kramer, Kaplan, & Thompkins, 1987). Though research regarding process variables of cognitive tests is limited, serial list learning tests have received the most attention with respect to analysis of process scores. Another type of memory test that is often administered is story or
paragraph recall such as the Wechsler Memory Scale Logical Memory subtest (LM; Wechsler, 1945); an exploratory error analysis of errors on LM revealed that increased volume around the temporal horns – an area of the brain typically affected in AD – was associated with greater number of correct responses, whereas an increase in the number of related errors, such as related confabulations, was associated with an increase in white matter hyperintensities (Libon et al., 2015). The association between brain region and errors suggests that error analysis does not only assist with the characterization of dementia, but also with identification of pre-clinical dementia when quantitative scores look stable (Libon et al., 2015).

Process scores also help with differential diagnosis in atypical presentations of dementia. Typically, AD is distinguishable from FTD due to the frank episodic memory impairment that is characteristic of the former etiology (Perry & Hodges, 2000). However, some research suggests that frontal executive difficulties may be the earliest feature of AD, typically in a younger AD cohort, thereby complicating its differentiation from FTD (Nedjam, Devouche, & Della Barba, 2004; Perry & Hodges, 2000). Individuals with both AD and FTD exhibit difficulties on tasks of working memory and attention; however, the differences in their process scores aid in differential diagnosis. For example, individuals with FTD tend to exhibit sequencing errors on digit span, a test that requires individuals to recite presented digits forwards and backwards, and show reduced performance on a working memory task only when distracted; in comparison, greater omission errors are seen on digit span in AD along with impaired performance on working memory even without delay or distraction (Stopford, Thompson, Neary, Richardson, & Snowden, 2012). Similarly, process variables aid in the differentiation
between AD and DLB; research suggests that both groups show intrusion errors, though individuals with DLB tend to exhibit a higher frequency of errors and more environmentally-cued intrusions than AD (Doubleday, Snowden, Varma, & Neary, 2002).

Overall, it is evident that an analysis of frequency and type of errors assists in the early detection of dementia and in differential diagnosis. Information about cognitive profiles of various dementing illnesses allows for further characterization of cognitive deficits with process variables and error analysis improving detection and determination of level of severity and etiology earlier in the disease process. Knowledge about precise cognitive abilities also helps in the generation of deficit-specific treatment recommendations for better management of symptoms and improvement of overall quality of life.

Development of the Colorado Cognitive Assessment (CoCA)

The CoCA was developed to address gaps in current cognitive screens. As described above, current cognitive screens are lacking in one or more of the following: coverage of a breadth of domains, coverage of various frontal/executive abilities, and influence of sociodemographic factors on performance. To the best of the author’s knowledge, no cognitive screening test provides a way to identify and score process variables. In addition, while performance on cognitive screening tests may correlate with functional outcomes, tasks assessing functional abilities are not included as part of any known cognitive screens. Furthermore, most cognitive screens are scored using CTT, which makes important untested assumptions. The dependency of item and person statistics in CTT poses challenges in estimating a person’s true ability (MacDonald &
Thus, the CoCA was developed to serve as a comprehensive cognitive screen that can be used in primary care settings, general practice, research studies, memory clinic evaluations, and neuropsychological evaluations to help with accurate detection of dementia, possibly earlier in the disease process; to assist with differential diagnosis; to be utilized in longitudinal research designs; to aid in the identification of cognitive strengths and weaknesses; and to be minimally influenced by systematic biases related to sex, education, age, or culture. Based on the limitations of current cognitive screens, tasks in the CoCA were chosen and developed to meet the needs identified above.

Tasks in the CoCA by cognitive domain. There is general consensus that only a handful of current cognitive screening instruments sufficiently assess a wide range of cognitive abilities, with most tools examining frontal/executive dysfunction to a limited extent (Shulman et al., 2006). The MoCA (Nasreddine et al., 2005), ACE-R (Mioshi et al., 2006), the Alzheimer’s Disease Assessment Scale – Cognitive section (ADAS-Cog; Rosen, Mohs, & Davis, 1984), and the Cambridge Assessment of Memory and Cognition (CAMDEX; Roth et al. 1986), are some of the current cognitive screening tools that cover the five cognitive domains of orientation, memory, executive functioning, visuospatial ability, and language, with varying emphases, but with more depth than the MMSE. Given that dementia is characterized by deterioration in global cognition, with individual and etiological differences in specific cognitive domains, an effort was made to include tasks assessing cognitive abilities affected in the major dementia syndromes in the CoCA. Overall, tasks in each domain were chosen bearing in mind maximization of the attainment of process scores and minimization of the influence of culture, sex, and
education on performance. With respect to the former, process scores allow for examination of unique variance in neuropsychological test performance, with recent research suggesting that it is the unique variance across a battery – rather than shared variance – that contributes best to differential diagnosis (John et al., 2016).

**Memory.** Memory disturbance is considered to be the most common symptom of AD, and a common symptom of other dementia etiologies as well; thus, most cognitive screening instruments assess memory functioning in some way, though not all of them adequately examine different aspects of memory (Salmon & Bondi, 2009). For example, to the best of the author’s knowledge, no cognitive screening instrument assesses visual learning. Deficits in visual memory are evident in dementias such as AD and DLB, and help with differential diagnosis, for example, between DLB and PDD (Kawas et al., 2003; Mondon et al., 2007). Visual memory deficits are also evident in those with lateralized lesions, with right hemispheric lesions resulting in impaired visual memory and relatively preserved verbal memory in comparison with those with left hemispheric lesions (Trahan, Larabee, & Quinatana, 1990). In addition, differences exist with respect to individuals verbal and visual memory abilities (Mayer & Massa, 2003); thus, a visual memory task was included in the CoCA in order to maximize information obtained about the patient’s memory. Like visual memory, prospective memory is not assessed in any known cognitive screening test at the present time; due to its high ecological validity and vulnerability to early stage dementia pathology (Huppert, Johnson, & Nickson, 2001; Spooner & Pachana, 2006), a test of prospective memory was included in the CoCA.

List learning tests assess multiple important aspects of memory, including encoding, forgetting, retention, and retrieval, all of which are impacted to varying degrees
in different dementia etiologies. List learning tests have consistently been shown to
differentiate between AD, MCI, and controls with adequate sensitivity and specificity (de
Jager, Hogervorst, Combrinck, & Budge, 2003; Karrasch, Sinerva, Gronholm, Rinne, &
Laine, 2005; Salmon et al., 2002). Though list-learning tests are included in most
cognitive screens, tests differ with respect to their length and the number of learning
trials; for example, the MMSE and the ACE-R provide three words with one learning
trial, whereas the MoCA and ADAS-COG provide five and ten words with two and three
learning trials, respectively (Folstein et al., 1975; Mioshi et al., 2006; Nasreddine et al.,
2005; Rosen et al., 1984). The length of the serial list-learning test in the CoCA was
chosen based on the average word span for normal adults, which is considered to be
between five and nine words (Kausler & Kausler, 2001; Miller, 1956). Three learning
trials were included since research suggests that the first trial is typically more reflective
of attentional rather than memory processes, allowing for two additional learning trials
(Stepanov, Abramson, Wolf, & Convit, 2010). List-learning tests can hold several biases
associated with education and culture. In fact, the second revision of the CVLT (CVLT-
II) included a new word list that was easier and less biased by geographical and
sociodemographic factors (Baños & Martin, 2002).

When developing the list-learning task for the CoCA, words that were thought to
be minimally influenced by geographical and sociodemographic backgrounds, and
literacy and education levels were chosen. Moreover, tests differ in the type and number
of delayed recall and recognition trials administered. In order to maximize information
obtained from process scores, free delayed recall, cued recall, and multiple-choice
recognition, as well as forced-choice recognition, were included in the CoCA; the forced-
choice recognition task may serve as an embedded measure of performance validity (Ashendorf & Sugarman, 2016), another potentially unique feature of the CoCA.

**Language.** Word-finding difficulty, comprehension deficits, and non-fluent speech are common features of language often affected in individuals with neurodegenerative disorders, stroke, traumatic brain injury, and epilepsy (Kempler & Goral, 2008; Yochim et al., 2015). To assess language production, tasks of semantic fluency and confrontation naming were included in the CoCA. A test unique to the CoCA is the non-visual assessment of confrontation naming; visual impairment is common in older adults, but tests of naming typically involve the visual identification of pictures. Thus, the validity of assessment of naming ability in a substantial portion of older adults is often affected (Ryskulova et al., 2008; Yochim et al., 2015). While a task of semantic fluency is relatively less biased, naming tests are often biased by cultural variables (e.g., the Boston Naming Test; Barker-Collo, 2000). Thus, a non-visual test of confrontation naming drawn from Yochim et al.’s (2015) Verbal Naming Test (VNT), was included in the CoCA. Though 15 items that were considered less influenced by education and culture were chosen from the VNT, the goal was to identify three to five items with the least bias in the CoCA for the final version of this test.

**Visuospatial.** Visuospatial deficits are often early markers of dementia, with a wide range of visual processes being impacted in various dementing illnesses (Possin, Laluz, Alcanter, Miler, & Kramer, 2011). Most cognitive screens include a figure construction task to assess visuospatial abilities; clock construction is included in the MoCA, SLUMS, and ACE-R (Nasreddine et al., 2005; Tariq et al., 2006), whereas the MMSE and ACE-R include the copy of a simple geometric figure (Folstein et al., 1975;
Mioshi et al., 2005). In order to assess visuoconstructional and visuospatial ability and examine change from construction to copy of the same object, the construction and copy of a clock were both included in the CoCA. As described earlier, the CDT has been consistently shown to inform important aspects of cognitive functioning and was chosen due to richness in process information obtained from this test. Comparison of the copy and construction versions of this test also provides important process variables that guide differential diagnosis (Libon, Swenson, Barnoski, & Sands, 1993; Libon, Malamut, Swenson, Sands, & Cloud, 1996; Royall, Cordes, & Polk, 1998). Rouleau, Salmon, Butters, Kennedy, and McGuire (1992) were the first to compare CDT related process variables between different dementia etiologies; for example, errors demonstrated by those with AD were typically stimulus-bound in nature and reflective of conceptual deficits whereas individuals with HD made errors that were reflective of poor planning. Also, in comparison to individuals with subcortical deficits, such as those with VaD and HD, individuals with AD demonstrate improvement from clock construction to clock copy (Libon et al., 1993; 1996). The inclusion of a copy trial in the CoCA is thus an advantage over the MoCA and SLUMS, which only consist of clock construction. In addition, a unique feature of the CoCA is the inclusion of a novel complex figure to assess visuospatial and visuoperceptual abilities; the development of this figure is based on the widely used Rey Osterreith Complex Figure Test, performance on which has been shown to be vulnerable to various dementia etiologies (ROCF; Possin et al., 2011; Rey, 1941). Specific features of the figure in the CoCA were developed to obtain information about visuospatial skills, such as visual perception and integration that may not be assessed in other tasks of visuospatial functioning. In addition, the complexity and details
of the figure were developed with the aim of capturing frontally-mediated executive skills. Though some studies have shown an individual’s cultural background can influence the perception of the structure of the figure (Bossuroy, Wallon, Falissard, & Moro, 2014), this primarily affects the approach and strategies used for figure copy, further reinforcing the case for the examination of process scores.

**Executive Function.** Executive functioning is commonly affected in MCI and has been found to predict conversion to dementia (Brandt et al., 2009). Aspects of executive functioning are assessed to varying degrees in cognitive screens; for example, vigilance, focused attention, working memory, lexical fluency, and abstract reasoning are assessed in the MoCA, whereas working memory is the only aspect of executive functioning assessed in the MMSE (Folstein et al., 1975; Nasreddine et al., 2005). In general, cognitive screens do not assess a broad range of executive functioning abilities (Shulman et al., 2006). In an attempt to ensure that the CoCA assesses a breadth of executive functioning abilities, tasks assessing cognitive flexibility, psychomotor speed, working memory, focused attention, abstract reasoning, and lexical fluency were included. In general, tasks on the CoCA were also chosen to maximize process scores that primarily provide information about executive dysfunction.

Digit span forward is often used as a measure of focused attention in various cognitive screens (e.g., MoCA); however, many of the process scores obtained from this test are only attainable after the administration of several trials, which is not conducive to the brief administration time desired in a cognitive screening tool. Additionally, research suggests that the clinical utility of digit span forward is questionable, especially for differential diagnosis (Groth-Marnat & Baker, 2003; John et al., 2016). In order to
overcome these limitations, another test of everyday attention and vigilance based on the Stop and Go Switch Task of the BTACT (Tun & Lachman, 2006) was developed for the CoCA. The Stop and Go Switch Task is considered to be relatively free of cultural- and education-related bias and requires the examinee to respond with “go” and “stop” to the words green and red, respectively.

Working memory deficits are common in early stages of dementia, though commonly used tasks of working memory such as digit span backward and arithmetic, have been found to load on to the language factor rather than executive functioning or memory (Stopford et al., 2008). As with digit span forward, administering only a couple of trials of digit span backward provides inadequate information about the underlying cognitive ability. Thus, another test of working memory was developed for the CoCA; this task was developed to be comparatively more difficult than other tests of working memory so as to capture cognitive deficits early in the disease process. Specifically, this task uses both letters and numbers and requires the examinee to recite only the numbers in the backwards order, thereby assessing selective attention and working memory. Lamar et al. (2007) showed that serial order digit recall and digit recall in any order are associated with different levels of subcortical pathology. Thus, the scoring of this item allows for partial credit, thereby allowing for a large variance in cognitive abilities based on work by Lamar et al. (2007), who have shown that recalling digits on span tasks in any order versus serial order represents distinct abilities of executive functioning.

As a test of psychomotor speed, a task similar to the Symbol Digit Modalities Test (SDMT; Smith, 1968; 1982) and Digit Symbol Substitution Test (DSST; Wechsler, 1981) was developed. The SDMT and DSST ask examinees to fill in numbers associated
with specific symbols or vice versa using a provided key within a short time frame. In a recent meta-analysis that examined the utility of various neuropsychological tests in detecting pathologically confirmed AD from normal aging, the DSST was shown to have the largest effect size in distinguishing between AD and cognitively healthy individuals (Gurnani & Gavett, 2016).

Similar to other cognitive screens, a test of abstract reasoning and verbal concept formation was included in the CoCA. In this task, the examinee is asked to explain how two non-obviously related items are similar with the goal of being able to identify the most pertinent features of the objects or concepts (Giovanetti et al., 2001). In addition to assessing knowledge about objects/concepts, this task also requires the examinee to process multiple pieces of information at the same time (Giovanetti et al., 2001). The items for the CoCA were developed to capture process variables that would provide more information about frontal-executive abilities.

Another test used frequently in conjunction with and included as part of cognitive screens is the Trail Making Test – B (TMT-B; Reitan & Wolfson, 1985). The TMT-B requires the examinee to connect circled numbers and letters in an alternating number to letter sequence as fast as possible. The TMT-B provides information about skills that are typically affected in AD and FTD, such as visual scanning, processing speed, cognitive flexibility, set-shifting, and set maintenance (Albert, Moss, Tanzi, & Jones, 2001; Perry & Hodges, 2000), and is one of the few neuropsychological tests consistently shown to be associated with driving ability (Reitan & Wolfson, 1985; Silva, Laks, & Engelhardt, 2009). However, its utility is limited in cross-cultural contexts due to the use of the English alphabet. Though a more culturally fair version of the TMT known as the Color
Trails Test (CTT) has been developed, its version that is similar to TMT-B (CTT-2) has not been found to be equivalent to TMT-B (Dugbartey, Towns, & Mahurin, 2000). In order to minimize the effect of education and culture, a more culturally appropriate version of this test, adapted from Zhao et al. (2013), was used in the CoCA.

Another commonly used test of executive functioning is lexical fluency, which assesses cognitive speed, strategy utilization, and response inhibition, abilities that are often affected in dementia (Abwender, Swan, Bowerman, & Connolly, 2001). This task asks the examinee to recite words beginning with a specific letter in one minute. Though similar in structure to animal fluency, this task provides the examinee with certain rules that make it more challenging to come up with words that may be easier to generate such as proper nouns. This test was chosen due to its ability to provide qualitatively rich information about executive control and to obtain a discrepancy score between lexical and semantic fluency, which assists in differential diagnosis (Abwender et al., 2001).

Orientation. Depending on dementia severity and level of engagement, individuals may not be oriented to their surroundings. Thus, a task assessing orientation to situation, time, and place was included in the CoCA.

Functional Ability. A recent advancement in the field of early detection of dementia involves the development of the latent dementia phenotype, “$\delta$”, by Royall and colleagues, who conceptualize the process of development of dementia as the “cognitive correlates of functional status” (Royall & Palmer, 2012; Royall, Palmer, Vidoni, Honea, & Burns, 2012). The concomitant presence of early and subtle cognitive and functional changes is the most distinguishing feature of dementia (Gavett et al., 2015; Royall & Palmer, 2012). $\delta$ is derived from Spearman’s general intelligence factor, “$g$,” which can
be understood as the construct representing the shared variance across observed psychometric performance (Royall et al., 2012; Spearman, 1904). Spearman’s \( g \) is further subdivided into two independent constructs, \( g' \) and \( \delta \). \( \delta \) is the shared variance between cognitive and functional measures, whereas \( g' \) is the variance in cognitive ability that is not associated with functional decline that occurs as a result of neurodegenerative illness (Royall et al., 2007). Though \( \delta \) and \( g' \) are similar in that they are both latent constructs representing psychometric performance, they differ slightly in that only \( \delta \) is thought to underlie one’s ability to perform instrumental activities of daily living (IADL).

\( \delta \) has been consistently shown to have an independent and strong association with a measure of dementia severity, the Clinical Dementia Rating – Sum of Boxes (CDR-SOB; Gavett et al., 2015; Royall et al., 2012). \( \delta \) has consistently distinguished between those with and without clinically diagnosed dementia, with varying neuropsychological tests, and has been validated in various samples cross-sectionally and longitudinally, including healthy controls and those with different dementia etiologies (Gavett et al., 2015; Gurnani, John, & Gavett, 2015; Gurnani et al., 2015; Royall & Palmer, 2012; Royall et al., 2012; Royall, Palmer, Vidonu, & Honea, 2013). Given the strong support surrounding \( \delta \) as an aid in the early identification of dementia, a test of functional ability was developed for the CoCA. In this way, the global score on the CoCA would be representative of simultaneous performance in cognitive and functional abilities. The task assessing functional ability in the CoCA was developed to mirror an instrumental activity of daily living. Scoring of this item allows for partial credit, allowing for increased variance in functional abilities; this type of scoring also helps with serial assessments to measure changes in functional and cognitive ability simultaneously. To the best of the
author’s knowledge, no cognitive screen at the present time includes a measure of functional ability, making this a unique feature of the CoCA.

Psychometric Properties of the CoCA

As briefly discussed earlier, most tests are scored using CTT. The total score for a test using CTT is usually the summation of correct responses (Gavett, Crane, & Dams O’Connor, 2013). The total score in CTT is referred to as the observed score, which reflects the unobserved variable of interest (true score) plus error due to several factors that could influence the observed score (DeVellis, 2006). CTT poses several advantages such as general familiarity with its basic concepts and methodology, adequate model fit for certain types of instruments, and modest relationships between items and the underlying variable of interest (DeVellis, 2006). Disadvantages include frequent violation of restrictive and untested assumptions about the relationship between test items and the overall construct being measured, and the lack of laborious scrutiny of item characteristics that may be available through other statistical techniques (DeVellis, 2006; Gavett et al., 2013). Typically in CTT, each item contributes equally to the total score irrespective of the level of item difficulty. In addition, CTT does not account for the ability level of the participant and assumes that the test is equally precise in measuring the variable of interest for people with differing abilities (DeVellis, 2006). Total scores arrived at using CTT often assume linear measurement properties, meaning that a two point difference at the center of the scale is interpreted to mean the same thing about the true score difference as a two point difference at either the high or low ends of the scale (DeVellis, 2006). However, tests developed using CTT rarely, if ever, possess linear measurement properties in reality. Linear measurement properties are desirable when
considering scoring of cognitive screening tests that are likely used in serial assessments (Mungas & Reed, 2000). Given the impact of the disadvantages described above on assessing true cognitive performance, it would be beneficial to score the CoCA using modern psychometrics to offset some of the primary limitations of CTT.

**Confirmatory factor analysis (CFA).** In contrast with CTT, confirmatory factor analysis (CFA) takes into account item difficulty when using factor scores to estimate an examinee’s ability level. CFA allows for the modeling of the relationship between observed measures such as test items or test scores and the latent variables (e.g., memory) believed to underlie performance on tests (Kay, 2004). An important feature of CFA is that measurement error is considered to be a property of the observed variables, not the latent variables; therefore, scores produced using this method are only affected by error in estimation, not by error in measurement.

One way to apply CFA is through the use of a bi-factor model, which usually includes a general factor that is representative of all the cognitive abilities measured by the test along with a secondary factor or group of factors that either reflect subdomains of cognition or method effects, such as the same list of words used to generate multiple test scores (Gavett et al., 2013; Jennrich & Bentler, 2011). A bi-factor model will be ideal for the CoCA given the various cognitive subdomains and method effects between tests. Validating the CoCA using a bi-factor model (Figure 1) remains a long-term goal; however, given the preliminary nature of the present study, a unidimensional CFA model was chosen (Figure 2). While not ideal, the use of unidimensional CFA is still a more sophisticated and psychometrically sound approach to scale validation than CTT.
Figure 1. Model to be validated in the future using confirmatory factor analysis (CFA). $g$ is the latent variable representing general cognitive ability. Cognitive domains of memory, language, visuospatial ability, and executive function are latent variables modeled as uncorrelated factors. Cognitive tests are the observed variables. VLL-IR = Verbal List Learning Immediate Recall; VLL-DR = Verbal List Learning Delayed Recall; VLL-Y/N = Verbal List Learning Yes/No Recognition; PMT = Prospective Memory Test; C-Command = Clock Command; C-Copy = Clock Copy; VNT = Verbal Naming Test; Animals = Animal Fluency; STT = Shape Trails Test; WM = Test of Working Memory; FA = Test of Focused Attention; NSMT = Number Symbol Matching Test; “S” Words = Lexical Fluency; TFA = Test of Functional Abilities
Figure 2. Proposed model of the CoCA that was planned for the current study and examined using confirmatory factor analysis (CFA). $g$ is the latent variable representing general cognitive ability. Cognitive tests are the observed variables. VLL = Verbal List Learning Immediate Recall + Delayed Recall + Yes/No Recognition; PMT = Prospective Memory Test; CDT = Clock Command + Copy; VNT = Verbal Naming Test; Fluency = Animal + Lexical Fluency; STT = Shape Trails Test; WM = Test of Working Memory; FA = Test of Focused Attention; NSMT = Number Symbol Matching Test; TFA = Test of Functional Abilities

**Measurement Invariance (MI) and Differential Item Functioning (DIF).** In order to justify that a scale or cognitive screening tool is measuring the same construct in all groups, one needs to establish that the test items measure the same construct in the same way across groups, (e.g., male vs. female) such that any differences or similarities between groups are both meaningful and interpretable (Milfont & Fischer, 2010). Lack of MI occurs when a test item functions differently for different groups with the same ability
level, suggesting that the item has distinct characteristics based on group membership (Hortensius, 2012). In order to compare the CoCA across different groups, each observed item in the CFA model must relate to $g$ in the same way. The CoCA will be considered to be measurement invariant when all the test items indicated in the CFA model of the CoCA relate to the latent factor of CoCA global ability ($g$) in the same way (Milfont & Fischer, 2010). MI analyses are used to assess differential item functioning (DIF), which occurs when members belonging to different groups with the same ability level have unequal probabilities of responding to an item or items on a measure or test (Brown, 2012).

There are two types of DIF: uniform and non-uniform. Uniform DIF occurs when one group has a higher probability of responding correctly to an item over another group across all ability levels (Swaminathan & Rogers, 1990). Non-uniform DIF occurs when the probability of one group providing a correct response to an item over another group is different at different ability levels (Swaminathan & Rogers, 1990). There is no interaction between group membership and ability level in uniform DIF, whereas group membership and ability level interact in non-uniform DIF (Swaminathan & Rogers, 1990). Thus, differential functioning of an item based on group membership results in the interpretations and conclusions made based on the test to be potentially artifactual and misleading.

In order for the CoCA to be a reliable and valid measure of cognitive ability, it should be able to produce precise measurements of cognitive functioning that can be meaningfully compared across various groups. The CoCA would be biased if one group (e.g., males) performed better or worse than another group (e.g., females) under certain
circumstances resulting in differences in total CoCA scores for males and females not truly reflecting differences in cognitive ability, but rather reflecting the effects of test bias (Bauer, 2016). One of the proposed unique characteristics of items on the CoCA is the consideration given to include tasks that are minimally influenced by sociodemographic variables including sex, education, age, race, and ethnicity. Thus, it would be imperative to establish MI for all items on the CoCA with respect to the variables mentioned above.

The Present Study

There is an imminent need for a cognitive screening instrument that can aid in the accurate and early detection of dementia and its etiologies. Though several cognitive screens have been developed over the past few decades, there still exists a gap for a tool that has the potential to overcome limitations of current screening instruments. Thus, the CoCA was developed to serve as a comprehensive cognitive screening instrument that differs from existing screens by: (1) assessing a wide range of cognitive abilities; (2) using tasks that are relatively less biased by education, sex, and culture; (3) providing a method to identify and record process variables; and (4) using modern psychometrics for scoring, thereby assisting in the early identification of dementia and in differential diagnosis. Although investigating the process scoring mechanism is an important aspect of the CoCA, it was not explored in the present study and is considered an important next-step in the development process of the CoCA. The primary purpose of the present study was to develop the CoCA and validate it in a sample of community-dwelling older adults using modern psychometrics (i.e., CFA). A second, related goal of the present study was to establish MI for the items of the CoCA with respect to education, age, and sex. As part of the validation process, another goal was to establish convergent validity of
the CoCA with the MoCA and Neuropsychological Assessment Battery (NAB) Judgment subtest, and divergent validity with respect to mood. Though examining the utility of the CoCA in differential diagnosis is outside the scope of the current study, it is considered to be the next step in the validation process.
CHAPTER II

METHOD

Participants

One hundred and fifty one older adults were recruited through UCCS and the larger Colorado Springs community using methods like newspaper advertisements and flyers. Inclusion criteria for study eligibility included (a) participant was 50 years or older; (b) participant had not had a concussion or traumatic brain injury in the past three months; (c) participant was not diagnosed with major depressive disorder or other mental health disorder; (d) participant was not diagnosed with a known neurological condition or neurodegenerative disease such as dementia; (e) participant could read and write in English; and (f) participant did not have any major problems with vision or hearing based on self-report. Thus, individuals below the age of 50 or with self-reported neurological or mental health illnesses were excluded from the present study. The Institutional Review Board (IRB) at UCCS approved the study for research with human subjects.

Measures

The CoCA was developed for the present study. The tasks used in the CoCA, along with the order of the tests and scoring of the CoCA, are described below. Other measures, including the Geriatric Depression Scale-15 (GDS-15), Geriatric Anxiety Scale -10 (GAS-10), Neuropsychological Assessment Battery (NAB) Judgment Subtest, the Montreal Cognitive Assessment (MoCA), and a demographics questionnaire were administered along with the CoCA, and are briefly described below.
CoCA. The CoCA was developed to identify the early stages of dementia with improved accuracy in comparison to commonly used cognitive screens such as the MMSE. Another aim of the CoCA was to assess relevant cognitive abilities using tasks that are free of or minimally influenced by culture, sex, and education. The CoCA was designed to assess memory, visuospatial ability, language, executive functioning, orientation, and functional ability. Though performance on each task typically involves utilization of skills belonging to various cognitive domains, the test was developed with the goal of having one primary domain underlying performance on each specific task. The tasks used in the CoCA are divided into primary domains as described below. The CoCA is available in Appendix A. Each task on the CoCA receives a total score. Each item on the CoCA along with scoring procedures is described below. In addition, process variables are identified for each item; a template for scoring process variables is available in Appendix B. The CoCA instruction and scoring manual is available in Appendix C.

Memory. The memory domain is made up of the following tasks:

1. **Prospective Memory Test (PMT):** This item assesses the examinee’s ability to remember to perform an action after the occurrence of a target event. In the CoCA, the examinee is asked to remind the examiner that he or she has to look at his or her bank statement when informed that the test has been completed. A total score is assigned for the delayed recall of this task to provide credit for remembering to make a payment. A maximum of one point can be attained on this task. There are no process scores for this test.

2. **Verbal List Learning (VLL) Immediate Recall:** This test assesses the examinee’s ability to recall seven words belonging to different categories immediately
after presentation. The words are administered over three learning trials. The total score for this task is a sum of all correctly recalled words across all three learning trials, with one point credited for every accurately recalled word. Maximum possible score on this task is 21. Process scores include number and type of intrusions (i.e., incorrect words recalled that are novel, and semantically-related or phonemically-related to a target word), and number of repetitions (recalling the same word more than once per trial).

3. **VLL Delayed Recall**: This task examines the examinee’s ability to retrieve previously presented words on the list after a brief delay. For words that were not freely recalled, the examinee is provided with category cues. The total score on this test is the sum of all accurately recalled words with and without a category cue, with one point given for every correct word. Maximum possible score on this task is 14. Process scores include number and type of intrusions (incorrect words recalled that were novel, and either semantically or phonemically related to the target word), intrusions on immediate recall that were repeated on delayed recall, and repetitions (recalling the same word more than once per trial).

4. **VLL Yes/No Recognition**: In this task, the examinee has to identify target words and deny the presence of distractor words when presented with a list of target and semantically-related distractor words in a yes/no format. The total score is the sum of all accurate responses, with one point credited for every correct response. A total of 14 points can be attained on this task. There are no process scores for this test.

5. **VLL Forced Choice Recognition**: In this task, the examinee has to choose the target word from a pair of words. The distractor words in this task are not related to the target word. The purpose of this item is to serve as an embedded performance validity
measure (this is similar to the procedure used by the California Verbal Learning Test-II [CVLT-II] to identify performance validity; see Delis, Kramer, Kaplan, & Ober, 2000; Moore & Donders, 2004; Root, Robbins, Chang, & Van Gorp, 2006); poor performance on this test might be suggestive of poor effort due to the minimal memory and learning demands of this task. The score from this task is not added towards the CoCA global score. A maximum of seven points can be obtained on this task. There are no process scores for this task.

6. **Figure Delayed Recall:** In this task, the examinee is instructed to draw a previously copied complex figure from memory. Performance on this task provides a composite score of visual memory; comparing performance on this test with **Figure Copy** (described below) provides further information about memory processes, including forgetting and retrieval. The figure is made up of nine components; the total score represents the aggregate of accuracy and placement scores for each component. A half point or full point can be earned for the accuracy of each component of the figure, and one full point can be earned for the placement of each figure component. A maximum of 18 points can be obtained on this task. Process scores include evidence for perseveration (repetitive behavior in the absence of a stimulus), poor planning, spatial errors, right-to-left organization, visual neglect, micrographia (small and cramped style of writing or drawing), tremors, and accuracy-to-placement ratio (which may be useful for identifying the location of a brain lesion, for example, as seen in dorsal vs. ventral stream impairment). Percent of information retained from the copy trial is also included as a process score for this test.
**Visuospatial ability:** The visuospatial domain consists of the following tasks:

1. **Clock Drawing Command and Copy:** This test assesses visuoconstructional, visuospatial, and visuoperceptual abilities. The examinee is first asked to draw the face of a clock and to set the hands to “ten after eleven,” followed by copying the drawing of a clock also set to the same time. The total score for this item is based on Rouleau et al.’s (1992) scoring system, which is available in Appendix C. A maximum of 10 points can be obtained on the clock copy and clock command. Process scores include evidence for inappropriate size of the clock, graphical errors, stimulus bound errors, conceptual errors, spatial or planning deficits, perseveration, neglect, self-correction, rotation of paper while drawing, change in performance from command to copy, copy-to-command ratio, and requesting a reminder regarding the time of the clock. Process errors for this test are explained in the instruction and scoring manual available in Appendix C.

2. **Figure Copy:** This test assesses visuoconstructional, visuospatial, and visuoperceptual abilities by asking the examinee to copy a more complex figure within a limited time (2 minutes). The total score is the summation of accuracy and placement scores for each of the nine components, with one point each given for accuracy and placement of each component. A half point or full point can be earned for the accuracy component of the figure, resulting in a maximum score of 18. Process scores include evidence for perseveration (repeating parts of the figure), poor planning, spatial errors, right to left organization, visual neglect, micrographia (small and cramped handwriting or drawing), motor issues, tremors, and accuracy/placement ratio.
**Language:** The language domain consists of the following tasks:

1. **Semantic Fluency (Animal Fluency):** In this task, the examinee is asked to generate as many animal words as possible in one minute. This test assesses generativity that is primarily dependent on intact semantic memory. The total score is the total number of correct words produced in one minute, with each accurate response counting for one point. Process scores include number of words generated every 15 seconds, perseverations (repeating the same word more than once), intrusions (incorrect words belonging to another category), and self-corrections. Percent set loss, semantic index (animal fluency/[animal + lexical fluency]), and percent perseveration are other process scores included in this test.

2. **Verbal Naming Test (VNT):** This test is a non-visual alternative to confrontation naming, requiring the examinee to identify the object being described by the examiner. Thus, this task is a verbal and non-visual way of assessing word-finding. Like confrontation naming, the examinee is provided with a phonemic cue if he or she is unable to recall a word within a specific amount of time. With the author’s permission, 15 items from the 60-item VNT (Yochim et al., 2015) were selected for the CoCA. Since the VNT is relatively new, research regarding the effects of education, sex, and culture on performance is limited. Thus, 15 items were selected for the purpose of this study, with the goal of using the data collected to identify three to five items that have the least education-, sex-, and cultural-related bias for inclusion on the final version of the CoCA. The total score on this test is the sum of all correctly identified words without phonemic cues, with each accurate response obtaining one point. A maximum of 15 points can be attained on this task. Process scores include number of phonemic cues given and the
number of correct words generated with phonemic cues. Other process variables include
evidence for phonemic (substitution of a target word with a word that contains some of
the syllables of the intended word) and semantic paraphasias (substitution of a target
word with another word that is from the same category as that of the intended word),
circumlocution (describing or explaining the target word without actually stating it), and
perseveration.

Executive function. The executive function domain consists of the following
tasks

1. Shape Trails Test (STT; Zhao et al., 2013): This test is a culturally fair
alternative to the Trail Making Test – Part B (TMT-B), which assesses set-shifting,
cognitive flexibility, visual scanning, and psychomotor speed (Spreen & Strauss, 1998).
In this variant of the TMT-B, the examinee is presented with squares and circles
consisting of numbers 1 through 8; the examinee has to draw lines alternating between
square and circle in ascending order of the numbers while disregarding numbers of the
alternate shape. In the CoCA, the examinee is given two minutes to complete this test and
is not provided with any cuing or prompting to help with completion of the test. The total
score on this test is the time taken to complete the test divided by the number of correctly
drawn lines (maximum number of correctly drawn lines is 7). Process scores include
number of sequencing errors (incorrect sequence of numbers), set loss errors (connecting
the same shapes instead of alternating shapes), and capture errors (errors that occur due to
the physical proximity of a distractor item), as well as evidence for visuospatial neglect
and/or difficulties, motor difficulties, tremors, and losing track and starting from the
wrong target.
2. **Lexical Fluency:** In this task, the examinee is given one minute to generate as many words as possible that begin with the letter “S.” This test is reliant on the mental lexicon and requires the examinee to rapidly generate solutions to a given problem. Though similar in structure to the animal fluency task, the ability to retrieve grammatical representations and sound forms of words while meeting certain constraints is less reliant on semantic memory and more reliant on executive functioning ability (Shao, Janse, Visser, & Meyer, 2014). The letter “S” was chosen for this task based on the most common initial letter of a word found after analyzing 5000 of the most frequently used words retrieved from www.wordfrequency.info (Davies, 2016). The total score on this test is the total of all correctly produced words in one minute, with one point being credited for every accurate response. Process scores include number of words generated every 15 seconds, perseverations, intrusions, rule violations, and self-corrections. Percent set loss and percent perseverations are other process scores included in this test.

3. **Working Memory:** This test examines working memory, which is the ability of an individual to hold, process, and manipulate information at the same time. In this task, the examinee is required to repeat digits in the reverse order of their presentation when read a string of letters and numbers. The examinee is given a sample, followed by two trials that increase in span length. Credit is given for each correctly identified digit and for recognizing accurate positioning of digits in the string. The total score is derived by assigning one-half point for each correctly identified digit and one-half point for each digit’s correct placement in the sequence across both test trials. A maximum of 3 points and 5 points can be achieved on Trials 1 and 2, respectively. Process scores include number of omissions, commissions (additions), perseverative
errors (repeating a number or letter consecutively), and set loss errors (stating letters instead of numbers).

4. **Focused Attention**: In this test, participants are instructed to say “see” and “smell” upon hearing the words “eyes” and “nose,” respectively. Participants hear a series of words consisting of target words, words beginning with a sound similar to that of a target word, and words that begin with phonemic sounds similar to accurate responses given by the examinee. The words are presented at a relatively fast pace (1 word per second), thereby assessing vigilance, focused attention, and disinhibition. The task’s total score refers to the number of correct responses by the examinee, including the lack of a response when appropriate (i.e. when not presented with a target word). Maximum possible score on this task is 42. Process scores include number of omissions, commissions (additions), perseverative errors (repeating a response despite change in stimulus), and set loss errors (providing an incorrect response to the target word).

5. **Similarities**: This test assesses concept formation and abstract reasoning by asking the examinee to describe the similarity between two objects or concepts. The examinee is provided with a sample followed by two test trials. Credit is given for concrete and abstract responses. The total score is the sum of points obtained on the test trials, with a concrete response receiving one point and an abstract response obtaining two points. Maximum possible score on this task is four. Process scores include in-set and out-of-set errors. Process errors and examples of concrete and abstract responses are explained in the instruction and scoring manual found in Appendix C.

6. **Number Symbol Matching Test (NSMT) – Written and Oral Version**: The examinee is presented with nine different symbols corresponding to numbers 1 through 9,
followed by presentation of the symbols in a random order with empty boxes below each symbol. The examinee has to fill the empty box under each symbol with the corresponding number, with the goal of completing the maximum number of boxes possible in 30 seconds. This task assesses visual scanning and psychomotor speed and is based on the Symbol Digit Modalities Test (SDMT) task created by Smith (1982). Symbols chosen for this task were picked to resemble the shape of certain numbers, for example, a Z resembles a 2. Discordant looking symbols and numbers were assigned to each other (e.g., Z was assigned to 8) such that examinees with executive function difficulties may be drawn to exhibit stimulus bound errors. The total score refers to the sum of correctly drawn numbers, with every accurately matched symbol receiving one point. A maximum of 63 points can be obtained on this task.

For the purpose of the present study, the written version of the test was first administered, immediately followed by the oral version. The score on the oral version was not planned for inclusion in the total score of the CoCA; rather, it was designed to be a useful alternative for patients with upper limb movement difficulties. The assessment of manual motor speed and agility as assessed in the written version of the NSMT is considered to be integral to cognitive screening, thus, including the written version of the NSMT in a model of the CoCA was considered to be imperative. In the future, the oral version of the NSMT is recommended for use instead of the written version when individuals’ motor difficulties affect performance. Process scores include total number of errors and stimulus bound errors. Perseveration of symbols (repeating numbers consecutively despite change in stimulus/symbol), overwriting, micrographia, tremors,
set loss errors (copying symbols rather than using the key), and expansion of symbols are other process variables included.

**Orientation.** This task assesses the examinee’s orientation to the present; the examinee is asked to provide the date, year, day of the week, place of evaluation, state, city, and president of the country. The total score refers to an aggregate of all accurately provided responses, with every correct answer receiving one point. A maximum of eight points can be obtained on this task. There are no process scores for this test.

**Test of functional abilities.** This test assesses the examinee’s understanding of a bank statement and the ability to think through the consequences of a financial decision. Specifically, the examinee is asked to determine whether there is enough money in his or her account to make a certain payment based on a provided bank statement. The examinee is asked to describe the consequences of making a payment when there are insufficient funds in the bank account. Credit is based on the thoroughness of the participant’s response. The total score refers to the sum of responses on both items. A maximum of four points can be obtained on this task. On the first item, stating that there are not sufficient funds in the account receives one point. On the second item, a maximum of three points can be achieved based on the thoroughness of the response. Detailed information about scoring of this item is available in Appendix C. There are no process scores for this test.

**Criterion Measures.** The following tests were chosen for criterion validation of the CoCA as a measure of judgment and decision-making capabilities and global cognitive ability, as well as to account for the influence of mood symptoms on CoCA scores.
Neuropsychological Assessment Battery (NAB) Executive Function Module:

**Judgment Subtest** (NAB Judgment [Form 1]; Stern & White, 2003b). The NAB is a nationally standardized compilation of neuropsychological tests that assesses five cognitive domains, including executive functioning. The Judgment subtest is part of the Executive Functions Module of the NAB; it consists of ten questions that ask about an individual’s judgment about health, medical, and safety decisions (Stern & White, 2003 a,b). Poor performance on this test possibly indicates reduced judgment and decision-making abilities about one’s health and safety. A cut-off score of ≤ 12 on this test has been shown to result in the best sensitivity (.61) and specificity (.88) values (Gavett et al., 2012) for differentiating MCI from AD. In an assisted living sample, this test demonstrated adequate internal consistency reliability (.83; MacDougall & Mansbach, 2013), though its reliability in the older adult standardization sample of the NAB was low (α = .45; White & Stern, 2003). However, the NAB Judgment subtest demonstrates adequate inter-rater reliability (.85; White & Stern, 2003). Given the ease of assessment, relatively short time for administration, and ability to assess knowledge and decision-making capacity around health and safety, the NAB Judgment subtest was chosen as a measure of overall judgment. Research has shown a correlation of .3 between the MoCA and the Dementia Rating Scale (DRS; Lam et al., 2013). As mentioned previously, dementia is defined by concurrent decline in cognitive ability and everyday functioning. Since the NAB judgment subtest provides information about an individual’s daily functioning, adequate convergent validity, as judged by a correlation coefficient of .30 or greater between the NAB judgment subtest and the CoCA, would indicate that performance on the CoCA is reflective of everyday judgment ability.
Montreal Cognitive Assessment (MoCA; Nassreddine et al., 2005; see Appendix D). The MoCA is a 30-point cognitive assessment tool that assesses short-term memory, language, orientation, visuospatial ability, and aspects of executive functioning including attention, working memory, abstraction, and lexical fluency (Nasreddine et al., 2005). The MoCA demonstrates good internal consistency (.83) and is able to detect MCI and AD with a sensitivity of 90% and 100% respectively, and with an overall specificity of 87% (Nasreddine et al., 2005). Given the psychometric properties of the MoCA and its ability to assess several cognitive domains, the MoCA was chosen to serve as a measure of convergent validity for the CoCA. As above, a correlation of .60 between the MoCA and the CoCA will be used to identify acceptable levels of convergent validity.

Geriatric Depression Scale – 15 (GDS-15; Sheikh & Yesavage, 1986; see Appendix E). The GDS-15 is a screening instrument used to identify depression in older adults; it is sensitive (sensitivity = .87) in detecting depression in individuals with mild to moderate levels of dementia (Korner et al., 2006; Sheikh & Yesavage, 1986). This assessment was chosen in order to examine the effect of mood on cognitive performance on the CoCA. A score greater than 5 on the GDS-15 is suggestive of possible depression, whereas a score greater than 10 is almost always indicative of depressive (Sheikh & Yesavage, 1986). The sample for the present study is supposed to consist of individuals that are not depressed based on self-report, thus, a correlation of .30 or less between the CoCA and GDS-15 will be used to identify adequate levels of divergent validity.

Geriatric Anxiety Scale – 10 (GAS-10; Mueller et al., 2015; Segal et al., 2010; see Appendix F). The GAS-10 is a brief screening instrument used to identify anxiety in older adults (Mueller et al., 2015). It was derived from the GAS, which is made up of 30
items and demonstrates high internal consistency (Segal et al., 2010). The GAS-10 also has excellent internal consistent reliability (Cronbach’s α = .89), and was significantly and positively correlated with the GAS total score (r = .96, p < .01). The GAS-10 was chosen to examine the impact of anxiety on cognitive performance. Scoring of the GAS-10 is as follows: <7 = minimal anxiety; 7-9 = mild anxiety; 10 = moderate anxiety; and >12 = severe anxiety (Mueller et al., 2015; Segal et al., 2010). The sample for the present study is supposed to consist of individuals that are not significantly anxious based on self-report, thus, a correlation of .30 or less between the CoCA and GAS-10 will be used to identify adequate levels of divergent validity.

Demographics Questionnaire (see Appendix G). This questionnaire asks about the participant’s date of birth, age, sex, marital status, years of education, race, ethnicity, handedness, primary occupation, and first language. A question about living situation (independent housing, assisted living or caregiver in home, and nursing home) is also included in the questionnaire.

Procedure

As described above, the CoCA consists of a verbal and visual memory test, both of which have delayed recall components. To ensure that there was a long enough (at least eight minutes) delay between the immediate and delayed recall portions of the memory tasks, the learning trials of both tasks were ordered at the beginning of the test and the delayed recall trials were placed towards the end of the test. Tasks that are expected to take roughly the same amount of time for completion irrespective of the examinee were sequenced between the immediate and recall portions of the memory measures so that all examinees would roughly have the same delay period. In general, the
order of the tests was chosen to reflect variability in use of visual, verbal, and motor skills so that presented material from tasks does not interfere with performance on remaining tasks. In addition, tests assessing the same ability were not placed in consecutive order; repeatedly testing someone in an area of difficulty may influence effort and mood, which can likely impact test performance. Along the same lines, the tasks in the CoCA were ordered so that tasks of varying levels of difficulty were interspersed through the cognitive assessment in order to prevent fatigue and frustration, which could impact performance.

After the consent form, a demographics questionnaire was administered. Following this, either the CoCA or the MoCA was administered; the administration of the CoCA and the MoCA was counterbalanced such that half the participants were administered the CoCA at the beginning and the remaining half were administered the MoCA at the start of testing. Participants that were administered the CoCA at the beginning were given the MoCA as the last test of the evaluation, and vice versa. Between the CoCA and the MoCA, NAB Judgment, GDS-15, and GAS-10 were administered in that order. Participants were debriefed at the end of the evaluation. To ensure the validity of the statistical analyses (described below), participants who scored lower than six on the forced choice recognition task of the VLL were excluded.

**Statistical Analyses**

All analyses were conducted using R version 3.4.0 (R Core Team, 2017), including *lavaan* version 0.5-23.1097 and *semTools* version 0.4-14 packages (Rosseel, 2012; semTools Contributors, 2016).
Scoring and missing data. For the purpose of the present study, CFA was used to model the latent structure of the CoCA. In order to account for any missing data, full information maximum likelihood (FIML) was used, as this method uses all available data to estimate parameter values that are the most likely to produce the sample data. In this way, FIML will produce unbiased estimates with good precision and type 1 error control (Baraldi & Enders, 2009).

Validation of the CoCA. Based on clinical judgment, a hypothesized bi-factor CFA model of the CoCA was developed (Figure 1). All the cognitive test scores obtained from the CoCA were the manifest variables. General cognitive ability, $g$, and the domains of memory, language, visuospatial, and executive function were hypothesized to be the latent variables underlying performance. All latent variables were modeled as uncorrelated factors, as any residual variance across domains should be accounted for by the global factor. Residual correlations representing method effects were estimated between the following manifest variables: immediate recall, multiple choice recall, and delayed recall of VLL; figure copy and figure delay; clock copy and clock command; semantic and lexical fluency. This model has 18 observed variables, five latent variables, 29 free path coefficients, and six residual variances, resulting in 76 free parameters; the rule of thumb is to have a minimum of five observations per parameter (Bentler & Chou, 1987), requiring a minimum sample size of 380 participants for the validation of this model, which is not feasible for the current project. Though this model is considered to be an aspirational model for validation in the future, it was simplified for the purposes of this study (see Figure 2).
The simplified CFA model in Figure 2 consists of one latent global factor, \( g \). In order to decrease the number of parameters to be freely estimated, the number of observed variables in the model was reduced. Verbal list learning scores of immediate recall, delayed recall, and recognition were combined into one manifest variable. Clock drawing and clock copy were combined to obtain one variable for the clock drawing test. Animal and lexical fluency were combined to form one variable of fluency. Thus, the model proposed for the current study consists of 14 observed variables, one latent variable, thirteen free path coefficients, and one residual correlation, resulting in 56 free parameters. Based on the rule-of-thumb of 5 participants per parameter estimate, validation of this model will require a sample size of 280 participants. This sample size too, is not feasible for the current project. Some research suggests that a sample size of 100 - 200 participants is typically adequate for CFA (Boomsma, 1982; 1985). In order to provide preliminary validation of a theoretically sound model for the CoCA in a feasible manner, a sample size of 150 participants was considered to be adequate for the present study.

In order to find the best fitting model of the CoCA, several models with various composites and combinations of observed variables were run using CFA. The first model consisted of all the observed variables in the CoCA; subsequent models were derived iteratively based on examination of model fit indices and factor loadings of previous models. Generally, items with standardized factor loadings less than .4 were removed (Stevens, 1992). A total of 20 models of the CoCA were examined. Model fit was determined on the basis of \( \chi^2 \), root mean square error of approximation (RMSEA) and its 90% confidence intervals, standardized root mean square residual (SRMR), Tucker-
Lewis Index (TLI), and comparative fit index (CFI). A non-significant $\chi^2$, CFI and TLI values $\geq 0.95$, RMSEA $\leq 0.06$, and SRMR $\leq 0.08$ were considered indicators of good model fit (Hu & Bentler, 1999).

The variance of the latent variable in the CFA model was freely estimated. The factor loading of the STT was fixed to 1 in order to identify the model. Factor scores were generated using a linear regression approach. In this method, the standardized observed values of the indicator items serve as the predictor variables that are weighted by regression coefficients to produce global factor scores (DiStefano, Zhu, & Mîndrilă, 2009). Factor scores were re-scaled so that they had a mean of 0 and a standard deviation of 1.

**Measurement Invariance (MI).** Another goal of the present study was to establish measurement invariance with respect to age, sex, education, and mood (total scores on GDS-15 and GAS-10). Since a large majority of participants were Caucasian/White, there was not enough racial or ethnic diversity to examine the measurement invariance of the CoCA items with respect to these demographic variables. Thus, MI analyses were conducted with respect to sex, age, education, GDS-15, and GAS-10. MI was evaluated using the multiple-indicators multiple-causes modeling (MIMIC) approach, which uses latent variable modeling to examine MI (Hauser & Goldberger, 1971; Jöreskog & Goldberger, 1975). In this analysis, age, sex, education, GDS-15, and GAS-10 were covariates. The use of a MIMIC model allows the latent factor (i.e., the global factor score of the CoCA) to be measured by multiple observed items (i.e., observed test scores on the CoCA) and regressed onto multiple covariates, such as age, sex, education, GDS-15, and GAS-10 (Bauer, 2016). As such, MIMIC
modeling allows for examination of the effects of exogenous variables on observed items (e.g., differences in an item’s score on the CoCA based on sex) while holding equal their standing on the latent global factor (Bauer, 2016). This approach was chosen over other methods of evaluating MI because MIMIC modeling may have greater power in detecting measurement non-invariance (bias) when group sizes are small while accounting for multiple categorical and continuous predictors in the same model (Bauer, 2016; Muthén, 1989). With MIMIC modeling, a minimum sample size of 100 is needed to obtain adequate power and reasonably accurate parameter estimations (Woods, 2000). However, it should be noted that only uniform MI/DIF, not non-uniform MI/DIF (difference in item intercepts and not factor loadings) was examined in the present study.

To facilitate interpretation of the regression coefficients in the MIMIC model, age and education were centered at 70 years and 12 years, respectively. GDS-15 and GAS-10 were centered at their respective mean scores. In the first step, a baseline CFA model was analyzed, based on the best fitting model for the CoCA in which the observed variables were treated as being unbiased by sex, age, education, GDS-15, and GAS-10. In the baseline model, the covariates (i.e., age, sex, education, GDS-15, and GAS-10) were allowed to freely predict the latent variable (i.e., global CoCA factor score), whereas the regression coefficients between the covariates and multiple indicators (i.e., observed test scores on the CoCA) were fixed at 0. In the second step, modification indices for the specific regressions that were fixed to 0 were examined and the largest modification index was identified as the most obvious source of measurement non-invariance. In the third step, the regression coefficient between one of the covariates (i.e., the covariate with the largest modification index) and an indicator variable was freely estimated instead of
being fixed to 0. A $\chi^2$ difference test was then conducted between the models in the first and third steps to examine the presence of item bias or measurement non-invariance. A non-significant difference between the two models would suggest the presence of measurement invariance (lack of bias) for the item examined. A significant difference between the two models would suggest measurement non-invariance, or item bias. Subsequently, the model freely estimating the regression of the biased item onto the relevant covariate would serve as the baseline model and steps 1-3 would be repeated by freeing new regression parameters one at a time until the $\chi^2$ difference test revealed no significant differences.

**Reliability (convergent and divergent) and validity.** Pearson’s $r$ correlations were conducted to examine convergent validity between the CoCA factor scores and the MoCA, and the CoCA factor scores and the NAB Judgment subtest. Divergent validity was examined between the CoCA factor scores and GDS-15, and the CoCA factor scores and GAS-10. Adequate convergent and divergent validities were judged by values of $r \geq .6$ and $r \leq .3$, respectively. The global factor score reliability of the CoCA was also estimated using Mplus version 8 (Muthén & Muthén, 1998-2017). For basic and applied research, reliabilities greater than .80 are considered acceptable (Lance, Butts, & Michels, 2006; Nunnaly, 1978).

**Comparison to the MoCA.** Similar to the CoCA, a CFA model of the MoCA based on its current scoring methodology was developed (Figure 3). This model consisted of one latent variable, $g$. Observed variables included the following: sum of trails test, cube, and clock (visuospatial/executive functioning); sum of the three naming trials (naming); sum of digits forwards, digits backwards, tapping, and serial subtraction
(attention); sum of sentence repetition and fluency (language); abstraction; delayed free recall; and orientation. Model fit was examined using the same fit indices as described above with respect to the CoCA. As with the CoCA, MI analyses were conducted with the above-described model of the MoCA. The reliability of the global factor score of the MoCA was also estimated as described above.

Figure 3. CFA model of The Montreal Cognitive Assessment. $g$ is the latent variable representing general cognitive ability. Cognitive tests are the observed variables. Visuospatial/Executive Functioning = Trails + cube + clock; Naming = Trials 1-3 of Naming; Language = Sentence repetition + fluency; Abstraction = Trials 1 and 2 of Abstraction; Attention = Digits forward and backward + tapping + serial subtraction.
CHAPTER III
RESULTS

Participants

Participants were 151 community dwelling older adults. Participants were predominantly female (71.5%), White/Caucasian (93.1%), and not of Hispanic, Latino, or Spanish origin (93.5%). 98% of the sample considered English to be their primary language (2% identified Spanish as their primary language) and all participants lived in independent housing. Participants were fairly well educated ($M = 15.82$, $SD = 2.35$) with a mean age of 71.21 ($SD = 8.05$). The study sample on average was cognitively healthy ($M_{MoCA} = 25.87$; $SD_{MoCA} = 2.85$), without any significant anxiety ($M_{GAS} = 3.85$; $SD_{GAS} = 3.26$) or depression ($M_{GDS} = 1.56$; $SD_{GAS} = 2.17$). Additional demographic characteristics of the participants are listed in Table 1.

Modeling

In order to obtain the best fitting model of the CoCA for the data, 20 CFA models with various indicator combinations were examined. The models were developed using an iterative process by which the models were incrementally adjusted based on their fit statistics and factor loadings until the most suitable model was identified. Model suitability was judged by a combination of fit statistics meeting the thresholds described above and standardized factor loadings $> .4$. Model fit indices for each of the models examined are presented in Table 2. Indicator variables included in each CFA model of the CoCA are presented in Table 3.
### Sample Demographic Characteristics (n = 151)

<table>
<thead>
<tr>
<th>Demographics</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>n (Percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>71.21</td>
<td>8.05</td>
<td>52-93</td>
<td>--</td>
</tr>
<tr>
<td>Years of Education</td>
<td>15.82</td>
<td>2.35</td>
<td>9-20</td>
<td>--</td>
</tr>
<tr>
<td>MoCA</td>
<td>25.87</td>
<td>2.85</td>
<td>18-30</td>
<td>--</td>
</tr>
<tr>
<td>NAB Judgment</td>
<td>15.25</td>
<td>2.29</td>
<td>9-20</td>
<td>--</td>
</tr>
<tr>
<td>GDS-15</td>
<td>1.56</td>
<td>2.17</td>
<td>0-12</td>
<td>--</td>
</tr>
<tr>
<td>GAS-10</td>
<td>3.85</td>
<td>3.26</td>
<td>0-19</td>
<td>--</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>108 (71.5%)</td>
</tr>
<tr>
<td>Male</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>43 (28.5%)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Hispanic, Latino, or Spanish Origin</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>129 (93.5%)</td>
</tr>
<tr>
<td>Hispanic, Latino, or Spanish Origin</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>8 (5.8%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (0.7%)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White/Caucasian</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>135 (93.1%)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>3 (2.1%)</td>
</tr>
<tr>
<td>Asian</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (0.7%)</td>
</tr>
<tr>
<td>Other/Unknown</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>6 (4.1%)</td>
</tr>
</tbody>
</table>

*Note. n = sample size; M = Mean; SD = Standard deviation; MoCA = Total score on the Montreal Cognitive Assessment; Judgment = Total score on the NAB Judgment subtest; GDS-15 = Total score on the Geriatric Depression Scale – 15; GAS-10 = Total score on the Geriatric Anxiety Scale – 10.*

The originally hypothesized model (Figure 2; Table 2 [Model 13]) revealed inadequate fit: \( \chi^2(74) = 99.491, p = .037; \) CFI = 0.934; TLI = 0.921; RMSEA = 0.045 (90% CI [0.012, 0.068]); SRMR = 0.062, which necessitated the examination of alternative models. The model with the best fit (Figure 4; Table 2 [Model 20]) consisted of the following observed variables: STT; sum of immediate and delayed free recall of the VLL test; sum of figure copy and figure recall; sum of lexical and semantic fluency; working memory; focused attention; oral version of the NSMT; VNT; orientation; PMT; and test of functional abilities. This model produced the following fit indices: \( \chi^2(44) = 38.243, p = .716; \) CFI = 1.000; TLI = 1.029; RMSEA = 0.000 (90% CI [0.000, 0.043]);
Table 2

Model Fit Indices for Each Model of the CoCA

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$</th>
<th>CFI</th>
<th>TLI</th>
<th>AIC</th>
<th>BIC</th>
<th>RMSEA</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>130.244</td>
<td>101</td>
<td>.027</td>
<td>0.933</td>
<td>0.921</td>
<td>8892.555</td>
<td>9046.097</td>
<td>0.044</td>
<td>0.061</td>
</tr>
<tr>
<td>Model 2</td>
<td>132.857</td>
<td>101</td>
<td>.018</td>
<td>0.927</td>
<td>0.914</td>
<td>9037.819</td>
<td>9191.362</td>
<td>0.046</td>
<td>0.062</td>
</tr>
<tr>
<td>Model 3</td>
<td>132.087</td>
<td>101</td>
<td>.021</td>
<td>0.928</td>
<td>0.914</td>
<td>8328.925</td>
<td>8482.468</td>
<td>0.045</td>
<td>0.062</td>
</tr>
<tr>
<td>Model 4</td>
<td>97.414</td>
<td>87</td>
<td>.209</td>
<td>0.975</td>
<td>0.969</td>
<td>8384.237</td>
<td>8528.747</td>
<td>0.028</td>
<td>0.053</td>
</tr>
<tr>
<td>Model 5</td>
<td>120.922</td>
<td>101</td>
<td>.086</td>
<td>0.952</td>
<td>0.943</td>
<td>8833.919</td>
<td>8987.461</td>
<td>0.036</td>
<td>0.058</td>
</tr>
<tr>
<td>Model 6</td>
<td>127.208</td>
<td>101</td>
<td>.040</td>
<td>0.939</td>
<td>0.928</td>
<td>8856.151</td>
<td>8989.694</td>
<td>0.042</td>
<td>0.060</td>
</tr>
<tr>
<td>Model 7</td>
<td>123.005</td>
<td>101</td>
<td>.068</td>
<td>0.949</td>
<td>0.939</td>
<td>8612.396</td>
<td>8765.938</td>
<td>0.038</td>
<td>0.058</td>
</tr>
<tr>
<td>Model 8</td>
<td>105.419</td>
<td>87</td>
<td>.087</td>
<td>0.956</td>
<td>0.947</td>
<td>8203.661</td>
<td>8348.172</td>
<td>0.038</td>
<td>0.056</td>
</tr>
<tr>
<td>Model 9</td>
<td>95.335</td>
<td>74</td>
<td>.048</td>
<td>0.949</td>
<td>0.937</td>
<td>8080.278</td>
<td>8215.756</td>
<td>0.044</td>
<td>0.057</td>
</tr>
<tr>
<td>Model 10</td>
<td>92.082</td>
<td>74</td>
<td>.076</td>
<td>0.956</td>
<td>0.946</td>
<td>8361.111</td>
<td>8496.590</td>
<td>0.040</td>
<td>0.059</td>
</tr>
<tr>
<td>Model 11</td>
<td>76.981</td>
<td>62</td>
<td>.095</td>
<td>0.963</td>
<td>0.953</td>
<td>7850.675</td>
<td>7977.121</td>
<td>0.040</td>
<td>0.053</td>
</tr>
<tr>
<td>Model 12</td>
<td>82.603</td>
<td>74</td>
<td>.231</td>
<td>0.979</td>
<td>0.974</td>
<td>7974.789</td>
<td>8110.268</td>
<td>0.028</td>
<td>0.051</td>
</tr>
<tr>
<td>Model 13</td>
<td>99.491</td>
<td>76</td>
<td>.037</td>
<td>0.934</td>
<td>0.921</td>
<td>7789.480</td>
<td>7918.937</td>
<td>0.045</td>
<td>0.062</td>
</tr>
<tr>
<td>Model 14</td>
<td>71.240</td>
<td>64</td>
<td>.250</td>
<td>0.978</td>
<td>0.973</td>
<td>7283.161</td>
<td>7403.586</td>
<td>0.027</td>
<td>0.053</td>
</tr>
<tr>
<td>Model 15</td>
<td>69.227</td>
<td>64</td>
<td>.306</td>
<td>0.984</td>
<td>0.981</td>
<td>7194.080</td>
<td>7314.505</td>
<td>0.023</td>
<td>0.052</td>
</tr>
<tr>
<td>Model 16</td>
<td>60.838</td>
<td>54</td>
<td>.243</td>
<td>0.975</td>
<td>0.969</td>
<td>6718.404</td>
<td>6826.787</td>
<td>0.029</td>
<td>0.051</td>
</tr>
<tr>
<td>Model 17</td>
<td>43.631</td>
<td>35</td>
<td>.150</td>
<td>0.967</td>
<td>0.958</td>
<td>6185.165</td>
<td>6275.484</td>
<td>0.041</td>
<td>0.050</td>
</tr>
<tr>
<td>Model 18</td>
<td>55.760</td>
<td>44</td>
<td>.110</td>
<td>0.956</td>
<td>0.946</td>
<td>6594.553</td>
<td>6693.904</td>
<td>0.042</td>
<td>0.053</td>
</tr>
<tr>
<td>Model 19</td>
<td>47.506</td>
<td>44</td>
<td>.332</td>
<td>0.987</td>
<td>0.983</td>
<td>6308.878</td>
<td>6408.229</td>
<td>0.023</td>
<td>0.048</td>
</tr>
<tr>
<td>Model 20</td>
<td>38.243</td>
<td>44</td>
<td>.716</td>
<td>1.000</td>
<td>1.029</td>
<td>6319.290</td>
<td>6418.641</td>
<td>0.000</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Note. df = Degree of freedom; CFI = Comparative fit index; TLI = Tucker Lewis index; AIC = Akaike information criterion; BIC = Bayesian information criterion; RMSEA = Root mean square error of approximation; SRMR = Standardized root mean square residual

SRMR = 0.045. However, this model used the oral version of the NSMT, which, as discussed above, was not preferred for the final version of the test.

The next best fitting model of the CoCA (Figure 5; Table 2 [Model 19]) used the written version of the NSMT instead of the oral version; this model also yielded excellent fit: $\chi^2(44) = 47.506$, $p = .332$; CFI = 0.987; TLI = 0.983; RMSEA = 0.023 (90% CI [0.000, 0.061]); SRMR = 0.048. As such, Model 19 (Figure 5) was considered to be the most desirable model for the CoCA and will be used in further analyses as described below.
Table 3

*Indicators Used in Each CFA Model of the CoCA*

<table>
<thead>
<tr>
<th>Model</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock + Similarities + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 2</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock + Similarities + NSMT-Written + VNT + OR + VLL-Delayed Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 3</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock + Similarities + NSMT-Written + VNT + OR + VLL-Retention + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 4</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Similarities + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 5</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock Improvement + Similarities + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 6</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock Command + Similarities + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 7</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock Copy + Similarities + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 8</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock Copy + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 9</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock Copy + NSMT-Written + VNT + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 10</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + Clock + NSMT-Written + VNT + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 11</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + NTM-Written + VNT + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 12</td>
<td>STT + VLL-Immediate + FC + Animals + S Words + WM + Attention + NSMT-Written + VNT + OR + VLL-Delayed Free Recall + FR + PMT + TFA</td>
</tr>
</tbody>
</table>
Table 3 (continued)

<table>
<thead>
<tr>
<th>Model</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 13</td>
<td>STT + VLL-Total + FC + Fluency + WM + Attention + Clock + Similarities +</td>
</tr>
<tr>
<td></td>
<td>NSMT-Written + VNT + OR + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 14</td>
<td>STT + VLL-Total + FC + Fluency + WM + Attention + Similarities + NSMT-Written + VNT + OR + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 15</td>
<td>STT + VLL-Immediate and Delayed Free Recall + FC + Fluency + WM + Attention + Similarities + NSMT-Written + VNT + OR + FR + PMT + TFA</td>
</tr>
<tr>
<td>Model 16</td>
<td>STT + VLL-Immediate and Delayed Free Recall + Figure Total + Fluency + WM + Attention + Similarities + NSMT-Written + VNT + OR + PMT + TFA</td>
</tr>
<tr>
<td>Model 17</td>
<td>STT + VLL-Immediate and Delayed Free Recall + Figure Total + Fluency + WM + Attention + NSMT-Written + VNT + PMT + TFA</td>
</tr>
<tr>
<td>Model 18</td>
<td>STT + VLL-Immediate and Delayed Free Recall + Figure Total + Fluency + WM + Attention + Similarities + NSMT-Written + VNT + PMT + TFA</td>
</tr>
<tr>
<td>Model 19</td>
<td>STT + VLL-Immediate and Delayed Free Recall + Figure Total + Fluency + WM + Attention + NSMT-Oral + OR + VNT + PMT + TFA</td>
</tr>
<tr>
<td>Model 20</td>
<td>STT + VLL-Immediate and Delayed Free Recall + Figure Total + Fluency + WM + Attention + NSMT-Oral + OR + VNT + PMT + TFA</td>
</tr>
</tbody>
</table>

Note. Total scores on the following tests: STT = Shape Trails Test; VLL-Immediate = Verbal List Learning Test – Immediate Recall; FC = Figure Copy; WM = Test of Working Memory; Attention = Test of Focused Attention; Clock = Clock Command + Clock Copy; NSMT-Written = Number Symbol Matching Test – Written Version; VNT = Verbal Naming Test; OR = Orientation; VLL – Delayed Recall = VLL Delayed Free Recall + Category Cue; FR = Figure Recall; PMT = Prospective Memory Test; TFA = Test of Functional Abilities; VLL – Retention = VLL Delayed Free Recall/ VLL Immediate Recall Trial 3; Clock Improvement = Clock Copy – Clock Command; VLL-Total = VLL Immediate Recall + VLL Delayed Free Recall + VLL Recognition; Fluency = Animals + S Words; Figure Total = Figure Copy + Figure Recall; NSMT-Oral = Number Symbol Matching Test – Oral Version

The factor loading for the orientation test in this model was below the threshold of .40. However, this item was retained for two reasons: removal of this item from the model reduced model fit and the similarity between this item with one of the items on the MoCA can help link the CoCA and the MoCA psychometrically. The parameter estimates for this model are presented in Table 4.
Figure 4. Best fitting model of the CoCA using confirmatory factor analysis (CFA). $g$ is the latent variable representing general cognitive ability. Cognitive tests are the observed variables. VLL Immediate + Free Recall = Verbal List Learning Immediate Recall + Delayed Free Recall; PMT = Prospective Memory Test; VNT = Verbal Naming Test; Fluency = Animal + S words; STT = Shape Trails Test; WM = Test of Working Memory; FA = Test of Focused Attention; NSMT - Oral = Number Symbol Matching Test – Oral Version; TFA = Test of Functional Abilities
Figure 5. Model of the CoCA using confirmatory factor analysis (CFA). $g$ is the latent variable representing general cognitive ability. Cognitive tests are the observed variables. VLL Immediate + Free Recall = Verbal List Learning Immediate Recall + Delayed Free Recall; PMT = Prospective Memory Test; VNT = Verbal Naming Test; Fluency = Animal + S words; STT = Shape Trails Test; WM = Test of Working Memory; FA = Test of Focused Attention; NSMT - Written = Number Symbol Matching Test – Written Version; TFA = Test of Functional Abilities
### Table 4

*Unstandardized and Standardized Parameter Estimates for the CFA Model of the CoCA Depicted in Figure 5*

<table>
<thead>
<tr>
<th>Variable</th>
<th>(U)</th>
<th>SEE</th>
<th>(S)</th>
<th>(p)</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape Trails Test</td>
<td>1.000</td>
<td>--</td>
<td>0.495</td>
<td>--</td>
<td>0.245</td>
</tr>
<tr>
<td>VLL – Immediate + Delayed Free Recall</td>
<td>2.312</td>
<td>0.579</td>
<td>0.558</td>
<td>&lt;.01</td>
<td>0.311</td>
</tr>
<tr>
<td>Figure Copy + Figure Recall</td>
<td>3.560</td>
<td>1.022</td>
<td>0.561</td>
<td>&lt;.01</td>
<td>0.315</td>
</tr>
<tr>
<td>Fluency</td>
<td>7.593</td>
<td>1.843</td>
<td>0.600</td>
<td>&lt;.01</td>
<td>0.360</td>
</tr>
<tr>
<td>Test of Working Memory</td>
<td>1.382</td>
<td>0.366</td>
<td>0.642</td>
<td>&lt;.01</td>
<td>0.412</td>
</tr>
<tr>
<td>Test of Focused Attention</td>
<td>1.395</td>
<td>0.444</td>
<td>0.428</td>
<td>0.002</td>
<td>0.183</td>
</tr>
<tr>
<td>NSMT-Written</td>
<td>2.924</td>
<td>0.699</td>
<td>0.645</td>
<td>&lt;.01</td>
<td>0.415</td>
</tr>
<tr>
<td>Verbal Naming Test</td>
<td>0.705</td>
<td>0.207</td>
<td>0.466</td>
<td>&lt;.01</td>
<td>0.217</td>
</tr>
<tr>
<td>Orientation</td>
<td>0.107</td>
<td>0.049</td>
<td>0.224</td>
<td>0.028</td>
<td>0.050</td>
</tr>
<tr>
<td>Prospective Memory Test</td>
<td>0.242</td>
<td>0.070</td>
<td>0.407</td>
<td>&lt;.01</td>
<td>0.166</td>
</tr>
<tr>
<td>Test of Functional Abilities</td>
<td>0.439</td>
<td>0.138</td>
<td>0.457</td>
<td>&lt;.01</td>
<td>0.209</td>
</tr>
</tbody>
</table>

*Note.* \(U\) = unstandardized parameter estimate; \(S\) = standardized parameter estimate; \(SEE\) = standard error of the unstandardized estimate; VLL Immediate + Free Recall = Verbal List Learning Immediate Recall + Delayed Free Recall; Fluency = Animal + S words; NSMT - Written = Number Symbol Matching Test – Written Version

### CoCA Descriptive Statistics

In comparison to the MoCA, which took an average of eight minutes to complete (SD = 1.41), the mean completion time of the CoCA was approximately 21 minutes (SD = 2.69). The CoCA global factor scores were negatively skewed, with a skewness of -0.77 (SE = 0.08) and a kurtosis of 0.53. Regression of completion time of the CoCA onto the CoCA global factor score revealed an unstandardized estimate of -0.177 (S.E. = 0.044; \(p < .01\)). Descriptive data for each item on the CoCA is presented in Table 5. Distribution of the global factor scores of the CoCA is presented in Figure 6.
Table 5

Descriptive Data for Each Subtest of the Colorado Cognitive Assessment (CoCA)

<table>
<thead>
<tr>
<th>Subtest</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>STT - Completion Time</td>
<td>17.17</td>
<td>8.71</td>
<td>6</td>
<td>56</td>
</tr>
<tr>
<td>STT (# of correct lines)</td>
<td>6.84</td>
<td>0.54</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>STT – Total</td>
<td>2.56</td>
<td>1.55</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>VLL – Immediate Recall</td>
<td>18.38</td>
<td>1.91</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>VLL – Delay Free Recall</td>
<td>3.77</td>
<td>2.00</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>VLL- Yes/No Recognition</td>
<td>12.56</td>
<td>1.69</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>VLL – Forced Choice</td>
<td>6.91</td>
<td>0.35</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Figure Copy</td>
<td>15.99</td>
<td>2.13</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Figure Recall</td>
<td>12.40</td>
<td>3.28</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Fluency – Animals</td>
<td>21.07</td>
<td>6.19</td>
<td>1</td>
<td>42</td>
</tr>
<tr>
<td>Fluency – “S” Words</td>
<td>14.38</td>
<td>4.99</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>WM – Trial 1</td>
<td>2.63</td>
<td>0.64</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>WM- Trial 2</td>
<td>3.41</td>
<td>1.28</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Focused Attention</td>
<td>40.35</td>
<td>2.49</td>
<td>26</td>
<td>42</td>
</tr>
<tr>
<td>Clock – Command</td>
<td>9.26</td>
<td>1.11</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Clock – Copy</td>
<td>9.76</td>
<td>0.53</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>NSMT - Written</td>
<td>15.43</td>
<td>3.45</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>NSMT – Oral</td>
<td>17.73</td>
<td>3.59</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>VNT</td>
<td>13.95</td>
<td>1.17</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Orientation</td>
<td>7.86</td>
<td>0.37</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>PMT</td>
<td>0.70</td>
<td>0.46</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. STT = Shape Trails Test; VLL = Verbal List Learning Test; WM = Working Memory Test; NSMT = Number Symbol Matching Test; VNT = Verbal Naming Test; PMT = Prospective Memory Test

Measurement Invariance (MI)

Model 19 (Figure 5, Table 2) was used to generate the baseline model. In the first step, the model was run with the assumption that the global factor score (g) was not biased by age, education, sex, GDS-15, or GAS-10 (see Figure 7).
Figure 6. Distribution of global factor scores of the CoCA. (A) Histogram showing the distribution of global factor scores of the CoCA. \( M = \text{mean}, \ SD = \text{standard deviation} \). (B) Distribution of global factor scores of the CoCA by years of education and sex. (C) Distribution of global factor scores of the CoCA by age and sex. (D) Scatterplot showing the correlation between the CoCA global factor score and CoCA completion time.
Figure 7. Step 1 of multiple-indicators multiple-causes (MIMIC) modeling. Geriatric Anxiety Scale-10 (GAS-10), Geriatric Depression Scale-15 (GDS-15), age, sex, and education are the covariates. \( g \) is the latent variable representing general cognitive ability. Cognitive tests are the observed variables. VLL Immediate + Free Recall = Verbal List Learning Immediate Recall + Delayed Free Recall; PMT = Prospective Memory Test; VNT = Verbal Naming Test; Fluency = Animal + S words; STT = Shape Trails Test; WM = Test of Working Memory; FA = Test of Focused Attention; NSMT - Written = Number Symbol Matching Test – Written Version; TFA = Test of Functional Abilities. The latent variable is freely regressed onto all covariates and the regression coefficients between the covariates and observed variables are fixed at zero.

As expected, the fit of the proposed baseline model was adequate, \( \chi^2(94) = 109.261, p = .134 \); CFI = 0.955; TLI = 0.947; RMSEA = 0.033 (90% CI [0.000, 0.057]); SRMR = 0.053. The unstandardized regression estimates of the CoCA global factor on each of the covariates in the baseline model were as follows: GDS -15 = -0.098 (S.E. =
0.046), \( p = .032 \); GAS-10 = -0.004 (S.E. = 0.021), \( p = .848 \); Sex = -0.115 (S.E. = 0.139), \( p = .409 \); Age = -0.044 (S.E. = 0.011), \( p < .01 \); Education = 0.095 (S.E. = 0.032), \( p = .003 \). An examination of the modification indices revealed the largest modification index was the regression of the test of functional abilities on years of education; modification index (ModI) = 3.951; expected unstandardized parameter change (EPC) = -0.052; standardized expected parameter change (SEPC) = -0.166. Thus, allowing years of education to exert a non-zero direct effect on the test of functional abilities is expected to reduce the \( \chi^2 \) by 3.951.

The alternative model (see Figure 8) that freely estimated the regression of the functional item on years of education produced slightly better fit statistics than the model that fixed this regression parameter to 0; CFI = 0.964; TLI = 0.957; RMSEA = 0.030 (90% CI [0.000, 0.055]); SRMR = 0.052. However, a \( \chi^2 \) difference test comparing the model with the bias adjustment (\( \chi^2 (93) = 105.174 \)) to the model that did not account for the bias (\( \chi^2 (94) = 109.261 \)) was not significant; \( \chi^2 \) difference (1) = 3.377, \( p = .066 \). Therefore, the CoCA items were determined to be free from bias at this point, obviating the need for additional MI testing.

**Reliability and Validity**

There was a strong correlation between the global factor score on the CoCA and total score on the MoCA; \( r = .625, 95 \% \text{ CI} [.517, .714], p < .01 \). The CoCA global factor score was also moderately correlated with the total score on the Judgment subtest of the NAB; \( r = .350 95\% \text{ CI} [.201, .483], p < .01 \). There was a negative moderate correlation between the CoCA global factor score and total score on the GDS-15; \( r = -.323 95\% \text{ CI} [-.460, -.172], p < .01 \). The CoCA global factor score had a small negative correlation
with the total score on the GAS-10; $r = -0.149$ 95% CI [-.303, -.011], $p = .068$. Scatterplots depicting the above-mentioned relationships are presented in Figure 9. A correlogram showing the intercorrelations between all measures is presented in Figure 10.

Figure 8. Step 3 of multiple-indicators multiple-causes (MIMIC) modeling. Geriatric Anxiety Scale-10 (GAS-10), Geriatric Depression Scale-15 (GDS-15), age, sex, and education are the covariates. $g$ is the latent variable representing general cognitive ability. Cognitive tests are the observed variables. VLL Immediate + Free Recall = Verbal List Learning Immediate Recall + Delayed Free Recall; PMT = Prospective Memory Test; VNT = Verbal Naming Test; Fluency = Animal + S words; STT = Shape Trails Test; WM = Test of Working Memory; FA = Test of Focused Attention; NSMT - Written = Number Symbol Matching Test – Written Version; TFA = Test of Functional Abilities. The latent variable is freely regressed onto all covariates. Regression coefficients between the covariates and observed variables (except TFA) are fixed at zero. Regression coefficients of TFA on age, sex, GDS-15, and GAS-10 are fixed at zero, whereas the effect of education on TFA score is freely estimated.
Figure 9. Scatterplot showing the relationships between the global factor score on the CoCA and total scores on the Montreal Cognitive Assessment (MoCA; A), NAB Judgment subtest (B), Geriatric Depression Scale-15 (GDS-15; C), and Geriatric Anxiety Scale-10 (GS-10; D). CoCA global factor score is on the X-axis and total scores on the MoCA, NAB Judgment, GDS-15, and GAS-10 are on the Y-axis.
Figure 10. Correlogram showing correlations between global factor score on the CoCA, age, education, total Montreal Cognitive Assessment (MoCA) score, total Geriatric Depression Scale-15 (GDS-15) score, total Geriatric Anxiety Scale-10 (GAS-10) score, and total NAB Judgment score

Reliability of the CoCA. The ratio of the global factor variance to the total (global factor and error) variance of the CoCA was .84, indicating a high factor score reliability of the CoCA global factor.

Comparison to the MoCA

Model fit. The MoCA model (Figure 3) had poor fit, $\chi^2(14) = 28.536$, $p = .012$; CFI = 0.817, TLI = 0.725; RMSEA = 0.083 (90% CI [0.038, 0.127]); SRMR = 0.062.

Measurement invariance. The baseline model of the MoCA is presented in Figure 3. In the first step, the model was run with the assumption that the global factor
score (g) was not biased by age, education, sex, GDS-15, and GAS-10. The fit of the proposed baseline model was poor, $\chi^2(44) = 64.347$, $p = .024$; CFI = 0.807; TLI = 0.754; RMSEA = 0.056 (90% CI [0.021, 0.083]); SRMR = 0.060. An examination of the modification indices revealed the largest modification index was the regression of abstraction on age; $\text{ModI} = 4.622$; $\text{EPC} = 0.012$; $\text{SEPC} = 0.179$. Thus, allowing abstraction to freely vary based on age is expected to reduce the $\chi^2$ by 4.622. The alternative model that freely estimated the regression of abstraction on age produced similar fit statistics; CFI = 0.843; TLI = 0.796; RMSEA = 0.051 (90% CI [0.005, 0.080]); SRMR = 0.056. A $\chi^2$ difference test comparing the model with the bias adjustment ($\chi^2(43) = 59.505$, $p = .048$) to the model that did not account for the bias ($\chi^2(44) = 64.347$, $p = .024$) was significant; $\chi^2$ difference (1) = 4.939, $p = .026$. Thus, at least one item on the MoCA was considered biased at this time, necessitating the examination of other potential sources of bias.

The next obvious potential source of measurement non-invariance was the regression of naming on years of education; $\text{ModI} = 4.854$; $\text{EPC} = -0.031$; $\text{SEPC} = -0.204$. The alternative model that freely estimated the regression of naming on years of education and of abstraction on age produced slightly better fit statistics than the model with the regression of abstraction on age alone; CFI = 0.884; TLI = 0.846; RMSEA = 0.044 (90% CI [0.000, 0.075]); SRMR = 0.052. A $\chi^2$ difference test comparing the model with the bias adjustment ($\chi^2(42) = 54.186$, $p = .098$) to the model that did not account for the bias ($\chi^2(43) = 59.505$, $p = .048$) was significant; $\chi^2$ difference (1) = 5.547, $p = .018$, necessitating the examination of additional potential sources of non-invariance.
The third most obvious source of bias was the regression of visuospatial and executive functioning (VSEF) abilities on GDS-15; ModI = 3.224; EPC = 0.067; SEPC = 0.15. The alternative model that freely estimated the regression of VSEF on GDS-15 along with the previous sources of non-invariance produced slightly better fit statistics than the model regressing naming on years of education and abstraction on age; CFI = 0.908; TLI = 0.874; RMSEA = 0.040 (90% CI [0.000, 0.072]); SRMR = 0.052. A $\chi^2$ difference test comparing the model with the bias adjustment ($\chi^2(41) = 50.742, p = .142$) to the model that did not account for the bias ($\chi^2(43) = 59.505, p = .048$) was significant; $\chi^2$ difference (1) = 19.968, $p < .01$, necessitating the examination of additional potential sources of non-invariance.

The fourth most obvious source of measurement non-invariance was the regression of delayed recall score on sex; ModI = 3.082; EPC = -0.47; SEPC = -0.144. The alternative model that freely estimated the regression of delayed recall score on sex in addition to the three other sources of bias produced better fit statistics than the model regressing VSEF on GDS-15, naming on years of education, and abstraction on age; CFI = 0.928; TLI = 0.899; RMSEA = 0.036 (90% CI [0.000, 0.070]); SRMR = 0.049. A $\chi^2$ difference test comparing the model with the bias adjustment ($\chi^2(40) = 47.614, p = .191$) to the model that did not account for the bias ($\chi^2(41) = 50.742, p = .142$) was not significant; $\chi^2$ difference (1) = 3.235, $p = .072$, thereby obviating the need to examine any other potential sources of non-invariance. Overall, the MoCA was considered to have three sources of bias; the effect of age on abstraction, the effect of years of education on naming, and the effect of GDS-15 on VSEF.
**Reliability.** Reliability of the MoCA was 0.74, indicating a lower factor reliability than the CoCA.
CHAPTER IV

DISCUSSION

The CoCA was developed to serve as a comprehensive cognitive screen to identify dementia and assist with differential diagnosis early in the disease process. The development of the CoCA also aimed to improve on current cognitive screens by: (i) assessing a wider range of cognitive abilities; (ii) utilizing tasks that are potentially less biased by age, education, and sex; (iii) providing a template to identify and record process scores for each task; and (iv) scoring using modern psychometrics. The primary goal of the present study was to validate the CoCA in a sample of community dwelling older adults. Multiple secondary related goals were examined including, (i) establishing the measurement invariance of the CoCA with respect to age, education, sex, and mood; (ii) establishing the convergent validity of the CoCA with the MoCA and NAB Judgment subtest; (iii) establishing the divergent validity of the CoCA with the GDS-15 and GAS-10; (iv) establishing the CoCA’s global factor score reliability; (v) comparing the CoCA and the MoCA with respect to measurement invariance and reliability.

Validation of the CoCA

Most cognitive tests are scored using CTT, which poses several limitations to test reliability and validity as described earlier. Primarily, the lack of accounting for item difficulty and examinees’ level of ability make CTT a less informative and less sophisticated approach to test development and scoring, yet most commonly used cognitive tests rely on CTT due to familiarity with the process (Brown, 2015). One of the
goals of the present study was to use modern psychometrics to score and validate the CoCA. This was achieved by the use of CFA, which verifies the number of factors underlying the CoCA and also examines item-factor relationships, thereby informing the scoring of the CoCA by accounting for item difficulty and examinees’ ability level (Brown, 2015). Further, CFA offers a robust analytical framework to evaluate construct validity, reliability, and measurement equivalence of the CoCA (Brown, 2015). Additionally, CFA removes measurement error by allowing for the use of latent variables instead of observed variables (e.g., raw score on the MoCA). Models developed using CFA can be incorporated into larger structural models to test hypotheses about factors that influence cognition.

Several CFA models of the CoCA with indicator variables in different combinations were explored. Since the originally hypothesized model (Model 13) of the CoCA had inadequate fit (Figure 2; Table 2 $\chi^2(74) = 99.491, p = .037; \text{CFI} = 0.934; \text{TLI} = 0.921; \text{RMSEA} = 0.045 (90\% \text{CI} [0.012, 0.068]); \text{SRMR} = 0.062$, theoretical considerations and examination of factor loadings were used to develop and test additional models. Initially, different scores from the same subtests were included in the model separately (e.g., VLL immediate recall and delayed free recall); however, composite variables consisting of tasks belonging to the same overall test appeared to improve model fit. In general, clock command and copy, similarities, and orientation did not have adequate factor loadings, indicating that $g$ (global factor of the CoCA) did not explain enough of the variance in these test scores. Though several studies have demonstrated the clock drawing test’s utility in detecting and predicting cognitive change, especially with examination of process variables, some research also suggests
that its ability to detect dementia in its earliest stages is limited (Cahn et al., 1996; Powlishta et al., 2002; Seigerschmidt, Mosch, Siemen, Forstl, & Bickel, 2002; Shulman, 2000).

Both the clock command and copy were included in the CoCA based on literature suggesting that improvement from command to copy assists with differential diagnosis of dementia (Libon et al., 1993; 1996). The CDT in the CoCA was scored using Rouleau et al.’s (1992) ten point scoring system. As mentioned previously, several scoring systems that emphasize different cognitive abilities have been created for the CDT (Hubbard et al., 2008; Shulman, 2000). Though the scoring systems are generally well-correlated, one study found that the Shulman method (Shulman et al., 1993) was better than two other scoring systems at identifying mild to moderate dementia (Brodaty & Moore, 1997), while another study found the the Cahn protocol (Cahn et al., 1996) to be superior at differentiating MCI from normal aging (Yamamoto et al., 2008). Since the literature surrounding the efficacy of scoring systems in comparison to each other is inconsistent, it is possible that scoring the clock using a different method could have resulted in the CDT having a higher factor loading.

Similarly, the low factor loading for the similarities test in a cognitively healthy sample may be explained by research indicating that the analysis of the types of responses and errors is better at detecting the earliest cognitive changes than the total quantitative score on this test (Giovanetti et al., 2001; Jacobs et al., 1997). The inclusion of only two items on the similarities test in the CoCA could have also contributed to the CoCA global factor not explaining a significant amount of variance in this test score. Although the prospective memory test met the cutoff of .4 with respect to its standardized
factor loading in the final model of the MoCA, the scoring did not allow for the capturing of variance in responses. Thus, adding varying levels of cuing on this test may help increase its factor loading and should be explored in future models of the CoCA. Though the orientation test had a low factor loading, its addition in the CFA model appeared to improve model fit (see Tables 2 & 3 [Models 19 & 20]). Items asking about physical and temporal orientation have been included in most cognitive screens, with individuals with different dementia etiologies exhibiting differential impairment on orientation questions (Jefferson et al., 2002). Thus, despite the low factor loading, the orientation item was retained in the final model of the CoCA as it improved model fit and can serve as a way to psychometrically link the CoCA to other cognitive screens. One possible reason for orientation having a low factor loading is that the sample in the present study consistent of primarily non-demented individuals. Thus, most participants likely scored very well on this item resulting in ceiling effects, which may have contributed to its low factor loading in the present study. As such, it is important to examine the factor loading of orientation in a sample of demented individuals who are less likely to fare well on this item. Further validation of the CoCA in larger samples with different population characteristics is needed to determine if items with low factor loadings will remain in the final version of the CoCA.

**Model of the CoCA used in the present study.** The best fitting model (Model 20) of the CoCA (Tables 2 & 3; Figure 4) had excellent fit; however, it used the oral version of the NSMT, which serves as an optional test in the event that an examinee is unable to perform the written version of the NSMT. Thus, the next best fitting model (Tables 2 & 3 [Model 19]; Figure 5) included the written version of the NSMT; because
this model also had excellent fit, it was considered the most desirable model; $\chi^2(44) = 47.506, p = .332$; CFI = 0.987; TLI = 0.983; RMSEA = 0.023 (90% CI [0.000, 0.061]); SRMR = 0.048. All standardized factor loadings for the final model (Model 19) of the CoCA (Tables 2 & 3; Figure 5) were above the threshold of .40, with the exception of two tests (Orientation and PMT; Table 4). Excellent overall model fit with generally high factor loadings indicate that the items on the CoCA measure global cognitive ability well (Brown, 2015). Items with higher factor loadings are better at measuring $g$; amongst all the items on the CoCA, the test of working memory and NSMT-Written were the best at measuring $g$, whereas orientation and PMT were the least effective at measuring $g$. At this time, an individual’s global cognitive functioning can be estimated by the CoCA using the CFA model in Model 19. Interpretation of an individual’s CoCA score will be based on norms; the development of norms is dependent on the future validation of the CoCA in diverse clinical and non-clinical samples.

**Comparison to the literature on other cognitive screens.** The fit of the CoCA cannot be directly compared to cognitive screens used in other studies due to different sample characteristics, tests, and models that affect fit. CFA models have been used to examine other cognitive screens, including the ACE-R (CFI = .884 - .994; RMSEA = .054 - .143; McGrory, Starr, Shenkin, Austen, & Hodges, 2015), MoCA (CFI = .956 - .981, RMSEA = .026 - .104; Duro, Simoes, Ponciano, & Santana, 2009; Freitas, Simoes, Maroco, Alves, & Santana, 2012), and the SLUMS (CFI = .99; RMSEA = .03; Stern, 2014). In general, models of the above mentioned cognitive screens were validated only in clinical samples, with multi-factor solutions yielding better fit than one-factor solutions. Overall, literature examining the factor structures of cognitive screening tools
90

is limited. For example, despite the widespread use of the MoCA and validation in various international samples, its psychometric properties and normative data within the U.S. has been studied to a limited extent (Freitas et al., 2012; Rossetti et al., 2011). Two studies have examined the factor structure of the MoCA; both have occurred in a Portuguese population (Duro et al., 2009; Freitas et al., 2012).

Though Nasreddine et al. (2005) proposed a six-factor model of the MoCA, the fit of the MoCA in the present study was based on the seven cognitive domains outlined in the MoCA (i.e., the seven domain scores contribute to a global factor, $g$). This model revealed poor fit; $\chi^2(14) = 28.536, p = .012; CFI = 0.817, TLI = 0.725; RMSEA = 0.083 (90\% CI [0.038, 0.127]); SRMR = 0.062$, which is in contrast to the fit obtained for a six-factor model, a two-factor model, and a one-factor second-order model of the MoCA (i.e., the six proposed factors contribute to an underlying factor of global cognition) by Freitas et al. (2012). Freitas et al. (2012) showed that the original six-factor model of the MoCA as conceptualized by Nasreddine et al. (2005) had best fit; $\chi^2(448) = 708.877, p < .01; CFI = 0.981, TLI = 0.978; RMSEA = 0.026$, similar to the model fit of the CoCA. Overall, the fit of the CoCA was comparable if not better than that of other cognitive screens including the MoCA.

**Measurement Invariance of the CoCA**

One of the considerations in the development of the CoCA was the inclusion of tasks that were unbiased by age, sex, education, and racial/ethnic background. Since the present study sample was primarily Caucasian, examining whether the CoCA was biased by race/ethnicity was not possible. After the validation of the CoCA, it was necessary to establish that items on the CoCA measure the same construct irrespective of an
examinee’s age, level of education, and sex. In other words, it was important to establish that the CoCA is invariant to these factors so that scores between individuals of different ages, education levels, and sex are directly comparable (Bauer, 2016). During the study, we also decided to investigate MI based on depression and anxiety, as that is another potential source of bias that we could test.

One way of establishing measurement invariance using CFA is through the use of MIMIC models that allow latent and indicator variables to be regressed onto covariates such as sex, age, education, and mood (Brown, 2015). In the present study, MIMIC modeling was used to examine whether sources of measurement non-invariance – or bias – exist within the items of the CoCA. Results indicated that the CoCA is invariant to age, sex, education, and GDS-15 and GAS-10 scores in the current sample, indicating that any individual differences in the CoCA global factor score are reflective of differences in global cognitive ability and not reflective of the confounding effects of age, sex, education, or mood.

Though none of the items on the CoCA were biased, there were significant direct effects of education, GDS-15, and age on the global factor score of the CoCA. It is important to emphasize that education, depression, and age affected performance on the CoCA as described below, but did not bias performance on the CoCA. It was found that for every one additional year of education above 12 years (centering point), the CoCA global score increased by 0.137 SD units. For every one year increase in an individual’s age beyond 70 years (centering point), the CoCA global factor score increased by 0.059 SD units. And for every one additional point on the GDS-15 above the sample mean (centering point), the CoCA global factor score decreased by 0.132 SD units. Though
there were no direct effects of sex on the CoCA global factor, Figures 6b and 6c allow for
the visual inspection of CoCA global factor score with respect to sex and education, and
sex and age, respectively. Consistent with the direct effect of education on CoCA global
factor score, Figure 6b shows that the global factor score of the CoCA increases at about
the same rate both males and females with increasing years of education. Consistent with
the direct effect of age on global factor score of the CoCA, Figure 6c shows that CoCA
global factor scores decrease with increasing age for both males and females, though
males appear to be showing a steeper decline than females.

**Comparison to other cognitive screens.** The lack of measurement invariance in
the CoCA is in contrast with presence of biases observed in most other cognitive screens.
For example, the MoCA has been shown to be biased by sex and years of education
(Benitez, Gross, Apostolova, & ADNI, 2015; Rossetti et al., 2011), the ADAS-COG has
been shown to be biased by gender (Verma, Beretvas, Pascual, Masdeu, & Markey,
2015), the CASI has demonstrated DIF with respect to age and education, and the MMSE
has exhibited DIF with respect to sex, education, and language of administration (Jones,
2006; Jones & Gallo, 2002; Ramirez, Teresi, Holmes, Gurland, & Lantigua, 2006). As a
result, these cognitive screens have provided adjustments for the existing DIF, even
though that does not necessarily result in elimination of the bias itself (Jones & Gallo,
2002).

To the best of the author’s knowledge, DIF has been examined with various
cognitive screens, though MIMIC modeling has only been utilized with the MMSE
(Jones, 2006). DIF has been well explored, especially with respect to the MoCA. In
various clinical samples and languages of MoCA assessment, studies have shown that
DIF exists within items of the MoCA with respect to education, sex, language, pathology, education, and gender (e.g., Benitez et al., 2015; Freitas, Prieto, Simoes, & Santana, 2014; Koski, Xie, & Finch, 2009; Zhou et al., 2015).

In addition to validating a CFA model of the MoCA, the present study examined the measurement equivalence of the MoCA. Results from the present study revealed that the MoCA had three sources of bias or non-invariance; the effect of age on abstraction, the effect of years of education on naming, and the effect of GDS-15 on VSEF. Some of these results are in contrast with findings from other studies. For example, some studies have found no effect of age on any of the MoCA items (Freitas et al., 2014; Koski, Xie, & Konsztowicz, 2011). With respect to education-related DIF, one study found an effect of education on several items of the MoCA (cube, abstraction, word recall, and orientation; Freitas et al., 2014), while another study found education-related DIF on two items of the MoCA (serial 7s and fluency; Kotwal et al., 2015). Though no sex related bias on the MoCA was observed in the present study, Kotwal et al. (2015) found that two items, serial 7s and delayed word recall, were biased by sex. In contrast, another study found no sex, language, or education related DIF with respect to any of the MoCA items (Koski et al., 2011). Thus, DIF with respect to the MoCA may be dependent on sample characteristics; for example, Koski et al. (2009) found two items on the MoCA to be biased by testing sample even though there were no significant differences in demographics across the two samples. While the literature regarding DIF with respect to the MoCA is inconsistent, the finding from the present study regarding an item on the MoCA being biased by depression appears to be a novel finding.
The presence of bias or non-invariance in a test prevents one from making valid inferences about observed differences between individuals, thereby making measurement invariance a prerequisite for unbiased testing (Borsboom, 2006; Meredith & Teresi, 2006). In addition, the presence of bias negatively impacts construct validity of an instrument because it indicates that the test is not purely measuring what it is supposed to measure; rather, it is also measuring the effects of confounding variables. The measurement invariance of the CoCA provides further preliminary support for its construct validity and applicability as an unbiased measure of global cognitive ability. Though the CoCA appears to be unbiased by a number of factors at this time, the rise in racial and ethnic diversity in the U.S. makes the MI analysis of the CoCA with respect to different racial and ethnic groups an increasingly important step to be taken in the process of validation of the instrument. Since longitudinal assessment is a cornerstone of cognitive aging research and clinical practice, future studies should also evaluate MI of the CoCA longitudinally.

Reliability and Validity of the CoCA

Factor score reliability. Examining the psychometric properties of the CoCA was another step in continuing to establish its utility as a screening tool. The CoCA had a high global factor score reliability ($r = .84$), suggesting that global cognitive ability can be estimated with acceptable precision. As mentioned earlier, there is limited research regarding the factor structure of cognitive screens, resulting in almost no data on their factor score reliability, though literature on internal consistency reliability is available. The internal consistency reliability (Cronbach’s $\alpha$) of the MoCA and ACE-R was found to be .90 and .80, respectively, in clinical samples (Frietas et al., 2012; Mioshi et al.,
2006), whereas the reliability of the MoCA has been found to be low (Cronbach’s \( \alpha = .50 - .60 \)) in non-clinical samples (Bernstein, Lacritz, Barlow, Weiner, & DeFina, 2010). The threshold of .80 is indicative of good reliability, particularly when decisions are made based on a group of tests with similar levels of reliability. For the purposes of decision making based on a single test score, a reliability coefficient of .90 or higher is considered suitable (Frary, n.d.). The factor score reliability of the CoCA is above the threshold of .80 (Nunnally, 1978) and comparable to the internal consistency reliabilities of other cognitive screens. In contrast, the factor score reliability of the MoCA obtained in the present study was .74, which was below the .80 threshold and lower than the CoCA, adding to preliminary evidence for the CoCA as providing a more reliable measure of global cognition than the MoCA.

**Convergent validity.** In order to assess convergent validity, the CoCA was compared to the MoCA and NAB Judgment subtests. Pearson’s \( r \) values of .10, .30, and .50 were judged as small, medium, and large correlations (Cohen, 1988). Results revealed that the CoCA had a strong and significant correlation with the MoCA (\( r = .625 \)) and a moderate and significant correlation with NAB Judgment subtest (\( r = .350 \)). The strong correlation between the CoCA and the MoCA, two cognitive screening tools, indicates that like the MoCA, the CoCA also assesses global cognitive ability. This association between the CoCA and the MoCA was expected given the similarity in structure and content of some of the items in both instruments. The size of the correlation supports the convergent validity between the CoCA and the MoCA; at the same time, the correlation is not too strong so as to render it redundant with the MoCA. The strong correlation between the CoCA and MoCA along with other results from the current study,
especially the direct comparison between the CoCA and the MoCA, addresses the value added by the CoCA. As such, the convergent validity between the CoCA and the MoCA as found in the present study provides further confirmation of the CoCA as a test of global cognition.

Consistent with the proposed cutoff for adequate validity, the moderate correlation between the CoCA and a test of everyday decision-making is similar to those found in the literature between the MoCA and the MMSE (Lam et al., 2013). Thus, the correlation between the NAB Judgment subtest and the CoCA is indicative of the CoCA assessing functional loss, confirming its convergent validity and making it suitable for dementia assessment. These findings suggest that the CoCA correlates well with commonly used tests of cognition and judgment, further confirming its construct validity as a measure of overall cognitive ability.

**Divergent validity.** Divergent validity was assessed by comparing the CoCA to the GDS-15 and the GAS-10. The CoCA had a moderate correlation with total score on the GDS-15 ($r = -.323$), indicating that higher GDS-15 scores are associated with reduced performance on the CoCA. The CoCA had a minimal correlation with scores on the GAS-10 ($r = -.149$), indicating that scores on the CoCA are weakly related to an individual’s level of anxiety. To the best of the author’s knowledge, literature regarding the assessment of divergent validity with respect to cognitive screens is sparse. One study found that low overall scores on the MMSE are associated with level of depressive symptoms (van Ojen et al., 1995); however, the effect of mood on individual items has not been tested. Research has consistently shown that increasing levels of depression are associated with deficits in executive functioning, particularly attention and processing
speed, as well as in memory (Boone et al., 1995; Levin, Heller, Mohanty, Herrington, & Miller, 2007; Nitschke, Heller, Imig, McDonald, & Miller, 2004; Snyder, 2013). The strength of the association between the CoCA global factor score and GDS-15 in the present study is similar to findings from a meta-analysis that found that severity of depression is related to reduced cognitive performance in episodic memory ($r = -.32$), processing speed ($r = -.16$), and executive functioning ($r = -.31$; McDermott & Ebmeier, 2009). The negative association between level of depression and CoCA global factor score is further confirmed by the results of the regression of CoCA global factor scores onto GDS-15 scores, which indicates that as an individual’s score on the GDS-15 increases by one point above the sample mean (centering point), the CoCA global factor score decreases by 0.132 SD units. It is important to note though that negative association between the CoCA global factor score and level of depression is not indicative of any of the items of the CoCA being biased by depressive symptomatology, as confirmed by the MI analysis conducted in the present study. Also, these results only apply to individuals who did not report a clinical diagnosis of depression and thus, cannot be applied to clinically depressed groups.

In contrast to the large amount of literature supporting the impact of depression on neuropsychological test performance, research regarding the association between anxiety and cognitive performance is rather limited. One study found that mild levels of anxiety affected performance on tests of memory and executive functioning only in males (Martin & Franzen, 1989), while another study has shown that increased anxiety is associated with reduced performance on tests of short term and delayed memory (Mantella et al., 2007). One study also showed that increased anxiety is associated with
poorer processing speed and inhibition (Beaudreau & O’Hara, 2009). Thus, the findings regarding the impact of anxiety on cognitive performance are inconsistent. The small and non-significant correlation between the GAS-10 and CoCA global factor scores demonstrates the divergent validity of the CoCA with respect to anxiety. Though the regression of CoCA global factor score onto GAS-10 scores was non-significant, results indicate that for every one point increase in GAS-10 scores beyond the sample mean (centering point), CoCA global factor scores decrease by 0.007 SD units. Visual inspection of the scatterplot of GAS-10 score and CoCA global factor score (Figure 9d) suggests that an increase in an individual’s level of anxiety is accompanied by a slight reduction in their CoCA global factor score. As with the GDS-15, MI analysis conducted in the present study confirmed that the items on the CoCA are not biased by an individual’s level of anxiety.

It is evident from a visual comparison of Figures 9c and 9d that the slope of the trendline showing the association between GDS-15 scores and CoCA global factor score (Figure 9c) is much steeper than that of GAS-10 scores and CoCA global factor score (Figure 9d). This indicates that in the current sample, depressive symptomatology affects one’s overall performance on the CoCA more than anxiety symptoms. Since the effect of mood on cognitive screens has not been examined, this study is unique in its assessment and establishment of divergent validity with respect to mood.

**Advantages of the CoCA**

The CoCA was developed to provide more thorough and more accurate information about an individual’s cognition in a relatively short time frame by combating some of the weaknesses of current cognitive screens. Overall, the CoCA samples from all
major cognitive domains, including verbal and visual memory, visuospatial ability, executive functioning, and language. While this is true for many other cognitive instruments as well, the CoCA attempts to assess a breadth of abilities underlying each cognitive domain with relatively equal emphasis on each domain. The present study found that the best fitting model of the CoCA was one that placed the most emphasis on attention and executive functioning, and least emphasis on visuospatial ability. Examination of future models of the CoCA will help determine whether all cognitive domains are relatively equally important to the model.

Other cognitive screens have ceiling effects, which can prevent the early detection of mild cognitive changes. For example, the MoCA has been criticized for consisting of items that are too easy for those having mild memory complaints (Koski, Xie, & Finch, 2009). Thus, including items of varying levels of difficulty was a key feature of the CoCA so as to be able to better capture cognitive heterogeneity and detect cognitive decline in higher functioning individuals. The absence of ceiling effects is evident from the distribution of the global factor scores of the CoCA (Figure 6a), which came from a fairly well educated cognitively healthy sample. The distribution of scores was negatively skewed (Figure 6a); most individuals demonstrated an average level of performance, while some performed above average and more performed below average. Given that the current sample consisted of fairly well-educated cognitively healthy individuals, floor effects could not be assessed, making future research with demented participants imperative.

The use of CFA to model and score the CoCA is not only advantageous in that it accounts for item difficulty and person’s ability level, but it also possesses linear
measurement properties, which make the CoCA psychometrically ideal for longitudinal assessment. By possessing linear measurement properties, a change in CoCA factor scores has the same meaning regardless of the initial score. In other words, a 1-SD change (for example) in total CoCA score in a cognitively impaired individual is reflective of the same degree of change in a cognitively intact individual (Mungas & Reed, 2000). This aspect of the CoCA fills an important gap in modern psychometric test development, which is in need of a test of global functioning that is more psychometrically suitable for tracking dementia progression (Mungas & Reed, 2000). But longitudinal research is needed before we can be confident in the psychometrics for assessing change.

The CoCA is also unique in that it assesses a wide range of executive function abilities, another characteristic that is not present in many cognitive screens. Research suggests that executive functioning can be broken into three subcomponents: information monitoring and updating (working memory), mental set shifting, and inhibition (Miyake et al., 2000). The CoCA may be unique among cognitive screens in that it assesses all three executive functioning subdomains to some extent. The inclusion of tasks that try to equally emphasize verbal and visual skills can provide more comprehensive information about an individual’s strengths and weaknesses. The inclusion of a forced choice recognition task is hypothesized to serve as an embedded test of performance validity and needs to be examined in the future. As discussed earlier, the recording of process variables through the CoCA can potentially assist with differential diagnosis early in the disease process making it a useful tool in the administration of briefer neuropsychological batteries (Au & Devine, 2013). Given that the purpose of the CoCA is to quickly detect
cognitive changes to inform further assessments or treatment recommendations, the fact that CoCA takes only 10 minutes longer than the MoCA makes it a feasible tool that has applicability in various healthcare settings.

**Limitations and Future Directions**

Results from the present study demonstrate the utility of the CoCA as a reliable cognitive screening tool. However, several limitations exist with the present study. The sample size was adequate, yet smaller than desired for CFA based analyses, which also prevented the examination of more complex CFA models of the CoCA. The sample was primarily female, well-educated, and Caucasian, thereby reducing the generalizability of results from this study. Participants in this study may have participated in numerous other psychology studies at UCCS, resulting in a selection bias, and not providing an accurate representation of the Colorado Springs community of older adults. Participants’ inclusion in the present study was based on criteria assessed through self-report and not thorough examination of their medical records or results from screening measures, thus increasing the potential for inclusion of participants who potentially met criteria for various mental health diagnoses including clinical depression and anxiety, and dementia. Additionally, the principal investigator administered and scored all the tests in the present study, thereby increasing the possibility of examiner bias. Another source of possible error variance is from the examiner improving in administration of the CoCA over time that may have influenced results. The tests used for the assessment of convergent and divergent validity may have some limitations as well, but were chosen based on overall content assessed and administration time. For example, a meta-analysis found that the GDS-15 was clinically “poor” at depression case-finding, with some research suggesting
that it assesses distress versus depressive symptomatology (Mitchell, Bird, Rizzo, & Meader, 2010; Smarr & Keefer, 2011). Additionally, the Independent Living Scales (ILS) Health and Safety subtest may be better suited as a measure of convergent validity since it more thoroughly assesses decisions around everyday health and safety and accounts for greater cognitive heterogeneity in scores than the NAB Judgment test (Loeb, 1996).

Even then, the results from the current study are promising, though several steps need to be taken in the near future to further examine and validate the utility of the CoCA for cognitive screening in older adults. First, the CoCA needs to be further validated in various samples: a larger sample of community dwelling older adults, an ethnically and racially diverse sample, and in different clinical populations (e.g., MCI, AD). Results obtained from these additional samples can be used for the development of norms for the CoCA by examining various CFA models, including the originally hypothesized bi-factor model of the CoCA (Figure 1). In order to allow for CFA based scoring of the CoCA, results from these studies will be used to develop an online and manual scoring mechanism for the global factor score of the CoCA. At this time, the global factor score of the CoCA can be calculated using an online scoring tool (https://neuropsych.shinyapps.io/CoCA/). Second, the process scoring mechanism of the CoCA needs to be examined and validated and its utility in differential diagnoses needs to be assessed. Third, measurement invariance of the CoCA with respect to race and ethnicity needs to be explored. Fourth, the ability of the CoCA to track change over time needs to be established since it is likely that it will be used in tracking dementia progression. As such, adequate test-retest reliability and inter-rater reliability are important psychometric properties that need to be established as part of the validation
process of the CoCA. After cross-sectional validation of the CoCA in clinical samples, its ability to predict conversion from MCI to dementia should be investigated. Longitudinal examination of MI of the CoCA and assessment of its convergent and divergent validity over time also needs to be investigated. Once the CoCA has been validated in various samples, another future direction involves examining its utility cross-sectionally and longitudinally in various cultures by translating the CoCA to other languages and revising items based on the values prescribed by each culture.

There are various aspects of the CoCA that can be further improved. For example, the results of the present study suggest that the clock drawing test and similarities test can be eliminated from the CoCA while maintaining its validity as a measure of global cognitive ability. Further analysis of specific items (e.g., the 15 VNT items) on this test will help determine whether any items can be discarded to shorten the time required without a loss of information. Some of the items on the CoCA may be too difficult for someone with moderate to severe dementia. Thus, it is important to examine its use with individuals with advanced dementia and possibly develop a shorter and/or easier version of the test for those with increased levels of cognitive impairment.

**Conclusion**

The CoCA was developed to accomplish two goals that were identified as inadequately met by current cognitive screens: First, accurate and early estimation of global cognitive ability and cognitive change, particularly dementia. Second, early and accurate differential diagnosis. The approach to the content, structure, and scoring of the items included in the CoCA attempted to mitigate some of the weaknesses or limitations of current cognitive screens. The results indicated that in a sample of community
dwelling older adults, the CoCA measures global cognition well in individuals with differing levels of ability while maintaining excellent psychometric properties for a screening measure. The items of the CoCA are also not biased by age, education, sex, or symptoms of depression and anxiety. The CoCA CFA models (even those that were the poorest fitting of the CoCA models) fit the data much better than a CFA model using the MoCA, one of the most widely used screening tests of global cognition. The model fit of the one-factor solution of the CoCA illustrates the construct validity and unidimensionality of the CoCA, and provides support for its use as a measure of global cognition. The use of modern psychometrics, inclusion of items with varying levels of difficulty that assess a wide range of cognitive abilities, potential to examine process scores, and use of items that were not biased by demographic factors and mood, make the case for its use as a reliable and comprehensive screening tool for measuring global cognitive ability. Future research is needed to investigate the process scoring mechanism of the CoCA in addition to cross-validating the results from the present study in various samples and exploring other questions associated with longitudinal validation and measurement invariance.
REFERENCES


Alzheimer’s Society (2007). Dementia, UK. London: Alzheimer’s Society

Aprahamian, I., Martinelli, J. E., Neri, A. L., & Yassuda, M. S. (2010). The accuracy of the Clock Drawing Test compared to that of standard screening tests for Alzheimer's disease: Results from a study of Brazilian elderly with heterogeneous educational backgrounds. International Psychogeriatrics, 22, 64-71. doi:10.1017/S1041610209991141


Gavett, B. E., & Stern, R. A. (2012). Dementia has a categorical, not dimensional, latent structure. Psychology and Aging, 27, 791-797. doi: 0.1037/a0027687


Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variables in

cerebral small vessel disease and Alzheimer's disease. *Journal of the International
Neuropsychological Society, 20*, 413-421. doi:10.1017/S1355617714000101

The impact of verbal and visuospatial ability. *Neuropsychology, 13*, 590-597. doi:
0.1037/0894-4105.13.4.590


Holsinger, T., Deveau, J., Boustani, M., & Williams, J. W. (2007). Does this patient have

Hoops, S., Nazem, S., Siderowf, A. D., Duda, J. E., Xie, S. X., Stern, M. B., &
doi:10.1212/WNL.0b013e3181c34b47

Retrieved from http://www.tc.umn.edu/

Anatomic, clinical, and molecular correlates. *Canadian Journal of Psychiatry,

Houston, W. S., Delis, D. C., Lansing, A., Jacobson, M. W., Cobell, K. R., Salmon, D. P.,
at-risk for Alzheimer's disease: Verbal versus design fluency. *Journal of the
International Neuropsychological Society, 11*, 863-870.
doi:10.1017/S1355617705051015

Addenbrooke's Cognitive Examination III in frontotemporal dementia and
doi: 0.1159/000351671

analysis: Conventional criteria versus new alternatives. *Structural Equation
Modeling: A Multidisciplinary Journal, 6*, 1-55. doi:
10.1080/10705519909540118


Smarr, K. L., & Keefer, A. L. (2011). Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). *Arthritis Care & Research, 63*, S454-S466. doi: 10.1002/acr.20556


APPENDIX A

CoCA INSTRUMENT

COLORADO COGNITIVE ASSESSMENT (CoCA)  Name:  Age:
DOB:  Sex:  Education:  Date:  Points

1. Prospective Memory Test: Suppose the examinee has to make a payment. When you are done with the
test, instruct examinee to remind you to show him the account balance and make a payment.

2. Shape Trail Test (Refer to Supplement 1): Alternate between square and circle in ascending order of
numbers. Discontinue after 2 mins (Zhao et al., 2013). \( \frac{\text{sec}}{\text{correct lines}} \) =

3. Verbal List Learning: Read list of words. Administer 3 trials. Examinee can repeat words in any order.
Please record examinee's answers verbatim. Score = 1 pt for each accurately recalled word in each trial.

| Trial 1 |  |  |  |  |  |  |
|---------|  |  |  |  |  |  |
| Trial 2 |  |  |  |  |  |  |
| Trial 3 |  |  |  |  |  |  |

/21

4. Figure Copy (Refer to Supplement 2): Copy figure. Discontinue after 2 minutes. Score = 1 pt each for
Accuracy and Placement (see below)

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer Rectangle</td>
<td></td>
</tr>
<tr>
<td>Inner Plus</td>
<td></td>
</tr>
<tr>
<td>Inner Diagonals</td>
<td></td>
</tr>
<tr>
<td>Concentric Rectangles</td>
<td></td>
</tr>
<tr>
<td>Hoops</td>
<td></td>
</tr>
<tr>
<td>Inner Triangle</td>
<td></td>
</tr>
<tr>
<td>Inner Vertical Lines</td>
<td></td>
</tr>
<tr>
<td>Inner Circles</td>
<td></td>
</tr>
<tr>
<td>Ramparts</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

/18

5. Animal Fluency (Refer to Supplement 3): Discontinue after 1 minute. Score = 1 pt for each correct word

6. Lexical Fluency "5" (Refer to
Supplement 3): Discontinue after 1 minute.
Score = 1 pt for each correct word

7. Working Memory: Examinee says
numbers in reverse order of presentation.
Score: 0.5 pt for correct digit, 0.5 for
correct placement of digit

Sample:

| 5 | Q | 0 | L | 3 |

| Digit |  |
| Order |  |

/3

Trial 1:

| 2 | F | 7 | C | S | 9 | N |

| Digit |  |
| Order |  |

/5

Trial 2:

| 3 | T | S | 7 | V | 0 | B | K | 1 | Y | 6 | U |

| Digit |  |
| Order |  |

Assessor

CoCA Total Page 1:
COLORADO COGNITIVE ASSESSMENT (CoCA)

Name:  
DOB:  
Sex:  
Education:  
Age:  
Date:  

8. **Focused Attention:** When you say *eyes,* examinee says *see.* When you say *nose,* examinee says *smell.* Place a checkmark over all correct responses including the lack of a response when appropriate.

Score: 1 pt for all correct responses.

Eyes; aisle; note; iron; nose; small; eyes; seat; note; nose; hose; nod; aisle; eyes; seat; nod; nose; eyes; aisle; nose; small; eyes; nod; iron; note; eyes; seat; lies; nose; aisle; iron; note; lies; nose; small; hose; eyes; nose; aisle; lies; hose; note

9. **Clock Command (Refer to Supplement 4):** Draw clock with all the numbers. Show time at 10 past 11. Refer to scoring in manual. Score is out of 10 points.

/10

10. **Clock Copy (Refer to Supplement 5):** Copy clock. Same scoring criteria as item 9 above.

/10

11. **Similarities:** State how two things are similar. Score: 0 = no similarity, 1 = concrete response, 2 = abstract response

Sample: How are a carrot and cucumber alike (vegetables)

a. How are television and radio alike (modes of information/entertainment) /2

b. How are tall and short alike? (ways of describing height) /2

12. **Number Symbol Matching Test (Refer to Supplement 6):** Discontinue after 30 sec. Score = 1 pt for all correct responses.

/81

13. **Verbal Naming:** Examinee has to identify what you are describing. Place a check mark if cue was given. Score = 1 pt for all correct responses without cue. (Adapted from Yochim et al., 2015).

<table>
<thead>
<tr>
<th>Item</th>
<th>Answer</th>
<th>Cue</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What ice does when it gets hot</td>
<td>Mist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. What you use to measure how many inches something is</td>
<td>Ruler</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. A long, severe snowstorm</td>
<td>Blizzard</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. What you put your head on to sleep at night</td>
<td>Pillow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. The part of the shirt that covers your arms</td>
<td>Sleeves</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. A person who works at a drug store to fill prescriptions</td>
<td>Pharmacist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. A moving set of stairs</td>
<td>Escalator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. A device that measures temperature</td>
<td>Thermometer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. What do you do to a pencil or knife that becomes dull</td>
<td>Sharpen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. A place people go to gamble money</td>
<td>Casino</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. When you take a deep breath when your sleepy or bored</td>
<td>Yawn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. A baby cow</td>
<td>Calf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. A desert plant that has spikes</td>
<td>Cactus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. A toy that has a string and floats in the air when it is windy</td>
<td>Kite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. The poison a snake uses to kill its prey</td>
<td>Venom</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
COLORADO COGNITIVE ASSESSMENT (CoCA)

Name:  
DOB:  
Sex:  
Education:  
Age:  
Date:  
Points:  

14. Orientation: **Score = 1 pt for each correct response**

<table>
<thead>
<tr>
<th>Date</th>
<th>Month</th>
<th>Year</th>
<th>Day</th>
<th>Place</th>
<th>City</th>
<th>State</th>
<th>President</th>
</tr>
</thead>
</table>

/8

15. Verbal List Learning Delayed Recall: Ask examinee to state all words remembered from the list. Provide category cue for all words not recalled. Record responses verbatim. **Score: 1 pt for each correct response**

<table>
<thead>
<tr>
<th>Free Recall</th>
<th>Car</th>
<th>River</th>
<th>Market</th>
<th>Desk</th>
<th>Foot</th>
<th>Square</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category Cue</td>
<td>(automobile)</td>
<td>(water body)</td>
<td>(place where you buy something)</td>
<td>(furniture)</td>
<td>(part of the body)</td>
<td>(geometric figure)</td>
<td>(type of job)</td>
</tr>
</tbody>
</table>

/14

16. Verbal List Learning Yes/No Recognition: Read each word listed below and ask examinee to answer in either yes or no. **Score = 1 pt for each correct answer.**

1. Foot Y/N  
2. Table Y/N  
3. River Y/N  
4. Store Y/N  
5. Square Y/N  
6. Desk Y/N  
7. Hand Y/N  
8. Car Y/N  
9. Market Y/N  
10. Artist Y/N  
11. Rectangle Y/N  
12. Stream Y/N  
13. Bus Y/N  
14. Teacher Y/N  

/14

17. Verbal List Learning Forced Choice Recognition: Ask examinee to pick correct word.

1. Seven or Car  
2. Fight or Teacher  
3. Television or Foot  
4. Market or Voice  
5. Desk or Country  
6. Game or Square  
7. River or Education  

/7

18. Figure Delayed Recall (Refer to Supplement 7): Ask examinee to draw figure from memory. **Score = 1 pt each for Accuracy and Placement (see below)**

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer Rectangle</td>
<td></td>
</tr>
<tr>
<td>Inner Plus</td>
<td></td>
</tr>
<tr>
<td>Inner Diagonals</td>
<td></td>
</tr>
<tr>
<td>Concentric Rectangles</td>
<td></td>
</tr>
<tr>
<td>Hoops</td>
<td></td>
</tr>
<tr>
<td>Inner Triangle</td>
<td></td>
</tr>
<tr>
<td>Inner Vertical Lines</td>
<td></td>
</tr>
<tr>
<td>Inner Circles</td>
<td></td>
</tr>
<tr>
<td>Ramparts</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

/18

19. Prospective Memory Test Delayed Recall: Tell the examinee that you are done with the test. **Score = 1 pt if examinee states that he or she has to make a payment**

/1

20. Test of Functional Abilities: Show examinee Supplement 8.

20A. Do you have enough money in your account to make your utilities payment of $189.75? 
Yes  
No  
If No, continue to item 8. If Yes, tell them that the answer is No. Then ask, item 8 below.  

/1

20B. What will happen if you make this payment from your account anyway? (Score = 1 pt for each of the following: still have to pay bill, overdrawn account, and fee charged)  

/3

CoCA Total Page 3:
### Animal and Lexical Fluency

<table>
<thead>
<tr>
<th>Animals</th>
<th>“S” Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 secs</td>
<td></td>
</tr>
<tr>
<td>30 secs</td>
<td></td>
</tr>
<tr>
<td>45 secs</td>
<td></td>
</tr>
<tr>
<td>1 min</td>
<td></td>
</tr>
</tbody>
</table>
Supplement 6  
Number Symbol Matching Test  

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Name:  
Date:  

| |海水|海水|海水|海水|海水|海水|海水|海水|
|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|

|海水|海水|海水|海水|海水|海水|海水|海水|
Supplement 7
Figure Delayed Recall

Name:
Date:
XYZ Bank Combined Statement of Accounts
Primary account number: 8887779999 ■ May 18, 2016 - June 17, 2016 ■ Page 1 of 3

Happy Smile
123 Happy Lane Apt 555
Sample City ZZ 88888-7777

You and XYZ Bank
Thank you for being a loyal XYZ Bank customer. We value your trust in our company and look forward to continuing to serve you with your financial needs.

IMPORTANT ACCOUNT INFORMATION
When can your account be closed?

We can close your account at any time. If the account is closed, we may send the remaining balance on deposit in your account by traditional mail or credit it to another account you maintain with us.

Summary of Accounts
Checking/Prepaid

<table>
<thead>
<tr>
<th>Account</th>
<th>Page</th>
<th>Account number</th>
<th>Ending balance last statement</th>
<th>Ending balance this statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>XYZ Checking</td>
<td>1</td>
<td>8887779999</td>
<td>$1,442.41</td>
<td>$172.06</td>
</tr>
</tbody>
</table>

Activity Summary
Beginning balance on 05/18 $1,442.41

Deposits/Additions $500.00

Withdrawals/Subtractions 1,770.35

Ending balance on 06/16 $172.06

Transaction History

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
<th>Deposits/Additions</th>
<th>Withdrawals/Subtractions</th>
<th>Ending Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/19</td>
<td>McDonald's #F21389 Parker</td>
<td>$8.41</td>
<td></td>
<td>$1434.00</td>
</tr>
<tr>
<td>05/25</td>
<td>Deposit 15859375</td>
<td>$500</td>
<td></td>
<td>$1934.00</td>
</tr>
<tr>
<td>05/28</td>
<td>123 Apartments, Colorado Springs, CO</td>
<td>$1000.00</td>
<td></td>
<td>$934.00</td>
</tr>
<tr>
<td>06/02</td>
<td>King Soopers, Colorado Springs, CO</td>
<td>$52.70</td>
<td></td>
<td>$881.30</td>
</tr>
<tr>
<td>06/03</td>
<td>United Airlines</td>
<td>$487.65</td>
<td></td>
<td>$393.65</td>
</tr>
<tr>
<td>06/05</td>
<td>Colorado Springs Utilities</td>
<td>$98.37</td>
<td></td>
<td>$295.28</td>
</tr>
<tr>
<td>06/09</td>
<td>Applebees, Colorado Springs, CO</td>
<td>$36.89</td>
<td></td>
<td>$258.39</td>
</tr>
<tr>
<td>06/10</td>
<td>Lowes Hardware, Colorado Springs, CO</td>
<td>$76.14</td>
<td></td>
<td>$182.25</td>
</tr>
<tr>
<td>06/13</td>
<td>Starbucks, Colorado Springs, CO</td>
<td>$10.19</td>
<td></td>
<td>$172.06</td>
</tr>
</tbody>
</table>
## APPENDIX B

### CoCA PROCESS VARIABLES RECORD FORM

#### 2. Shape Trails Test

<table>
<thead>
<tr>
<th>Process Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing Error</td>
</tr>
<tr>
<td>Set Loss Error</td>
</tr>
<tr>
<td>Capture Error</td>
</tr>
<tr>
<td>Visuospatial Neglect</td>
</tr>
<tr>
<td>Visual Difficulties</td>
</tr>
<tr>
<td>Motor Difficulties/Tremors</td>
</tr>
<tr>
<td>Losing track and starting from wrong target</td>
</tr>
</tbody>
</table>

#### 4. Figure Copy & 18. Figure Recall

<table>
<thead>
<tr>
<th>Process Variables</th>
<th>Copy</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perseverations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorly Planned</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right to Left Organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neglect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Micrographia/Motor issues/Tremors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy/Placement Ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent retained</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 3, 15, & 16. Verbal List Learning Test

<table>
<thead>
<tr>
<th>Process Variables</th>
<th>IR</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic Intrusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonemic Intrusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novel Intrusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat Intrusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetitions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 5. Animal Fluency & 6. Lexical Fluency

<table>
<thead>
<tr>
<th>Process Variables</th>
<th>Animals</th>
<th>S Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epoch 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epoch 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epoch 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epoch 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perseverations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Correction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent set loss</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% perseveration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule Violations</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 7. Test of Working Memory

<table>
<thead>
<tr>
<th>Process Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omissions</td>
</tr>
<tr>
<td>Additions/Commissions</td>
</tr>
<tr>
<td>Perseverative Errors</td>
</tr>
<tr>
<td>Set Loss Errors</td>
</tr>
<tr>
<td>Sequencing Errors</td>
</tr>
</tbody>
</table>

### 8. Test of Focused Attention

<table>
<thead>
<tr>
<th>Process Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omissions</td>
</tr>
<tr>
<td>Additions/Commissions</td>
</tr>
<tr>
<td>Perseverative Errors</td>
</tr>
<tr>
<td>Set Loss Errors</td>
</tr>
</tbody>
</table>

### 9. Clock Command

<table>
<thead>
<tr>
<th>Process Variables</th>
<th>Clock Command</th>
<th>Clock Copy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of Clock</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphical Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Bound Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conceptual Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial/Planning Deficits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perseveration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neglect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotated Paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-correction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requested Reminder of Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Command to Copy Ratio</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplement 9
Process Scoring Sheet

### 11. Similarities

| Process Variables |  
|-------------------|---
| In-Set Errors     |  
| Vague Response    |  
| Subordinate Response |  
| Out-of-Set Errors |  
| One object Response |  
| Juxtaposition Response |  
| Different Responses |  

### 12. Number Symbol Matching Test

| Process Variables |  
|-------------------|---
| Number of Errors  |  
| Stimulus Pull Errors |  
| Perseveration of Symbols |  
| Overwriting/ Micrographia/ Tremors |  
| Expansion of Symbols |  

### 13. Verbal Naming

| Process Variables |  
|-------------------|---
| # of phonemic cues given |  
| # correct with phonemic cues |  
| Phonemic Paraphasia |  
| Semantic Paraphasia |  
| Circumlocution |  
| Perseveration |  

**Behavioral Observations Checklist:** Please indicate the presence and rate the severity of any of the below mentioned behavioral features:

<table>
<thead>
<tr>
<th></th>
<th>None</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Distractibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Impulsivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Reduced/Lack of Affect and Social Engagement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Socially Inappropriate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Decreased Initiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Agitation/Aggression/ Irritation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Other:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The CoCA is made of 20 items. Items have a total score and various process scores. The total score is indicated on the main pages of the CoCA; total scores are added together using a CFA model to arrive at the global score of the CoCA. Based on the CFA model described in this manuscript, an online scoring program is available to obtain the global factor score of the CoCA (https://neuropsych.shinyapps.io/CoCA/). There is an additional supplement to assist with scoring of process variables. Process variables are not reflected in the global score of the CoCA. Specific process variables are identified for items; the examiner can use the process variables supplement to indicate the presence and number of item-specific qualitative features.

In general, each correct response receives one point unless otherwise specified. Items that are not scored are indicated as such in this manual and their scoring sections are greyed out on the CoCA.

Examiners are encouraged not to affirm or respond to examinee’s responses. If the examinee makes a mistake, do not correct them, unless otherwise specified.

No additional materials besides a pen and a stopwatch are needed for the CoCA. All stimulus materials needed for administration of the CoCA are available as supplemental pages. Please try and record examinee’s responses verbatim as much as possible.

1. Prospective Memory Test

Materials Needed: None

Instructions: Suppose you have to look at a bank statement in a little while. Only when I say that the test is over, I want you to remind me that you have to look a bank statement.

Scoring: This item is not scored

* If examinee reminds you about looking at bank statement after the shape trail test, restate instructions and emphasize that the reminder only has to occur when the examiner states that the test is over.
2. Shape Trail Test *(Adapted from Zhao et al., 2013)*

**Materials Needed:** Supplement 2 and Stopwatch

**Instructions:** Here *(point to stimulus)*, you see numbers in circles and squares. I want you to draw a line from 1 in the square to the 2 in the circle to the 3 in the square *(keep pointing)* and just like that alternating between square and circle in ascending order of the numbers till you reach the end here *(point)*. You will have two minutes. Ready? Go!

![Shape Trail Test Diagram](image)

**Scoring:**

- **Global:**
  - Total time (seconds)/ no of correct lines – maximum number of correct lines is 7
- **Process:**
  - Sequencing error
  - Set loss error
  - Capture error
  - Visuospatial neglect
  - Visual difficulties
  - Motor difficulties/tremors
  - Losing track and starting from wrong target

3. Verbal List Learning Test

**Materials Needed:** None.

**Instructions:** I am going to read you a list of words. When I am done, I want you to tell me all the words you remember, in any order. Ready? *(read list at 1 word per second).* Now, tell me all the words you remember.

**Repeat twice:** I am going to read you the same list of words one more time. When I am done, tell me all the words you remember, including the ones you have already told me before. You can tell me the words in any order.
At the end of the test: Try and keep these words in mind because you may be asked about them at the end of the test.

Scoring:
- Global:
  - 1 point for each word correctly recalled in each trial for a maximum of 21 points across all three learning trials
- Process:
  - Intrusions
    - Semantic
    - Phonemic
    - Novel
  - Repetitions

4. Figure Copy

Materials Needed: Supplement 3 and Stopwatch

Instructions: I want you to copy this figure as accurately as you can in the space below. You will have two minutes. Ready? Go!
• 1 point or half point for each accuracy component and one point for each placement component described below
• Maximum of 9 points for accuracy and 9 points for placement for a total of 18 points for the test.

Components:
• Outer rectangle
  o Accuracy:
    0.5 pt: not accurately drawn but resembles a rectangle even if lines are wavy and not straight. Rectangle may be much smaller or larger than that of the actual figure.
    1 pt: all four lines are present; lines are approximately straight (slight waves and breaks are acceptable); angles formed are approximately 90 degrees; width of rectangle is obviously longer than height.
  o Placement: reasonably close to center of blank space provided, not rotated more than 15 degrees.
• Inner plus
  o Accuracy:
    0.5 pt: not accurately drawn but resembles a plus within the rectangle even if lines are wavy, not straight, and do not bisect each other or touch the outer rectangle.
    1 pt: all four lines are present; lines are approximately straight (slight waves and breaks are acceptable); should approximately bisect each other.
  o Placement: lines should approximately bisect all four sides of the rectangles; should touch outer rectangle, and not end before or extend outer rectangle by more than 1/4 inch; lines should intersect the inner diagonals approximately at the point of intersection of inner diagonals.
• Inner diagonals
  o Accuracy:
    0.5 pt: not accurately drawn but resembles two intersecting diagonals within the rectangle even if lines are wavy, not straight, and do not intersect each other.
    1 pt: both lines are present; lines are approximately straight (slight waves and breaks are acceptable); should approximately bisect each other.
  o Placement: ends of lines should meet corners of the rectangle without significant overlap or measurable distance between the ends of the lines and corners; should touch outer rectangle, or not end before or extend outer rectangle by more than 1/4 inch; lines should intersect the inner plus approximately at the point of intersection of inner plus.
• Concentric rectangles
  o Accuracy:
0.5 pt: not accurately drawn but resembles a single or concentric (with three or four lines) rectangles even if lines are wavy, not straight, and touch each other.

1 pt: three concentric rectangles with lines that do not touch each other; the three rectangles decrease in size by approximately the same amount; horizontal sides of rectangles are obviously longer than vertical sides; lines are continuous or only have only slight waves and/or breaks; the size of the largest concentric rectangle is proportionate to the size of the outer rectangle (height and width do not exceed 50% of height and width of outer rectangle).

- **Placement:** placed above the outer rectangle on the left side; the base of the concentric rectangles touches the top left horizontal segment of the outer rectangle; not rotated more than 15 degrees. The right vertical segment of the inner concentric rectangle is roughly aligned with the right edge of the inner triangle.

- **Hoops**
  - **Accuracy:**
    - 0.5 pt: not accurately drawn but resembles any number of vertical semicircles of varying sizes (looping allowed) that do not continue beyond the bottom of the left hand corner.
    - 1 pt: Approximately nine (±2) small vertical semicircles that overlap with each other; semicircles are relatively equal in size; touch the left lower vertical side of the rectangle. Should be drawn separately rather than in a continuous motion (i.e., no looping on the side closest to the outer rectangle). Hoops do not continue beyond the bottom left hand corner.

- **Placement:** placed on the outside of left vertical side of the rectangle; hoops are placed from the point of horizontal bisection of the height of the outer rectangle to the base of the rectangle.

- **Inner triangle**
  - **Accuracy:**
    - 0.5 pt: resembles a triangle of any size; any figure resembling a triangle (even if without third side).
    - 1 pt: angle formed by three sides of triangle are all less than 90 degrees; sides are approximately straight (slight waves and breaks are acceptable).

- **Placement:** vertical side of triangle touches or almost touches left inner top and bottom diagonals; vertex of triangle touches or almost touches the point of bisection of the left vertical side of rectangle.

- **Inner vertical lines**
  - **Accuracy:**
0.5 pt: Three vertical lines are present, even if they are of varying lengths and do not decrease progressively in length.
1 pt: three vertical lines are continuous and straight (slight waves and breaks are acceptable); the three lines visibly decrease in length from left to right; spacing between lines is roughly equal and is proportionate to the figure.

- Placement: in appropriate segment (upper half of top right quadrant); touching the top right horizontal side of the outer rectangle and the right upper diagonal.

- Inner circles
  - Accuracy:
    - 0.5 pt: five dots or circle that are present in an obvious arrangement (does not have to match arrangement in figure) that do not touch each other or any other part of the figure.
    - 1 pt: five circles that are round and closed (filled dots are not acceptable); are continuous and any breaks are only due to mild sloppiness; circles roughly equal size; in the depicted arrangement and not rotated more than 15 degrees; not touching each other or any other part of the figure.
  - Placement: in appropriate segment (upper half of bottom right quadrant).

- Ramparts
  - Accuracy:
    - 0.5: an arrangement of any number of figures that do not have look like Ms and triangles; the arrangement may be sloppy/messy looking
    - 1 pt: three Ms and three right angled triangles (with unclosed bases) in a horizontal arrangement; Ms and triangles are arranged in an alternating manner; Ms and triangles are connected to each other by short horizontal lines; Ms and triangles are equal in height; all Ms are roughly equal in width; all triangles are roughly equal in width. Height of ramparts arrangement does not exceed 25% of length of the right vertical side of the outer rectangle.
  - Placement: ramparts arrangement protrudes from right lower corner of the rectangle such that it appears to be a continuation of the bottom horizontal side of the rectangle

- Process:
  - Perseverations
• Spatial errors
• Poorly planned
• Right to left organization
• Neglect
• Micrographia/motor issues/tremors
• Accuracy/placement ratio

5. Animal Fluency

Materials Needed: Supplement 3 and Stopwatch

Instructions: Now, I am going to give you one minute to name as many animals as you can. Ready? Go!

Scoring:
• Global:
  • 1 point for each correctly generated word
• Process:
  • Number of words produced every 15 seconds (Epochs 15, 30, 35, and 60)
  • Perseverations
  • Intrusions
  • Self correction
  • Percent set loss
  • Percent perseveration
  • Semantic index = animal/(animal + letter)

6. Lexical Fluency

Materials Needed: Supplement 3 and Stopwatch

Instructions: Now, we are going to do something a little different. I am going to give you one minute to name as many words as you can think of that belong to a particular letter of the alphabet that I will tell you in a moment. You can say any kind of word you like except for proper nouns (like Bob or Boston), numbers, or words that have the same beginning but different endings (like play, player, playing). Ok, now you have one minute to name as many words as you can think of that begin with the letter “S”

Scoring:
• Global:
• 1 point for each correctly generated word
• Process:
  • Number of words produced every 15 seconds (Epochs 15, 30, 35, and 60)
  • Perseverations
  • Intrusions
  • Self correction
  • Percent set loss
  • Percent perseveration
  • Rule violations

7. Test of Working Memory

Materials Needed: None.

Instructions: I am going to read you some numbers and letters. When I am done, I want you tell me all the numbers, in the reverse order of their presentation. Lets start with a sample (read one letter/digit per second):

• Sample: 5Q0L3 (Answer: 305)

If correct, move on to the test items.
If incorrect, say - I said 5Q0L3; the order of the presented numbers was 503, so backwards it would be 305. Understood? Ok, I am going to move on to the test item.

• Trial 1: 2F7CS9N (Answer: 972)
• Trial 2: 3TS7V0BK1Y6U (Answer: 61073)

Scoring:
• Global:
  • 0.5 pt for correct digit
  • 0.5 pt for correct order of digit
  • Maximum of 3 points on Trial 1 and 5 points on Trial 2
  • The addition of a number at the end of a sequence, for example, 9720 instead of 972, results in a total score of 2.5, with credit given for the three correct digits (9, 7, & 2) and credit given for the correct order of 9 and 7. In this case, credit is not given for the correct order of 2
  • The addition of an incorrect number in the sequence, for example, 9702 instead of 972, results in a total score of 2.5, with credit given for the three correct digits (9, 7, & 2) and credit given for the correct order of 9 and 7. In this case, credit is not given for the correct order of 2

• Process:
  • Omissions
8. Test of Focused Attention

Materials Needed: None

Instructions: I am going to read you a list of words. When I say eyes, you say see. When I say nose, you say smell. If I say a different word, do not say anything. Understand? Ready? (Read words below in order at the rate of 1 word/second).

Eyes; aisle; note; iron; nose; small; eyes; seat; note; nose; hose; nod; aisle; eyes; seat; nod; nose; eyes; aisle; nose; small; eyes; nod; iron; note; eyes; seat; lies; nose; aisle; iron; note; lies; nose; small; hose; eyes; nose; aisle; lies; hose; note

* If the participant forgets instructions during the task, remind once quickly, but keep going with the task and do not restart the task.

Scoring:
- Global:
  - 1 point for every correct response including not responding when appropriate (for example, not responding when the examiner says aisle), for a maximum of 42 points
  - If an examinee makes an error and then self-corrects, do not give credit for the correct. Instead, count the error and make a note of the number of self-corrects.

- Process
  - Omissions
  - Additions/commissions
  - Set loss errors
  - Perseverative errors

9. Clock Command

Materials Needed: Supplement 4 and Stopwatch

Instructions: I want you to draw the face of a clock with all the numbers in it and have the time showing ten past 11.

Scoring: see below*

10. Clock Copy
Materials Needed: Supplement 5 and Stopwatch

Instructions: Now, I want you to copy this clock as accurately as possible in the space below.

*Scoring for clock copy and command:
- Global: Use scoring system below (Rouleau et al., 1992), for a maximum of 10 points for each (command and copy)

**Scoring System for Clock Drawing Command and Copy**

- **Integrity of the clock face (maximum: 2 points)**
  2: Present without gross distortion
  1: Incomplete or some distortion
  0: Absent or totally inappropriate

- **Presence and sequencing of the numbers (maximum: 4 points)**
  4: All present in the right order and at most minimal error in the spatial arrangement
  3: All present but errors in spatial arrangement
  2: Numbers missing or added but no gross distortions of the remaining numbers; numbers placed in counterclockwise direction; numbers all present but gross distortion in spatial layout (i.e., hemineglect, numbers outside the clock)
  1: Missing or added
  0: Absence or poor representation of numbers

- **Presence and placement of the hands (maximum: 4 points)**
  4: Hands are in correct position and the size difference is respected.
  3: Slight errors in the placement of the hands or no representation of size difference between the hands.
  2: Major errors in the placement of the hands (significantly out of course including 10 to 11)
  1: Only one hand or poor representation of two hands
  0: No hands or perseveration on hands

- Process:
  - **Size of clock**: small if its measures less than 1.5 inches and large if it measures more than 5 inches
  - **Graphical errors**: lines are not precise resulting in distortions of the clock face or making the numbers difficult to read. The hands are not straight and sometimes fail to connect in the middle. The overall performance appears inaccurate and clumsy.
  - **Stimulus-bound errors**: hands are set for 10 to 11 instead of 10 past 11; time is written in numbers/letters besides 10 and/or 11 or between 10 and 11; hands are either absent or pointed to 10 and/or 11.
• **Conceptual errors:** clock face without numbers or inappropriate use of numbers; time is misrepresented on the clock either by writing it on the clockface or the hands are absent or inadequately represented
• **Spatial/planning deficits:** neglect of the left hemisphere; planning difficulties with inappropriate gap before 12 or 3, 6, or 9; disorganization in the layout of the numbers, numbers written outside the clockface, numbers written counter clockwise
• **Perseveration:** presence of more than 2 hands; ongoing tracing of hands; ongoing prolongation of numbers; repetition of numbers
• **Neglect**
• Rotated paper while placing numerals/numeral substitutes
• Attempt to self-correct significant error
• Requested reminder of time for hand-setting
• Copy to command ratio

11. **Similarities**

**Materials Needed:** None

**Instructions:** *Now I am going to ask you some questions about how two things are alike.*  
(Read sample item below)

Sample:
- How are a potato and an onion alike?

Answer: *vegetables.*
If correct, say: *yes, they are vegetables (Move on to test items)*
No answer or incorrect answer: *Can you think of another way in which they are alike? They are vegetables.*

Test:
1. How are television and radio alike?
   - 2 point: *modes of entertainment/information/communication/media*
   - 1 point: *electric/electrical; play/give out/broadcast signals, sound, news, or music; receive/transmit signals electronically*
   - 0 point: *send words/visions*

2. How are tall and short alike?
   - 2 point: *ways of describing height/size/measurement/length/persons stature*
   - 1 point: *variations/opposites of a dimension/scale*
   - 0 point: *not alike; opposites; antonyms; descriptors; adjectives*

**Scoring:**
- Global:
  - 0 = for no similarity or pointing out differences
• 1 = concrete response
• 2 = abstract response

• Process (Giovanetti et al., 2001):
  • In set errors:
    • Vague response: provision of a superordinate, but superficial categorical response (e.g., dog – lion: “they eat”)
    • Subordinate response: provision of responses that relate to shared concrete traits (e.g., orange - banana: they have skin) or specific properties about the items that may not be correct in all instances (e.g., boat-car: they have motors)
  • Out of set errors:
    • One object response: provision of a response that only addresses one word from the pair (e.g., train-bicycle: one is faster than the other)
    • Juxtaposition responses: provision of a response that describes the interaction of one member of the word pair with the other (e.g., cat-mouse: the cat can eat the mouse)
    • Different responses: provision of a response that accurately describes how the two words of the pair are different (e.g., tongue-nose: you smell with your nose and taste with your tongue).

12. Number Symbol Matching Test

Materials Needed: Supplement 6A and 6B and Stopwatch

Instructions:

Written version (Supplement 6A):

Here, (point to the key), you see two rows; there are symbols in the top row and the bottom row consists of corresponding numbers. Each symbol has its own number.

Now, here (point), you see rows that have symbols in the top part and are empty in the bottom part. I want you fill in the empty boxes with their corresponding numbers.

Look at the first symbol (point to first symbol on Sample). I look at this symbol, then look at the key. The symbol matches the number 3. So I go ahead and put that in the empty box below the symbol (write 3 below the first symbol on the Sample)

Look at the next symbol (point to the second symbol on Sample). I look at this symbol, then look at the key. The symbol matches the number 7. So I go ahead and put that in the empty box below the symbol (write 7 below the second symbol on the Sample)

Look at the next symbol (point to the third symbol on Sample). I look at this symbol, then look at the key. The symbol matches the number 9. So I go ahead and put that in the empty box below the symbol (write 9 below the third symbol on the Sample)
Now, I want to you to fill in the boxes until this heavily shaded line.

Now, I am going to give you 30 seconds to fill in as many boxes as you can. Begin here (point) and fill in as many boxes as possible, without skipping any. When you finish one row move on to the next row. Continue working until I ask you to stop. Ready? Go!

**Oral Version (Supplement 6B):**

Now, we are going to do something the same thing a little differently. I want you verbally state the numbers that match the presented symbols.

For example, look at the first symbol (point to first symbol on Sample). I look at this symbol, then look at the key. The symbol matches the number 3. So I want you to say 3.

Now, I am going to give you 30 seconds to state as many numbers matching the presented symbols as fast as you can. When you finish one row move on to the next row. Continue working until I ask you to stop. Ready? Go!

**Key:**

<table>
<thead>
<tr>
<th>)</th>
<th>S</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Z</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

**Answers:**

| Sample: 1 7 9 8 7 1 5 2 6 |
| 1 6 9 8 7 3 2 5 3 |
| 9 4 6 9 1 6 8 2 7 |
| 3 6 8 7 9 2 1 5 4 |
| 5 6 7 1 3 2 8 7 3 |
| 5 2 9 7 1 4 6 3 8 |
| 9 4 2 5 1 7 4 6 8 |
| 5 4 9 2 6 1 7 9 5 |
| 6 3 1 7 8 2 5 4 9 |
| 7 1 3 5 9 1 7 8 3 |

**Scoring:**

- Global:
  - 1 point for each correct answer
- Process:
- Number of errors
- Stimulus pull errors
- Perseveration of symbols
- Overwriting/ micrographia/ tremors
- Expansion of symbols beyond boxes
- Set loss errors

13. **Verbal Naming Test** *(Adapted from Yochim et al., 2015).*

**Materials Needed:** Stopwatch

**Instructions:** Now, we are going to do something different. I am going to describe an object or a verb and I want you to tell me the name of what I am describing. *(Read test item as presented)*

Allow examinee 10 sec to respond. If examinee gives an incorrect response, say “No, its something else.”

If incorrect or no response is provided in 10 seconds, provide the phonemic cue by saying, “It starts with the sound...(state the underlined part of the answer).”

If incorrect or response is provided after 10 seconds from the phonemic cue, proceed to next item

**Items:**

1. What ice does when it gets hot – **Melt**
2. What you use to measure how many inches something is – **Ruler**
3. A long, sever snowstorm - **Blizzard**
4. What you put your head on to sleep at night – **Pillow**
5. The part of your shirt that covers your arms – **Sleeves**
6. A person who works at a drug store to fill prescriptions – **Pharmacist**
7. A moving set of stairs – **Escalator**
8. A device that measures temperature – **Thermometer**
9. What do you do to a pencil or knife that becomes dull – **Sharpen**
10. A place people go to gamble money – **Casino**
11. When you take a deep breath when your sleepy or bored – **Yawn**
12. A baby cow – **Calf**
13. The desert plant that has spikes – **Cactus**
14. A toy that has a string and floats in the air when it is windy – **Kite**
   * If examinee says balloon, state: “can you think of another word?”
15. The poison a snake uses to kill its prey – **Venom**

**Scoring:**

- Global
  - 1 point for every correct answer without phonemic cue for a maximum of 15 points
• Process:
  • Number of phonemic cues given
  • Number correct with phonemic cues
  • Phonemic paraphasia
  • Semantic paraphasia
  • Circumlocution
  • Perseverations

14. Orientation:

Materials Needed: None

Instructions: Ask questions in order presented below. What is the...
1. Date
2. Month
3. Year
4. Day of week
5. Place
6. City
7. State
8. President of the country

Scoring:
• Global:
  • 1 point for every correct response for a maximum of eight points
  • There are no process scores for this test.

15. Verbal List Learning Delayed Recall

Materials Needed: None

Instructions: A little while ago, I read you a list of words. Please tell me all the words you remember from that list. (Write examinee’s responses verbatim, including any additional words that the examinee may state).

Scoring: See below +

Instructions: For the words not recalled on Delayed Recall, state: One of the words I read to you was a type of (eg. automobile). Do you remember that word?(See below for categories)

1. Car (automobile)
2. River (water body)
3. Market (place where you buy something)
4. Desk (furniture)  
5. Foot (part of the body)  
6. Square (geometric figure)  
7. Teacher (type of job)  

**Scoring:** See below.

16. **Verbal List Learning Yes/No Recognition**

**Materials Needed:** None

**Instructions:** Now I am going to read you some more words. Some of these words were on the list I read to you earlier, and others are new words that you have not heard before. For each word, I want you to tell me in Yes or No, whether you think the word was on the list. Was (word) on the list? Yes or No?

1. Foot  
2. Table  
3. River  
4. Store  
5. Square  
6. Desk  
7. Hand  
8. Car  
9. Market  
10. Artist  
11. Rectangle  
12. Stream  
13. Bus  
14. Teacher  

**Scoring:** See below.

17. **Verbal List Learning Forced Choice Recognition**

**Materials Needed:** None
**Instructions:**
1. Was it seven or car
2. Was it fight or teacher
3. Was it television or foot
4. Was it market or voice
5. Was it desk or country
6. Was it game or square
7. Was it river or education

**Scoring:** See below +

*Scoring for all Verbal List Learning Items:*
- **Global:**
  - 1 point for each correct response on Delayed Recall, Yes/No Recognition, and Forced Recognition for a maximum of 14, 14, and 7 points respectively.
  - The Global score from Forced Recognition does not count towards the CoCA total score.
- **Process (applies to Delayed Recall only):**
  - Repetitions
  - Intrusions
  - Semantic
  - Phonemic
  - Novel
  - Repeat intrusions from immediate recall

**18. Figure Delayed Recall**

**Stimulus:** Supplement 7 and Stopwatch

**Instructions:** A little while ago, you copied a figure. Now, I would like you draw as much of that figure as you remember now. If you remember a part and are not sure where it goes, put it anywhere. Try and draw as much of it as you remember until I ask you to stop. You will have two minutes. Ready? Go!

*Note:* If examinee starts to draw a clock, remind him/her that he/she copied another figure as well.

**Scoring:**
- **Global:**
  - 1 point or half point for each accuracy component and one point for each placement component in the Figure Copy test
• Maximum of 9 points for accuracy and 9 points for placement for a total of 18
  points for the test.

• Process:
  • Perseverations
  • Spatial errors
  • Poorly Planned
  • Right to left organization
  • Neglect
  • Micrographia/ motor issues/ tremors
  • Accuracy/placement ratio
  • Percent retained

19. Prospective Memory Test Delayed Recall

**Stimulus:** None

**Instructions:** The test is over. (Wait to see if the examinee responds by saying that he or she has to look at a bank statement; do not provide any hints or look expectantly).

**Scoring:**
• Global:
  • 1 point if examinee states that he or she has to look at a bank statement or bank account. No points are given if examinee remembers that they have to do something else, for example, make an appointment or balance a check book.
  • There are no process scores for this test.

20. Test of Functional Abilities

**Stimulus:** Supplement 8

**20A. Instructions:** (Show examinee bank statement and say) – Suppose you have to pay your utilities bill of $189.75. This is your bank statement. Do you have enough money in your account to make this payment?

If the examinee says No, move on to item 20B.
If the examinee says Yes, say – You do not have enough money in your account to make this payment. Now, move on to item 20B.

**Scoring:**
• Global:
  • 1 point if the examinee’s answer is No.
• There are no process scores for this test.

20B. Instructions: *What will happen if you make this payment from your account anyway?*
After the examinee’s response, state “*anything else?*” and record any additional responses.

Scoring:
• Global:
  1. 1 point for any of the following for a total of three points:
     1. Any response suggesting that the utilities bill still has to be paid
     2. Any response indicating overdrawn/overdraft account or that money will be deducted from reserve/savings
     3. Any response indicated that a fee/ fine will be charged or that a penalty will be applied
     4. Any response that indicates that the check will bounce or that the check will be returned or that the check will not clear
     5. Any response suggesting that the examinee’s credit rating will be negatively impacted
     6. Any response suggesting that the utilities/lights will be shut off
     7. A response suggesting that the account balance is negative or that there is not enough money in the account or that $15 - $17 is owed in conjunction with responses #2 or #4 does not obtain an additional point

• The following responses receive 0 points:
  1. Being arrested for a bad check
  2. The account will be closed
APPENDIX D

MONTREAL COGNITIVE ASSESSMENT (MoCA)

MONTREAL COGNITIVE ASSESSMENT (MOCA)
Version 7.1 Original Version

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

VISUOSPATIAL / EXECUTIVE

Copy cube

Draw CLOCK (Ten past eleven)
(3 points)

POINTS

MEMO RY

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

1st trial

2nd trial

ATTENTION

Read list of digits (1 digit/sec.).
Subject has to repeat them in the forward order
Subject has to repeat them in the backward order

FACE VELVET CHURCH DAISY RED

No points

POINTS

LANGUAGE

Fluency / Name maximum number of words in one minute that begin with the letter F

Serial 7 subtraction starting at 100

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

POINTS

ABSTRACTION

Similarity between e.g. banana - orange = fruit

POINTS

DELAYED RECALL

Has to recall words WITH NO QUE

FACE VELVET CHURCH DAISY RED

Points for UNCUED recall only

Optional

Category cue

Multiple choice cue

POINTS

ORIENTATION

Date Month Year Day Place City

POINTS

TOTAL

© Z.Nasreddine MD

www.mocatest.org

Administered by:______________________________

Normal 26 / 30

Add 1 point if 5+ 1yr old
## APPENDIX E

### GERIATRIC DEPRESSION SCALE (SHORT FORM)

**Geriatric Depression Scale (Short Form)**

*Self-Rated Version*

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
<th>Answer</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Are you basically satisfied with your life?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Have you dropped many of your activities and interests?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Do you feel that your life is empty?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Do you often get bored?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Are you in good spirits most of the time?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Are you afraid that something bad is going to happen to you?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Do you feel happy most of the time?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Do you often feel helpless?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Do you prefer to stay at home, rather than going out and doing new things?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Do you feel you have more problems with memory than most people?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Do you think it is wonderful to be alive?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Do you feel pretty worthless the way you are now?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Do you feel full of energy?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>Do you feel that your situation is hopeless?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>Do you think that most people are better off than you are?</td>
<td>YES / NO</td>
<td></td>
</tr>
</tbody>
</table>

TOTAL

*(Sheikh & Yesavage, 1986)*
APPENDIX F

GERIATRIC ANXIETY SCALE – 10

Geriatric Anxiety Scale – 10 Item Version (GAS-10)
© Daniel L. Segal, Ph.D., 2015

Below is a list of common symptoms of anxiety or stress. Please read each item in the list carefully. Indicate how often you have experienced each symptom during the PAST WEEK, INCLUDING TODAY by checking under the corresponding answer.

<table>
<thead>
<tr>
<th>Item</th>
<th>Not at all (0)</th>
<th>Sometimes (1)</th>
<th>Most of the time (2)</th>
<th>All of the time (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I was irritable.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I felt detached or isolated from others.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. I felt like I was in a daze.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I had a hard time sitting still.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. I could not control my worry.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I felt restless, keyed up, or on edge.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. I felt tired.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. My muscles were tense.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. I felt like I had no control over my life.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. I felt like something terrible was going to happen to me.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G

DEMOGRAPHICS QUESTIONNAIRE

Participant ID: ____________________ Examiner: ____________________
Date: ______________ For experimenter use only Condition: __________/________

Date of birth: ____________________ Age: ___________

Years of education (circle one)
Less than 9th grade
Finished 9th grade (9)
Finished 10th grade (10)
Finished 11th grade (11)
High School Diploma (12)
One year of college (13)
Two years of college or Associate’s degree (14) Three years of college (15)
Bachelor’s degree or equivalent (16)
Post-graduate work (17) Master’s degree or equivalent (18)
All but dissertation (19) Doctoral degree or equivalent (20)

Sex (circle one)
Female Male

Ethnicity (circle one)
Hispanic, Latino, or Spanish origin Not Hispanic, Latino, or Spanish origin Unknown

Race (circle all that apply)
American Indian or Alaska Native
Asian
Black or African American
Native Hawaiian or Other Pacific Islander
White or Caucasian
Other /Unknown

What is your current living situation? (circle one)
Independent housing
Assisted Living Facility/caregiver (professional or family) in home
Nursing home

Primary Occupation

Is English your first language? (circle one)
Yes
No, second language (First language: ____________________)
No, third or other language (First language: ____________________)
APPENDIX H

IRB APPROVAL

University of Colorado
Colorado Springs
Institutional Review Board (IRB) for the Protection of Human Subjects

Date: 7/26/2016

IRB Review

Approved

IRB PROTOCOL NO.: 17-410
Protocol Title: The Colorado Cognitive Assessment (CoCA): Development of a screening tool that provides qualitative information and is scored using modern psychometrics
Principal Investigator: Ahsit Parmiani
Faculty Advisor if Applicable: Brandon Gavett
Application: New Application
Type of Review: Expedited 7
Risk Level: No more than Minimal Risk
Renewal Review Level (If changed from original approval) if Applicable: N/A No Change
This Protocol involves a Vulnerable Population: N/A (No Vulnerable Population)
Expires: 25 July 2017

[Text continues with additional details]

OSP #: [Blank]

Sponsor: [Blank]

Thank you for submitting your Request for IRB Review. The protocol identified above has been reviewed according to the policies of this institution and the provisions of applicable federal regulations. The review category is noted above, along with the expiration date, if applicable.

Once human participant research has been approved, it is the Principal Investigator’s (PI) responsibility to report any changes in research activity related to the project:
- The PI must provide the IRB with all protocol and consent form amendments and revisions.
- The IRB must approve these changes prior to implementation.
- All amendments requiring study subjects must be reviewed prior to approval by the IRB.
- The PI must promptly inform the IRB of all unanticipated serious adverse events (within 24 hours). All unanticipated adverse events must be reported to the IRB within 1 week (see 45CFR46.103a(b)(2)). Failure to comply with these federally mandated responsibilities may result in suspension or termination of the project.
- Notify study subjects of the IRB prior to expiration.
- Notify the IRB when the study is complete.

If you have any questions, please contact Research Compliance Specialist in the Office of Sponsored Programs at 719-255-3903 or orc@uccs.edu.

Thank you for your concern about human subject protection issues, and good luck with your research.

Sincerely yours,

Michele Olson, PhD
IRB Reviewer