Validation of Smartphone-Based Cognitive Assessments for
Individuals with Major Depressive Disorder

by

Therond, Alexandra

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Master of Arts

in

Psychology

Carleton University
Ottawa, Canada

© 2021
Therond, Alexandra

Abstract

Cognitive deficits are often present in major depressive disorder (MDD) and negatively impact functional outcomes. However, it remains challenging to assess these impairments in clinical and research settings. Smartphone applications provide the opportunity to measure cognitive impairments in an accessible way. In this study, 24 individuals with MDD and 34 healthy controls (HC) completed the Trail Making Tests (TMT), and the smartphone-based versions, named the Jewels Trail Tests (JTT). Significant positive relationships between the JTT and TMT were observed with a moderate concurrent validity for Parts A and strong concurrent validity for Parts B. The intraclass correlations showed moderate test-retest reliability for Part A of the JTT and good reliability for Part B. This study did not find significant differences between the MDD and HC groups completion time. Lastly, higher sleep quality was associated with a faster completion time on the speed processing task over a period of three months.

Keywords: depression, cognitive deficits, smartphone application, cognitive assessments, speed of processing, executive function

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. Synthia Guimond. Her enthusiasm in research and endless support in my studies has made this final year of my degree a great success. Without her guidance and persistent help, this thesis would not have been possible. Thank you for always being available, with an open door (or video call) when I needed last minute support and encouragement. I would also like to thank my secondary supervisor, Dr. Andrea Howard, who supported me throughout the process and provided insightful advice that aided me in the completion of my thesis. I would like to acknowledge all the participants who took part in this study. It was a great enjoyment to experience their excitement in this research. In addition, I would like to thank Feng Gu and Cecelia Shvetz for their help with the project, and Dr. John Torous of The Beth Israel Deaconess Medical Center at Harvard Medical School. A thank you to Rayya Kaidbey for always providing unconditional support during those years. Financial support was provided by the Natalia Mental Health Foundation.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vi
List of Figures.....	vii
List of Appendices.....	viii
Introduction.....	1
<i>Cognitive Deficits in Major Depressive Disorder.....</i>	4
<i>Cognitive Dysfunction Related to Functional Decline in Major Depressive</i>	
<i>Disorder .....</i>	6
<i>Sleep Disturbances and Cognitive Functioning.....</i>	8
<i>Limitations of Cognitive Assessments.....</i>	11
<i>Advantages of Using a Smartphone Application to Assess Cognition and</i>	
<i>Behavior.....</i>	13
<i>MindLAMP Application.....</i>	17
The Present Study.....	19
Method.....	20
<i>Participants.....</i>	20
<i>Procedure.....</i>	21
<i>Measures.....</i>	24
<i>Statistical Analysis.....</i>	30
Results.....	36

<i>Outlier Detection Analysis</i>	36
<i>Demographic and Clinical Data</i>	37
<i>Concurrent Validity</i>	38
<i>Test-Retest Reliability</i>	40
<i>Between-Group Differences in JTT performance</i>	40
<i>Self-Reported Sleep Quality</i>	47
<i>Practice Effects</i>	50
Discussion.....	52
<i>The Validity of a Smartphone Application to Assess Cognition</i>	54
<i>Comparison of Patients and Healthy Control</i>	56
<i>Towards Remote Cognitive Assessments</i>	59
<i>Monitoring Cognitive Functioning and Symptomatology</i>	
<i>Longitudinally</i>	61
<i>Strengths and Limitations</i>	63
<i>Implications</i>	64
<i>Conclusion</i>	66
References.....	69

List of Tables

Table 1: <i>Baseline Medication Information for MDD group.....</i>	21
Table 2: <i>Final Sample Size of Each Aim and Statistical Analysis After Removing Missing Values and Erroneous Datapoints.....</i>	37
Table 3: <i>Baseline Demographic and Clinical Data for Patients and Controls.....</i>	38
Table 4: <i>Means and Standard Deviations of the MDD and HC groups for TMT Completion Time in the Laboratory Setting.....</i>	41
Table 5: <i>Estimates of the Linear Regression Model of MDD Indicator on the TMT Completion Time in the Laboratory Setting.....</i>	42
Table 6: <i>Means and Standard Deviations of the MDD and HC Groups for JTT Completion Time in the Laboratory Setting.....</i>	43
Table 7: <i>Estimates of the Linear Regression Model of MDD Indicator on the JTT Completion Time in the Laboratory Setting.....</i>	44
Table 8: <i>Means and Standard Deviations of the MDD and HC Groups for JTT Completion Time in a Remote Setting.....</i>	45
Table 9: <i>Fixed Effect Estimates of the Mixed Effects Model of MDD Indicator on the JTT Completion Time in a Remote Setting.....</i>	46
Table 10: <i>Fixed Effect Estimates of the Mixed Effects Model of Sleep Scale on the JTT Completion Time.....</i>	48
Table 11: <i>Fixed Effect Estimates of the Mixed Effects Model of the Practice Effects on the JTT Completion Time.....</i>	50

List of Figures

Figure 1: <i>Smartphone-Based Cognitive Assessments on MindLAMP A) JTT A: Participant taps on numbers in chronological order; B) JTT B: Participant taps on numbers in chronological order while alternating between the two shapes.....</i>	29
Figure 2: <i>Relationship Between Completion Time of Participants on the Pen-and-Paper TMT and the Smartphone Application JTT.....</i>	39
Figure 3: <i>Longitudinal Mixed Effect Model Prediction of the Impact of Self-Reported Sleep Quality Measured with the Sleep Scale on Cognition Over Time, Measured as the JTT Completion Time. A) Part A of the test; B) Part B of the test.....</i>	49
Figure 4: <i>Longitudinal Mixed Effect Model Prediction of the Impact of Practice Effects, Measured as the Number of Previous Trials, on the JTT Completion Time. A) Part A of the test; B) Part B of the test.....</i>	51

List of Appendices

Appendix A: <i>Informed Consent Form</i>	104
Appendix B: <i>Publications and Conference Presentations</i>	112

Introduction

Major depressive disorder (MDD) is characterized by a decreased interest or pleasure in daily activities, depressed mood, weight changes, sleep and psychomotor disturbances, fatigue, feelings of worthlessness, difficulties concentrating, and suicidal ideation (American Psychiatric Association, 2013). MDD represents a major public health concern, with global estimates indicating that about 10.8% of individuals are affected by this condition (Lim et al., 2018). MDD is a leading cause of disability worldwide and accounts for 40.5% of the global diseases caused by psychiatric disorders, in terms of years of disability and years of life lost due to premature mortality (Whiteford et al., 2013; World Health Organization, 1992). Costs of MDD include mortality predominantly from suicide, diminished functioning, increased clinical morbidity, and significant reduction of quality of life (Klerman, 1992).

MDD is also related to decreased work productivity and poor social functioning (Rizvi et al., 2015; Stewart et al., 2003). Wilhelm and colleagues (2003) analyzed data from an Australian National Survey and observed that unemployment was the strongest correlate of MDD. Moreover, Dooley and colleagues (2000) found that MDD symptoms amongst youth with employment in their early thirties predicted a loss of employment two years later. The economic burden of MDD in 2000 in the U.S.A was estimated at \$83.1 billion (Greenberg et al., 2015). Moreover, dysfunctional social interactions are correlated with decreased work performance, unemployment, disability (Rizvi et al., 2015) and remain persistent even after three years of remission (Rhebergen et al., 2010). Psychosocial impairments related with MDD were higher in the social role domain, leading to reduced marital and interpersonal functioning, than

in the work role domain (Fink & Shapiro, 2013; Hirschfeld et al., 2002; Kessler et al., 2003).

Amongst the symptoms that MDD patients most frequently reported, cognitive deficits were predominant (Albert et al., 2016). These dysfunctions are meaningful predictors of relapse or recurrence of the illness despite pharmacotherapy and psychological interventions (Majer et al., 2004). Cognitive symptoms impact several domains, namely executive functioning, processing speed, attention, and memory (Rock et al., 2014). Cognitive disorders tend to persist even during remission of clinical symptoms. Indeed, they are present 80-90% of the time during episodic phases and 40-45% of the time during remission phases (Conradi et al., 2011). They have also emerged as predictors of overall functioning (Jaeger et al., 2006). Processing speed and executive functioning have been found to specifically impact daily functioning in MDD (Nakano et al., 2008). Indeed, poor cognitive health such as deficits in processing speed and executive functioning may be partly responsible for patients' failure to return to work (Baune et al., 2010; Woo et al., 2016). Several factors may exacerbate cognitive dysfunctions in patients with MDD. Accumulating evidence has supported that sleep is fundamental to healthy cognitive functioning and lack of sleep may impact several cognitive domains (Alhola & Polo-Kantola, 2007). Nevertheless, limited research has explored the relationship between MDD, sleep and cognitive functioning (Wu et al., 2019).

Cognitive symptoms should be regarded as critical targets of assessment. However, no clinical guidelines have been provided to clinicians and healthcare providers in terms of cognitive evaluations for MDD (Trivedi & Greer, 2014). The

current neuropsychological batteries present numerous challenges such as involving long hours of assessments and entailing extensive training for the clinical assessors (Dagum, 2018; Howieson, 2019; Torous & Powell, 2015; Torous et al., 2019). Moreover, subjective reports of cognitive deficits in MDD are not correlated with objective reports of cognitive health (Naismith et al., 2007).

Utilizing smartphone-based assessments may be a feasible and cost-effective option to monitor the cognitive functioning and behavioral patterns in patients with MDD (Torous et al., 2019). These transportable tests can assess cognition in numerous environments. The mobile assessments can allow continuous tracking of cognitive functioning, reduce retrospective recall with real-time data collection and assess daily fluctuations (Torous et al., 2014). The goal of the technology is to augment the ability of healthcare providers to appropriately assess cognition in individuals with depressive symptoms (Torous et al., 2014). The MindLAMP smartphone application is a free application that has been developed by our team at the Beth Israel Deaconess Medical Center at Harvard Medical School (Torous et al., 2019). The MindLAMP is a digital mental health application used by clinicians and researchers to collect, store, and share data from individuals with mental health disorders. The application was designed to be flexible and customizable to the patient's needs and preferences. It provides the opportunity for researchers and clinicians to utilize an engaging and usable platform for patients with MDD (Torous et al., 2019). Numerous smartphone applications utilized for mental health purposes cannot be used by individuals with serious psychiatric disorders due to the lack of knowledge and confidence with the technology (Sarkar et al., 2016). The MindLAMP application also supported the development of

two cognitive assessments called the Jewels-A (JTT A) and Jewels-B (JTT B) tests, which stem from the validated Trail Making Test A and B (Reitan & Wolfson, 1985; Torous et al., 2019). Like the TMT, the test involves a Part A assessing processing speed and participants are asked to tap the jewel-shaped numbers on their screen in sequential order from 1 to 25 as fast as possible. In Part B, which measures executive functioning, participants are asked to alternate between two different shapes of diamond of the same number while continuing to tap the numbers in sequential order from 1 to 25. Preliminary evidence shows that it is feasible to use these MindLAMP cognitive tests with other mental disorders such as schizophrenia (Liu et al., 2019). However, the tests remain to be validated to assess cognitive functioning in individuals with MDD.

This study examined whether the MindLAMP smartphone-based cognitive assessments can monitor processing speed and executive function in individuals with MDD.

Cognitive Deficits in Major Depressive Disorder

Investigation of cognitive symptoms of MDD was for a long time secondary after mood symptoms. It is only in the past decade that interest has heightened about neuropsychological dysfunctions relating to MDD. Traditionally, the physical and emotional domains have been emphasized in the assessment and treatment of the illness (Richardson & Adams, 2018). However, up to two-thirds of acutely depressed people experience neurocognitive deficits (Krol et al., 2011; Olesen et al., 2012; Rock et al., 2014; Thomas & Morris, 2003). A systematic review and meta-analysis on cognitive impairment during episodic phases of MDD observed significant moderate

deficits in attention, executive function and memory when compared to healthy controls. Furthermore, the meta-analysis observed that patients remitted from MDD also exhibited cognitive deficits compared to a healthy control group (Rock et al., 2014). Neuropsychological dysfunctions have been identified at the onset of the disorder during the first episode and in young adults who developed MDD later in life (Lee et al., 2012). Tsourtos and colleagues (2002) observed evidence of early information processing speed deficit in younger patients with MDD. Indeed, patients showed cognitive slowing compared to healthy control groups. These findings infer that deficits in cognition can be present at the early stages of the disease.

In addition to the well-established presence of more severe cognitive deficits in active phases of MDD, evidence indicates that cognitive impairment may persist after clinical remission. Cognitive symptoms in MDD last longer than the acute episodes (Conradi et al., 2011). Over 30% of patients whose clinical symptoms have resolved, reported residual cognitive problems, including reduced processing speed, reduced attentional mechanisms, apathy, word-finding difficulty, and mental slowness (Harrison et al., 2018). Residual symptoms have also been observed 44% of the time in remitted primary-care patients with MDD (Conradi et al., 2011). Shimizu and colleagues (2013) examined the link between neuropsychological dysfunction and quality of life in individuals with remitted MDD. Patients in remission had poorer cognitive performances compared to age and education matched controls, specifically for attention, verbal memory and psychomotor speed. Cognitive deficits can occur during, before or after a depressive episode and worsen after each consecutive episode (Keefe et al., 2014; Semkovska et al., 2019). Therefore, repeated episodes of MDD

may increase vulnerability for additional cognitive deficits and require constant monitoring (Ragguett et al., 2016).

Cognitive Dysfunction Related to Functional Decline in Major Depressive Disorder

Cognitive symptoms are important contributors to functional decline in MDD (Conradi et al., 2011). A wide range of cognitive domains namely, executive functions, memory, processing speed, attention, and language, are strongly affected and impact productivity performance (Burt et al., 1995; Caligiuri & Ellwanger, 2000; Evans et al., 2014; Ravnkilde et al., 2002; Sobin & Sackeim, 1997). Decreased concentration and deficits in executive functioning are present in over 70% of medicated patients (McClintock et al., 2011). Godard and colleagues (2011) found that individual diagnosed with MDD were the most impaired in information processing, with a majority of patients exhibiting impairments in executive function. Indeed, all patients experienced negative effects on work functioning (Godard et al., 2011). Accumulating evidence suggests that lasting cognitive symptoms in MDD play an important role in patients' functional recovery (Cambridge et al., 2018; Jaeger et al., 2006). A body of research supports the notion that cognitive dysfunctions are related to psychosocial deficits, particularly in the domains of professional functioning, interpersonal functioning, and daily obligations (Baune et al., 2010; Evans et al., 2014; Gotlib & Joormann, 2010).

Cognitive dysfunctions related to MDD may play a role in the patients' ability to perform in the workplace (Woo et al., 2016). Due to the early age of onset and the chronic course of illness, MDD has a large influence on work efficiency (Druss et al., 2000). Indeed, cognitive dysfunction is associated with unemployment status (Baune et

al., 2010). Poor cognitive health may be partly responsible for patient's failure to return to work. Baune and colleagues (2010) revealed that only cognitively recovered patients were able to successfully restart their employment, indicating that cognition is an important factor that contributes to the burden of disease. Godard and colleagues (2012) aimed to characterize psychosocial and cognitive profiles in patients experiencing a major depressive episode. All participants had psychosocial impairments, characterized by occupational dysfunctions. At a 12-month time point follow-up, work-related, psychosocial impairments, and neurocognitive impairments, were present and impacted the daily functioning of patients (Godard et al., 2012). Furthermore, a study also evaluated the effect of cognitive symptoms on daily functioning in MDD including occupational functioning (McIntyre, 2013). The researchers observed that dysfunctions in psychomotor processing speed, executive function, attention and working memory speed were the largest influence on work performance.

Cognitive dysfunction, particularly in the domains of processing speed and executive functioning have been related to functional disability in MDD. Processing speed is defined as the amount of time it takes a person to complete a task requiring cognitive effort (Tsourtos et al., 2002). This domain is impaired in individuals with MDD and affects daily functioning (Caligiuri & Ellwanger, 2000; Evans et al., 2014; Sobin & Sackeim, 1997). A systematic review observed that processing speed was one of the cognitive domains engendering deficits in psychosocial functioning (Evans et al., 2014). Patients will often report that their mental speed is slow (O'Connor et al., 1990) and are less able to carry out tasks in a timely manner (Godard et al., 2011). Moreover,

Kalb and colleagues (2006) found that reaction time to take decisions is slower in individuals with MDD compared to control participants. Therefore, processing speed affects work productivity reducing decision making abilities (Clark et al., 2016).

Executive functioning pertains to a broad range of higher order cognitive functions, including working memory, cognitive flexibility, impulse inhibition, and planning. Dysfunctions in that area is considered to be an indicator of cognitive impairment in MDD (Miyake et al., 2000). Kiosses and Alexopoulos (2005) suggested that executive functioning deficits had a mediating effect on the association between MDD and impaired daily functioning. Executive dysfunction can also be a predictive factor of suicidality (Dombrovski et al., 2008). This deficit was found among depressed patients with suicide attempts but not among depressed patients with no history of suicidality, it appears that patients with suicidal ideations have impaired executive functioning (Keilp et al., 2013, Keilp et al., 2001). Executive functioning has been related to occupational functioning in people with MDD (Godard et al., 2011). Previous research that examined the relationship between cognition and psychosocial functioning found that performance in executive functioning and psychomotor speed was associated with functional outcomes (Withall et al., 2009).

Sleep Disturbances and Cognitive Functioning

Sleep is essential in the consolidation of neuronal connections and to process new information (Chambers, 2017). A wealth of research supports the necessity of sleep for healthy brain functioning. Sleep deprivation can affect a wide array of cognitive functions, and these can include attention, memory and decision-making (Alhola & Polo-Kantola, 2007). Also, MDD has been recognized as a significant risk factor for

sleep dysfunctions (Nutt et al., 2008). Although sleep is a symptom of MDD, research has shown that lack of sleep can also play a role in worsening other depressive symptoms (Riemann et al., 2001). It can affect the relapse, the severity, and the duration of the mood disorder (Franzen & Buysse, 2008). Sleep symptoms in MDD have been found in both clinical and epidemiological studies (Argyropoulos & Wilson, 2005). In a longitudinal study of adults with MDD, sleep symptoms occurred in more than 50% aged between 21 and 30 (Breslau et al., 1996). Approximately 75% of patients with MDD will report difficulties in initiating or maintaining sleep (Yates et al., 2007). A national survey study based in the United Kingdom interviewed 8580 depressed individuals and found that 80% had at least one insomnia symptom. These numbers were higher for patients aged between 55-to-64-years with an average of 90% of the sample reporting insomnia (Stewart et al., 2006). Sleep disturbances create an important distress and impact the quality of life of those affected (Nutt et al., 2008). One study observed that 59% of patients with MDD indicated that poor sleep affected their quality of life and a large number reported that the difficulties were present at the onset of the disorder (Paterson et al., 2009). Indeed, it is well documented that patients with MDD present physiological changes in their sleep architecture (Benca et al., 1992). An impaired hypothalamic–pituitary–adrenal axis and abnormal neurotransmitter levels may be responsible for this relationship (Murphy & Peterson, 2015). An imbalance of interaction between cholinergic and adrenergic mechanisms promotes sleep abnormalities in depressed individuals.

From a neuropsychological perspective, it appears that poor sleep impacts the prefrontal cortex and interferes with higher functions including executive functioning

(Drummond et al., 1999; Horne, 1993; McCrae et al., 2012). Wimmer and colleagues (1992) examined the impact of sleep deprivation on cognitive flexibility and processing speed. A sample of university students were sleep-deprived and completed cognitive testing the following day. The participants with less sleep demonstrated a lower performance on the Trail Making Test compared to control participants. Therefore, it appears that sleep maintains cognitive functions and restores attentional mechanisms. Furthermore, a study observed that poor sleepers exhibited worse executive functioning compared to good sleepers (Nebes et al., 2009). Additionally, measures of sleep fragmentation, sleep onset latency and sleep efficiency were associated with a reduced cognitive ability. The adverse effects of sleep disturbances observed in laboratory designs have also been confirmed in ecological settings (Philibert, 2005; Samkoff & Jacques, 1991). Healthcare workers engaged in more errors during repetitive tasks when sleep deprived (Otmani et al., 2005). Furthermore, professional drivers and military personnel who lack sleep have also shown a deterioration in attention and alertness (Russo et al., 2005).

Evidence from observational studies support the relationship between sleep disturbances and cognition (Chen et al., 2012; Cricco et al., 2001; Foley et al., 2004; Killgore, 2010; Merlino et al., 2010). Yaffe and colleagues (2014) suggested that biological mechanisms such as sleep wake cycles may accelerate brain ageing and the development of cognitive impairments in populations at risk for dementia. Moreover, sleep and MDD are strongly associated with cognitive dysfunction. A recent study sought to examine the relationship between depressive symptoms, poor sleep and cognitive impairments in community-dwelling participants (Wu et al., 2019). The

results indicated that difficulty in initiating sleep was a complete mediator of the adverse effect of MMD on cognitive impairment. These findings suggest that assessing sleep symptoms is necessary in order to manage MDD and cognitive deficits. Furthermore, de Almondes and colleagues (2016) investigated the association between sleep symptoms, executive functioning and subjective depressive symptoms in 95 older adults. MDD and sleep complaints interaction would determine worse executive functioning and working memory performances than these two conditions alone. These data are important as executive functions are predictors of neurodegenerative diseases and a history of MDD may increase the risk factor for the development of dementia (Geerlings et al., 2008; Mulyala & Varghese, 2010; Speck et al., 1995). Nevertheless, the extent to which sleep quality is related to depression-related cognitive impairments such as processing speed and executive functioning remains unknown. Therefore, additional research is warranted to explore the impact of sleep quality on cognitive functioning in individuals with MDD.

Limitations of Cognitive Assessments

Given the persistence of cognitive symptoms during both active and remitted phases of the illness and its relationship with daily functioning, continuous monitoring of cognition is consequential (McIntyre, 2013; Ragguett et al., 2016). Nevertheless, no guidelines have been developed to provide appropriate assessments of neurocognitive deficits in clinical settings (Trivedi & Greer, 2014). The evaluation of cognition in individuals with MDD comes with numerous practical challenges. In clinical practice, detailed in-depth objective assessments of cognition are time consuming. Additionally, they require the participant to interrupt their daily routine. Utilizing a complete

neuropsychological battery in patients with MDD taxes sparse resources including staff and space (Casaletto & Heaton, 2017). Therefore, the cognitive tests are impractical as they are episodic, time consuming and only provide a limited amount of data (Dagum, 2018). Furthermore, there is currently no consensus on what constitutes an appropriate cognitive assessment in MDD (Bakkour et al., 2014). Some patients tend to over-report cognitive symptoms when explicitly asked, while others rarely report cognitive problems to health care professionals due to embarrassment or fear of being diagnosed with dementia (Iverson & Lam, 2013).

Self-report questionnaires and interviews are the backbones of clinical and cognitive assessments (Baumeister et al., 2007). A recurrent limitation is the collection of retrospective instead of real-time reports (Csikszentmihalyi & Larson, 1987). Subjective reports of cognitive deficits in MDD are not correlated with objective reports of cognitive health (Naismith et al., 2007). Indeed, there is a high degree of discordance between subjective cognitive scales and objective performance, which necessitates the systematic assessment of both (Popkin et al., 1982). It is common for individuals with MDD to assume that their cognitive deficits will spontaneously resolve because the clinical symptoms improve (Antikainen et al., 2001). Numerous questionnaires measuring the severity of depressive symptoms solely include a limited number of items regarding cognitive symptoms (McIntyre, 2013). For example, the Montgomery–Åsberg Depression Rating Scale (Montgomery & Åsberg, 1979) is comprised of 10 items and only two address cognitive symptoms. Furthermore, validated screening assessments for cognitive impairments such as the Montreal Cognitive Assessment and the Mini-Mental Status Exam may not be sensitive enough

to detect subtle changes in cognition amongst young adults with MDD (McIntyre, 2013; Ragguett et al., 2016). Sensitivity is defined as the proportion of positive observations (e.g., patients with MDD) correctly identified, while specificity is referred to the proportion of negative observations (e.g. healthy controls) correctly identified. It follows, that the higher the specificity and sensitivity, the better the performance of the tests at distinguishing between healthy controls and patients with MDD (Parikh et al., 2008). A short cognitive assessment, sensitive to a change in functioning, suitable as a repeated measure and with a broad conceptual coverage does not currently exist (McIntyre et al., 2015; Ragguett et al., 2016). This highlights the necessity to conduct more research investigating whether the systematic assessment of cognitive function improves outcomes in MDD. To date, several clinician-friendly assessments are in the process of being validated for MDD like the THINC-it screening tool, but brief and feasible options are limited for a straightforward implementation in clinical settings (Harrison et al., 2018).

Advantages of Using a Smartphone Application to Assess Cognition and Behavior

Traditionally, neuropsychological assessments have been completed in-person typically involving an administrator (Strauss et al., 2006). Health providers require years of training to practice neuropsychological testing and produce complete reports on patient's cognitive impairments. The resources in clinical settings such as psychiatric hospitals are limited, therefore it is not practical for clinicians to evaluate all MDD patients with cognitive testing. Moreover, burdens of neuropsychological test administration and practice effects hinder the possibility of continuous sampling to examine longitudinal assessments (Bauer et al., 2012). Smartphone-based assessments

may be a feasible intermediate option between self-reports and comprehensive neuropsychological batteries to provide cognitive monitoring. Mobile technology can support evaluations of cognitive functioning in the patient's home or workplace. Ecological momentary assessment is one method that involves repeated assessments in a naturalistic setting. First, ecological momentary assessments can monitor cognitive health in real-time and enhances our understanding of cognitive dynamics. Furthermore, this method can reduce retrospective recall bias by decreasing the reliance on episodic memory which can be fallacious. Third, it allows to capture brief symptomatic experiences and daily fluctuations (Torous et al., 2014). The possibility to assess changes in a patients' cognitive ability represents a strength of ecological momentary assessments (Moore et al., 2016; Moore et al., 2017). These transportable tests can be completed at any point in time as well as in situations of daily life encompassing a wide array of distractions and therefore better characterizing cognitive functioning. Data collection can occur through two different methods: active collection and passive collection. Active data collection requests the user to enter values or perform a measurement, for example a smartphone-based cognitive test that prompts a patient to complete a Trail Making Test (TMT) (Reitan & Wolfson, 1985). On the other hand, passive data collection continuously acquires data without the knowledge of the user, for example a mobile application estimating the number of steps (Kourtis et al., 2019). The goal of technology is not to replace clinicians or human care providers but rather augment the ability of clinicians to monitor cognitive health in a non-intrusive and cost-effective manner (Torous et al., 2016).

Utilizing smartphone applications may represent a feasible opportunity to monitor

mental health conditions and cognitive health. Previous literature indicates that 97% of individuals with mental health disorders own a phone (Torous et al., 2014). Iliescu and colleagues (2021) investigated smartphone ownership in 101 psychiatric inpatients and found that 82.8% indicated owning a smartphone. Overall, 59.4% of inpatients from all age groups expressed an interest in trying an application for clinical purposes.

Smartphones offer the potential of capturing real-time information in a person's life in contrast to current clinical interviews relying on retrospective recall (Torous et al., 2014). Applications allow live daily tracking and monitoring of patient's mental health and cognitive function (Koo & Vizer, 2019). Ecological momentary assessment methods are "time-stamped" and the precision of when and how entries are recorded adds clinical validity. Hence, they could benefit the development of clinical interventions for MDD (Torous & Powell, 2015). Cognitive functioning might also predict treatment response (Rush et al., 2009). Therefore, continuous assessment of cognition could be a cost-effective and personalized approach to prevent the selection of ineffective treatments or pharmacotherapy (Simon & Perlis, 2010).

Descriptive and rich data from an active collection can yield a plethora of information on cognitive domains such as processing speed or executive functioning. Touch screens can provide detailed information about cognitive health while the user is swiping or typing (Kourtis et al., 2019). Most studies utilizing devices to assess cognition have examined healthy individuals (Dagum, 2018; Moore et al., 2017). However, Bouvard and colleagues (2018) successfully assessed executive functioning in patients with substance-related addictions over a one-week period. They found that repetitive cognitive monitoring can determine changes in cognitive function due to fatigue and

environment. Moreover, participants were not burdened by the active data collection and the assessments became part of their normal phone usage. A growing body of evidence also highlights the use of mobile devices to assess cognition in individuals with dementia or patients at risk for neurodegenerative disorders such as MDD patients (Kourtis et al., 2019). Deficits in fine motor skills and processing speed are possible signs of Alzheimer's and can be examined with tracing accuracy and finger tapping. Rabinowitz and Lavner (2014) found that the contact time for the tapping test was slower for individuals with a mild cognitive impairment compared to healthy controls. In a recent study, researchers found that text keystrokes per minute and the number of pauses can distinguish between individuals with and without a cognitive deficit (Stringer et al., 2018). Early detection of those impairments may help clinicians better tackle cognitive dysfunctions in at-risk MDD patients. In addition, the use of mobile technology has been shown to enhance cognitive function and prevent depressive symptoms in older adults with MDD (Lin et al., 2020). Playing games or engaging in complex tasks on a smartphone demonstrated increased cognitive control abilities in older adults (Anguera et al., 2013). Despite the clear advantages of an ecological momentary assessment, it is not yet utilized in daily clinical interactions and decision-making (Torous et al., 2016). More research is needed as it is still unclear whether some cognitive domains can be better evaluated than others by ecological momentary assessments (Bouvard et al., 2018).

Cognition does not exist in isolation and is influenced by external behavioral factors. Therefore, it is relevant to assess sleep because it may negatively impact cognition and daily functioning (Lam et al., 2014; McIntyre, 2013; Murrough et al., 2011). Previous literature has examined the importance of sleep and MDD (Benca et

al., 1992). In clinical practice, sleep quality is captured through self-reported measures, however these clinical interviews are prone to retrospective recall bias (Popkin et al., 1982; Urban et al., 2018). Passive data collection may not be the optimal option to detect certain sleep dysfunctions. Kolla and colleagues (2016) concluded that sleep tracking devices underestimated sleep disturbances and overestimated sleep duration compared to an actigraphy. Therefore, self-reported sleep questionnaires using ecological momentary assessments may improve the model of prediction of sleep quality. This approach focuses on indicators and symptoms that can be utilized in a clinical context (Torous et al., 2016). Staples and colleagues (2017) have demonstrated that sleep monitoring through smartphone devices is feasible and cost-effective for patients with schizophrenia. There is currently a limited number of validated smartphone applications to support the tracking of sleep quality in mood disorders (Aledavood et al., 2019).

MindLAMP Application

MindLAMP (Learn, Assess, Manage and Prevent) is a free and sharable health smartphone application with functions such as clinical and cognitive assessments, self-management tools, and educational components (Torous et al., 2019). The smartphone application has been developed by our team at the Beth Israel Deaconess Medical Center (BIDMC), Harvard Medical School, Boston, United States of America. It has been utilized by 1000 individuals for research and augmented care in a “digital clinic” in Boston. The application was created to support healthcare professionals by collecting meaningful clinical information. Furthermore, it was built as an open and reusable platform to acquire data regarding the interactions between patients and their devices.

The application is highly customizable, flexible, and adaptable to the unique needs of the researcher or clinician (Torous & Vaidyam, 2020). It can be modified to the patient's preferences and goals, providing more control over what information can be shared. The clinician or researcher can modify the format of assessments, survey questions, frequency of notifications, language and color of the background. Indeed, the application can be altered through a portal for each research study. This novel type of data collection can strengthen the rapport between clinicians or researchers and the patients by gathering continuous data about their mental health. The application does not involve any data sharing or user agreements with technology companies to guarantee the data privacy of participants and clinical research teams. It is available on iOS and Android platforms to ensure it does not discriminate based on technology. Moreover, it requires low power and connectivity, enabling those with older phones and troublesome access to the internet to use the application. MindLAMP was built with the general purpose of monitoring, managing and preventing chronic mental illnesses (Torous & Vaidyam, 2020).

The MindLAMP application also enabled the development of two cognitive assessments called the Jewels Trail Tests A (JTT A) and Jewels Trail Tests B (JTT B), which are modelled off the validated TMT A and B (Reitan & Wolfson, 1985). The TMT measures processing speed and executive functioning, two cognitive domains that are frequently impaired in MDD (Withall et al., 2009). The main difference between the smartphone and standard pen-and-paper versions is that in order to make the tests usable and engaging on a smartphone, users tap the numbers instead of drawing a line between them. They also have to alternate between different shapes instead of letters and numbers for Part B. Preliminary evidence shows that it is feasible to use these MindLAMP

cognitive tests with other mental disorders such as schizophrenia and that they may differentiate healthy controls from patients (Liu et al., 2019). Nevertheless, the tests remain to be validated to assess cognitive functioning in individuals with MDD.

The Present Study

In the present study, I intended to evaluate whether the MindLAMP smartphone-based cognitive assessments could monitor processing speed and executive function in individuals with MDD. The specific aims were as follows:

1. I aimed to determine whether the JTT are valid and reliable to measure cognition in MDD compared to the standard pen-and-paper TMT. I hypothesized that completion time collected on the smartphone application would show moderate concurrent validity when compared to the standard pen-and-paper TMT, as well as moderate test-retest reliability.

2. I aimed to assess the ability of the TMT and JTT at detecting a slower completion time in clinical populations, such as MDD. I hypothesized that significant differences would be observed between MDD patients and healthy control groups in both a laboratory setting and remote environment, with MDD individuals showing a slower completion time.

3. Finally, I aimed to assess predictive validity by exploring whether poor self-reported sleep quality impacted completion time over time measured by the JTT. I hypothesized that poor self-reported sleep quality would lead to a slower completion time on the JTT.

Method

Participants

Participants were included if they were aged between 18 and 75 years old inclusively for both patient and control groups. They were required to have the ability to speak and read English above the level of sixth grade. For the patient group, participants had to be diagnosed with MDD, confirmed by the Mini International Neuropsychiatric Interview (MINI) interview (Sheehan et al., 1998). We excluded those who had changed or stopped medication in the past month for their symptoms to be controlled. All participants in the current study were on stable medication for one month, information about the specific antidepressants of each patient is reported in Table 1. Exclusion criteria for all participants encompassed any significant neurological or medical disorders that may cause cognitive impairment. In addition, we excluded participants with a recent history of substance abuse or dependence within the past three months and decisional incapacity requiring a guardian. We also excluded participants who did not own a smartphone. Potential healthy controls (HC) were excluded from this study if they presented current or past psychiatric diagnoses or if they had a first-degree relative who received a diagnosis of MDD. Efforts were made to have each participant in each group matched in terms of age and sex.

The sample consisted of 38 participants in the HC group and 24 participants in the MDD group. Four participants were removed from the HC group: three for not meeting the inclusion/exclusion criteria as discovered during the clinical interview and one for not following the instructions during the assessment. As such, the final sample had 34 participants in the HC group and 24 participants in the MDD group. Due to

COVID-19, the data collection was paused hence this thesis reports an interim analysis and we anticipate recruiting a total of 40 participants in each group.

Missing data was observed for the JTT at baseline and during longitudinal data collection. For instance, some participants only completed one Part of the JTT, or failed to complete any trials in the remote setting. As such, missing timepoints were removed from the data during the statistical analyses.

Table 1

Baseline Medication Information for MDD group

Antidepressants	Number of MDD patients
Sertraline	4
Bupropion	4
Fluoxetine	5
Vortioxetine	1
Zopiclone	1
Venlafaxine	4
Duloxetine	1
Escitalopram	4

Procedure

The recruitment of research participants began in February 2019 and was discontinued in March 2020 due to COVID-19. We have recruited outpatients from the mood disorder units of the Royal Ottawa Mental Health Centre. Additionally, we have recruited from the local community, community services, including the Mood

Disorders Ottawa Mutual Support Group, websites and social media, including [ementalhealth.ca](http://mentalhealth.ca), Reddit and Twitter. Compensation was provided using cash. Participants received a one-time \$30.00 after the baseline in-laboratory assessment was completed.

A trained research assistant and I performed the clinical interviews and cognitive assessments for the research study in all participants in a quiet testing room at the Royal Ottawa Hospital. The totality of the baseline visit lasted approximately one hour and thirty minutes per person to be completed. At the beginning of the first visit, all participants were invited to read and sign the informed consent form. In addition, the Mini-International Neuropsychiatric Interview (MINI) and the Montgomery Asberg Depression Rating Scale (MADRS) were administered to ensure current depressive symptoms and no suicidal ideations. Four participants presented suicidal ideations and the research team assessed the risk of suicide using the suicide interview plan (the Columbia Suicide Severity Scale). In all participants, the risk of suicidal behavior was not considered high. Therefore, the research team recommended him/her to contact their physician and to call the suicide help hotline if needed.

In the second part of the in-laboratory baseline visit, participants underwent a cognitive assessment consisting of two separate tests. Instructions and short practice trials were given before each cognitive assessment to ensure that participants understood the task. First, participants completed the TMT (Reitan & Wolfson, 1985) using pen and paper in a quiet testing room in the laboratory. Participants were instructed to complete this task as quickly as possible without lifting the pen from the paper, and the amount of time for the participants to complete this task was scored as

their performance on this test.

Second, the participants were instructed to download the MindLAMP application on their smartphone in order to complete the JTT in the laboratory. Time was also used as a measure of performance however, on the application participants were limited to 90 seconds to complete each test. Mistakes were tracked by the application with an addition of two seconds to the participant's time, resulting in a poorer score. Performance on the JTT was automatically obtained by the application and was used as a measure of cognitive performance.

We invited the participants to keep the application on their phones for a period of three months. The MindLAMP application (Torous et al., 2019), sent pop-up notifications for participants to answer surveys and engage in the cognitive task once a week on either Mondays, Wednesdays, or Fridays. Nevertheless, participants were informed that they could complete the surveys as often as possible. The JTT and surveys took approximately 5 to 10 minutes to be completed each week. After three months, the application would no longer gather any data, and the participants were not able to use the application. In addition, participants were informed that they were able to delete the application on their phones, to cease the data collection, and to delete the data that have been collected if they wanted. If they dropped out of the study early or chose to withdraw at any time, they could also delete the application and it would stop collecting any data. Pen-and-paper and smartphone cognitive assessments were counterbalanced to avoid order effects. Lastly, a self-report questionnaire for sleep quality was filled in during the initial laboratory visit through the MindLAMP application on the participant's smartphones.

Measures

Demographic Information. Demographic information was collected by the research assistant or myself during the one-on-one interview process during the baseline in-lab visit. This information includes date of birth and sex.

Major Depressive Disorder Diagnosis. The Mini-International Neuropsychiatric Interview (MINI), based on the Diagnostic and Statistical Manual of Mental Disorders was administered during the baseline visit to confirm the diagnosis of people with MDD, and to ensure that HC participants did not have an undisclosed current or past mental illness (Sheehan et al., 1998). The MINI is a psychiatric structured diagnostic interview instrument where the patient requires to only answer “yes” or “no” to questions. Previous research supports the validity and reliability of the MINI in the assessment of psychiatric disorders (Mordal et al., 2010; Otsubo et al., 2005; Rossi et al., 2004). This clinical instrument has been translated into 43 languages (Boudrot et al., 2013). The validity and reliability of the tool has been explored for the original English and French versions. The MINI has demonstrated high accuracy for MDD patients in primary care and in psychiatric care (de Azevedo Marques & Zuardi, 2008; Pettersson et al., 2018). Excellent inter-rater agreement ($\kappa = 1.00$) of the English version of MINI have been reported (Lecrubier et al., 1997). According to Shrout and colleagues (1987), in test-retest analyses, kappa values ($\kappa = .84$) demonstrated excellent agreement for MDD diagnoses. In addition, an excellent sensitivity ($\kappa = .96$) and specificity ($\kappa = .88$) was observed for the MINI (Sheehan et al., 1997). A very good convergent validity was reported compared to the Composite International Diagnostic Interview (CIDI) (Lecrubier et al., 1997) and the Structured Clinical

Interview for Diagnostic and Statistical Manual (SCID) (Sheehan et al., 1997). The MINI is a time-efficient diagnostic interview that can be incorporated into routine clinical settings. It is acceptable to patients and improves diagnostic accuracy compared to other clinical assessments (Pinninti et al., 2003).

Depressive Symptoms. The MADRS was administered during the baseline visit to ensure current depressive symptoms and no suicidal ideations (Montgomery & Asberg, 1979). The MADRS is a diagnostic questionnaire that includes ten questions used to measure the severity of depressive symptoms in individuals with mood disorders (Montgomery & Asberg, 1979). The questionnaire incorporates items on the following clinical symptoms: apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulty, lassitude, inability to feel, pessimistic thoughts and suicidal thoughts. Each question produces a score of 0 to 6 and the total score ranges from 0 to 60. The levels of severity of MDD based on the total score are the following: normal or absent 0-6; mild 7-19; moderate 20-34; and severe 35-60 (Montgomery & Asberg, 1979). The MADRS has been widely used in research and given its advantages, has been translated from English into several different languages (Ahmadpanah et al., 2016; Dratcu et al., 1987; Lobo et al., 2002; Schmidtke et al., 1988; Soron, 2017). Numerous studies in different countries have demonstrated that the MADRS is sensitive to change, reliable and valid (Bondolfi et al., 2010; Davidson et al., 1986; Kearns et al., 1982). Paiva-Medeiros and colleagues (2015) examined the psychometric properties of the MADRS in severely obese patients ($N = 374$). The meant total score was 7.73 ($SD = 11.33$) with a Cronbach's alpha coefficient of internal consistency of .93, demonstrating the ability of the items to

assess homogeneously the target construct. Moreover, the results revealed a sensitivity of .81 and specificity of .85 (Paiva-Medeiros et al., 2015). The instrument has been validated in a Colombian sample presenting high internal consistency ($\alpha = .92$; Cano et al., 2016). In the same study, the findings of inter-rater reliability revealed that the tool was highly reliable ($r = .98$; Cano et al., 2016). A meta-analysis by Heo and colleagues (2007), identified that the MADRS and Hamilton Depression Rating Scale (HAM-D) resulted in comparable ratings in assessing changes in overall depressive symptom severity. Compared to the Hamilton Depression Rating Scale (HAM-D), the MADRS may have more precision in estimating MDD and a greater capacity to differentiate between responders and nonresponders to antidepressants (Bagby et al., 2004; Williams & Kobak, 2008).

Pen-and-paper cognitive assessments. The TMT (Reitan & Wolfson, 1985) was used during the baseline visit to assess the cognitive functioning in MDD and HC groups. Moreover, Part A of the TMT measures processing speed and required participants to draw lines connecting encircled numbers from 1-25 in ascending order. Similar to Part A, Part B measures executive functioning and required participants to draw lines connecting encircled numbers and letters in an alternating pattern (i.e., 1-A-2-B-3-C, etc.). The TMT is one of the most widely used instruments in neuropsychological assessment (Lezak, 1995; Mitrushina, 2005; Reitan, 1992). The tool was originally a part of the six subtests of the Army Individual Test Battery (1944) and was later incorporated into the Halstead-Reitan Battery (Reitan & Wolfson, 1985). It was used to assess processing speed and executive functions in all participants in this study. The TMT is a low cost, and low effort measure of cognition, which can be

quickly administered in 5 to 10 minutes as well as easily scored (Reitan & Wolfson, 1985). Previous research has shown that patients with MDD achieve significantly slower completion times than HC in both parts of the TMT (Talarowska et al., 2012). The TMT has been considered as a reliable, valid and sensitive measure of neurological function (Alvarez, 1962; Armitage, 1946; Lezak, 1995; Reitan, 1955, Reitan & Wolfson, 1985). The TMT has good psychometric properties (Gaudino et al., 1995) such as high interrater reliability (Carone, 2007). A study investigated the test-retest reliability in a sample of 150 patients with various neurological disorders and found that both Part A and B had a moderate but acceptable reliability ($r = .69$ and $r = .66$ respectively; Goldstein & Watson, 1989). However, the TMT may be susceptible to practice effects, similar to numerous neuropsychological assessments (Stuss et al., 1987). DesRosiers and Kavanagh (1987) evaluated the sensitivity of the TMT at detecting cognitive impairment in clinical populations. Both parts of the assessment clearly differentiated brain injured patients from control groups, TMT Part A ($F(1,62) = 9.01, p < .023$) and TMT Part B ($F(1,62) = 8.12, p < .034$). The TMT appears to be a particularly sensitive indicator of neurological impairment.

Smartphone-based cognitive assessments. A smartphone application version of the pen-and-paper TMT A and B named the JTT (Torous et al., 2019) was developed to assess cognition. Similar to the TMT, the test involves a Part A (Figure 1A) and a Part B measuring (Figure 1B) processing speed and executive functioning, respectively. In Part A, participants were asked to tap the jewel-shaped numbers on their screen in sequential order from 1 to 25 as fast as possible without making any mistakes. Indeed, Part A of this task required participants to tap on numbers enclosed

in 25 diamond shapes in ascending order as fast and as accurately as possible. Part B involved alternating between two different shapes of the same number while continuing to tap the numbers in sequential order (i.e., 1-shape A, 1- shape B, 2-shape A, 2-shape B...). Before starting the task, participants were told that if they made a mistake, two seconds would be added to their time, resulting in a slower time score. Further, completion time is the amount of time participants took to complete the cognitive tasks and was kept as their scores. Completion time on the JTT was automatically obtained by the application and was used as a measure of cognitive performance.

Figure 1

Smartphone-Based Cognitive Assessments on MindLAMP A) JTT A: Participant taps on numbers in chronological order; B) JTT B: Participant taps on numbers in chronological order while alternating between the two shapes.

Self-reported sleep quality survey. A self-report questionnaire for sleep quality was filled in during the initial laboratory visit through the MindLAMP application on the participant's smartphones. Indeed, sleep quality was assessed through a subjective 1-item questionnaire on the application. The question asked

participants to rate the quality of last night's sleep on a scale from 1- 10 (1= low; 10= excellent): *On a scale of 1-10, (1=very bad, and 10=excellent) how was your sleep last night?*. The self-reported questionnaire included on the MindLAMP application does not possess any psychometric properties.

Statistical Analysis

Outlier detection analysis. A detailed examination of outliers was conducted for each statistical analysis to remove data points that were likely to be erroneous. For the concurrent validity analyses, outliers were identified as datapoints that were more than three times the standard deviation away from the mean, independently for the TMT and JTT. Finally, a visual inspection of the outliers was conducted to determine which points were technical issues or distractions based on notes in the participant's file. These datapoints were subsequently removed from the concurrent validity analyses.

For the second aim, the outliers in the baseline TMT and JTT data that were identified and removed in aim one were also removed. An outlier analysis was also conducted to identify erroneous datapoints in the remote longitudinal data, computed at the participant level. Cook's D and DFBETAs were calculated and plotted to identify participants of high influence using the influence.ME package in R studio (Nieuwenhuis et al., 2016). An influential participant was denoted as having a Cook's D greater than 4 divided by the sample size, or a DFBETA greater than 4 divided by the square root of the sample size. This was followed by a visual inspection to identify participants that should be removed due to datapoints that were likely to be measurement errors. Finally, this analysis was also conducted for the remote longitudinal data in the third aim. Cook's D and DFBETAs were calculated and plotted, followed by a visual inspection.

Descriptive statistics. The descriptive statistics of the sample, including mean, median and range, were calculated. A two-sided t-test was performed to assess significant differences in the mean age of the MDD and HC groups. Similarly, a chi-square test was conducted to evaluate whether there was a difference in the proportion of females and smartphone usage within the two groups. A p -value of .05 was selected as the threshold for significance. These tests were conducted to ensure that both the MDD and HC groups were homogeneous in their characteristics, allowing for a more accurate comparison of cognitive performance between both groups.

Concurrent validity. The concurrent validity of the JTT was evaluated to determine whether the JTT measured cognition in participants similarly as the standard pen-and-paper TMT. The strength of the relationship of completion time between these tests, and the established TMT was evaluated. A Pearson's correlation was performed between the JTT A and TMT A for all participants, and an additional correlation was conducted for Part B. This analysis was conducted at three time points: in the laboratory setting, as well as each participant's first and last time point in the remote environment. Participants who only had one time point in the remote setting were only included in the data for the first time point. These computations allowed to assess whether the JTT completion time was similar to the TMT completion time, and whether it remained comparable over time when the JTT were completed in a remote environment. Scatter plots of these relationships were presented to further illustrate the strength of the relationship. Post-hoc general linear models were run to evaluate the between-group interactions of the MDD and HC group completion time on the TMT and JTT.

Test-retest reliability. Test-retest reliability was assessed by calculating the

intraclass correlation (ICC) between the longitudinal measurements of the JTT at three separate time points: the JTT completion time in the laboratory setting, as well as the first and last JTT completion time conducted in the remote setting for each participant. This analysis was performed to assess the reliability of the repeated measurements with the estimates arising from an analysis of variance (ANOVA) (Weir, 2005). As such, the test-retest reliability analysis examined the variability in the measurements taken by the same participants over time. The reliability analysis used an absolute agreement, 2-way mixed-effects model to calculate the ICC (Koo & Li, 2016). That is,

$$ICC(A, 1) = \frac{MS_R - MS_E}{MS_R + (k - 1)MS_E + \frac{k}{n}(MS_C - MS_E)}$$

where MS_R is the mean square for rows, MS_E is the mean square error, MS_C is the mean square for columns, k is the number of measurements and n is the number of subjects (McGraw & Wong, 1996).

The ICC was calculated independently for the JTT A and JTT B using the psych package in R (Revelle & Revelle, 2015). The confidence intervals of the ICC estimates were presented, and ICC values between 0.50 and 0.75 were denoted as having moderate reliability, while values between 0.75 and 0.90 indicated good reliability, and values greater than 0.90 were interpreted as having excellent reliability (Koo & Li, 2016).

Between-group differences in TMT and JTT completion time. An analysis was conducted to assess whether there were significant differences between the TMT completion time of participants in the HC group compared to the MDD group in data available from the laboratory setting. The means and standard deviations of the

laboratory data were compared between the MDD and the HC groups.

First, a linear model was conducted to evaluate the differences in the completion time of the groups in the laboratory setting for the TMT A and the TMT B. In both models, a binary indicator denoting participants in the MDD group was included. The model controlled for the age and the sex of the participants, with the dependent variable being the TMT completion time. The significance of the MDD indicator was evaluated, as it measures the difference in mean completion time between both groups.

Second, an analysis was conducted to assess whether there were significant differences between the JTT completion time of participants in the HC group compared to the MDD group in data available from the laboratory and remote settings. The means and standard deviations of the laboratory and remote data were compared between the MDD and the HC groups.

A linear model was conducted to evaluate the differences in the completion time of the groups in the laboratory setting for the JTT A and the JTT B. In both models, a binary indicator denoting participants in the MDD group was included. The model controlled for the age and the sex of the participants, with the dependent variable being the JTT completion time. The significance of the MDD indicator was evaluated, as it measures the difference in mean completion time between both groups.

Longitudinal mixed effects models were fitted to the longitudinal data to account for the dependency between the longitudinal measurements of the same participants. These longitudinal mixed effects models were conducted using the lmerTest package (Kuznetsova et al., 2017). This package provides a coefficient table

for the fixed effects that employs Satterthwaite's method to approximate the degrees of freedom to assess significance (Satterthwaite, 1941). The p -values reported for the longitudinal mixed effects model are those provided by this package.

Two longitudinal mixed effect models were produced. One model had the JTT A completion time as the dependent variable and the other had the JTT B completion time. In both models, a binary indicator denoting participants in the MDD group was included as the fixed independent variable, with the participant ID as the random-effect variable. Sex and age were included as control variables as these variables are known to impact cognitive functioning. That is, the model is defined as follows,

$$\text{JTT completion time} \sim \text{MDD indicator} + \text{Sex} + \text{Age} + (1 \mid \text{Participant ID})$$

The MDD indicator denotes the difference in completion time between the participants in the MDD and HC groups. It follows that I assessed the significance of the MDD indicator to determine whether there were significant differences in the completion time between participants in the MDD and HC groups.

Self-reported sleep quality. To explore whether sleep quality impacted cognition over time, two additional mixed effects models were created with the JTT A and JTT B completion time as the dependent variables. In these models, the sleep scale was included as a fixed independent variable and the age and the sex of the participant were added as control variables. The mixed effects model is defined as follows,

$$\text{JTT completion time} \sim \text{Sleep quality} + \text{Sex} + \text{Age} + (1 \mid \text{Participant ID}).$$

I aimed to explore the significance of the sleep scale by assessing whether self-

reported sleep quality significantly impacted the JTT completion time. Scatter plots were created to illustrate the relationship between the sleep scale and completion times as estimated by the longitudinal models. An additional mixed effects model was conducted with the inclusion of an interaction term, that is,

$$\text{JTT completion time} \sim \text{Sleep quality} * \text{MDD indicator} + \text{Sex} + \text{Age} + (1 \mid \text{Participant ID}).$$

The significance of the interaction term was evaluated to assess whether there were between-group differences on the relationship between self-reported sleep quality and JTT completion time.

Practice effects. The impact of practice effects, where a patient improves at the JTT due to practice over time, was evaluated. A model was run with the number of times the test has been previously completed as the fixed independent variable and the completion time of the JTT as the dependent variable. Age and sex were added as control variables. The participant ID was included as the random-effect variable. That is, the mixed effects model is as follows

$$\text{JTT completion time} \sim \text{Number of previous trials} + \text{Age} + \text{Sex} + (1 \mid \text{Participant ID}).$$

The significant effect of the number of trials on the JTT completion time was examined to quantify the impact of practice effects.

Results

Outlier detection analysis

For aim 1 in the baseline data, four participants were removed for Part A and one participant was removed for Part B. For the first timepoint in the remote setting, Part A saw two participants removed while one participant was deemed erroneous for Part B. Finally, for the last timepoint in the remote setting, two participants were removed in Part A and no participants were removed for Part B.

For the longitudinal data in aim 2, a total of two participants and 11 timepoints removed from Part A while two participants with 14 datapoints were removed from Part B.

Finally, in the longitudinal remote data for aim 3, two participants with 13 timepoints were removed from Part A and one participant with 8 timepoints was removed from Part B (see Table 2).

Table 2

Final Sample Size of Each Aim and Statistical Analysis After Removing Missing Values and Erroneous Datapoints.

Aim	JTT A completion time		JTT B completion time	
	Number of participants (HC/MDD)	Number of timepoints (HC/MDD)	Number of participants (HC/MDD)	Number of timepoints (HC/MDD)
Aim 1				
Baseline data	(34/18)	(34/18)	(34/21)	(34/21)
First timepoint in the remote setting	(24/21)	(24/21)	(22/20)	(22/20)
Last timepoint in the remote setting	(20/19)	(20/19)	(21/18)	(21/18)
Aim 2				
Baseline data	(34/18)	(34/18)	(34/21)	(34/21)
Longitudinal remote data	(25/21)	(153/153)	(22/22)	(145/143)
Aim 3				
Longitudinal remote data	(24/19)	(126/117)	(24/18)	(132/107)

Demographic and Clinical Data

The demographic data for the HC and the MDD group are presented in Table 3. The age range was 18 to 72 for the HC group ($M = 33.47$, $SD = 13.22$) and 19 to 75 for the MDD group ($M = 33.25$, $SD = 14.89$). The difference in age between both groups was not significant. The MDD sample consisted of 33.33% males ($n = 8$) and 66.67% females ($n = 16$), and the difference between the groups was not statistically significant. The

MADRS total score indicated that, on average, participants with MDD exhibited moderate depressive symptoms ($M = 29.14$, $SD = 5.52$) The cell phone use of participants was similar between the MDD and HC groups, with most participants using their smartphone daily.

Table 3

Baseline Demographic and Clinical Data for Patients and Controls

	MDD group	HC group	<i>p</i> -value
<i>N</i>	24	34	
Age (years)			.95
Mean (<i>SD</i>)	33.25 (14.89)	33.47 (13.22)	
Median [Min, Max]	27 [19, 75]	29 [18, 72]	
Sex			.06
Male	8 (33.33 %)	21 (61.76 %)	
Female	16 (66.67 %)	13 (38.24 %)	
MADRS total score (<i>SD</i>)	29.14 (5.52)		
Smartphone use			
Every day of the week	23 (95.83%)	34 (100%)	.86
6 or less days per week	1 (4.17%)	0 (0%)	

Concurrent Validity

The correlations between the JTT and the TMT in the laboratory baseline visit were positive, moderate, and significant for Part A ($r = .35$, $p = .01$) and strong and significant for Part B ($r = .58$, $p = <.001$) (see Figure 2 A and B). No significant

between-group interactions were observed on these associations for both Part A and Part B ($t = 0.80, p = .47$ and $t = 0.43, p = .87$, respectively).

Figure 2

Relationship Between Completion Time of Participants on the Pen-and-Paper TMT and Smartphone Application JTT

A second set of correlations was conducted to assess the concurrent validity of the TMT compared to the first JTT of each participant, completed on the smartphone application in a remote environment. The correlations between the first time point of the JTT in a remote environment and the pen and paper TMT were positive, moderate, and significant for Part A ($r = .30, p = .04$) and strong and significant for Part B ($r = .64, p = <.001$). The linear model showed that the interaction term was not significant for Part A and for Part B (Part A: $t = -0.72, p = .48$; Part B: $t = 1.56, p = .13$).

Finally, correlations between the TMT and each participant's last completion of the JTT in a remote environment were conducted to conclude the assessment of the concurrent validity of the JTT. The correlation between the last timepoint of the JTT and the TMT test was strong and significant for Part A ($r = .58, p < .001$) and for Part B ($r = .56, p = .002$). No significant between-group interactions were observed on these associations for both Part A and Part B (Part A: $t = -0.39, p = .70$; Part B: $t = 1.82, p = .08$).

Test-Retest Reliability

The ICC showed moderate reliability for Part A of the JTT ($ICC = .57, CI [0.43, 0.71]$) and good reliability for Part B ($ICC = .76, CI [0.64, 0.85]$).

Between-group differences in TMT and JTT completion time

Table 4 shows the mean completion time on the TMT for participants in the MDD and HC groups. The results demonstrated that participants in the MDD group had a similar completion time when compared to those in the HC group.

Table 4

Means and Standard Deviations of the MDD and HC Groups for TMT Completion Time in the Laboratory Setting

	MDD group	HC group
TMT A completion time		
<i>N</i>	18	34
Mean (<i>SD</i>)	21.89 (6.58)	20.38 (5.45)
TMT B completion time		
<i>N</i>	21	34
Mean (<i>SD</i>)	56.09 (26.25)	51.99 (21.12)

In the linear model with TMT as the dependent variable, controlling for age and sex, I observed that between-group differences in TMT completion time were not statistically significant (see Table 5).

Table 5

Estimates of the Linear Regression Model of MDD Indicator on the TMT Completion Time in the Laboratory Setting.

	Estimate	95% CI	<i>t</i> -value	<i>p</i> -value
TMT A completion time				
Intercept	17.47	(12.63, 22.32)	7.25	<.001
MDD indicator	1.84	(-1.75, 5.44)	1.03	.31
Age	0.08	(-0.07, 0.23)	1.05	.30
Sex	0.46	(-3.49, 4.41)	0.23	.82
TMT B completion time				
Intercept	38.72	(21.56, 55.88)	4.53	<.001
MDD indicator	3.49	(-9.45, 16.44)	0.54	.59
Age	0.48	(-0.03, 0.98)	1.89	.06
Sex	-4.38	(-18.48, 9.72)	-0.62	.54

Differences in mean completion time on the JTT for participants in the MDD and HC groups are available in Table 6. The results showed that participants in the MDD group had similar completion time when compared to those in the HC group.

Table 6

Means and Standard Deviations of the MDD and HC Groups for JTT Completion Time in the Laboratory Setting

	MDD group	HC group
JTT A completion time		
<i>N</i>	18	34
Mean (<i>SD</i>)	34.22 (9.82)	36.41 (9.36)
JTT B completion time		
<i>N</i>	21	34
Mean (<i>SD</i>)	46.48 (16.27)	45.29 (13.05)

In the linear model controlling for age and sex on JTT completion time, I observed that between-group differences in completion time were not statistically significant (see Table 7). The completion time for participants in the MDD group was, on average, similar to the HC group for both JTT A and the JTT B in the laboratory setting.

Table 7

Estimates of the Linear Regression Model of MDD Indicator on the JTT Completion Time in the Laboratory Setting

	Estimate	95% CI	<i>t</i> -value	<i>p</i> -value
JTT A completion time				
Intercept	22.89	(16.33, 29.44)	7.02	<.001
MDD indicator	-2.74	(-7.70, 2.21)	-1.11	.27
Age	0.48	(0.28, 0.69)	4.78	<.001
Sex	-4.27	(-9.64, 1.09)	-1.60	.12
JTT B completion time				
Intercept	27.25	(17.98, 36.51)	5.90	<.001
MDD indicator	1.17	(-5.93, 8.27)	0.33	.74
Age	0.58	(0.31, 0.86)	4.29	<.001
Sex	-2.37	(-10.05, 5.30)	-0.62	.54

Table 8 shows the comparison in mean completion time and standard deviations in the longitudinal data. Participants in the MDD group had a slower mean cognitive completion time in both JTT A and JTT B relative to those in the HC group.

Table 8

Means and Standard Deviations of the MDD and HC groups for JTT Completion Time in a Remote Setting

	MDD group	HC group
JTT A completion time		
<i>N</i>	153	153
Mean (<i>SD</i>)	39.76 (13.03)	35.20 (10.44)
JTT B completion time		
<i>N</i>	145	143
Mean (<i>SD</i>)	43.99 (15.69)	38.62 (10.82)

The results of the longitudinal mixed effects models evaluating the differences in completion time between the MDD and the HC groups are presented in Table 9. In both the JTT A and JTT B, participants in the MDD group had a slower mean completion time relative to participants in the HC group, but these differences were not statistically significant (see Table 9).

Table 9

Fixed Effect Estimates of the Mixed Effects Model of MDD Indicator on the JTT Completion Time in a Remote Setting

	Estimate	95% CI	<i>t</i> -value	<i>p</i> -value
JTT A completion time				
Intercept	21.18	(15.86, 26.51)	7.60	<.001
MDD indicator	2.75	(-1.66, 7.17)	1.19	.24
Age	0.51	(0.35, 0.68)	5.98	<.001
Sex	-1.87	(-6.82, 3.07)	-0.72	.47
JTT B completion time				
Intercept	20.09	(12.93, 27.25)	5.35	<.001
MDD indicator	4.66	(-1.06, 10.39)	1.55	.13
Age	0.60	(0.38, 0.81)	5.31	<.001
Sex	0.11	(-6.38, 6.61)	0.03	.97

Note. CI = Confidence Intervals.

Effect of Age

Table 7 and Table 9 also demonstrate that the control variable age was a significant predictor of the completion time for both Part A and Part B of the JTT ($p < .001$ and $p < .001$ for JTT A and JTT B, respectively in the laboratory setting; $p < .001$ and $p < .001$ for JTT A and JTT B, respectively in the remote environment). In the laboratory setting, a 40-year-old participant was, on average, 4.8 seconds slower than a 30-year-old participant on Part A, and 5.8 seconds on Part B. In the longitudinal data, a 40-year-old participant on Part A was 5.1 seconds slower on average than a 30-year-old

participant, and 6.0 seconds slower on Part B.

Self-Reported Sleep Quality

Table 10 shows the results of the longitudinal mixed effects models evaluating the impact of poor self-reported sleep quality on completion time over a three-months period, controlling for age and sex. In both models, the mean completion time was faster as scores on the sleep scale increased. For every unit increase on the sleep scale, indicating that a participant had a better night of sleep, the average completion time on the JTT A was significantly faster, while controlling for age and sex. Participants with better sleep quality appeared to have a faster average completion time on the JTT B, but the effect was not statistically significant (Figure 3).

Table 10*Fixed Effect Estimates of the Mixed Effects Model of Sleep Scale on the JTT Completion**Time*

	Estimate	95% CI	<i>t</i> -value	<i>p</i> -value
JTT A completion time				
Intercept	29.10	(22.34, 35.89)	8.26	<.001
Sleep scale	-0.82	(-1.52, -0.12)	-2.27	.02
Age	0.53	(0.37, 0.70)	6.35	<.001
Sex	-4.69	(-9.34, -0.05)	-1.94	.06
JTT B completion time				
Intercept	27.21	(17.31, 37.08)	5.31	<.001
Sleep scale	-0.22	(-1.01, 0.58)	-0.56	.58
Age	0.61	(0.35, 0.87)	4.44	<.001
Sex	-4.86	(-12.59, 2.86)	-1.21	.24

Note. *CI* = Confidence Intervals.

Figure 3

Longitudinal Mixed effect Model Prediction of the Impact of Self-Reported Sleep Quality Measured with the Sleep Scale on Cognition Over Time, Measured as the JTT Completion Time A) Part A of the test; B) Part B of the test

Note. HC: healthy control, MDD: Major depressive disorder. The black line represents the predicted values from the longitudinal mixed effects model, while the grey area represents the confidence interval.

The additional mixed effects model with the inclusion of an interaction term between the sleep scale and the MDD indicator was run. This interaction term was not statistically significant for JTT A ($t = 0.22, p = .82$) or JTT B ($t = 0.31, p = .76$). Therefore, the association between the self-reported sleep quality and the JTT completion time was similar between the MDD and HC groups.

Practice Effects

Practice effects, measured as the number of previous trials, were not significant for JTT A and JTT B (see Table 11 and Figure 4).

Table 11

Fixed Effect Estimates of the Mixed Effects Model of the Practice Effects on the JTT Completion Time

	Estimate	95% CI	<i>t</i> -value	<i>p</i> -value
JTT A completion time				
Intercept	24.37	(18.51, 30.21)	8.06	<.001
Number of previous trials	-0.11	(-0.37, 0.15)	-0.80	.42
Age	0.50	(0.32, 0.68)	5.33	<.001
Sex	-2.60	(-7.62, 2.44)	-1.00	.32
JTT B completion time				
Intercept	28.06	(19.28, 36.85)	6.17	<.001
Number of previous trials	-0.26	(-0.55, 0.02)	-1.80	.07
Age	0.58	(0.31, 0.86)	4.10	<.001
Sex	-4.77	(-12.58, 3.03)	-1.18	.25

Note. CI = Confidence Intervals.

Figure 4

Longitudinal Mixed Effect Model Prediction of the Impact of Practice Effects, Measured as the Number of Previous Trials, on the JTT Completion Time. A) Part A of the test; B) Part B of the test

Discussion

The purpose of this study was to examine whether the MindLAMP smartphone-based cognitive assessments could monitor processing speed and executive function in individuals with MDD. Participants with MDD and HC underwent baseline clinical interviews and cognitive assessments at the Royal Ottawa Hospital. In addition, participants were asked to keep the MindLAMP application on their phones for three months and to complete the JTT A and JTT B as well as the self-reported sleep survey during that time. Three research questions were examined. First, are the JTT valid and reliable to measure cognition in MDD compared to the standard pen-and-paper TMT? It was expected that completion time collected on the smartphone application would show moderate concurrent validity when compared to the standard pen-and-paper TMT, as well as moderate test-retest reliability. Second, can the TMT and JTT detect cognitive impairments in clinical populations, such as MDD? It was predicted that the TMT and JTT would significantly differentiate MDD patients and HC in a laboratory setting and in a remote environment. Lastly, does self-reported sleep quality impact completion time measured by the smartphone application over time and demonstrate predictive validity? It was expected that poor self-reported sleep quality would lead to a slower completion time on the smartphone application in all participants.

As expected, I observed a significant, positive, and moderate correlation between the completion time on the smartphone application and the standardized pen-and-paper TMT in the laboratory, which supports concurrent validity for Part A ($r = .35$) and a significant, positive, and strong correlation for Part B ($r = .58$). Therefore, a participant that has a faster completion time on the TMT is also expected to have a faster completion

time on the JTT, and vice versa. Similar results were found when examining the first completion time of participants on the JTT in a remote environment compared to the TMT in the laboratory setting. A moderate and significant correlation of .30 was observed for Part A of the tests, followed by a strong significant correlation of .64 for Part B. In addition, the last completion time of participants on the JTT in a remote environment significantly correlated with the TMT in the laboratory setting (Part A: $r = .58$; Part B: $r = .56$). These results emphasize the consistency of the participants completion time on the JTT, even when completed in a different environment to the TMT. This further supports the concurrent validity of the JTT.

There were no significant between group interactions between the MDD and HC groups, indicating that the relationship between the TMT and JTT was consistent across all participants, regardless of whether they had an MDD diagnosis or not. Moreover, the test-retest reliability coefficients exhibited a moderate reliability between the laboratory setting, as well as the first and the last timepoint in a remote setting for Part A of the test and good reliability for Part B. As such, these findings provide strong evidence of the consistency and the comparability of the JTT, both in laboratory and remote settings.

My second research question was not supported. I observed that between-group differences in completion time on the TMT and JTT in the laboratory setting were not statistically significant for both Part A and Part B of the tests. In addition, participants with MDD did not have a statistically significant slower completion time relative to participants in the HC group for the JTT Part A ($p = .24$) in the longitudinal data. The completion time of the participants in the MDD group was, on average, 5.37 seconds slower for the JTT B, though the difference between both groups was also not statistically

significant ($p = .13$). In addition, age was a significant predictor of the completion time for both Part A and Part B of the JTT.

The third research question was only partially supported. Indeed, self-reported sleep quality was a significant predictor of the completion time on the remote cognitive assessment Part A, demonstrating the predictive validity of the application. For every unit increase in the sleep scale, the average completion time on the JTT A significantly decreased by 0.82 seconds, while controlling for age and sex ($p = .02$). For the JTT B, every unit increase in the sleep scale decreased the completion time by an average of 0.22 seconds, but the effect was not significant ($p = .58$). In sum, the higher the self-reported sleep quality, the better the processing speed on the application for all participants. The interaction between sleep scale and the MDD indicator was not statistically significant for JTT A ($t = 0.22, p = .82$) and JTT B ($t = 0.34, p = .74$). This indicates that the relationship between the sleep scale and the JTT completion time was similar across both the MDD and HC groups.

Lastly, practice effects for JTT A and JTT B were not significant when controlling for the age and sex of participants. Therefore, the participants did not significantly improve their performance due to an increase in practice from completing previous trials.

The Validity of a Smartphone Application to Assess Cognition

The current study is the first aiming to validate an ecological momentary assessment assessing processing speed and executive functioning in MDD. The findings indicate that the MindLAMP application is a promising tool to assess cognitive functioning in individuals with MDD. The results highlight that the JTT work

in a similar way to the gold-standard TMT, which is widely used to measure processing speed and executive functioning (Bowie & Harvey, 2006). The results showed significant, positive, and moderate correlations between completion time on the smartphone application and the standardized pen-and-paper TMT when compared in a laboratory setting. In addition, the first remote JTT A and B completion time were moderately associated with the pen and paper TMT A and B completion time in the laboratory. Similarly, the last remote JTT A and B completion time were also moderately associated with the pen and paper TMT A and B completion time in the laboratory. This shows consistency across the various completion of the JTT in the remote setting.

These results are in line with previous studies intending to validate ambulant cognitive assessments on a smartphone. Jongstra and colleagues (2017) developed a digital version of the TMT on the iVitality application to assess cognition in healthy adults at increased risk for dementia. The authors compared the first cognitive completion time of 151 participants on the smartphone with their completion time on the existing neuropsychological test in a laboratory setting. They found a moderate correlation ($r = .38$) between the smartphone-based TMT and the conventional test. Further, the correlation ($r = .43$) increased when the pen-and-paper assessment was compared with the mean scores of all performed smartphone tests. Taken together, these results support the use of a digital version of the TMT to yield a valid representation of the neuropsychological functioning of individuals at risk for cognitive impairments. Further, one study from Liu and colleagues (2019) supported the feasibility of using these MindLAMP cognitive tests with other mental disorders such

as schizophrenia. However, it was warranted to assess the potential of the application with participants diagnosed with MDD. Overall, the current study provides initial validity for cognitive assessments on a smartphone application aiming to measure processing speed and executive functioning in MDD.

Furthermore, no significant between-group interactions were observed between the HC and MDD groups. Therefore, the relationship between the TMT and JTT did not significantly differ across both groups. This further supports the validity of the smartphone application, as the JTT completion time of participants in both groups was positively correlated to the TMT completion time, further demonstrating the consistency of the smartphone application.

Comparison of Patients and Healthy Controls

This study did not find significant differences between the MDD and HC groups completion time on both TMT and JTT during the baseline laboratory visit. In addition, the results indicated that there were no significant differences between the MDD and HC groups completion time on the JTT during the three-month period in a remote setting. This was an unexpected finding considering that cognitive deficits are predominant in individuals with MDD (Snyder, 2013). Previous research has shown that cognitive impairment during episodic phases of MDD observed significant moderate deficits in processing speed and executive function when compared to HC (Murrough et al., 2011; Rock et al., 2014).

The first possible consideration is that our participants were less cognitively impaired and not representative of the MDD population. The results showed that participants in the MDD group had a similar completion time on the TMT when

compared to those in the HC group. Further, the average MADRS score was 29.14, with 95% of our MDD patients experiencing mild or moderate clinical symptoms. Only one participant reported severe depressive symptoms on the scale. Self-selection bias might have been introduced in this research project as participants who voluntarily chose to participate could potentially differ in clinical symptoms from those who do not. It is plausible that patients with more severe clinical characteristics and lower cognitive abilities may not be drawn to take part in clinical research. Prior scientific literature has demonstrated that participants with poorer psychological health such as higher depressive symptoms were less likely willing to volunteer in research (Almeida et al., 2008; Donkin et al., 2012; van Heuvelen et al., 2005). In one study, Moritz and colleagues (2017) examined cognitive impairments in MDD. Patients with MDD exhibited more negative attitudes towards neuropsychological testing and lower performance motivation compared to HC. Additionally, other empirical work showed that higher depressive severity was related with more disengagement from complex cognitive tasks (Bowie et al., 2016). As 87% of our sample volunteered from the community and were not patients from the hospital, self-selection bias may have impacted the internal validity of the results. Therefore, the JTT are measuring cognition in a similar way as the TMT, and that the lack of significant differences is mostly due to characteristics of our current sample. Efforts will be made to recruit more symptomatic patients during the completion of the study and include a more representative sample of the MDD population

The second possible consideration is that cognitive dysfunction may fluctuate in patients with MDD (Murrough et al., 2011). Prior literature supports that cognitive deficits are associated with the number of previous depressive episodes in mood disorders

(Kessing, 1998). Indeed, Kessing (1998) found that patients were significantly more impaired if they experienced recurrent episodes compared to patients with a single episode. Xie and colleagues (2019) demonstrated that poor sleep quality and depressed mood predicted reduced cognitive functioning. Sleep disturbances may influence cognitive functioning in people with MDD. In addition, McIntyre and peers (2014) showed that specific antidepressant medications can also impact cognitive functioning. More specifically, they conducted a randomized controlled trial to evaluate the efficacy of vortioxetine on adults with MDD. The findings indicated that participants in the treatment group significantly improved on subjective and objective measures of cognition. Nevertheless, the study did not assess whether the improved cognitive functioning resulted in an increased daily functioning. In sum, a wide array of factors impact cognition in patients with MDD resulting in a dynamic cognitive functioning in this specific psychiatric population.

The third possible consideration is the size of the sample, with the data for the MDD and HC groups encompassing only 24 and 34 unique participants, respectively. As such, the analysis may lack the power to detect significant differences due to small sample sizes, especially for the MDD group. In the longitudinal data, on both the JTT A and JTT B, participants in the MDD group had, on average, a slower completion time relative to participants in the HC group, however these differences were not significant. Further recruitment may improve the statistical power of the analyses and could demonstrate the ability of the JTT to adequately assess differences in cognitive functioning between individuals with MDD and controls.

A fourth possible consideration is that age was not perfectly matched between

MDD and HC groups. Additionally, the current study found that age was a significant predictor of the completion time for both Part A and Part B of the JTT in the MD and HC groups. Specifically, as age increased, completion time on the JTT was slower, indicating poorer cognition. Since the sample ranged from ages 18-75, I observed a decline in cognitive performance on the JTT, which could be related to regular ageing effects on cognitive abilities. The results reflect previous scientific literature, which has emphasized the role of aging in MDD and cognitive impairments (Thomas et al., 2008). The TMT completion time has also been demonstrated to be affected by age (Corrigan & Hinkeldey, 1987). Indeed, TMT scores have generally been found to decline with age (Davies, 1968; Giovagnoli et al., 1996; Tombaugh, 2004). One study showed that with increasing age, TMT performances were slower for patients with traumatic brain injuries and control participants, and the validity of the TMT also decreased. There are also differences in age and smartphone usage, and habits. Youth may have more abilities with technology. Older adults experience difficulties in using smartphones due to a lack of knowledge (Mohadisudis & Ali, 2014). Zhou and colleagues (2014) noted that older adults were unable to appropriately utilize smartphone functionalities compared to younger adults.

Towards Remote Cognitive Assessments

The results demonstrate that the JTT holds promising psychometric properties for the long-term assessment of cognitive abilities. Notably, the test-retest reliability showed moderate reliability for Part A of the JTT and good reliability for Part B. That is, participants showed consistent cognitive functioning across the longitudinal scores of the JTT, supporting the reliability of the JTT across different completion times of the same

participant. These findings provide evidence that digital assessments can maintain standardized test procedures in a remote location. Manually administered tests are susceptible to variations in test procedures and test administrators and are limited to time and resources while remote digital assessments implement automated scoring, which could decrease time consumption for clinicians and patients (Torous & Powell, 2015; Woods et al., 2015). These transportable tests could be completed at any point in time as well as in situations of daily life encompassing a wide array of distractions and therefore better characterizing cognitive functioning (Torous et al., 2014).

Continuous and repeated monitoring of cognition is primordial in patients with MDD and therefore, warrants the examination of practice effects in the JTT (Conradi et al., 2011). Significant practice effects were not observed on either Part A or B of the JTT over the three-month longitudinal portion of the study. The absence of significant practice effects in the results suggests that the JTT could be used remotely for a period of numerous weeks. The pen-and-paper TMT has been reported to show practice effects, which have previously resulted in the development of alternate test forms (Wagner et al., 2011). To reduce the potential learning effect, variations of the organization of the jewels on the JTT were developed. The MindLAMP application delivers a different version of the JTT for every attempt of the cognitive test. Indeed, digital versions of cognitive tests provide the ideal format for generating an unlimited number of JTT instances. Jongstra and colleagues (2017) developed a digital version of the TMT without incorporating alternative versions or variations of the cognitive assessment. The authors noted a learning effect for their smartphone version of the TMT. As the testing was repeated, the test scores were significantly increasing with the number of tests performed. Therefore,

future research should consider providing variations of the digital cognitive assessments to evaluate users' cognitive functioning over time. Smartphone-based cognitive testing would facilitate the monitoring of cognitive assessment in a noninvasive manner, rapidly and repeatedly at a suitable time for the patient (Torous et al., 2014).

Monitoring Cognitive Functioning and Symptomatology Longitudinally

The results indicated that changes in self-reported sleep quality over time could influence processing speed as measured by the smartphone JTT A. The higher the quality of the self-reported sleep, the faster the completion time on the smartphone application over a period of three months. The results are in line with previous scientific literature examining sleep disturbances on cognitive variables. Specifically, Lim and Dignes (2010) conducted a meta-analysis examining the impact of short-term sleep deprivation in healthy adults across several cognitive domains. Significant differences were observed for processing speed with moderate effect sizes across the studies selected ($g = -0.30$, CI [0.46, 0.14]). Further, one study showed sleep–wake disturbances in older patients with MDD was related to neuropsychological functioning (Naismith et al., 2011). Sleep disturbance in late-life MDD was significantly related to performance on tests of processing speed ($r = -0.30$) and executive functioning ($r = -0.42$). Therefore, smartphones could be utilized to monitor these numerous manifestations of symptomatic behavior. Wisniewski and colleagues (2019) examined the relevance of the MindLAMP application in case report series for patients with psychotic illness. The findings demonstrated that a smartphone application can identify clinical and behavioral patterns. For example, one patient exhibited strong correlations between his anxiety, MDD, psychotic symptoms, sleep quality and cognitive scores. A benefit of utilizing a

subjective survey through a smartphone application, is that it can be paired with the JTT and used as an educational tool to provide key information to patients and their support teams such as tracking cognitive functioning and symptomatology in a longitudinal manner. Overall, the current findings provide evidence that cognition measured over time by the JTT may be susceptible to dynamic factors, such as sleep quality and support the predictive validity of the JTT A on the MindLAMP application.

Participants with higher sleep quality appeared to have a faster average completion time on the JTT B, however the effect was not significant. The JTT are still within the exploratory phase, and the subjective sleep quality survey has yet to be validated. Indeed, the subjective survey used on the MindLAMP app did not possess any psychometric properties. Furthermore, this research only focused on a small part of the data obtained by the sleep survey. Perhaps the current survey is not sensitive enough to detect meaningful sleep disturbances that impact executive functioning. It is possible that the other questions on the sleep survey may have been better predictors of sleep quality and should be considered in future research. These factors may have played a role in the results obtained in the present study. The small sample size of the study may have reduced the chances of detecting a true effect. Upon further review of the literature, it appears that sleep disturbances impact executive functioning in patients with depressive symptoms. Almondes and colleagues (2016) investigated the association between sleep symptoms, executive functioning, and subjective depressive symptoms in 95 older adults. MDD and sleep complaints interaction would determine worse executive functioning than these two conditions alone. Therefore, further investigation of the JTT B is warranted, as executive functions are predictors of neurodegenerative diseases and a

history of MDD may increase the risk factor for the development of dementia (Geerlings et al., 2008; Mulyala & Varghese, 2010; Speck et al., 1995).

Strengths and Limitations

This research has several strengths. The current study is the first to examine the validity and reliability of a smartphone application that assesses cognitive functioning in individuals with MDD. In addition, the current study provides initial evidence on the JTT acceptability, validity, and reliability, as well as critical insight on the potential of using such smartphone-based assessments to measure cognition in MDD.

Nevertheless, these results need to be interpreted cautiously, considering some limitations. Due to COVID-19, the study was halted, and the participant recruitment was not completed. Therefore, these preliminary results are limited to a small sample size. After removing the outliers, the baseline data contained only 18 unique participants for the MDD group, and 34 participants in the HC group. It is predicted that further recruitment and data collection will improve the statistical power of the analyses. For example, a larger sample size may show a stronger positive correlation between the TMT A and the JTT A or demonstrate significant completion time differences between the HC and MDD groups. A larger sample size may support the impact of sleep disturbances on executive functioning, as measured by the JTT B. Additionally, considering that the sample size is relatively small, it may affect the generalization of the observed effect. Therefore, further studies with greater sample sizes are needed.

Additionally, a few differences exist between the demographics of participants in the MDD and HC groups. While the sample was balanced in terms of age, it was not balanced in terms of sex. Indeed, MDD patients remain mostly female and the HC group

recruitment was mostly male. While the models did control for both age and sex, these differences may have introduced individual variabilities not captured in the control variables. I will actively seek to recruit participants to reduce these differences in demographics in the pursuit of this project.

An additional limitation of the present study involved technical difficulties regarding data collection. Smartphone application assessments are an increasingly new tool in the field of psychiatry, and difficulties are bound to occur as with the development of any other technology. Some occasional difficulties arose where the application failed to save completion times to the study database. This technical difficulty occurred for two patients, requiring these participants to be excluded from the data analyses. The team of application developers is continuously working to improve any technical issues occurring with the MindLAMP application.

In addition, participants were instructed to complete the JTT in a quiet environment however, these conditions could not be verified. The ideal testing environment for one participant may be too quiet or too noisy for another participant. Future research should implement a survey check-in after completing the JTT. It is important to ask if there were excessive noises in their surroundings or on the other hand, if they were focused on the assessments.

Implications

The increasingly high Internet and smartphone usage across all generations in Western populations indicates that such technology may benefit the current assessment of cognition, by providing daily and real-time assessments (Hollis et al., 2018). Daily tracking of symptoms and lifestyle habits could provide a closer look into the patient's

everyday life, and what mechanisms may influence their illness (Wisniewski et al., 2019). Preliminary results of the present study introduce this method of measuring cognitive functioning and symptomatology as a future direction in clinical practice and research development. Capturing data in this manner could set forth new methods of clinical assessments (Torous et al., 2016). While in-person assessments provide information that may not be conveniently captured through digital devices, remote cognitive technology represents a unique opportunity for more accurate data collection and enhanced clinical care (Wisniewski et al., 2019). Mobile technology can support evaluations of cognitive functioning in the patient's home or workplace. Participants from the current study had the potential to benefit from the ease of use, convenience, and accessibility of the smartphone application cognitive assessments and surveys.

Considering the steady rise of smartphone ownership and daily-usage, smartphone applications may be the key resource for reaching populations with limited access to mental health services. Previous empirical research has recognized the limited amount mental health services for special-needs communities, such as seniors in rural and remote areas and Indigenous communities within the boundaries of Canada. Indeed, a great number of these patients have difficulties accessing outpatient clinics or hospitals (Boksa et al., 2015; Morgan et al., 2002). Digital assessments could provide an alternative access to mental health care for these individuals at risk for cognitive impairments or experiencing cognitive impairments, and who do not typically have access to such evaluations (Morgan et al., 2009). Implementing remote cognitive testing could decrease the patient's burden to attend the clinic or hospital and increase their engagement resulting in potential benefits (Owens et al., 2020). Utilizing smartphone applications

may provide a continued tracking of a cognitive impairments progression in patients that would otherwise not have access to continued clinical care.

The COVID-19 pandemic has heightened the need for remote cognitive assessments (Nitkunan et al., 2020). The pandemic has led to a reduction in face-to-face clinical appointments. Further, one study based in the United States of America found that the risk of being diagnosed with COVID-19 was higher for patients with MDD (Wang et al., 2021). To reduce the risk of infection, social distancing measures were recommended, however, unintentional effects of such distancing lead to a reduction in social activity, physical activity and increased depressive symptoms, and these factors are associated with cognitive decline (De Pue et al., 2021). Implementing remote cognitive assessments to augment clinical care during COVID-19 is primordial. Nevertheless, there is a vital need for a framework to guide the remote cognitive and behavioral assessment (Geddes et al., 2020). Smartphone applications can provide an opportunity to examine how the conventional assessments could be adjusted for remote testing and used to monitor cognitive functioning in individuals with a serious mental illness (Larner, 2021; Wisniewski et al., 2019).

While the findings suggest that a smartphone application could be used to assess cognitive functioning in individuals with a mental illness, further research is needed to evaluate the validity of such assessments in wider mental illness populations.

Conclusion

The present study provides preliminary data indicating that utilizing a smartphone application is feasible to assess cognitive functioning in individuals with MDD and represents a novel approach to assessing cognition in the context of everyday life.

The results showed that scores from the pen-and-paper and smartphone application cognitive assessments positively correlated with each other. The comparison between Part A of the pen-and-paper and smartphone app assessments indicated a moderate positive correlation, and a strong positive significant correlation was observed between the Part B. No significant between-group interactions were observed between the HC and MDD groups. Therefore, the relationship between the TMT and JTT was similar across both groups.

This study did not find significant differences between the MDD and HC groups completion time on TMT and JTT during in the laboratory setting and on the JTT during the three-month period in a remote setting. Nevertheless, participants showed consistent completion time measurements across the longitudinal performance results of the JTT, supporting the reliability of the smartphone application across different scores of the same participant.

The current findings indicated that changes in self-reported sleep quality over time could influence processing speed as measured by the smartphone JTT A. In addition, participants with a higher sleep quality appeared to have a faster average completion time on the JTT B, but the effect was not statistically significant. Therefore, the results provide initial evidence that the novel MindLAMP app is an accessible, valid, and reliable tool to assess cognition in people with MDD.

Although MDD is a major public health concern, the benefit of smartphone application cognitive assessments could be seen in a range of cognitive disorders (Lim et al., 2018; Trivedi, 2006). Longitudinal research is needed to observe if smartphone application assessments have the potential to improve functioning in a variety of

cognitive domains (Liu et al., 2019). The current study can provide a template to further the development of additional smartphone application assessments. The smartphone Jewel Trail Tests appears to be useful to measure processing speed and executive functioning in patients with MDD, and other cognitive domains such as working memory or verbal memory remain to be tested.

References

- Albert, U., Brugnoli, R., Caraci, F., Dell'Osso, B., Di Sciascio, G., Tortorella, A., Vampini, C., Cataldo, N., & Pegoraro, V. (2016). Italian psychiatrists' perception on cognitive symptoms in major depressive disorder. *International journal of psychiatry in clinical practice*, 20(1), 2–9.
<https://doi.org/10.3109/13651501.2015.1093147>
- Aledavood, T., Torous, J., Triana Hoyos, A. M., Naslund, J. A., Onnela, J. P., & Keshavan, M. (2019). Smartphone-Based Tracking of Sleep in Depression, Anxiety, and Psychotic Disorders. *Current psychiatry reports*, 21(7), 49.
<https://doi.org/10.1007/s11920-019-1043-y>
- Alhola, P., & Polo-Kantola, P. (2007). Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric disease and treatment*, 3(5), 553–567.
- Almeida, L., Kashdan, T. B., Nunes, T., Coelho, R., Albino-Teixeira, A., & Soares-da-Silva, P. (2008). Who volunteers for phase I clinical trials? Influences of anxiety, social anxiety and depressive symptoms on self-selection and the reporting of adverse events. *European Journal of Clinical Pharmacology*, 64(6), 575–582. <https://doi.org/10.1007/s00228-008-0468-8>
- Alvarez, R. R. (1962). Comparison of Depressive and Brain-Injured Subjects on the Trail Making Test. *Perceptual and Motor Skills*, 14(1), 91–96.
<https://doi.org/10.2466/pms.1962.14.1.91>
- Ahmadpanah, M., Sheikhabaei, M., Haghighi, M., Roham, F., Jahangard, L., Akhondi, A., Sadeghi Bahmani, D., Bajoghli, H., Holsboer-Trachsler, E., & Brand, S. (2016). Validity and test-retest reliability of the Persian version of

- the Montgomery-Asberg Depression Rating Scale. *Neuropsychiatric disease and treatment*, 12, 603–607. <https://doi.org/10.2147/NDT.S103869>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publishing.
- Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., & Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature*, 501(7465), 97–101. <https://doi.org/10.1038/nature12486>
- Antikainen, R., Hänninen, T., Honkalampi, K., Hintikka, J., Koivumaa-Honkanen, H., Tanskanen, A., & Viinamäki, H. (2001). Mood improvement reduces memory complaints in depressed patients. *European archives of psychiatry and clinical neuroscience*, 251(1), 6–11. <https://doi.org/10.1007/s004060170060>
- Argyropoulos, S. V., & Wilson, S. J. (2005). Sleep disturbances in depression and the effects of antidepressants. *International review of psychiatry (Abingdon, England)*, 17(4), 237–245. <https://doi.org/10.1080/09540260500104458>
- Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs*, 60(1), i-48. <https://doi.org/10.1037/h0093567>
- Army Individual Test Battery. (1944). *PsycTESTS Dataset*. <https://doi.org/10.1037/t31500-000>
- Bakkour, N., Samp, J., Akhras, K., El Hammi, E., Soussi, I., Zahra, F., Duru, G., Kooli, A., & Toumi, M. (2014). Systematic review of appropriate cognitive assessment instruments used in clinical trials of schizophrenia, major depressive disorder

and bipolar disorder. *Psychiatry research*, 216(3), 291–302.

<https://doi.org/10.1016/j.psychres.2014.02.014>

Bagby, R. M., Ryder, A. G., Schuller, D. R., & Marshall, M. B. (2004). The Hamilton Depression Rating Scale: has the gold standard become a lead weight?. *The American journal of psychiatry*, 161(12), 2163–2177.

<https://doi.org/10.1176/appi.ajp.161.12.2163>

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 27(3), 362–373. <https://doi.org/10.1093/arclin/acs027>

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior?. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 2(4), 396–403.

<https://doi.org/10.1111/j.1745-6916.2007.00051.x>

Baune, B. T., Miller, R., Mcafoose, J., Johnson, M., Quirk, F., & Mitchell, D. (2010). The role of cognitive impairment in general functioning in major depression. *Psychiatry Research*, 176(2-3), 183–189.

<https://doi.org/10.1016/j.psychres.2008.12.001>

Benca, R. M., Obermeyer, W. H., Thisted, R. A., & Gillin, J. C. (1992). Sleep and psychiatric disorders. A meta-analysis. *Archives of general psychiatry*, 49(8),

651–670. <https://doi.org/10.1001/archpsyc.1992.01820080059010>

Boksa, P., Joobar, R., & Kirmayer, L. J. (2015). Mental wellness in Canada's Aboriginal communities: striving toward reconciliation. *Journal of psychiatry & neuroscience : JPN*, 40(6), 363–365. <https://doi.org/10.1503/jpn.150309>

Bondolfi, G., Jermann, F., Rouget, B. W., Gex-Fabry, M., McQuillan, A., Dupont-Willemin, A., Aubry, J. M., & Nguyen, C. (2010). Self- and clinician-rated Montgomery-Asberg Depression Rating Scale: evaluation in clinical practice. *Journal of affective disorders*, 121(3), 268–272. <https://doi.org/10.1016/j.jad.2009.06.037>

Boudrot, A., Sheehan, D., & Acquadro, C. (2013). Lost in Translation: Translatability of Psychiatric Terms – The Example of the Mini–International Neuropsychiatric Interview (M.I.N.I.). *Value in Health*, 16(7). <https://doi.org/10.1016/j.jval.2013.08.1692>

Bouvard, A., Dupuy, M., Schweitzer, P., Revranche, M., Fatseas, M., Serre, F., Misdrahi, D., Auriacombe, M., & Swendsen, J. (2018). Feasibility and validity of mobile cognitive testing in patients with substance use disorders and healthy controls. *The American journal on addictions*, 27(7), 553–556. <https://doi.org/10.1111/ajad.12804>

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature protocols*, 1(5), 2277–2281. <https://doi.org/10.1038/nprot.2006.390>

Bowie, C. R., Milanovic, M., Tran, T., & Cassidy, S. (2016). Disengagement from tasks as a function of cognitive load and depressive symptom severity.

Cognitive Neuropsychiatry, 22(1), 83–94.

<https://doi.org/10.1080/13546805.2016.1267617>

Breslau, N., Roth, T., Rosenthal, L., & Andreski, P. (1996). Sleep disturbance and psychiatric disorders: a longitudinal epidemiological study of young adults. *Biological psychiatry*, 39(6), 411–418. [https://doi.org/10.1016/0006-3223\(95\)00188-3](https://doi.org/10.1016/0006-3223(95)00188-3)

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychological bulletin*, 117(2), 285–305. <https://doi.org/10.1037/0033-2909.117.2.285>

Caligiuri, M. P., & Ellwanger, J. (2000). Motor and cognitive aspects of motor retardation in depression. *Journal of affective disorders*, 57(1-3), 83–93. [https://doi.org/10.1016/s0165-0327\(99\)00068-3](https://doi.org/10.1016/s0165-0327(99)00068-3)

Cambridge, O. R., Knight, M. J., Mills, N., & Baune, B. T. (2018). The clinical relationship between cognitive impairment and psychosocial functioning in major depressive disorder: A systematic review. *Psychiatry Research*, 269, 157–171. <https://doi.org/10.1016/j.psychres.2018.08.033>

Cano, J. F., Gomez Restrepo, C., & Rondón, M. (2016). Validación en Colombia del instrumento para evaluación de la depresión Montgomery-Åsberg Depression Rating Scale (MADRS) [Validation of the Montgomery-Åsberg Depression Rating Scale (MADRS) in Colombia]. *Revista colombiana de psiquiatria*, 45(3), 146–155. <https://doi.org/10.1016/j.rcp.2015.08.006>

Carone, D. A. (2007). E. Strauss, E. M. S. Sherman, & O. Spreen, A

Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. *Applied Neuropsychology*, 14(1), 62–63.

<https://doi.org/10.1080/09084280701280502>

Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological Assessment: Past and Future. *Journal of the International Neuropsychological Society : JINS*, 23(9-10), 778–790.

<https://doi.org/10.1017/S1355617717001060>

Chambers A. M. (2017). The role of sleep in cognitive processing: focusing on memory consolidation. *Wiley interdisciplinary reviews. Cognitive science*, 8(3), 10.1002/wcs.1433.

<https://doi.org/10.1002/wcs.1433>

Chen, P. L., Lee, W. J., Sun, W. Z., Oyang, Y. J., & Fuh, J. L. (2012). Risk of dementia in patients with insomnia and long-term use of hypnotics: a population-based retrospective cohort study. *PloS one*, 7(11), e49113. <https://doi.org/10.1371/journal.pone.0049113>

Clark, M., DiBenedetti, D., & Perez, V. (2016). Cognitive dysfunction and work productivity in major depressive disorder. *Expert review of pharmacoeconomics & outcomes research*, 16(4), 455–463.

<https://doi.org/10.1080/14737167.2016.1195688>

Conradi, H. J., Ormel, J., & de Jonge, P. (2011). Presence of individual (residual) symptoms during depressive episodes and periods of remission: a 3-year prospective study. *Psychological medicine*, 41(6), 1165–1174.

<https://doi.org/10.1017/S0033291710001911>

- Cricco, M., Simonsick, E. M., & Foley, D. J. (2001). The impact of insomnia on cognitive functioning in older adults. *Journal of the American Geriatrics Society*, 49(9), 1185–1189. <https://doi.org/10.1046/j.1532-5415.2001.49235.x>
- Corrigan, J. D., & Hinkeldey, N. S. (1987). Relationships between parts A and B of the Trail Making Test. *Journal of clinical psychology*, 43(4), 402–409. [https://doi.org/10.1002/1097-4679\(198707\)43:4<402::aid-jclp2270430411>3.0.co;2-e](https://doi.org/10.1002/1097-4679(198707)43:4<402::aid-jclp2270430411>3.0.co;2-e)
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of nervous and mental disease*, 175(9), 526–536. <https://doi.org/10.1097/00005053-198709000-00004>
- Dagum P. (2018). Digital biomarkers of cognitive function. *NPJ digital medicine*, 1, 10. <https://doi.org/10.1038/s41746-018-0018-4>
- Davidson, J., Turnbull, C. D., Strickland, R., Miller, R., & Graves, K. (1986). The Montgomery- Asberg Depression Scale: reliability and validity. *Acta psychiatrica Scandinavica*, 73(5), 544–548. <https://doi.org/10.1111/j.1600-0447.1986.tb02723.x>
- Davies A. D. (1968). The influence of age on trail making test performance. *Journal of clinical psychology*, 24(1), 96–98. [https://doi.org/10.1002/1097-4679\(196801\)24:1<96::aidjclp2270240131>3.0.co;2-u](https://doi.org/10.1002/1097-4679(196801)24:1<96::aidjclp2270240131>3.0.co;2-u)
- Donkin, L., Hickie, I. B., Christensen, H., Naismith, S. L., Neal, B., Cockayne, N.

- L., & Glozier, N. (2012). Sampling bias in an internet treatment trial for depression. *Translational Psychiatry*, 2(10).
<https://doi.org/10.1038/tp.2012.100>
- de Almondes, K. M., Costa, M. V., Malloy-Diniz, L. F., & Diniz, B. S. (2016). The Relationship between Sleep Complaints, Depression, and Executive Functions on Older Adults. *Frontiers in psychology*, 7, 1547.
<https://doi.org/10.3389/fpsyg.2016.01547>
- de Azevedo Marques, J. M., & Zuardi, A. W. (2008). Validity and applicability of the Mini International Neuropsychiatric Interview administered by family medicine residents in primary health care in Brazil. *General hospital psychiatry*, 30(4), 303–310.
<https://doi.org/10.1016/j.genhosppsych.2008.02.001>
- De Pue, S., Gillebert, C., Dierckx, E., Vanderhasselt, M. A., De Raedt, R., & Van den Bussche, E. (2021). The impact of the COVID-19 pandemic on wellbeing and cognitive functioning of older adults. *Scientific reports*, 11(1), 4636. <https://doi.org/10.1038/s41598-021-84127-7>
- DesRosiers, G., & Kavanagh, D. (1987). Cognitive assessment in closed head injury: Stability, validity and parallel forms for two neuropsychological measures of recovery. *International Journal of Clinical Neuropsychology*, 9(4), 162–173.
- Dooley, D., Prause, J., & Ham-Rowbottom, K. A. (2000). Underemployment and depression: longitudinal relationships. *Journal of health and social behavior*, 41(4), 421–436.
- Dombrovski, A. Y., Butters, M. A., Reynolds, C. F., 3rd, Houck, P. R., Clark, L.,

- Mazumdar, S., & Szanto, K. (2008). Cognitive performance in suicidal depressed elderly: preliminary report. *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry*, 16(2), 109–115. <https://doi.org/10.1097/JGP.0b013e3180f6338d>
- Dratcu, L., da Costa Ribeiro, L., & Calil, H. M. (1987). Depression assessment in Brazil. The first application of the Montgomery-Asberg Depression Rating Scale. *The British journal of psychiatry : the journal of mental science*, 150, 797–800. <https://doi.org/10.1192/bjp.150.6.797>
- Drummond, S. P., Brown, G. G., Stricker, J. L., Buxton, R. B., Wong, E. C., & Gillin, J. C. (1999). Sleep deprivation-induced reduction in cortical functional response to serial subtraction. *Neuroreport*, 10(18), 3745–3748. <https://doi.org/10.1097/00001756-199912160-00004>
- Druss, B. G., Rosenheck, R. A., & Sledge, W. H. (2000). Health and disability costs of depressive illness in a major U.S. corporation. *The American journal of psychiatry*, 157(8), 1274–1278. <https://doi.org/10.1176/appi.ajp.157.8.1274>
- Evans, V. C., Iverson, G. L., Yatham, L. N., & Lam, R. W. (2014). The Relationship Between Neurocognitive and Psychosocial Functioning in Major Depressive Disorder. *The Journal of Clinical Psychiatry*, 75(12), 1359–1370. <https://doi.org/10.4088/jcp.13r08939>
- Franzen, P. L., & Buysse, D. J. (2008). Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications. *Dialogues in clinical neuroscience*, 10(4), 473–481.
- Fink, B. C., & Shapiro, A. F. (2013). Coping Mediates the Association Between

Marital Instability and Depression, but Not Marital Satisfaction and Depression. *Couple & family psychology*, 2(1), 1–13.

<https://doi.org/10.1037/a0031763>

Foley, D., Ancoli-Israel, S., Britz, P., & Walsh, J. (2004). Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey. *Journal of psychosomatic research*, 56(5), 497–502.
<https://doi.org/10.1016/j.jpsychores.2004.02.010>

Gaudino, E. A., Geisler, M. W., & Squires, N. K. (1995). Construct validity in the Trail Making Test: what makes Part B harder?. *Journal of clinical and experimental neuropsychology*, 17(4), 529–535.
<https://doi.org/10.1080/01688639508405143>

Geddes, M. R., O'Connell, M. E., Fisk, J. D., Gauthier, S., Camicioli, R., Ismail, Z., & Alzheimer Society of Canada Task Force on Dementia Care Best Practices for COVID-19 (2020). Remote cognitive and behavioral assessment: Report of the Alzheimer Society of Canada Task Force on dementia care best practices for COVID-19. *Alzheimer's & dementia (Amsterdam, Netherlands)*, 12(1), e12111.
<https://doi.org/10.1002/dad2.12111>

Geerlings, M. I., den Heijer, T., Koudstaal, P. J., Hofman, A., & Breteler, M. M. (2008). History of depression, depressive symptoms, and medial temporal lobe atrophy and the risk of Alzheimer disease. *Neurology*, 70(15), 1258–1264.
<https://doi.org/10.1212/01.wnl.0000308937.30473.d1>

Giovagnoli, A. R., Del Pesce, M., Mascheroni, S., Simoncelli, M., Laiacona, M., & Capitani, E. (1996). Trail making test: normative values from 287 normal adult

controls. *Italian journal of neurological sciences*, 17(4), 305–309.

<https://doi.org/10.1007/BF01997792>

Godard, J., Baruch, P., Grondin, S., & Lafleur, M. F. (2012). Psychosocial and neurocognitive functioning in unipolar and bipolar depression: a 12-month prospective study. *Psychiatry research*, 196(1), 145–153.

<https://doi.org/10.1016/j.psychres.2011.09.013>

Godard, J., Grondin, S., Baruch, P., & Lafleur, M. F. (2011). Psychosocial and neurocognitive profiles in depressed patients with major depressive disorder and bipolar disorder. *Psychiatry research*, 190(2-3), 244–252.

<https://doi.org/10.1016/j.psychres.2011.06.014>

Goldstein, G., & Watson, J. R. (1989). Test-retest reliability of the halstead-reitan battery and the WAIS in a neuropsychiatric population. *Clinical Neuropsychologist*, 3(3), 265–272. <https://doi.org/10.1080/13854048908404088>

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual review of clinical psychology*, 6, 285–312.

<https://doi.org/10.1146/annurev.clinpsy.121208.131305>

Greenberg, P. E., Fournier, A. A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of clinical psychiatry*, 76(2), 155–162.

<https://doi.org/10.4088/JCP.14m09298>

Harrison, J. E., Barry, H., Baune, B. T., Best, M. W., Bowie, C. R., Cha, D. S., Culpepper, L., Fossati, P., Greer, T. L., Harmer, C., Klag, E., Lam, R. W., Lee, Y., Mansur, R. B., Wittchen, H. U., & McIntyre, R. S. (2018). Stability,

reliability, and validity of the THINC-it screening tool for cognitive impairment in depression: A psychometric exploration in healthy volunteers. *International journal of methods in psychiatric research*, 27(3), e1736. <https://doi.org/10.1002/mpr.1736>

Heo, M., Murphy, C. F., & Meyers, B. S. (2007). Relationship between the Hamilton Depression Rating Scale and the Montgomery-Asberg Depression Rating Scale in depressed elderly: a meta-analysis. *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry*, 15(10), 899–905. <https://doi.org/10.1097/JGP.0b013e318098614e>

Hirschfeld, R. M., Dunner, D. L., Keitner, G., Klein, D. N., Koran, L. M., Kornstein, S. G., Markowitz, J. C., Miller, I., Nemeroff, C. B., Ninan, P. T., Rush, A. J., Schatzberg, A. F., Thase, M. E., Trivedi, M. H., Borian, F. E., Crits-Christoph, P., & Keller, M. B. (2002). Does psychosocial functioning improve independent of depressive symptoms? A comparison of nefazodone, psychotherapy, and their combination. *Biological psychiatry*, 51(2), 123–133.
[https://doi.org/10.1016/s0006-3223\(01\)01291-4](https://doi.org/10.1016/s0006-3223(01)01291-4)

Hollis, C., Sampson, S., Simons, L., Davies, E. B., Churchill, R., Betton, V., Butler, D., Chapman, K., Easton, K., Gronlund, T. A., Kabir, T., Rawsthorne, M., Rye, E., & Tomlin, A. (2018). Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. *The lancet. Psychiatry*, 5(10), 845–854.
[https://doi.org/10.1016/S2215-0366\(18\)30296-7](https://doi.org/10.1016/S2215-0366(18)30296-7)

Horne J. A. (1993). Human sleep, sleep loss and behaviour. Implications for the

- prefrontal cortex and psychiatric disorder. *The British journal of psychiatry : the journal of mental science*, 162, 413–419. <https://doi.org/10.1192/bjp.162.3.413>
- Howieson D. (2019). Current limitations of neuropsychological tests and assessment procedures. *The Clinical neuropsychologist*, 33(2), 200–208.
<https://doi.org/10.1080/13854046.2018.1552762>
- Iliescu, R., Kumaravel, A., Smurawska, L., Torous, J., & Keshavan, M. (2021). Smartphone ownership and use of mental health applications by psychiatric inpatients. *Psychiatry research*, 299, 113806.
<https://doi.org/10.1016/j.psychres.2021.113806>
- Iverson, G. L., & Lam, R. W. (2013). Rapid screening for perceived cognitive impairment in major depressive disorder. *Annals of clinical psychiatry : official journal of the American Academy of Clinical Psychiatrists*, 25(2), 135–140.
- Jaeger, J., Berns, S., Uzelac, S., & Davis-Conway, S. (2006). Neurocognitive deficits and disability in major depressive disorder. *Psychiatry research*, 145(1), 39–48. <https://doi.org/10.1016/j.psychres.2005.11.011>
- Jongstra, S., Wijsman, L. W., Cachucho, R., Hoevenaer-Blom, M. P., Mooijart, S. P., & Richard, E. (2017). Cognitive Testing in People at Increased Risk of Dementia Using a Smartphone App: The iVitality Proof-of-Principle Study. *JMIR mHealth and uHealth*, 5(5), e68.
<https://doi.org/10.2196/mhealth.6939>
- Kalb, R., Dörner, M., & Kalb, S. (2006). Opposite effects of depression and antidepressants on processing speed and error rate. *Progress in neuro-psychopharmacology & biological psychiatry*, 30(2), 244–250.

<https://doi.org/10.1016/j.pnpbp.2005.10.009>

Kearns, N. P., Cruickshank, C. A., McGuigan, K. J., Riley, S. A., Shaw, S. P., & Snaith, R. P. (1982). A comparison of depression rating scales. *The British journal of psychiatry: the journal of mental science*, 141, 45–49.

<https://doi.org/10.1192/bjp.141.1.45>

Keefe, R. S., McClintock, S. M., Roth, R. M., Doraiswamy, P. M., Tiger, S., & Madhoo, M. (2014). Cognitive effects of pharmacotherapy for major depressive disorder: a systematic review. *The Journal of clinical psychiatry*, 75(8), 864–876. <https://doi.org/10.4088/JCP.13r08609>

Keilp, J. G., Gorlyn, M., Russell, M., Oquendo, M. A., Burke, A. K., Harkavy-Friedman, J., & Mann, J. J. (2013). Neuropsychological function and suicidal behavior: attention control, memory and executive dysfunction in suicide attempt. *Psychological medicine*, 43(3), 539–551.

<https://doi.org/10.1017/S0033291712001419>

Keilp, J. G., Sackeim, H. A., Brodsky, B. S., Oquendo, M. A., Malone, K. M., & Mann, J. J. (2001). Neuropsychological dysfunction in depressed suicide attempters. *The American journal of psychiatry*, 158(5), 735–741.

<https://doi.org/10.1176/appi.ajp.158.5.735>

Kessing L. V. (1998). Cognitive impairment in the euthymic phase of affective disorder. *Psychological medicine*, 28(5), 1027–1038.

<https://doi.org/10.1017/s0033291798006862>

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E., Wang, P. S., & National Comorbidity Survey Replication

- (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*, 289(23), 3095–3105. <https://doi.org/10.1001/jama.289.23.3095>
- Killgore W. D. (2010). Effects of sleep deprivation on cognition. *Progress in brain research*, 185, 105–129. <https://doi.org/10.1016/B978-0-444-53702-7.00007-5>
- Kiosses, D. N., & Alexopoulos, G. S. (2005). IADL functions, cognitive deficits, and severity of depression: a preliminary study. *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry*, 13(3), 244–249. <https://doi.org/10.1176/appi.ajgp.13.3.244>
- Klerman, G. L. (1992). The Course, Morbidity, and Costs of Depression. *Archives of General Psychiatry*, 49(10), 831. <https://doi.org/10.1001/archpsyc.1992.01820100075013>
- Kolla, B. P., Mansukhani, S., & Mansukhani, M. P. (2016). Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert review of medical devices*, 13(5), 497–506. <https://doi.org/10.1586/17434440.2016.1171708>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Koo, B. M., & Vizer, L. M. (2019). Mobile Technology for Cognitive Assessment of Older Adults: A Scoping Review. *Innovation in aging*, 3(1), igy038. <https://doi.org/10.1093/geroni/igy038>
- Kourtis, L. C., Regele, O. B., Wright, J. M., & Jones, G. B. (2019). Digital

biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. *NPJ digital medicine*, 2, 9. <https://doi.org/10.1038/s41746-019-0084-2>

Krol, M., Papenburg, J., Koopmanschap, M., & Brouwer, W. (2011). Do productivity costs matter?: the impact of including productivity costs on the incremental costs of interventions targeted at depressive disorders. *Pharmacoeconomics*, 29(7), 601–619. <https://doi.org/10.2165/11539970-000000000-00000>

Kuznetsova, A., Brockhoff, P., Christensen R. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>.

Lam, R. W., Kennedy, S. H., McIntyre, R. S., & Khullar, A. (2014). Cognitive dysfunction in major depressive disorder: effects on psychosocial functioning and implications for treatment. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, 59(12), 649–654. <https://doi.org/10.1177/070674371405901206>

Larner A. J. (2021). Cognitive testing in the COVID-19 era: can existing screeners be adapted for telephone use?. *Neurodegenerative disease management*, 11(1), 77–82. <https://doi.org/10.2217/nmt-2020-0040>

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., Janavs, J., & Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry*, 12(5), 224–231.

- [https://doi.org/10.1016/s0924-9338\(97\)83296-8](https://doi.org/10.1016/s0924-9338(97)83296-8) Lee, R. S., Hermens, D. F., Porter, M. A., & Redoblado-Hodge, M. A. (2012). A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. *Journal of affective disorders, 140*(2), 113–124. <https://doi.org/10.1016/j.jad.2011.10.023>
- Lezak, M.D. (1995). Executive functions and motor performance . In M.D. Lezak (Ed.), *Neuropsychological assessment* (pp. 650 – 685). New York : Oxford University Press.
- Lim, G. Y., Tam, W. W., Lu, Y., Ho, C. S., Zhang, M. W., & Ho, R. C. (2018). Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. *Scientific Reports, 8*(1). doi:10.1038/s41598-018-21243-x
- Lin, L., Jing, X. C., Lv, S. J., Liang, J. H., Tian, L., Li, H. L., Puts, M., & Xu, Y. (2020). Mobile device use and the cognitive function and depressive symptoms of older adults living in residential care homes. *BMC geriatrics, 20*(1), 41. <https://doi.org/10.1186/s12877-020-1427-1>
- Lobo, A., Chamorro, L., Luque, A., Dal-Ré, R., Badia, X., Baró, E., & Grupo de Validación en Español de Escalas Psicométricas (GVEEP) (2002). Validación de las versiones en español de la Montgomery-Asberg Depression Rating Scale y la Hamilton Anxiety Rating Scale para la evaluación de la depresión y de la ansiedad [Validation of the Spanish versions of the Montgomery-Asberg depression and Hamilton anxiety rating scales]. *Medicina clinica, 118*(13), 493–499. [https://doi.org/10.1016/s0025-7753\(02\)72429-9](https://doi.org/10.1016/s0025-7753(02)72429-9)
- Liu, G., Henson, P., Keshavan, M., Pekka-Onnela, J., & Torous, J. (2019). Assessing the potential of longitudinal smartphone based cognitive assessment in

- schizophrenia: A naturalistic pilot study. *Schizophrenia research*.
Cognition, 17, 100144. <https://doi.org/10.1016/j.scog.2019.100144>
- Majer, M., Ising, M., Künzel, H., Binder, E. B., Holsboer, F., Modell, S., & Zihl, J. (2004). Impaired divided attention predicts delayed response and risk to relapse in subjects with depressive disorders. *Psychological medicine*, 34(8), 1453–1463. <https://doi.org/10.1017/s0033291704002697>
- McClintock, S. M., Husain, M. M., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Trivedi, M. H., Cook, I., Morris, D., Warden, D., & Rush, A. J. (2011). Residual symptoms in depressed outpatients who respond by 50% but do not remit to antidepressant medication. *Journal of clinical psychopharmacology*, 31(2), 180–186.
<https://doi.org/10.1097/JCP.0b013e31820ebd2c>
- McCrae, C. S., Vathauer, K. E., Dzierzewski, J. M., & Marsiske, M. (2012). Habitual Sleep, Reasoning, and Processing Speed in Older Adults with Sleep Complaints. *Cognitive therapy and research*, 36(2), 156–164.
<https://doi.org/10.1007/s10608-011-9425-4>
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, 1(1), 30.
- McIntyre R. S. (2013). Using measurement strategies to identify and monitor residual symptoms. *The Journal of clinical psychiatry*, 74 Suppl 2, 14–18.
<https://doi.org/10.4088/JCP.12084su1c.03>
- McIntyre, R. S., Lophaven, S., & Olsen, C. K. (2014). A randomized, double-blind, placebo-controlled study of vortioxetine on cognitive function in depressed adults.

- The international journal of neuropsychopharmacology*, 17(10), 1557–1567.
<https://doi.org/10.1017/S1461145714000546>
- McIntyre, R. S., Xiao, H. X., Syeda, K., Vinberg, M., Carvalho, A. F., Mansur, R. B., Maruschak, N., & Cha, D. S. (2015). The prevalence, measurement, and treatment of the cognitive dimension/domain in major depressive disorder. *CNS drugs*, 29(7), 577–589. <https://doi.org/10.1007/s40263-015-0263-x>
- Merlino, G., Piani, A., Gigli, G. L., Cancelli, I., Rinaldi, A., Baroselli, A., Serafini, A., Zanchettin, B., & Valente, M. (2010). Daytime sleepiness is associated with dementia and cognitive decline in older Italian adults: a population-based study. *Sleep medicine*, 11(4), 372–377.
<https://doi.org/10.1016/j.sleep.2009.07.018>
- Mitrushina, M. N. (2005). *Handbook of normative data for neuropsychological assessment*. Oxford University Press.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Mohadisdudis, H.M., Ali, N.M. (2014). A study of smartphone usage and barriers among the elderly. In: *2014 3rd International Conference on User Science and Engineering (i-USEr)*, pp. 109–114. IEEE, Selangor
- Montgomery, S. A., & Asberg, M. (1979). Montgomery-Asberg Depression Rating Scale. *PsycTESTS Dataset*. doi:10.1037/t04111-000
- Moore, R. C., Depp, C. A., Wetherell, J. L., & Lenze, E. J. (2016). Ecological momentary

- assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *Journal of psychiatric research*, 75, 116–123. <https://doi.org/10.1016/j.jpsychires.2016.01.011>
- Moore, R. C., Swendsen, J., & Depp, C. A. (2017). Applications for self-administered mobile cognitive assessments in clinical research: A systematic review. *International journal of methods in psychiatric research*, 26(4), e1562. <https://doi.org/10.1002/mpr.1562>
- Mordal, J., Gundersen, O., & Bramness, J. G. (2010). Norwegian version of the Mini-International Neuropsychiatric Interview: feasibility, acceptability and test-retest reliability in an acute psychiatric ward. *European psychiatry : the journal of the Association of European Psychiatrists*, 25(3), 172–177. <https://doi.org/10.1016/j.eurpsy.2009.02.004>
- Morgan, D. G., Crossley, M., Kirk, A., D'Arcy, C., Stewart, N., Biem, J., Forbes, D., Harder, S., Basran, J., Dal Bello-Haas, V., & McBain, L. (2009). Improving access to dementia care: development and evaluation of a rural and remote memory clinic. *Aging & mental health*, 13(1), 17–30. <https://doi.org/10.1080/13607860802154432>
- Morgan, D. G., Semchuk, K. M., Stewart, N. J., & D'Arcy, C. (2002). Rural families caring for a relative with dementia: barriers to use of formal services. *Social science & medicine (1982)*, 55(7), 1129–1142. [https://doi.org/10.1016/s0277-9536\(01\)00255-6](https://doi.org/10.1016/s0277-9536(01)00255-6)
- Moritz, S., Stöckert, K., Hauschildt, M., Lill, H., Jelinek, L., Beblo, T., Diedrich, S., & Arlt, S. (2017). Are we exaggerating neuropsychological impairment in

- depression? Reopening a closed chapter. *Expert Review of Neurotherapeutics*, 17(8), 839–846. <https://doi.org/10.1080/14737175.2017.1347040>
- Muliyala, K. P., & Varghese, M. (2010). The complex relationship between depression and dementia. *Annals of Indian Academy of Neurology*, 13(Suppl 2), S69–S73. <https://doi.org/10.4103/0972-2327.74248>
- Murphy, M. J., & Peterson, M. J. (2015). Sleep Disturbances in Depression. *Sleep medicine clinics*, 10(1), 17–23. <https://doi.org/10.1016/j.jsmc.2014.11.009>
- Murrough, J. W., Iacoviello, B., Neumeister, A., Charney, D. S., & Iosifescu, D. V. (2011). Cognitive dysfunction in depression: neurocircuitry and new therapeutic strategies. *Neurobiology of learning and memory*, 96(4), 553–563. <https://doi.org/10.1016/j.nlm.2011.06.006>
- Naismith, S. L., Longley, W. A., Scott, E. M., & Hickie, I. B. (2007). Disability in major depression related to self-rated and objectively-measured cognitive deficits: a preliminary study. *BMC psychiatry*, 7, 32. <https://doi.org/10.1186/1471-244X-7-32>
- Naismith, S. L., Rogers, N. L., Lewis, S. J., Terpening, Z., Ip, T., Diamond, K., Norrie, L., & Hickie, I. B. (2011). Sleep disturbance relates to neuropsychological functioning in late-life depression. *Journal of affective disorders*, 132(1-2), 139–145. <https://doi.org/10.1016/j.jad.2011.02.027>
- Nakano, Y., Baba, H., Maeshima, H., Kitajima, A., Sakai, Y., Baba, K., Suzuki, T., Mimura, M., & Arai, H. (2008). Executive dysfunction in medicated, remitted state of major depression. *Journal of affective disorders*, 111(1), 46–51. <https://doi.org/10.1016/j.jad.2008.01.027>

Nebes, R. D., Buysse, D. J., Halligan, E. M., Houck, P. R., & Monk, T. H. (2009).

Self-reported sleep quality predicts poor cognitive performance in healthy older adults. *The journals of gerontology. Series B, Psychological sciences and social sciences*, 64(2), 180–187. <https://doi.org/10.1093/geronb/gbn037>

Nitkunan, A., Paviour, D., & Nitkunan, T. (2020). COVID-19: switching to remote neurology outpatient consultations. *Practical neurology*, 20(3), 222–224.

<https://doi.org/10.1136/practneurol-2020-002571>

Nieuwenhuis, R., Grotenhuis, M. T., & Pelzer, B. (2016). influence.ME: Tools for Detecting Influential Data in Mixed Effects Models.

<https://doi.org/10.31235/osf.io/a5w4u>

Nutt, D., Wilson, S., & Paterson, L. (2008). Sleep disorders as core symptoms of depression. *Dialogues in clinical neuroscience*, 10(3), 329–336.

<https://doi.org/10.31887/DCNS.2008.10.3/dnutt>

O'Connor, D. W., Pollitt, P. A., Roth, M., Brook, P. B., & Reiss, B. B. (1990).

Memory complaints and impairment in normal, depressed, and demented elderly persons identified in a community survey. *Archives of general psychiatry*, 47(3), 224–227. <https://doi.org/10.1001/archpsyc.1990.01810150024005>

Olesen, J., Gustavsson, A., Svensson, M., Wittchen, H. U., Jönsson, B., CDBE2010

study group, & European Brain Council (2012). The economic cost of brain disorders in Europe. *European journal of neurology*, 19(1), 155–162.

<https://doi.org/10.1111/j.1468-1331.2011.03590.x>

Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car

drivers. *Physiology & behavior*, 84(5), 715–724.

<https://doi.org/10.1016/j.physbeh.2005.02.021>

Otsubo, T., Tanaka, K., Koda, R., Shinoda, J., Sano, N., Tanaka, S., Aoyama, H., Mimura, M., & Kamijima, K. (2005). Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry and clinical neurosciences*, 59(5), 517–526. <https://doi.org/10.1111/j.1440-1819.2005.01408.x>

Owens, A. P., Ballard, C., Beigi, M., Kalafatis, C., Brooker, H., Lavelle, G., Brønnick, K. K., Sauer, J., Boddington, S., Velayudhan, L., & Aarsland, D. (2020). Implementing Remote Memory Clinics to Enhance Clinical Care During and After COVID-19. *Frontiers in psychiatry*, 11, 579934. <https://doi.org/10.3389/fpsy.2020.579934>

Paiva-Medeiros, P. F., Duarte-Guerra, L. S., Santo, M. A., Lotufo-Neto, F., & Wang, Y. P. (2015). Psychometric Properties of the Montgomery-Åsberg Depression Rating Scale in Severely Obese Patients. *The Spanish journal of psychology*, 18, E69. <https://doi.org/10.1017/sjp.2015.072>

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology*, 56(1), 45–50. <https://doi.org/10.4103/0301-4738.37595>

Paterson, L. M., Nutt, D. J., & Wilson, S. J. (2009). NAPSAQ-1: National Patient Sleep Assessment Questionnaire in depression. *International journal of psychiatry in clinical practice*, 13(1), 48–58.

<https://doi.org/10.1080/13651500802450498>

Pettersson, A., Modin, S., Wahlström, R., Af Winklerfelt Hammarberg, S., & Krakau, I. (2018). The Mini-International Neuropsychiatric Interview is useful and well accepted as part of the clinical assessment for depression and anxiety in primary care: a mixed-methods study. *BMC family practice*, 19(1), 19.

<https://doi.org/10.1186/s12875-017-0674-5>

Philibert I. (2005). Sleep loss and performance in residents and nonphysicians: a meta-analytic examination. *Sleep*, 28(11), 1392–1402.

<https://doi.org/10.1093/sleep/28.11.1392>

Pinninti, N. R., Madison, H., Musser, E., & Rissmiller, D. (2003). MINI International Neuropsychiatric Schedule: clinical utility and patient acceptance. *European psychiatry : the journal of the Association of European Psychiatrists*, 18(7), 361–364. <https://doi.org/10.1016/j.eurpsy.2003.03.004>

Popkin, S. J., Gallagher, D., Thompson, L. W., & Moore, M. (1982). Memory complaint and performance in normal and depressed older adults. *Experimental aging research*, 8(3-4), 141–145.

<https://doi.org/10.1080/03610738208260271>

Rabinowitz, I., & Lavner, Y. (2014). Association between finger tapping, attention, memory, and cognitive diagnosis in elderly patients. *Perceptual and motor skills*, 119(1), 259–278. <https://doi.org/10.2466/10.22.PMS.119c12z3>

Ragguett, R. M., Cha, D. S., Kakar, R., Rosenblat, J. D., Lee, Y., & McIntyre, R. S. (2016). Assessing and measuring cognitive function in major depressive disorder. *Evidence-based mental health*, 19(4), 106–109.

<https://doi.org/10.1136/eb-2016-102456>

Ravnkilde, B., Videbech, P., Clemmensen, K., Egander, A., Rasmussen, N. A., & Rosenberg, R. (2002). Cognitive deficits in major depression. *Scandinavian journal of psychology*, 43(3), 239–251. <https://doi.org/10.1111/1467-9450.00292>

Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology*, 19(5), 393–394. <https://doi.org/10.1037/h0044509>

Reitan, R.M. (1992). *Trail Making Test: Manual for administration and scoring*. Tucson, AZ : Reitan Neuropsychology Laboratory.

Reitan, R. M., & Wolfson, D. (1985). The Halstead–Reitan Neuropsychological Test Battery: Therapy and clinical interpretation. Tucson, AZ: Neuropsychological Press.

Revelle, W., & Revelle, M. W. (2015). Package ‘psych’. The comprehensive R archive network, 337, 338.

Rhebergen, D., Beekman, A. T., de Graaf, R., Nolen, W. A., Spijker, J., Hoogendijk, W. J., & Penninx, B. W. (2010). Trajectories of recovery of social and physical functioning in major depression, dysthymic disorder and double depression: a 3-year follow-up. *Journal of affective disorders*, 124(1-2), 148–156. <https://doi.org/10.1016/j.jad.2009.10.029>

Richardson, L., & Adams, S. (2018). Cognitive Deficits in Patients With Depression. *The Journal for Nurse Practitioners*, 14(6). <https://doi.org/10.1016/j.nurpra.2018.03.006>

- Riemann, D., Berger, M., & Voderholzer, U. (2001). Sleep and depression--results from psychobiological studies: an overview. *Biological psychology*, 57(1-3), 67–103. [https://doi.org/10.1016/s0301-0511\(01\)00090-4](https://doi.org/10.1016/s0301-0511(01)00090-4)
- Rizvi, S. J., Cyriac, A., Grima, E., Tan, M., Lin, P., Gallagher, L. A., McIntyre, R. S., & Kennedy, S. H. (2015). Depression and employment status in primary and tertiary care settings. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, 60(1), 14–22. <https://doi.org/10.1177/070674371506000105>
- Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological medicine*, 44(10), 2029–2040. <https://doi.org/10.1017/S0033291713002535>
- Rossi, A., Alberio, R., Porta, A., Sandri, M., Tansella, M., & Amaddeo, F. (2004). The reliability of the Mini-International Neuropsychiatric Interview--Italian version. *Journal of clinical psychopharmacology*, 24(5), 561–563. <https://doi.org/10.1097/01.jcp.0000139758.03834.ad>
- Rush, A. J., Warden, D., Wisniewski, S. R., Fava, M., Trivedi, M. H., Gaynes, B. N., & Nierenberg, A. A. (2009). STAR*D: revising conventional wisdom. *CNS drugs*, 23(8), 627–647. <https://doi.org/10.2165/00023210-200923080-00001>
- Russo, M. B., Kendall, A. P., Johnson, D. E., Sing, H. C., Thorne, D. R., Escolas, S. M., Santiago, S., Holland, D. A., Hall, S. W., & Redmond, D. P. (2005). Visual perception, psychomotor performance, and complex motor performance during an overnight air refueling simulated flight. *Aviation, space, and environmental*

medicine, 76(7 Suppl), C92–C103.

Sarkar, U., Gourley, G. I., Lyles, C. R., Tieu, L., Clarity, C., Newmark, L., Singh, K.,

& Bates, D. W. (2016). Usability of Commercially Available Mobile

Applications for Diverse Patients. *Journal of general internal medicine*, 31(12),

1417–1426. <https://doi.org/10.1007/s11606-016-3771-6>

Samkoff, J. S., & Jacques, C. H. (1991). A review of studies concerning effects of

sleep deprivation and fatigue on residents' performance. *Academic medicine :*

journal of the Association of American Medical Colleges, 66(11), 687–693.

<https://doi.org/10.1097/00001888-199111000-00013>

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika*, 6(5), 309–316.

Schmidtke, A., Fleckenstein, P., Moises, W., & Beckmann, H. (1988). Untersuchungen

zur Reliabilität und Validität einer deutschen Version der Montgomery-Asberg

Depression- Rating Scale (MADRS) [Studies of the reliability and validity of the

German version of the Montgomery-Asberg Depression Rating Scale

(MADRS)]. *Schweizer Archiv fur Neurologie und Psychiatrie (Zurich,*

Switzerland : 1985), 139(2), 51–65.

Semkovska, M., Quinlivan, L., O'Grady, T., Johnson, R., Collins, A., O'Connor, J.,

Knittle, H., Ahern, E., & Gload, T. (2019). Cognitive function following a major

depressive episode: a systematic review and meta-analysis. *The lancet.*

Psychiatry, 6(10), 851–861. [https://doi.org/10.1016/S2215-0366\(19\)30291-3](https://doi.org/10.1016/S2215-0366(19)30291-3)

Simon, G. E., & Perlis, R. H. (2010). Personalized medicine for depression: can we

match patients with treatments?. *The American journal of psychiatry*, 167(12),

1445–1455. <https://doi.org/10.1176/appi.ajp.2010.09111680>

- Sheehan, D., Lecrubier, Y., Sheehan, K., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, J., & Dunbar, G. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of clinical psychiatry*, 59 Suppl 20, 22- 57.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K. H., Janavs, J., Weiller, E., Keskiner, A., Schinka, J., Knapp, E., Sheehan, M. F., & Dunbar, G. C. (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*, 12(5), 232–241. [https://doi.org/10.1016/S0924-9338\(97\)83297-X](https://doi.org/10.1016/S0924-9338(97)83297-X)
- Shimizu, Y., Kitagawa, N., Mitsui, N., Fujii, Y., Toyomaki, A., Hashimoto, N., Kako, Y., Tanaka, T., Asakura, S., & Kusumi, I. (2013). Neurocognitive impairments and quality of life in unemployed patients with remitted major depressive disorder. *Psychiatry research*, 210(3), 913–918. <https://doi.org/10.1016/j.psychres.2013.08.030>
- Shrout, P. E., Spitzer, R. L., & Fleiss, J. L. (1987). Quantification of agreement in psychiatric diagnosis revisited. *Archives of general psychiatry*, 44(2), 172–177. <https://doi.org/10.1001/archpsyc.1987.01800140084013>
- Snyder H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychological bulletin*, 139(1), 81–132. <https://doi.org/10.1037/a0028727>
- Sobin, C., & Sackeim, H. A. (1997). Psychomotor symptoms of depression. *The*

American journal of psychiatry, 154(1), 4–17.

<https://doi.org/10.1176/ajp.154.1.4>

Soron, T. R. (2017). Validation of Bangla Montgomery Asberg Depression Rating Scale (MADRSB). *Asian journal of psychiatry*, 28, 41–46.

<https://doi.org/10.1016/j.ajp.2017.03.019>

Speck, C. E., Kukull, W. A., Brenner, D. E., Bowen, J. D., McCormick, W. C., Teri, L., Pfanschmidt, M. L., Thompson, J. D., & Larson, E. B. (1995). History of depression as a risk factor for Alzheimer's disease. *Epidemiology (Cambridge, Mass.)*, 6(4), 366–369.

<https://doi.org/10.1097/00001648-199507000-00006>

Staples, P., Torous, J., Barnett, I., Carlson, K., Sandoval, L., Keshavan, M., & Onnela, J. P. (2017). A comparison of passive and active estimates of sleep in a cohort with schizophrenia. *NPJ schizophrenia*, 3(1), 37.

<https://doi.org/10.1038/s41537-017-0038-0>

Stewart, R., Besset, A., Bebbington, P., Brugha, T., Lindesay, J., Jenkins, R., Singleton, N., & Meltzer, H. (2006). Insomnia comorbidity and impact and hypnotic use by age group in a national survey population aged 16 to 74 years. *Sleep*, 29(11), 1391–1397.

<https://doi.org/10.1093/sleep/29.11.1391>

Stewart, W. F., Ricci, J. A., Chee, E., Morganstein, D., & Lipton, R. (2003). Lost productive time and cost due to common pain conditions in the US workforce. *JAMA*, 290(18), 2443–2454.

<https://doi.org/10.1001/jama.290.18.2443>

- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press.
- Stringer, G., Couth, S., Brown, L., Montaldi, D., Gledson, A., Mellor, J., Sutcliffe, A., Sawyer, P., Keane, J., Bull, C., Zeng, X., Rayson, P., & Leroi, I. (2018). Can you detect early dementia from an email? A proof of principle study of daily computer use to detect cognitive and functional decline. *International journal of geriatric psychiatry*, 33(7), 867–874. <https://doi.org/10.1002/gps.4863>
- Stuss, D. T., Stethem, L. L., & Poirier, C. A. (1987). Comparison of three tests of attention and rapid information processing across six age groups. *Clinical Neuropsychologist*, 1(2), 139–152. <https://doi.org/10.1080/13854048708520046>
- Talarowska, M., Krzysztof, Z., Mossakowska-Wójcik, J., & Gałecki, P. (2012). Wykonanie Testu Łaczenia Punktów przez osoby z depresją i organicznymi zaburzeniami depresyjnymi [Results of the Trail Making Test among patients suffering from depressive disorders and organic depressive disorders]. *Psychiatria polska*, 46(2), 273–282.
- Thomas, A. J., Gallagher, P., Robinson, L. J., Porter, R. J., Young, A. H., Ferrier, I. N., O'Brien, J. T. (2008). A comparison of neurocognitive impairment in younger and older adults with major depression. *Psychological Medicine*, 39(5), 725–733. <https://doi.org/10.1017/s0033291708004042>
- Thomas, C. M., & Morris, S. (2003). Cost of depression among adults in England in

2000. *The British journal of psychiatry : the journal of mental science*, 183, 514–519. <https://doi.org/10.1192/bjp.183.6.514>
- Tombaugh T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 19(2), 203–214. [https://doi.org/10.1016/S0887-6177\(03\)00039-8](https://doi.org/10.1016/S0887-6177(03)00039-8)
- Torous, J. B., Chan, S. R., Yellowlees, P. M., & Boland, R. (2016). To Use or Not? Evaluating ASPECTS of Smartphone Apps and Mobile Technology for Clinical Care in Psychiatry. *The Journal of clinical psychiatry*, 77(6), e734–e738. <https://doi.org/10.4088/JCP.15com10619>
- Torous, J., Friedman, R., & Keshavan, M. (2014). Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth and uHealth*, 2(1), e2. <https://doi.org/10.2196/mhealth.2994>
- Torous, J., & Powell, A. C. (2015). Current research and trends in the use of smartphone applications for mood disorders. *Internet Interventions*, 2(2), 169–173. <https://doi.org/10.1016/j.invent.2015.03.002>
- Torous, J., & Vaidyam, A. (2020). Multiple uses of app instead of using multiple apps - a case for rethinking the digital health technology toolbox. *Epidemiology and psychiatric sciences*, 29, e100. <https://doi.org/10.1017/S2045796020000013>
- Torous, J., Wisniewski, H., Bird, B., Carpenter, E., David, G., Elejalde, E., Fulford, D., Guimond, S., Hays, R., Henson, P., Hoffman, L., Lim, C., Menon, M., Noel, V., Pearson, J., Peterson, R., Susheela, A., Troy, H., Vaidyam, A.,

- Weizenbaum, E., Naslund, J.A., Keshavan, M. (2019). Creating a Digital Health Smartphone App and Digital Phenotyping Platform for Mental Health and Diverse Healthcare Needs: An Interdisciplinary and Collaborative Approach. *Journal of Technology in Behavioral Science*, 4(2), 73–85.
<https://doi.org/10.1007/s41347-019-00095-w>
- Trivedi J. K. (2006). Cognitive deficits in psychiatric disorders: Current status. *Indian journal of psychiatry*, 48(1), 10–20. <https://doi.org/10.4103/0019-5545.31613>
- Trivedi, M. H., & Greer, T. L. (2014). Cognitive dysfunction in unipolar depression: implications for treatment. *Journal of affective disorders*, 152-154, 19–27.
<https://doi.org/10.1016/j.jad.2013.09.012>
- Tsourtos, G., Thompson, J. C., & Stough, C. (2002). Evidence of an early information processing speed deficit in unipolar major depression. *Psychological medicine*, 32(2), 259–265.
<https://doi.org/10.1017/s0033291701005001>
- Urban, E. J., Charles, S. T., Levine, L. J., & Almeida, D. M. (2018). Depression history and memory bias for specific daily emotions. *PloS one*, 13(9), e0203574. <https://doi.org/10.1371/journal.pone.0203574>
- van Heuvelen, M. J., Hochstenbach, J. B., Brouwer, W. H., de Greef, M. H., Zijlstra, G. A., van Jaarsveld, E., Kempen, G. I., van Sonderen, E., Ormel, J., & Mulder, T. (2005). Differences between participants and non-participants in an RCT on physical activity and psychological interventions for older persons. *Aging Clinical and Experimental Research*, 17(3), 236–245. <https://doi.org/10.1007/bf03324603>

- Wang, Q., Xu, R., & Volkow, N. D. (2021). Increased risk of COVID-19 infection and mortality in people with mental disorders: analysis from electronic health records in the United States. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, 20(1), 124–130. <https://doi.org/10.1002/wps.20806>
- Wagner, S., Helmreich, I., Dahmen, N., Lieb, K., & Tadic, A. (2011). Reliability of three alternate forms of the trail making tests a and B. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 26(4), 314–321. <https://doi.org/10.1093/arclin/acr024>
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, 19(1), 231–240. <https://doi.org/10.1519/15184.1>
- Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., Charlson, F. J., Norman, R. E., Flaxman, A. D., Johns, N., Burstein, R., Murray, C. J., & Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet (London, England)*, 382(9904), 1575–1586. [https://doi.org/10.1016/S0140-6736\(13\)61611-6](https://doi.org/10.1016/S0140-6736(13)61611-6)
- Wilhelm, K., Mitchell, P., Slade, T., Brownhill, S., & Andrews, G. (2003). Prevalence and correlates of DSM-IV major depression in an Australian national survey. *Journal of affective disorders*, 75(2), 155–162. [https://doi.org/10.1016/s0165-0327\(02\)00040-x](https://doi.org/10.1016/s0165-0327(02)00040-x)
- Williams, J. B., & Kobak, K. A. (2008). Development and reliability of a structured

- interview guide for the Montgomery Asberg Depression Rating Scale (SIGMA). *The British journal of psychiatry : the journal of mental science*, 192(1), 52–58. <https://doi.org/10.1192/bjp.bp.106.032532>
- Wimmer, F., Hoffmann, R. F., Bonato, R. A., & Moffitt, A. R. (1992). The effects of sleep deprivation on divergent thinking and attention processes. *Journal of sleep research*, 1(4), 223–230. <https://doi.org/10.1111/j.1365-2869.1992.tb00043.x>
- Wisniewski, H., Henson, P., & Torous, J. (2019). Using a Smartphone App to Identify Clinically Relevant Behavior Trends via Symptom Report, Cognition Scores, and Exercise Levels: A Case Series. *Frontiers in psychiatry*, 10, 652. <https://doi.org/10.3389/fpsyt.2019.00652>
- Withall, A., Harris, L. M., & Cumming, S. R. (2009). The relationship between cognitive function and clinical and functional outcomes in major depressive disorder. *Psychological medicine*, 39(3), 393–402. <https://doi.org/10.1017/S0033291708003620>
- World Health Organization. (1992). *The ICD-10 classification of mental and behavioral disorders: clinical descriptions and diagnostic guidelines*. World Health Organization. Available at <http://www.who.int/iris/handle/10665/37958>
- Woo, Y. S., Rosenblat, J. D., Kakar, R., Bahk, W. M., & McIntyre, R. S. (2016). Cognitive Deficits as a Mediator of Poor Occupational Function in Remitted Major Depressive Disorder Patients. *Clinical psychopharmacology and neuroscience : the official scientific journal of the Korean College of*

Neuropsychopharmacology, 14(1), 1–16.

<https://doi.org/10.9758/cpn.2016.14.1.1>

Woods, D. L., Wyma, J. M., Herron, T. J., & Yund, E. W. (2015). The Effects of Aging, Malingering, and Traumatic Brain Injury on Computerized Trail-Making Test Performance. *PloS one*, 10(6), e0124345.

<https://doi.org/10.1371/journal.pone.0124345>

Wu, C. R., Chen, P. Y., Hsieh, S. H., Huang, H. C., Chen, Y. T., Chen, T. J., & Chiu, H. Y. (2019). Sleep Mediates the Relationship Between Depression and Cognitive Impairment in Older Men. *American journal of men's health*, 13(1), 1557988319825765. <https://doi.org/10.1177/1557988319825765>

Yaffe, K., Falvey, C. M., & Hoang, T. (2014). Connections between sleep and cognition in older adults. *The Lancet. Neurology*, 13(10), 1017–1028.

[https://doi.org/10.1016/S1474-4422\(14\)70172-3](https://doi.org/10.1016/S1474-4422(14)70172-3)

Yates, W. R., Mitchell, J., John Rush, A., Trivedi, M., Wisniewski, S. R., Warden, D., Bryan, C., Fava, M., Husain, M. M., & Gaynes, B. N. (2007). Clinical features of depression in outpatients with and without co-occurring general medical conditions in STAR*D: confirmatory analysis. *Primary care companion to the Journal of clinical psychiatry*, 9(1), 7–15.

<https://doi.org/10.4088/pcc.v09n0102>

Appendix A: Informed Consent Form**Information & Consent Form to Participate in a Research Study**

(Patient participant version)

Protocol Title: An Observational Study of Digital Technology for Monitoring Cognition

Principal Investigator: Synthia Guimond, PhD, Scientist, The Royal's Institute of Mental Health Research

Phone: 613.722.6521 ext. 6586, Email:

Synthia.Guimond@theroyal.ca

Co-Investigators: John Torous, MD, Instructor in Psychiatry, Beth Israel Deaconess Medical Center (BIDMC), Harvard Medical School, USA, Email: jtorous@gmail.com

Cecelia Shvetz, Research Assistant, The Royal's Institute of Mental Health Research

Phone: 613.722.6521 ext. 6446, Email:

Cecelia.Shvetz@theroyal.ca

Study Coordinator: Feng Gu, MA, Research Assistant, The Royal's Institute of Mental Health Research

Phone: 613.722.6521 ext. 6599, Email: Feng.Gu@theroyal.ca

Sponsor: Dr. John Torous, BIDMC, Harvard Medical School
The Royal's IMHR eRIMh funding

Introduction

Before agreeing to take part in this project, it is important that you read the information in this research consent form. It includes details we think you need to know in order to decide if you wish to take part in the study. If you have any questions, ask one of the study researchers. You should not sign this form until you are sure you understand the information. A research assistant will always be available in the laboratory should you need any help.

All research is voluntary. You are free to withdraw from the study at any time. Withdrawing from this study will not affect your relationship with any members of the research team or any care providers at Royal Ottawa Mental Health Centre (ROHCG) now or in the future.

- You can discuss the study with your family doctor, a family member or close friends. If you decide to participate, it is important that you are completely truthful about your health history and any medications that you are taking.

Purpose of the Research

The purpose of this study is to learn if data collected through a smartphone application (app) from your smartphone may be helpful in understanding more about difficulties in thinking skills present in depression and schizophrenia. This study involves downloading an app to your smartphone. This app was developed by researchers at BIDMC and is used for research only. The app collects information in two different ways:

- The app will automatically collect location data stored in your phone if you decide to enable this function. The app will collect the location information gathered by your phone's location sensor/global positioning (GPS) sensors.
- The app will also collect data by asking you to complete cognitive tests and surveys (about your sleep quality and mood). This is called active data collection, and you can decide whether you want to consent to this data collection.

We plan to use the data we collect to study whether this information is useful in helping us better understand, predict and monitor your thinking skills. For example, does sleeping better lead to changes in your thinking skills the next day?

Description of the Research

If you qualify to take part in this research study, you will undergo these research procedures:

- At the lab visit, the research assistant will help you download the app to your smartphone and show you how the app works. We will ask you to complete all the pop-up tests during this visit to get a baseline and to ensure you understand the instructions. You will also be asked to answer questions about your mood, and perform a pen-and-paper test assessing your thinking skills.
- The visit will take about 2 hours 30 minutes. We will create a unique research ID, which will have no identifiable information as your login ID. This app will be password-protected using a password of your choice, so no one else will be able to access the app.
- You will be invited to keep the app on your phone for three months. If you enable the GPS recording, the app will be collecting the location data even if you are not using the phone.
- Once a week (on either Mondays, Wednesdays, or Fridays), the app will ask you to perform some brief thinking skills tests and take surveys that involve tapping on the screen of your phone. Completing the tests and surveys should last about 10 minutes.

- There are a number of thinking skills tests on the smartphone App, but you will only be required to perform 2 of them. Please do not complete other thinking skills tests. If you do so, a record will be kept and data will be saved. However, the data will not be used for analyses.
- It is important to know that none of this survey data or any app data will be monitored in real time.
- You are free to not take any tests or surveys when they pop up on your phone.
- You are free to disable the GPS recording at any time.
- The app will record data from the cognitive tests, surveys and GPS (if the GPS recording is enabled), and will not access any other information on your phone.
- All the data collected is immediately encrypted on the phone as soon as it is collected. When the smartphone is connected to wireless internet, all the de-identified and anonymized data will be encrypted again and sent to Amazon's secure servers for healthcare data. All these data are encrypted and totally anonymous.
- Whenever the app sends data off your phone, it deletes that data from your phone. If there is not a wifi connection, data will temporarily be stored on your phone in an encrypted manner.
- Your encrypted and anonymous data will also be downloaded to a ROHCG passport-protected computer in order to perform our study. No one will be monitoring this data during its actual collection.
- If you drop out of the study early or choose to withdraw at any time, you can delete the app and it will stop collecting any data. You can also delete all the data that have been collected from the app at any time by selecting “delete my data”.

This is a 90-day study in which you will first complete a series of questionnaires about your clinical symptoms and thinking skills in the laboratory, then use a smartphone app to monitor your mood, sleep quality, and thinking skills for 90 days. This study is not a treatment study and is only observing your activities, moods, sleep quality, and thinking skills – thus your treatment will not be affected. The study app collect two types of data that are outlined in detail below.

Passive Data Collection

The app will automatically collect location data stored in your phone if you decide to enable this function. The app will collect the location information gathered by your phone's location sensor/global positioning (GPS) sensors. GPS stands for Global

Positioning System and allows the phone to know where you are located and stores this data. GPS is usually accurate to within 10 feet, which means this app will be collecting data on where you are, although your personal location will not be used as data in the study. This app will collect GPS data constantly and send the coordinate to the server. The coordinate will then be converted to general information about where you are into broad summary metrics such as how far you have travelled in a single day or how many unique locations you have visited. This data is anonymized and will not be monitored in real time. You can disable the GPS recording at any time.

Active Data Collection

You can decide if you want to complete the weekly surveys and tests, which actively collect information from you. The app will offer you surveys and tests about your thinking skills, sleep, and mood. These surveys will pop up as reminders on your phone. You are free to either take these surveys and tests or choose to ignore them and not take them. The surveys will always pop up on a pre-set schedule regardless if you took or ignored the last survey.

These cognitive tests and surveys should take less than 10 minutes to complete.

All Data Collected

You are free to stop collecting data at any time by deleting the app from your phone. If you drop out of the study early or choose to withdraw at any time, you can delete the app and it will stop collecting any data. You can also delete all the data that has been collected from the app at any time by selecting “delete my data”.

As noted above, all data the app collects are not monitored in real time and no one will be able to respond to that data in real time. If at any time, and for any reason, you need help, then you can push the “help” button on the app. It will give you the opportunity to call the study coordinator (if you need help with the app), call an emergency number, or call any other preprogrammed phone numbers you have entered.

Potential Harms (Injury, Discomforts or Inconvenience)

Taking part in the project is considered to be of low risk.

The risks associated with participating are minor given the observational nature of this study. The topics discussed in the questionnaires and surveys may cause some emotional discomfort. We ask you to inform members of the research team if this is the case. The information gathered from these questionnaires is treated professionally and confidentially.

To prevent the risk of data insecurity you will need to add a PIN to your phone to prevent unwanted access. Security measures to protect privacy threats associated with users' devices include the following measures: password protection, users are automatically logged off after 5 minutes of no activity; any data stored locally is

automatically encrypted using 256-bit encryption based on the user authentication information and cannot be accessed without this information.

You should also never try to take surveys on the mobile application while driving.

You can contact the study coordinator if you need help with the app, and the coordinator may answer your call or call you back later. Moreover, you cannot expect that study staff will review the information you enter in the app in a timely matter. This warning will be reiterated to you when you launch the mobile app along with instructions that you should seek care if you experience a worsening of your condition from a health care professional, emergency room, or helpline. This screen will have single button links to automatically call any of these numbers.

Potential Benefits

There is no direct benefit to you from being in this study. However, your participation may help others in the future as a result of knowledge gained from the research. You will also receive \$30 in compensation for your time for your visit at The Royal's Institute of Mental Health Research.

Protecting Your Health Information

Staff from the Research Ethics Board may look at your research record to check that the study is following the proper laws and guidelines. You may be contacted by staff from the Research Ethics Board to answer questions about your experience in the study. This is done to improve the quality of our research work.

Your personal information, including results from the questionnaires and cognitive assessments will be kept strictly confidential except as required or permitted by law. Any information that would indicate that a child was being harmed or at risk of such harm, would not be kept confidential and instead be disclosed as appropriate to offset that risk.

The data produced from this study will be stored in locked filing cabinets in a secure office at the ROMHC. Electronic data will be stored in password-protected files on secured computers at the ROHCG. A log linking your name to your anonymous research data will be kept separately from the rest of your data. Only members of the research team will have access to the data. After the study completion, the data will be kept for 10 years after the last publication of this study. They will then be destroyed.

You will not be identified in any publication or presentation of this study.

A copy of the signed consent form will be provided to you.

Study Results

You will receive a copy of the final study results if you provide the required information in the signature pages below. Findings from this study are expected to be reported at scientific conferences and to be published in scientific journals.

Participation and Withdrawal

Participation in any research study is entirely voluntary. If you decide to participate in this study you can change your mind without giving a reason, and you may withdraw from the study at any time without penalty.

Research Ethics Board Contact

If you have any questions regarding your rights as a research participant, you may contact:

Research Ethics Board

613-722-6521 ext. 6214 or by e-mail: research.ethics@theroyal.ca

The study protocol and consent form have been reviewed by a committee called the Research Ethics Board at the ROHCG. The Research Ethics Board is a group of scientists, medical staff, people from other backgrounds (including law and ethics) as well as members from the community. The Board is established by the ROHCG to review studies for their scientific and ethical merit. The Board pays special attention to the potential harms and benefits involved in participation to the research participant, as well as the potential benefit to society. The Board is also required to do periodic review of ongoing research studies.

Signature Page

An Observational Study of Digital Technology for Monitoring Cognition

Principal Investigator: Synthia Guimond PhD, Scientist, The Royal's Institute of Mental Health Research

Phone: 613.722.6521 ext. 6586 Email: Synthia.guimond@theroyal.ca

The research project has been explained to me, and my questions have been answered to my satisfaction. I have the right not to participate and the right to withdraw without penalty. The potential harms and benefits (if any) of participating in this research study have been explained to me.

I have been told that I have not waived my legal rights nor released the investigators or involved institutions from their legal and professional responsibilities. I know that I may ask now, or in the future, any questions I have about the study. I have been told that records relating to me will be kept confidential and that no information will be disclosed without my permission unless required by law. I have been given sufficient time to read the above

information.

I also understand that all de-identified and anonymized data can accompany the publication of the findings in peer-reviewed scientific publication to promote open and transparent science. **Never your name or any personal information that could identify you** will be associated with the publication of the data.

I consent to participate to this study. I have been told I will be given a signed copy of this consent form.

- I agree to be contacted by IMHR Researchers to be offered other research opportunities
- I agree that all data collected in this study can be shared with other IMHR Researchers. The Principal Investigator (Dr. Synthia Guimond) will be the main person responsible for the data and its distribution.
- I am a patient receiving care at the Royal Mental Health Centre and I agree that the Principal Investigator can access my medical file to confirm my medication information after my first visit, as well as after 3 months to assess any medication change.

Signature of Participant	Name of Participant	Date

Signature of Person Obtaining Informed Consent	Name of Person Obtaining Informed Consent	Date

---

Principal Investigator Signature

I, Synthia Guimond, am the Principal Investigator responsible for the conduct of this study at the Royal Ottawa Mental Health Center, and I have delegated the explanation of this study to this participant to _____ (**name of person conducting the consent discussion**).

Signature of Investigator	Date

Request to be informed of final study findings

*I would like to receive summaries of the study findings

_____
Initials of Participant

Participant's email (or postal) address: _____

Appendix B: Publications and Conference Presentations

Peer-Reviewed Publications

Thérond, A., Pezzoli, P., Abbas, M., Howard, A., Bowie, C. R., & Guimond, S. (2021).

The Efficacy of Cognitive Remediation in Depression: A Systematic Literature Review and Meta-Analysis. *Journal of affective disorders*, 284, 238–246.

<https://doi.org/10.1016/j.jad.2021.02.009>

Conference Presentations

Thérond, A., Howard, A., Torous, T., Shvetz, C., Gu, Feng., & Guimond, S. (2021, June).

Smartphone-Based Assessments to Monitor Cognitive Functioning in Individuals with Major Depressive Disorder. Poster presented at the Brain Health Conference, online, Canada.

Thérond, A., Howard, A., Torous, T., Shvetz, C., Gu, Feng., & Guimond, S. (2021, June).

Smartphone-Based Assessments to Monitor Cognitive Functioning in Individuals with Major Depressive Disorder. Poster presented at the Canadian Psychological Association, online, Canada.

Thérond, A., Pezzoli, P., Abbas, M., Howard, A., Bowie, C., & Guimond, S. (2020, July).

The Efficacy of Cognitive Remediation Therapy in Depression: A Meta-Analysis. Poster presented at the Canadian Psychological Association, online, Canada.

Thérond, A., Pezzoli, P., Abbas, M., Howard, A., Bowie, C., & Guimond, S. (2020,

May). *The Efficacy of Cognitive Remediation Therapy in Depression: A Meta-Analysis*. Poster accepted at the Cognitive Remediation in Psychiatry Conference, New York City, U.S.A. Cancelled due to COVID-19.