The Montreal Cognitive Assessment (MoCA) as a cognitive screen in addiction health care: A validation study for clinical practice

Carolien J. W. H. Bruijnen, Mandy Jansen, Boukje A. G. Dijkstra, Serge J. W. Walvoort, Selma Lugtmeijer, Wiebren Markus, Cor A. J. De Jong, and Roy P. C. Kessels

Abstract

Objective: The current study assessed the criterion validity of the Montreal Cognitive Assessment (MoCA) as a short cognitive screen for use in addiction health care. Method: Eighty-two patients were assessed with two parallel versions of the MoCA; at intake (baseline) and directly preceding an extensive neuropsychological assessment (NPA) approximately 8 weeks later (follow-up). Results: Of all included patients, 54.9% were classified as having substance-induced neurocognitive disorder. The most common primary substance of abuse was alcohol (70.7%). The criterion validity was determined predictively and concurrently, and sensitivities of .56 and .67 and specificities of .62 and .73 were found, respectively. Conclusion: While the MoCA is an adequate screen when administered at the same time as the NPA, the predictive validity of administering this cognitive screen at intake is limited. Furthermore, the relation between MoCA domain scores and the performance on their corresponding cognitive domain in the NPA is more reliable when the MoCA is administered at the same time as the NPA. While the MoCA can be used to screen for cognitive impairments in patients in addiction health care, the instrument’s sensitivity is not optimal, which should be taken into account when interpreting results.

Introduction

About 0.6% of the adult population worldwide (an estimated 29.5 million) suffer from substance use disorder (SUD; United Nations Office on Drugs and Crime, 2017). SUD affects the individual in social, physical and economical ways (Laudet, Savage, & Mahmoud, 2002) and may result in cognitive impairments interfering with treatment (Aharonovich et al., 2006; Bates, Pawlak, Tonigan, & Buckman, 2006; Copersino et al., 2012). The Diagnostic and Statistical Manual of Mental Disorders (5th ed.; DSM-5; American Psychiatric Association, 2013) introduced the term “neurocognitive disorder” (NCD) in which the subtype substance-induced NCD can be classified as either major or mild, based on severity and everyday limitations. Cognitive impairments in patients with SUD have an estimated prevalence of 30–80% (Copersino et al., 2009). The exact prevalence of substance-induced NCD is, however, difficult to establish based on the existing literature (Toledo-Fernández et al., 2017).

The effects of chronic substance use on cognitive functioning are both acute and chronic and vary across substances, resulting in decreased treatment adherence, lower self-efficacy and less treatment retention (Aharonovich et al., 2006; Bates et al., 2006; Copersino et al., 2012). Therefore, insight into an individual’s cognitive functioning is crucial, as it enables to personalize and optimize treatment effectiveness (Allen, Goldstein, & Seaton, 1997; Bates, Buckman, & Nguyen, 2013; Sofuoglu, Sugarman, & Carroll, 2010). Often patients with SUD lack insight into their NCD, as indicated by a lack of correlation between objectively measured and subjectively experienced cognitive deficits (Horner, Harvey, & Denier, 1999; Walvoort, van der Heijden, Kessels, & Egger, 2016). Although neurocognitive assessment can accurately detect the pattern and severity of cognitive impairment in patients with SUD, the administration of such an extensive neuropsychological assessment (NPA) is not always feasible. Therefore, this study investigated the validity of a short and easy-to-administer cognitive screen, the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), in patients with SUD.

Originally developed to detect mild cognitive impairment (MCI; Nasreddine et al., 2005), the MoCA has also been found to be valid in SUD for detecting cognitive deficits (Copersino et al., 2009; Ridley et al., 2018; Rojo-Mota, Pedrero-Pérez, Ruiz-Sánchez de León, Llanero-Luque, & Puerta-García, 2013). However, only one study in patients with alcohol use disorders (AUD) has correlated the MoCA...
to the gold standard NPA performance (Ewert et al., 2018), with no studies in users of other substances.

The current prospective study assessed the criterion validity of the MoCA as a screen for cognitive impairment in a sample of patients with SUD. First, the optimal cutoff for use in addiction health care was established. Next, the criterion validity was assessed by comparing MoCA results with an extensive NPA. Additionally, the interaction of substance type, abstinence duration and MoCA performance has been examined. In order to maximize the external validity of the design, this study was designed to apply as much as possible with treatment as usual in all participating institutions.

Method
Design
A prospective study was performed with two time points of assessment using two authorized and validated parallel versions of the MoCA (Costa et al., 2012). MoCA version 7.1 was administered at intake (baseline) and MoCA version 7.2 directly preceding the NPA at follow-up approximately 8 weeks later. Data were collected between August 2012 and March 2015. The study was approved by the internal review boards of all participating health-care centers and the research board of the Nijmegen Institute for Scientist-Practitioners in Addiction.

Participants
The aim was to recruit a total of 100 outpatients seeking treatment for SUD from four participating addiction health-care centers in the Netherlands (IrisZorg, Novadic-Kentron, Tactus and Vincent van Gogh for Mental Health). This study is part of a larger study (N = 691), for which the inclusion criteria were (1) dependency or abuse of a substance (excluding nicotine) or behavior; (2) age 18–75; and (3) signed informed consent for participation at baseline and/or follow-up. The exclusion criterion was an inability to administer the MoCA, due to for instance: a neurological (e.g., stroke, dementia, traumatic brain injury) or very unstable acute psychiatric disorder, severe lack of motivation, or insufficient Dutch language skills. Patients were included regardless of abstinence to increase the external validity of the design.

Materials
Montreal Cognitive Assessment
The MoCA (Nasreddine et al., 2005) consists of twelve elements measuring seven cognitive domains. These domains include executive functioning; visuospatial abilities; attention, concentration and working memory (called “attention” from now on); language; abstract reasoning; memory; and orientation. A total score is calculated, with a maximum of 30 points. Nasreddine et al. (2005) found a score of ≤25 to be indicative of MCI. However, several studies have identified different cutoff scores for different populations (Julayanont, Phillips, Chertkow, & Nasreddine, 2013). The validity of the MoCA, including both alternate forms, has been established in detecting MCI, with sensitivities and specificities ranging from .90 to 1.00 and .57 to .62, respectively. Alternate-form reliability for healthy controls ranged from .52 to .69 and all versions were found to be equivalent in previous research (Costa et al., 2012; Nasreddine & Patel, 2016). The authorized Dutch translations of two parallel versions were used (see www.mocatest.org for materials and instructions).

Neuropsychological assessment
The tests included in the NPA, that was administered at follow-up, were selected based on the cognitive domains targeted by the MoCA, thus assessing a broad range of cognitive functions. The allocation of the NPA (sub)tests and MoCA elements to each cognitive domain was based on DSM-5 criteria for NCD (pp. 593–595; American Psychiatric Association, 2013) and is summarized in Appendix 1.

Measurements in the addictions for triage and evaluation (MATE 2.1)
The MATE 2.1 (Schippers, Broekman, & Buchholz, 2011) is part of the intake procedure and consists of an interview and self-report questionnaires for collecting relevant patient characteristics. In this study, Section 1 ‘Substance use’ and Section 3 ‘History of treatment for substance use disorders’ were used.

Procedure
At baseline, the MATE 2.1 was administered and written informed consent was obtained. MoCA version 7.1 was administered by trained professionals during intake procedure. For administration of the NPA (follow-up) an appointment approximately 6–8 weeks later was made. The NPA procedure was fixed for all institutions and administration was done by trained psychologists. All professionals were trained in MoCA and/or NPA administration in accordance with the test manuals, by the coordinator of this study.

Recruitment for follow-up was based on random selection (i.e., one in eight patients of the large study were randomly selected for a follow-up), indication (i.e., based on care as usual), or both. Three patients with a behavioral disorder without substance use were excluded for this study. At follow-up, MoCA version 7.2 was administered preceding the NPA. Prior to both baseline and follow-up, a self-reported estimation of substance use in the week before administration, or abstinence duration (if >7 days) was obtained.

Analyses
Patient characteristics
For determining the presence of NCD, criteria for substance-induced NCD of the DSM-5 (American Psychiatric Association, 2013) were combined with the “cognitive impairment, no dementia” criteria as outlined in van den Berg, Kessels, de Haan, Kappelle, and Biessels (2005). All raw scores of the NPA were transformed into standard z-scores, according to the normative data. These standard z-scores were classified as: 0 = average (≥–1.00); −1 = below average (between −1.00 and −1.65); −2 = impaired (≤–1.65). An average score for a domain of ≤–1.00 was considered to
be impaired. If at least two out of all seven NPA domains were impaired, the patient was classified as having substance-induced NCD. Nine patients had missing data in one or more NPA domain, but the remaining results were sufficient to validly classify each patient. Mean NPA domain scores and patient characteristics were compared between patients with and without NCD by using independent t-tests, chi-square tests or Mann–Whitney U tests (non-normal variables).

**Criterion validity**

Level of education was classified on a seven-point scale ranging from 1 = less than primary school to 7 = university degree or higher (Duits & Kessels, 2014), a classification system comparable to the International Standard Classification of Education (United Nations Educational Scientific and Cultural Organization, 2012). As it was found that years of education affects performance on the MoCA (Chertkow, Nasreddine, Johns, Phillips, & McHenry, 2011; Nasreddine et al., 2005), Spearman’s rho correlations were used to relate the unadjusted MoCA total scores at baseline and follow-up to this level of education. Based on the studies by Chertkow et al. (2011) and van der Elst, van Boxtel, van Breukelen, and Jolles (2005), the MoCA total score was then adjusted for education (low level of education, classifications 1–3: two additional points; average level of education, classifications 4 and 5: one additional point; and highly educated patients, classifications 6 and 7: no additional points), with the maximum MoCA total score remaining 30 in all cases. MoCA results were then explored and differences between patients with and without NCD were computed using independent t-tests. Furthermore, MoCA domain scores were correlated with mean z-scores on the corresponding NPA domain, and systematic differences between MoCA domain and total scores at baseline and at follow-up were assessed with paired t-tests.

The predictive validity was assessed by computing a receiver operating characteristic curve with the corresponding area under the curve (AUC) for the MoCA total score at baseline, with NPA classification (NCD or no-NCD) at follow-up as a criterion. The cutoff point was determined by the optimal sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). After applying the cutoff to the data, overall agreement and chance-adjusted agreement (Cohen’s kappa) were determined. The concurrent validity was assessed in the same way as the predictive validity by using MoCA and NPA results at follow-up.

**Substance type and abstinence duration**

The influence of substance type and abstinence duration on MoCA performance at follow-up was estimated using logistic regression with abstinence duration, substance type (alcohol versus other drugs), MoCA total score, and interactions between each as predictors, with NCD classification (NCD versus no-NCD) as the dependent variable. The Outlier Labelling Rule (Hoaglin & Iglewicz, 1987) was used to exclude outliers, leading to the exclusion of one outlier for abstinence duration. All data were computed and analyzed with IBM SPSS version 24.0.

**Results**

**Patient characteristics**

A total of 82 patients were included in this study, 54.9% of whom were classified as having NCD based on the NPA results. Mean NPA-domain scores differed significantly between patients with and without NCD for all cognitive domains except orientation (Table 1).

The overall mean age was 44.1 (SD = 13.77) and 68.3% were men. The most prevalent primary problem substance of abuse was alcohol (70.7%). More patients were abstinent at follow-up than at baseline (an increase of 28.0%). The majority (42.7%) was not abstinent at either time point. Except for age and marital status, both patient groups were comparable (Table 2).

**Criterion validity**

**Montreal Cognitive Assessment**

Level of education correlated with the unadjusted MoCA total scores both at baseline (ρ = .421, p < .001) and at follow-up (ρ = .550, p < .001).

Patients with NCD performed significantly worse on the MoCA than patients without NCD, both at baseline and at follow-up. The same was true for the MoCA domain executive functioning. At baseline, the difference in performance on the MoCA domain memory was also significant, while performance on the MoCA domains language and abstract reasoning differed significantly between both patient groups at follow-up. Performance on the other MoCA domains did not differ significantly between both patient groups (Table 3).

At baseline, correlations between performance on the MoCA and NPA were significant for the domains executive functioning (r = .238, p = .032), abstract reasoning (r = .300, p = .006), and memory (r = .423, p < .001). At follow-up, there was an almost perfect correspondence between MoCA and NPA performance: all MoCA domain scores were significantly correlated to the corresponding NPA domain (executive functioning: r = .328, p = .003; visuospatial abilities: r = .241, p = .029; attention: r = .396, p < .001; abstract reasoning: r = .542, p < .001; memory: r = .455, p < .001; orientation: r = .229, p = .043).

**Predictive validity at baseline**

An AUC value of .676 was found (ρ = .006; Figure 1) and a cutoff score of 24 yielded the most optimal sensitivity (.56), specificity (.62), PPV (64.1%), and NPV (53.5%), using the NPA as gold standard. Applying this cutoff score, 39 out of

---

**Table 1.** Mean (SD) performance in z-scores for each domain of the neuropsychological assessment (NPA) for patients with and without neurocognitive disorders (NCD) and in the total sample.

<table>
<thead>
<tr>
<th>NPA domain</th>
<th>Total (n = 82)</th>
<th>NCD (n = 45)</th>
<th>No-NCD (n = 37)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive functioning</td>
<td>–0.10 (0.89)</td>
<td>–0.32 (1.02)</td>
<td>0.16 (0.62)</td>
<td>.011**</td>
</tr>
<tr>
<td>Visuospatial abilities</td>
<td>–0.83 (0.99)</td>
<td>–1.21 (0.88)</td>
<td>–0.37 (0.92)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Attention</td>
<td>–0.78 (0.90)</td>
<td>–1.16 (0.71)</td>
<td>–0.32 (0.89)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Abstract reasoning</td>
<td>–0.71 (0.83)</td>
<td>–1.00 (0.85)</td>
<td>–0.35 (0.64)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Memory</td>
<td>–0.67 (1.01)</td>
<td>–1.00 (1.02)</td>
<td>–0.27 (0.86)</td>
<td>.001**</td>
</tr>
<tr>
<td>Orientation</td>
<td>4.20 (1.11)</td>
<td>4.05 (1.32)</td>
<td>4.36 (0.80)</td>
<td>.211</td>
</tr>
<tr>
<td>Processing speed</td>
<td>–0.61 (0.89)</td>
<td>–1.03 (0.80)</td>
<td>–0.10 (0.71)</td>
<td>&lt;.001***</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01; ***p < .001.
82 patients were classified as having NCD, while 45 out of 82 patients were classified as having NCD based on the NPA. The overall agreement with the NPA was 58.5% and the chance-adjusted agreement was 17.5% (Table 4).

**Concurrent validity at follow-up**
An AUC of .745 was found ($p < .001$; Figure 1) and a cutoff score of 25 yielded the most optimal sensitivity (.67), specificity (.73), PPV (75.0%), and NPV (64.3%), using the NPA as
Applying this cutoff score, 40 out of 82 patients were classified as having NCD. The overall agreement with the NPA was 69.5% and the chance-adjusted agreement was 39.2% (Table 4).

Systematic differences between MoCA versions 7.1 and 7.2
Paired t-tests between MoCA results for version 7.1 (baseline) and 7.2 (follow-up) showed that only the domains abstract reasoning, memory, and orientation did not differ significantly. All other domains and the MoCA total score differed significantly between both versions. For all domains except language, mean scores on MoCA version 7.2 were higher (Table 3).

Substance type and abstinence duration
The logistic regression model was statistically significant ($\chi^2 (7) = 16.58, p = .020$), correctly classifying 69.0% of cases. However, neither the predictors nor any interaction between the predictors in the model were statistically significant.

Discussion
This study is the first to examine the MoCA as a short cognitive screen in a sample of patients with SUD, using an extensive NPA as benchmark. The results show that administration of the MoCA at baseline resulted in a worse validity than the MoCA administered at follow-up. Also, while at follow-up all MoCA cognitive domains correlated with the corresponding domain of the NPA, at baseline only the MoCA domains executive functioning, abstract reasoning, and memory significantly predicted NPA performance 8 weeks later.

These findings are partly in line with other MoCA studies (Alarcon, Nalpas, Pelletier, & Perney, 2015; Oudman et al., 2014; Wester, Westhoff, Kessels, & Egger, 2013) where only Alarcon et al. (2015) reported a higher predictive validity. There are, however, important differences between these studies and the current. First, only homogeneous groups of patients with AUD were included in previous studies, limiting their external validity. Second, patients in those studies were abstinent for at least...
1 week (Alarcon et al., 2015) to a minimum of 6 months (Oudman et al., 2014), while in clinical practice patients are often not abstinent at intake. To date, only one study in AUD related MoCA performance directly to an NPA (Ewert et al., 2018), which is considered to be the gold standard for the assessment of cognitive impairments (Lezak, Howieson, Bigler, & Tranel, 2012). Ewert et al. (2018) found a higher education-adjusted cutoff score than was currently found to be indicative of cognitive impairment, using a homogeneous group of hospitalized patients with AUD. The only study in a heterogeneous group of patients with SUD (Copersino et al., 2009) was more in line with the present findings – albeit that slightly better psychometric properties were found.

Regarding the relation between MoCA domain scores at baseline and NPA domain performance at follow-up, caution should be taken when interpreting the MoCA domain scores. This is in line with a previous study also showing MoCA domain scores to be poor predictors of impairment on neuropsychological tests (Moafmashhadi & Koski, 2013). The difference in findings between the predictive and concurrent validity can be explained by the interval between baseline and follow-up. Abstinence could also be an explanation as cognitive recovery is likely to occur with sustained abstinence in AUD (Stavro, Pelletier, & Potvin, 2013; van Holst & Schilt, 2011), cannabis (Lyons et al., 2004), and stimulants (ludicello et al., 2010; Vonmoos et al., 2014; Wood, Sage, Shuman, & Anagnostaras, 2014; Zhong et al., 2016). Although more patients were abstinent at follow-up, we did not find a significant effect of abstinence on MoCA performance in our statistical model.

There are several strengths to the current study. First, the heterogeneity of the sample largely represents clinical practice, which makes the results generalizable to addiction health care. Second, the used adjustment method for level of education (based on Chertkow et al., 2011) is more fine-grained than the original adjustment method by Nasreddine et al. (2005). Third, the extensive gold standard NPA, using widely used, valid and reliable tests, made analysis of specific domains and comparisons between patients with or without NCD possible, which has not been done before in a heterogeneous group of patients with SUD. Fourth, parallel MoCA versions were administered at two time points, which made it possible to assess validity predictively and concurrently. Finally, the effects of substance type and abstinence duration on MoCA performance were taken into account.

Although a moderate concurrent validity of the MoCA as compared to the NPA was found, it should be stressed that using a MoCA cutoff score of 25 results in only 66.7% of patients with NCD being classified correctly. Therefore, we underscore the fact that the MoCA as a screen can never substitute an extensive NPA. Therefore, a subsequent extensive NPA is recommended, especially in patients who perform above or at the cutoff point, given the low sensitivity.

**Acknowledgments**

Many thanks go to Dr Arie Wester who had a special role in this project and Dr Lydia Jonker for her contributions. Unfortunately, both died unexpectedly before they could see the finish line. Also, we would like to thank all institutions and professionals that participated in the data collection.

**Funding**

This research was funded by the Nijmegen Institute for Scientist-Practitioners in Addiction (NISPA), Nijmegen, The Netherlands, and Vincent van Gogh for Mental Health, Venray, The Netherlands.

**References**


# Appendix 1

Cognitive domains with the corresponding elements of the Montreal Cognitive Assessment (MoCA) and (sub)tests of the neuropsychological assessment (NPA)

<table>
<thead>
<tr>
<th>Cognitive domain</th>
<th>MoCA element (score range)</th>
<th>NPA element (used score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive functioning</td>
<td>Alternating trail making (0–1) Verbal fluency (0–1)</td>
<td>D-KEFS TMT Letter-number switching (scaled score) Number sequencing (scaled contrast score*) Letter sequencing (scaled contrast score*) Stroop CWT Interference score (t-score)</td>
</tr>
<tr>
<td>Visuospatial abilities</td>
<td>Figure copy (0–1) Clock drawing (0–3)</td>
<td>WAIS-IV-NL Block design (scaled score) RCFT Copy (t-score)</td>
</tr>
<tr>
<td>Attention</td>
<td>Digit span (0–2) Sustained attention (0–1) Serial subtraction (0–3)</td>
<td>WAIS-IV-NL Digit span forward (scaled score) Digit span backward (scaled score)</td>
</tr>
<tr>
<td>Language</td>
<td>Naming (0–3) Sentence repetition (0–2)</td>
<td>WAIS-IV-NL Similarities (scaled score) DART (deviation IQ)</td>
</tr>
<tr>
<td>Abstract reasoning</td>
<td>Abstraction (0–2)</td>
<td>WAIS-IV-NL</td>
</tr>
<tr>
<td>Memory</td>
<td>Delayed recall (0–5)</td>
<td>RAVLT Total correct (t-score) Delayed recall (t-score) RCFT Immediate reproduction (t-score)</td>
</tr>
<tr>
<td>Orientation</td>
<td>Orientation (0–6)</td>
<td>CST-14 (raw score) Stroop CWT Word reading (t-score) Color naming (t-score) D-KEFS TMT Motor speed (scaled score)</td>
</tr>
</tbody>
</table>

D-KEFS TMT = Delis-Kaplan Executive Function System Trail Making Test (Delis, Kaplan, & Kramer, 2007); Stroop CWT = Stroop Color Word Test (Hammes, 1971); WAIS-IV-NL = Wechsler Adult Intelligence Scale–IV–Dutch (Wechsler, 2012); RAVLT = Rey-Osterrieth Complex Figure Test (Meyers & Meyers, 1995); DART = Dutch Adult Reading Test (Schmand, Lindeboom, & van Harskamp, 1992); CST-14 = Cognitive Screening Test–14 (de Graaf & Deelman, 1991); * = Scaled score obtained by contrasting performance on the number sequencing and letter sequencing conditions against performance on the number-letter switching condition.