HHS Public Access

Author manuscript

J Int Neuropsychol Soc. Author manuscript; available in PMC 2023 September 02.

Published in final edited form as:

J Int Neuropsychol Soc. 2022 April ; 28(4): 401–413. doi:10.1017/S1355617721000503.

Smartphone-Based Neuropsychological Assessment in Parkinson's Disease: Feasibility, Validity, and Contextually Driven Variability in Cognition

Emma L. Weizenbaum[1], **Daniel Fulford**[1,2], **John Torous**[3], **Emma Pinsky**[1,4], **Vijaya B. Kolachalama**[5,6], **Alice Cronin-Golomb**[1,*]

[1]Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA

[2]Department of Occupational Therapy and Rehabilitation Sciences, Boston University, Boston, MA, USA

[3]Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

[4]Department of Psychology, Bryn Mawr College, Bryn Mawr, PA, USA

[5]Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA

[6]Department of Computer Science, and Faculty of Computing and Data Sciences, Boston University Alzheimer's Disease Center; Boston University, Boston, MA, USA

Abstract

Objectives: The prevalence of neurodegenerative disorders demands methods of accessible assessment that reliably captures cognition in daily life contexts. We investigated the feasibility of smartphone cognitive assessment in people with Parkinson's disease (PD), who may have cognitive impairment in addition to motor-related problems that limit attending in-person clinics. We examined how daily-life factors predicted smartphone cognitive performance and examined the convergent validity of smartphone assessment with traditional neuropsychological tests.

Methods: Twenty-seven nondemented individuals with mild–moderate PD attended one in-lab session and responded to smartphone notifications over 10 days. The smartphone app queried participants 5x/day about their location, mood, alertness, exercise, and medication state and administered mobile games of working memory and executive function.

Results: Response rate to prompts was high, demonstrating feasibility of the approach. Between-subject reliability was high on both cognitive games. Within-subject variability was higher for working memory than executive function. Strong convergent validity was seen between traditional tests and smartphone working memory but not executive function, reflecting the latter's ceiling effects. Participants performed better on mobile working memory tasks when at home and after

[*]Correspondence and reprint requests to: Alice Cronin-Golomb, Ph.D., Department of Psychological and Brain Sciences, Boston University, 900 Commonwealth Ave., 2nd floor, Boston, MA 02215, USA. alicecg@bu.edu.

CONFLICTS OF INTEREST

The authors have nothing to disclose.

recent exercise. Less self-reported daytime sleepiness and lower PD symptom burden predicted a stronger association between later time of day and higher smartphone test performance.

Conclusions: These findings support feasibility and validity of repeat smartphone assessments of cognition and provide preliminary evidence of the effects of context on cognitive variability in PD. Further development of this accessible assessment method could increase sensitivity and specificity regarding daily cognitive dysfunction for PD and other clinical populations.

Keywords

Parkinson's disease; Cognition; Smartphone; Mobile phone; Ecological sampling method; Precision medicine

INTRODUCTION

The rates of neurodegenerative disorders are rising as the population ages, and costs of assessment are a barrier to timely diagnosis and treatment (Hurd, Martorell, Delavande, Mullen, & Langa, 2013). Persons with Parkinson's disease (PD) experience common barriers to assessment found with other neurodegenerative disorders (Bradford, Kunik, Schulz, Williams, & Singh, 2009), but motor impairment is an additional impediment to clinic and lab visits (e.g., slow gait, risk of falls, challenges to driving or using public transit). These limitations result in delayed diagnosis for some individuals long after symptoms emerge (Breen, Evans, Farrell, Brayne, & Barker, 2013) as well as suboptimal treatment (Hogan et al., 2008). Nonetheless, clinics and labs are required for current methods of assessment of PD symptoms, including non-motor symptoms such as cognitive impairment.

Neuropsychological testing to characterize cognitive deficits requires hours of time from persons with PD, who often experience daytime fatigue (Friedman et al., 2007), and current methods of testing capture cognitive performance only at a single time point. This is particularly problematic for persons with PD, which is a disorder marked by fluctuations in symptoms over time including changes in cognition, mood, motivation, and arousal level (Cronin-Golomb, Reynolds, Salazar, & Saint-Hilaire, 2019).

Digital technology has the potential to address limitations (e.g., access, time, and cost) inherent to neuropsychological assessment in clinical populations such as PD (Parsey & Schmitter-Edgecombe, 2013). Emerging digital assessments use computerized versions of the lengthy analog tests that only provide a snapshot of cognition in a single session (Bauer et al., 2012; Rentz et al., 2016); these tests can now be completed at home for tracking cognitive function. An emerging literature indicates that smartphone assessments of cognition in clinical samples show strong convergent validity with their analog traditional neuropsychological assessments (Dagum, 2018; Moore et al., 2020; Pal et al., 2016; Schuster, Mermelstein, & Hedeker, 2016; Sliwinski et al., 2018; Timmers et al., 2014).

Another benefit of remote digital assessment is the ability to easily assess cognitive variability with a higher volume of trials than one-time testing allows. Intraindividual variability is a sensitive marker of function and risk for decline in neurodegenerative

disorders (Christ, Combrinck, & Thomas, 2018; Haynes, Bauermeister, & Bunce, 2017). Specifically in PD, we know that greater intraindividual variability in processing speed, beyond mean changes, is associated with later stages of the disease (de Frias, Dixon, Fisher, & Camicioli, 2007). Repeated testing of cognition and state-based variables via a mobile assessment device, such as a smartphone, allows us to identify changes in disease symptoms, which can subsequently predict more subtle changes in disease state or progression. Furthermore, compared to the invasive and costly nature of biomarkers, within-person variability as measured via intensive longitudinal data could produce a sensitive metric of disease stage and diagnosis without undue burden to individuals. The ubiquity of smartphones allows for evidence-based assessment in the hands of those who may most benefit from it. The potential of this technology is becoming increasingly apparent given the ever-shrinking digital divide between younger and older adults as smartphone ownership among seniors continues to rise (Anderson & Perrin, 2017).

Few studies to date have used ecological momentary assessment (EMA) in PD, which refers to frequent, typically brief assessments delivered in real-world environments (Csikszentmihalyi & Larson, 1987). Our lab demonstrated the feasibility of conducting a smartphone EMA study of mood and subjective sleep quality in PD (Wu & Cronin-Golomb, 2020). Further, a single-case EMA study established proof-of-concept using smartphone technology to track intraindividual fluctuations in PD motor symptoms as they related to mood and then modeled these fluctuations using network analysis; they found that anxiety was prospectively associated with more tremor, and cheerfulness with less tremor (van der Velden, Mulders, Drukker, Kuijf, & Leentjens, 2018). The mPower Study used a brief repeated-measures smart-phone design to assess PD motor symptoms including motor and reaction time over six months (Bot et al., 2016; Lipsmeier et al., 2018). All elements of the assessment related to at least one component of the Unified Parkinson's Disease Rating Scale (UPDRS); e.g., finger tapping (motor speed) was significantly correlated with UPDRS ability to independently dress oneself without difficulty (Lipsmeier et al., 2018). Another study used smartphone test data to develop an independent and valid disease severity score using a supervised and weighted machine learning algorithm. The score was significantly associated with the UPDRS and other standard PD motor tests such as Timed Up and Go (Zhan et al., 2018). A limitation of this latter study was that all assessments occurred within the clinic setting and no measures of neuropsychological function beyond reaction time were included. To date, no studies have examined patterns of neuropsychological testing using smartphone technology in PD, although research has found cognitive correlates in PD (Lo et al., 2019).

The first purpose of the present study was to use a smartphone EMA design to assess executive function and working memory, which are among the earliest cognitive dysfunctions to arise in PD (reviewed in Cronin-Golomb et al., 2019; Dirnberger & Jahanshahi, 2013; Miller, Neargarder, Risi, & Cronin-Golomb, 2013) and are related to a wide range of other symptoms, including anxiety (Reynolds, Hanna, Neargarder, & Cronin-Golomb, 2017), depression and apathy (reviewed in Dirnberger & Jahanshahi, 2013), and impairments in spatial judgment (Salazar, Moon, Neargarder, & Cronin-Golomb, 2019) and gait (Morris et al., 2019). While the cognitive measures we included are similar to tests used in other app-based measures of cognition (Moore, Swendsen, & Depp, 2017), performance

on the app-based tests have generally not been measured alongside performance on the traditional neuropsychological tests (Spatial Span [working memory] and Trail Making Test [executive function]) upon which they are based. A recent smartphone-based study of performance on a Stroop-based task by persons aged 50 years and above with HIV, a disorder that is like PD in that it affects the basal ganglia, showed good feasibility and convergent validity with respect to traditional neuropsychological measures of executive function and working memory (Moore et al., 2020). We predicted a significant association in PD between scores on our traditional neuropsychological tests and overall mean score (across time points) on the respective smartphone cognitive assessments.

An additional objective was to examine how PD symptoms and individual differences relate to design feasibility and to intraindividual variability in cognitive performance. We predicted that the majority of participants, who were nondemented, would exhibit a completion rate similar to that seen in other studies of mobile cognitive assessment, roughly 70%–80% completion of surveys (Moore et al., 2017). We hypothesized that those with lower baseline scores on a global cognitive screening measure, the Montreal Cognitive Assessment (MoCA), would have lower rates of response due to difficulty navigating the smartphone app interface without the assistance of a researcher or a clinician. We expected that those with greater symptom burden (higher UPDRS score) would have lower rates of response due to motor or other impairments (e.g., pain, fatigue).

Because the literature has demonstrated associations between specific contextual variables and within-person fluctuations in cognition (Weizenbaum, Torous, & Fuford, 2020), we developed exploratory hypotheses regarding contextually driven within-person fluctuations. Specifically, we examined whether lower mood and higher anxiety would be associated with worse performance on smartphone tasks, as prior research has shown negative mood (Brose, Schmiedek, Lövdén, & Lindenberger, 2012; Jefferies, Smilek, Eich, & Enns, 2008) and stress (Sliwinski, Smyth, Hofer, & Stawski, 2006) to be associated with poorer momentary performance on attention and working memory tasks, especially in older adults. Similarly, we assessed whether lower motivation related to worse smartphone performance due to poor task engagement, as shown in another EMA study of working memory (Brose et al., 2012). As older adults typically perform best on working memory in the morning (West, Murphy, Armilio, Craik, & Stuss, 2002), we assessed whether time of day would be a significant predictor of performance. Subjective alertness can also be predictive of cognitive performance on attention accuracy tasks (Manly, Lewis, Robertson, Watson, & Datta, 2002); hence, we hypothesized that alertness would be positively associated with performance. Although social activity has been associated with improved cognitive performance in certain domains, including memory and processing speed (Bielak, Mogle, & Sliwinski, 2017), the visual and auditory distraction in these settings may be less conducive to optimal performance, and variable noise in the surrounding environment is associated with worse working memory (Lange, 2005). We examined whether being at home (versus in other environments) would produce a more controlled and familiar sensory environment and hypothesized that it may lead to improved cognitive performance.

We also considered whether smartphone cognitive performance would be related to self-reported on-off periods of medication, as off-periods are associated with an increase in

motor symptoms including tremor as well as nonmotor symptoms such as pain and anxiety (Cheon, Park, Kim, & Kim, 2009). In those without medication-related fluctuations, it was predicted that off-periods would be associated with poorer cognitive performance at that time point. The final of the secondary contextual hypotheses was whether factors such as recent exercise may be associated with improved cognitive performance, as previous work suggests widespread benefit of exercise to cognition in people with PD (Murray et al., 2014; Reynolds et al., 2017).

METHODS

Participants

We initially recruited 30 community-dwelling participants who met the clinical criteria for mild to moderate idiopathic PD, following the United Kingdom Parkinson's Disease Society Brain Bank diagnostic criteria (Hughes, Daniel, Kilford, & Lees, 1992). They were recruited from the Boston University Center for Neurorehabilitation and the Parkinson's Disease and Movement Disorders Center at Boston Medical Center. Study protocols were approved by the Boston University Institutional Review Board, with consent obtained according to the Declaration of Helsinki. All participants were native-English speaking adults who owned a smartphone with iOS or Android operating systems. Exclusion criteria included a score of less than 20 on the Montreal Cognitive Assessment (MoCA) or a motor impairment (e.g., significant tremor or dyskinesia) that prohibited them from regularly using a smartphone. Out of the 30 enrolled, 27 completed the remote smartphone assessments in addition to the in-lab initial assessment. One dropped out due to other time commitments, and the other two experienced technical failures with their phone and/or the study app. All but one of the 27 participants were taking medication to treat PD. We calculated levodopa equivalent daily dosages (LEDD) based on convention (Tomlinson et al., 2010). See Table 1 for descriptive statistics of the final sample.

Study Procedure

After enrollment by phone, participants were emailed a link to online questionnaires to complete before the in-lab assessment. In-lab (Study Day 0), after the MoCA confirmed eligibility, the UPDRS was administered by a trained examiner as well as two brief traditional neuropsychological measures: the WMS-III Spatial Span Test and the Trail Making Test A & B. Participants then downloaded the study app onto their own smartphone. They were guided through a slideshow presentation on how to use the app and practiced completing surveys and games with the researcher present. Immediately following the visit, participants were emailed the slideshow of instructions and the structure of smartphone assessment administration was described to them. On Study Days 1–10, smartphone assessment notifications appeared consistently at 9am, 12pm, 3pm, 6pm, and 9pm. The smartphone assessments consisted of a brief survey and two games (described below) and typically required five minutes to complete. Participants had the opportunity to respond to the assessments within two hours of the initial notification before the assessment window would close until the next time prompt (Table 2). Participants were encouraged to contact researchers if questions arose and were also emailed on Day 2 and Day 6 to check-in and troubleshoot any problems encountered using the study app. Participants were compensated

at the rate of $15/hour for time spent in-lab and $1 for every smartphone assessment, with the opportunity to earn a $15 bonus if they completed at least 80% (40/50) of all assessments.

Study Measures

Online pre-study survey—The following questionnaires were completed online before the lab visit. *Mood and Anxiety Symptoms Questionnaire (MASQ-short)* assesses symptoms of depression and anxiety (Wardenaar et al., 2010). *Positive and Negative Affect Scale* assesses trait positive and negative affect (Watson, Clark, & Tellegen, 1988). *Motivation and Pleasure Scale* assesses motivation and pleasure-seeking behavior (Llerena et al., 2013). *Behavior Rating Inventory of Executive Functioning* assesses self-perceived difficulties in executive functioning (Roth, Gioia, & Isquith, 2005). *Epworth Sleepiness Scale (ESS)* assesses daytime sleepiness (Johns, 1991).

In-lab baseline assessment—The following tests were administered in-lab: *UPDRS*, a standard assessment of PD symptoms including in-person interview and motor examination and Hoehn and Yahr (H&Y) stage rating (Fahn, Elton, & UPDRS program members, 1987); the neuropsychological measures, *Spatial Span* (Wechsler Memory Scale-III, Wechsler, 1997) for visual working memory as the analog of the smartphone Backwards Spatial Span, and *Trail Making Test A & B* (Tombaugh, 2004), for processing speed and complex attention as the analog of to the smartphone Trails-B test.

mindLAMP smartphone assessments—The app-based assessment included questions of context, mood state, and cognition using the Mind Learn-Assess-Manage-Prevent app developed at the Department of Digital Psychiatry of Beth Israel Deaconess Medical Center, Boston (Torous et al., 2019). See screenshots of the app interface in Figure 1a–d. Code for the smartphone assessments is publicly available at: https://github.com/BIDMCDigitalPsychiatry

Location & Social Context Survey: Location & Social Context Survey assessed type of current location and social context (Figure 1a). Location options included home, work/school, someone else's home, public place indoors, public place outdoors, and transportation. Social context survey options included alone, strangers, classmates/coworkers, acquaintances, close friends, family members, and romantic partner. *Mood & Alertness Survey* included five questions relating to mood (happy/sad), anxiety, alertness, and level of motivation to try one's hardest on subsequent cognitive games (Figure 1b).

As noted above, executive function and working memory are among the earliest cognitive domains affected by PD and also are domains affected by contextual variables and for this reason we included the following tests:

Backwards Spatial Span Test: Backwards Spatial Span Test consisted of four increasingly challenging trials requiring one to tap the sequence of boxes that appeared on the screen during that trial but in reverse order (Figure 1c). *Trails-B Test* consisted of four increasingly lengthy trials requiring one to tap numbers and letters on the screen in a sequential and alternating manner (Figure 1d).

Data Analysis

EMA response rate was calculated as the number of completed smartphone assessments divided by the total possible number of assessments for each person (50), and then within-person means were averaged to create a sample response rate mean. To determine whether there was reliable and systematic between-person variability that was greater than within-person variability in participants' cognitive tests scores (i.e., Backwards Spatial Span and Trails-B), a multilevel modeling approach was taken using the software MPlus Version 8.1.7 that uses between- and within-person variance to calculate between-person reliability. The formula for reliability is where Var(BP) is the total variance in scores that is between persons, Var(WP) is the total variance in scores that is within persons, and n is the number of assessments (Sliwinski et al., 2018).

$$BP\ \text{Reliablity} = \frac{Var(\text{BP})}{\left(Var(\text{BP}) + \frac{Var(\text{WP})}{n}\right)}$$

Scores were aggregated across time points to find the sample's average within-person means and standard deviations across the full study period and by time point and day. In addition to whole-sample analyses, within-person means were calculated along with within-person standard deviation to identify whether a person's level of variability in score was associated with the overall performance on the task. Pearson correlations were used to assess potential convergence between within-person smartphone cognitive test means and self-reported executive dysfunction (BRIEF-A). Because of the repeated nature of smartphone tests compared to the one-time BRIEF-A questionnaire, a multilevel model maximum likelihood regression was used with the BRIEF-A as a Level 2 fixed effect predictor of smartphone test performance, which was entered as a Level 1 random effect outcome variable. To assess convergent validity with traditional neuropsychological measures, in-lab test scores were entered as Level 2 fixed effect predictors of Backwards Spatial Span and Trails-B performance (the Level 1 outcome variable) each in their own separate model.

We also conducted multilevel model regressions to assess the extent to which smartphone test performance on the Backwards Spatial Span or Trails-B task related to context and state variables collected at the same time point. In this model, each contextual variable was entered as an independent Level-1 predictor of the Level-1 outcome: Backwards Spatial Span or Trails-B. Survey time point (1–50) and day of study were independently analyzed as Level-1 predictors to determine whether smartphone cognitive test performance improved over the course of the study period as a result of practice and familiarity with the task. A two-level regression model was used to determine whether Level-2 fixed effects (between-person measures) moderated the within-person relations between context and smartphone score. These fixed effects included demographics, PD-specific measures (UPDRS Total Score (Parts I-IV), UPDRS Motor Subscale (Part III), UPDRS tremor item total (items 20 and 21), H&Y stage), baseline cognitive score (MoCA), and online questionnaires of trait affect, anxiety, sleepiness tendency, and motivation.

RESULTS

Association with Sample Characteristics

LEDD was not significantly correlated with any cognitive measures (in-lab or smartphone) (r's < .20, p's > .33). Age was negatively associated with the smartphone Backwards Spatial Span Task (r = −.59, p < .01), and UPDRS total score was negatively associated with the smartphone Trails-B task (r = −.53, p < .01). No other participant characteristics were associated with smartphone or in-lab cognitive measures.

Feasibility

Adherence was 73% (SD = 10%); on average, participants responded to roughly 37 out of the 50 possible assessments over the course of the 10-day study period. Age, disease severity (UPDRS total), and baseline MoCA score were not significantly correlated with response rate to EMA prompts. Responses were captured in a variety of social contexts. Participants reported being alone 44% of the time when completing the surveys and games, followed by 32% with family members. Overall, responses were evenly distributed across the course of the day, with most responses in the afternoon (40%) and evening (37%), which is to be expected as two prompts occurred in the afternoon and two in the evening. Responses were most likely to occur within the physical context of being home (71%), with the remaining 29% relatively evenly distributed across other locations including public places indoors, outdoors, work, someone else's home, and during transit (Figure 2).

Smartphone Cognitive Test Psychometrics

Backward spatial span—Within-person mean accuracy was 85.9% (SD = 5.2) (Table 3). Within-person means were significantly negatively correlated with standard deviation (r = −.70, p < .001), meaning that high accuracy was associated with less intraindividual variability. Between-person reliability was high (.89). Day of the study (1–10) was significantly related to higher accuracy (B = .33, SE = .13, p = .01), indicating a practice effect; however, no significant residual variance indicated no individual differences in this effect. Specifically, when scores were aggregated by day, the greatest amount of variability occurred on Day 1 (mean score = 84.0 [SD = 2.1]), after which SD rapidly narrowed to the study average of 85.9 (SD = 5.2).

Trails-B accuracy—When scores were collapsed across time points within individuals, Trails-B accuracy was 96.2% (SD = 3.3), indicating a ceiling effect with little variability around score from moment to moment within individuals (Table 3). Within-person means were significantly negatively correlated with standard deviation (r = −.93, p < .001), meaning that high accuracy was associated with less intraindividual variability around the mean. Between-person reliability was .87. There was no association between day of study (B = .13, SE = .08, p = .12) and accuracy, nor was there significant residual variance between persons. This indicates the absence of a practice effect, but likely reflects the high and unvaried score accuracy across participants and across the study period.

Smartphone Cognitive Test Convergent Validity

The smartphone Backwards Spatial Span was significantly predicted by performance on the MoCA, WMS-III Spatial Span Test (including total score backward and forward and subtest scores), and the Trail Making Test (time on A & B) (Table 4). By contrast, smartphone Trails-B task accuracy was not predicted by any of the traditional neuropsychological measures, including baseline cognitive screen (MoCA), WMS-III Spatial Span Test, or Trail Making Test A & B Time (Table 4), potentially owing to the limited variability in scores on the smartphone measure. Neither smartphone test score was predicted by the BRIEF-A questionnaire score nor was BRIEF-A (subjective self-report) correlated with any of the traditional in-lab test scores, all from objective task-based measures.

Individual differences as independent predictors of smartphone test scores—Using multilevel regression models of between-person (Level 2) predictors of repeated (Level 1) measurements, we found that higher scores on the UPDRS tremor items ($B = -.77$, $SE = .27$, $p = .005$), UPDRS Part III motor sub-scale ($B = -.25$, $SE = .07$, $p = .001$), and UPDRS Part I-IV total score ($B = -.13$, $SE = .05$, $p = .009$) predicted lower overall smartphone Trails-B accuracy scores, but not Backwards Spatial Span scores. Older age predicted lower overall Backwards Spatial Span scores, but not Trails-B scores. Neither the UPDRS Part I mood-mentation-behavior subscale nor the UPDRS Part II ADL subscale significantly predicted smartphone test scores, although there was a trend for an association between the ADL subscale score and Trails-B accuracy ($B = -.280$, $SE = .149$, $p = .060$) (Table 5).

EMA contextual variables and smartphone test performance—For Backwards Spatial Span, being at home (vs. other location) predicted a higher score at the same time point ($B = 7.19$, $SE = 3.01$, $p = .02$); there was no significant between-person residual variance in this association. Having reported exercising in the last 3 h also predicted higher Backwards Spatial Span score at the same time point ($B = 2.16$, $SE = .78$, $p = .01$); there was significant between person residual variance associated with this effect ($B = 1.26$, $SE \leq .001$, $p \leq .001$), meaning that individual differences existed in the strength of the exercise and Span score association (Table 6). When entered as a Level 2 predictor, a person's mean exercise frequency score (the percent of times a person endorsed having exercised in the past 3 h) predicted the strength of the relation between exercise and spatial span score. Specifically, when those who generally exercised less frequently endorsed having exercised in the last 3 h, Backwards Spatial Span score tended to be higher ($B = -2.42$, $SE = .10$, $p < .001$). The question remained as to whether specific individual differences may predict which participants exhibited variability in their accuracy scores in relation to context variables. Based on exploratory correlations between individual differences and contextual variables, we entered individual differences as Level 2 predictors in multilevel models. No individual differences examined predicted the exercise and spatial span association (Table 7).

For Trails-B, there were no significant associations between contextual EMA variables and accuracy. There was a significant level of between-person variability, however, given the residual variance values in the relation between time of day and Trails-B score ($B = 3.75$, SE

= 1.65, p = .02) and being at home (versus another location) and Trails-B score (B = 57.76, SE = 23.49, p = .01). Despite the lack of sample-level relations between Trails-B score and time-of-day, this association did vary significantly from person to person in the study, as illustrated by the examples in Figure 3: Participant 40011 tended to be more accurate as the day progressed, whereas Participant 40008's accuracy did not differ across the morning, afternoon, and evening.

Individual difference variables that were significantly correlated with Trails-B accuracy were entered as Level 2 or between-person predictors in individual multilevel models of context and performance (Table 7). Those who reported lower baseline daytime sleepiness tended to perform better on Trails-B later in the day. No individual difference variables predicted the strength of a participant's association between Trails-B accuracy and being at home.

DISCUSSION

This study examined the utility and feasibility of repeated smartphone assessment of cognition and context in persons with PD. There is a paucity of data on smartphone cognitive assessment in PD, and while a visual working memory task has been incorporated into other apps used to measure symptoms in PD (Bot et al., 2016), to our knowledge this is the first study to report on remote repeated measurement of working memory and executive functioning, cognitive domains that are affected early in PD.

On both smartphone cognitive measures, higher within-person mean accuracy was associated with lower within-person variability. This indicates that those who were performing more poorly were likely to be less consistent. An implication is that traditional one-time testing may not always be sufficient to capture the extent of variability (the good days and bad days) that people with cognitive impairment experience in day-to-day life. On Backwards Spatial Span, participants' scores increased over time, primarily on the first day of the study. We found that average performance was very high on the smartphone Trails-B task, indicating a ceiling effect. Increasing the difficulty or scoring of this measure would improve psychometrics. Despite this limitation, the measure was found to be reliable in differentiating between participants when taking into account within-person fluctuation of repeated measurements.

Convergent Validity

Performance on the smartphone Backwards Spatial Span task was related to a number of the traditional in-lab neuropsychological measures. Most notably, in-lab WMS-III Spatial Span (including total, forward, and backward scores) strongly predicted an individual's performance on the smartphone task. This provides evidence that this smartphone task likely measures a similar construct as the traditional in-lab neuropsycho-logical measure and is the first to validate it against traditional measures. The present study supports use of smartphone cognitive assessment as a valid adjunct to traditional neuropsychological measures, especially because smartphone tasks can be completed quickly, administered remotely, and repeated across time points with high reliability.

Performance on Trails-B was not related to performance on any of the traditional neuropsychological tests administered in-lab, likely due to ceiling effects. Further, the smartphone Trails-B was a measure of switching accuracy versus speed, whereas completion time, or speed, is the primary metric of performance in the traditional Trail Making Test B. Our Trails-B task accuracy reflects inhibition and planning, components of executive function, not processing speed. Our version measured total time per Trails-B game, but the number of trials within a game varied based on accuracy; hence, response time was not comparable within or between people. Updated time measurement and scores based on speed or screen tap latency could make this task a more sensitive measure of processing speed and executive function. For example, it may be that the latency or response time in between alternating numbers and letters in the task is what differentiates people more than the accuracy of switches, and this kind of sensitive timing ability is what differentiates digital assessment from current analog versions of this task.

When considering the use of smartphone cognitive assessment in clinical populations with motor impairments, such as PD, it is essential to understand relations between disease symptoms and performance on the tasks themselves. In our study, the Trails-B task required quick and spatially precise taps on small on-screen circles. We found that higher UPDRS motor and tremor subscale scores were associated with lower scores on the task itself. If the objective of the assessment is to measure changes in processing speed and motor accuracy as a result of disease, this could be valuable to capture in a smartphone cognitive measure. When the goal is to understand complex attention, switching, and executive function; however, it is clear that a participant's motor impairments may obscure cognitive ability. By contrast, the Backwards Spatial Span task was not associated with PD symptoms as measured by the UPDRS, which signals its utility in measuring cognitive performance via smartphone in those with motor speed impairments.

EMA Contextual Variables and Smartphone Test Performance

An additional aim of this study was to explore the extent to which cognitive performance varied across different contexts. We found that being home rather than in another environment, and endorsing recent exercise, predicted higher Backwards Spatial Span scores. On this measure of working memory, it seems logical that being in a familiar and perhaps less distracting environment such as one's home would be conducive to stronger performance.

In interpreting the direct effect of exercise on visual working memory performance, it is important to note that this is an observational finding only. One possibility is that the physiological effects of exercise could contribute to improved cognitive performance in PD. Indeed, a recent systematic review showed that a variety of exercise programs can lead to improvements in global cognitive function, processing speed, attention, and mental flexibility in people with PD (da Silva et al., 2018). A number of studies point toward increased levels of blood-derived neurotrophic factor as the mechanism of action in the relation between exercise and higher cognitive performance in PD (Hirsch, van Wegen, Newman, & Heyn, 2018). It is also possible that on days when people are feeling better, broadly, they may be more likely to perform better on a working memory task and be

more motivated to exercise. We found significant residual variance in this association indicating heterogeneous individual differences in how much recent exercise associated with Backwards Spatial Span scores. Specifically, those who least frequently endorsed exercise were the ones most likely to have a higher score when they *did* report recent exercise. This finding may further support the correlational interpretation that people with PD who exercise more frequently may be exercising unconditionally on “good” days and “bad” days or despite other factors of variability in their daily lives. By contrast, those who endorse exercising less frequently may have a more conditional relationship with exercise, such that choosing to exercise may be associated with other factors (e.g., less physical discomfort, a less busy day, etc.) that could also be conducive to better working memory.

Whereas there were no direct effects between contextual predictors and Trails-B, there were significant individual differences in these effects; specifically, those with less symptom burden (UPDRS total score) and less daytime sleepiness (ESS score) showed a stronger effect between Trails-B performance and later time of day. While there is reason to be cautious about these findings given the ceiling effect and motor confounds seen with Trails-B, it is worth considering that those with greater disease burden may show *less* sensitivity to varying contexts; alternatively, they may be more homebound and so have less exposure to varying physical and social contexts. One potential future question is how variability in context and one’s response to different contexts relate to stage of a disorder. For example, could a decrease in the variety of contexts or one’s differential response to them as measured by repeated mobile testing help to predict advancing disease state? Alternatively, mobile testing may be most useful for those in early stages of a disease. It may also help identify contexts in which performance is lower and targeted interventions could be of benefit for people across the spectrum of disease severity.

Results are in line with other studies regarding engagement. Participants completed 73% of the EMA prompts on average. This response rate is similar to other studies of mobile cognitive assessment in a variety of clinical samples, which found average study response to be 79% (Moore et al., 2017) and is higher than the 61% response rate in a study of smartphone assessment of PD motor symptoms (Lipsmeier et al., 2018).

Study Limitations

This study examined a sample of individuals with mild to moderate PD severity without dementia and cannot generalize to individuals with more severe motor and cognitive symptoms. It is also important to note that participants in this sample were self-selected and had either previously participated in research or had expressed interest in participating. This type of sample often leads to narrow demographics in terms of socioeconomic status, education, and race, which poses barriers to community-wide generalization. A limitation in our methods was that one of our smartphone cognitive measures elicited very high accuracy and was not able to capture processing speed; nevertheless, the data produced were useful in regard to assessment of the relation between accuracy and performance variability.

Future Directions

In our study design, smartphone assessments were prompted at fixed times during the day. As such, participants could anticipate the exact time of an assessment and may have altered their routine to accommodate test-taking, which could have restricted the range of locations and contexts in which tests were completed. Further investigation is needed to compare performance under conditions of random vs. fixed-time prompts within daytime intervals. A second avenue of investigation would be to intentionally time test administration to coincide with participant-specific activities and routines. Future studies should also aim to measure cognitive domains beyond visual working memory and the inhibition and planning components of executive function, in order to compare results with those presented here; investigators wishing to use our tests may do so through this link https://github.com/BIDMCDigitalPsychiatry. We also suggest expanding the scope of studies of smartphone-based assessments of cognition to include measures of activities of daily living (ADLs) beyond the UPDRS ADL subscale used in our study where a trend was seen between this score and smartphone Trails-B accuracy. This line of future investigation is supported by an extensive literature relating performance on traditional neuropsychological tests of executive function to ADLs in older adults with and without cognitive impairment (e.g., Bell-McGinty, Podell, Franzen, Baird, & Williams, 2002; see Overdorp, Kessels, Claassen, & Oosterman, 2016 for a review), including in PD (Higginson, Lanni, Sigvardt, & Disbrow, 2013; Kudlicka, Hindle, Spencer, & Clare, 2018).

Conclusion

This study demonstrated that repeated smartphone assessment of cognitive performance in persons with PD is a feasible and useful method of neuropsychological assessment. Our smartphone measures of executive functioning and visual working memory both displayed strong between-person reliability. There was convergent validity for the smartphone test of visual working memory, with its results predicted by traditional neuropsychological tests administered in-lab. Performance on this task was also unrelated to PD motor impairment and tremor score, which demonstrates feasibility even in those with notable motor-based symptoms of PD. Further, we found that certain contexts and conditions (being at home and recent exercise) differentially predicted performance on this task at that same time point. Together these findings provide support for further investigation of smart-phone cognitive assessment in PD as a means of understanding idiographic patterns of symptoms, with the potential to better predict disease progression and provide targeted and individualized interventions.

ACKNOWLEDGMENTS

Anup Dupaguntla, BA, provided valuable support on data entry and data management. Aditya Vaidyam, MS, and Phillip Henson, MS, assisted greatly with technical support for the mindLAMP and data formatting and compilation. Tairmae Kangarloo, BS, provided assistance with initial data visualization and analysis planning. We are grateful to Terry Ellis, PhD, and Cristina Colon-Semenza, PhD, for referring individuals with PD to this study from the Boston University Center for Neurorehabilitation and to Olivier Barthelemy, PhD, for referring participants from other lab studies. We acknowledge with gratitude the efforts of these colleagues and especially the efforts of all of the study participants.

FUNDING

This work was funded by a Boston University Digital Health Initiative Research Incubation Award (#2018-02-006) with the support of the Institute for Health System Innovation & Policy and the Hariri Institute for Computing and Computational Science & Engineering at Boston University.

REFERENCES

Anderson M & Perrin A (2017). Technology Use among Seniors. Washington, DC: Pew Research Center for Internet & Technology.

Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, & Naugle RI (2012). Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. The Clinical Neuropsychologist, 26(2), 177–196. 10.1080/13854046.2012.663001 [PubMed: 22394228]

Bell-McGinty S, Podell K, Franzen M, Baird AD, & Williams MJ (2002). Standard measures of executive function in predicting instrumental activities of daily living in older adults. International Journal of Geriatric Psychiatry, 17(9), 828–834. 10.1002/gps.646 [PubMed: 12221656]

Bielak AAM, Mogle J, & Sliwinski MJ (2017). What did you do today? Variability in daily activities is related to variability in daily cognitive performance. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences 10.1093/geronb/gbx145

Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, …Trister AD(2016). The mPower study, Parkinson disease mobile data collected using ResearchKit. Scientific Data, 3(1), 160011. 10.1038/sdata.2016.11 [PubMed: 26938265]

Bradford A, Kunik ME, Schulz P, Williams S, & Singh H (2009). Missed and delayed diagnosis of dementia in primary care prevalence and contributing factors. Alzheimer Disease & Associated Disorders, 23(4), 306–314. [PubMed: 19568149]

Breen D, Evans J, Farrell K, Brayne C, & Barker R (2013). Determinants of delayed diagnosis in Parkinson's disease. Journal of Neurology, 260(8), 1978–1981. 10.1007/s00415-013-6905-3 [PubMed: 23572347]

Brose A, Schmiedek F, Lövdén M, & Lindenberger U (2012). Daily variability in working memory is coupled with negative affect: the role of attention and motivation. Emotion, 12(3), 605–617. 10.1037/a0024436 [PubMed: 21787075]

Cheon S-M, Park MJ, Kim W-J, & Kim JW (2009). Non-motor off symptoms in Parkinson's disease. Journal of Korean Medical Science, 24(2), 311. 10.3346/jkms.2009.24.2.311 [PubMed: 19399276]

Christ BU, Combrinck MI, & Thomas KGF (2018). Both reaction time and accuracy measures of intraindividual variability predict cognitive performance in Alzheimer's disease. Frontiers in Human Neuroscience, 12, 124. 10.3389/fnhum.2018.00124 [PubMed: 29686610]

Cronin-Golomb A, Reynolds GO, Salazar RD, & Saint-Hilaire M-H (2019). Parkinson's disease and Parkinson-plus syndromes. In Alosco ML& Stern RA(Eds.), The Oxford handbook of adult cognitive disorders (pp. 599–630; By Cronin-Golomb A, Reynolds GO, Salazar RD, & Saint-Hilaire M-H. Oxford University Press. 10.1093/oxfordhb/9780190664121.013.28

Csikszentmihalyi M & Larson R (1987). Validity and reliability of the experience-sampling method. Journal of Nervous and Mental Disease, 175(9), 526–536. [PubMed: 3655778]

da Silva FC, da Iop RR, de Oliveira LC, Boll AM, de Alvarenga JGS, Gutierres Filho PJB, … da Silva R (2018). Effects of physical exercise programs on cognitive function in Parkinson's disease patients: a systematic review of randomized controlled trials of the last 10 years. PLOS ONE, 13(2), e0193113. 10.1371/journal.pone.0193113 [PubMed: 29486000]

Dagum P (2018). Digital biomarkers of cognitive function. NPJ Digital Medicine, 1(1), 10. 10.1038/s41746-018-0018-4 [PubMed: 31304295]

de Frias CM, Dixon RA, Fisher N, & Camicioli R (2007). Intraindividual variability in neurocognitive speed: a comparison of Parkinson's disease and normal older adults. Neuropsychologia, 45(11), 2499–2507. 10.1016/j.neuropsychologia.2007.03.022 [PubMed: 17507058]

Dirnberger G & Jahanshahi M (2013). Executive dysfunction in Parkinson's disease: a review. Journal of Neuropsychology, 7(2), 193–224. 10.1111/jnp.12028 [PubMed: 24007368]

Fahn S, Elton R, & UPDRS Program Members. (1987). Unified Parkinsons disease rating scale. In Fahn S, Marsden CD, Goldstein M, & Calne D(Eds.), Recent developments in Parkinsons disease (*Vol.* 2, pp. 153–163). Florham Park, NJ: Macmillan Healthcare Information.

Friedman JH, Brown RG, Comella C, Garber CE, Krupp LB, Lou J-S, … Taylor CB (2007). Fatigue in Parkinson's disease: a review. Movement Disorders, 22(3), 297–308. 10.1002/mds.21240 [PubMed: 17133511]

Haynes BI, Bauermeister S, & Bunce D (2017). A systematic review of longitudinal associations between reaction time intraindividual variability and age-related cognitive decline or impairment, dementia, and mortality. Journal of the International Neuropsychological Society: JINS, 23(5), 431–445. 10.1017/S1355617717000236 [PubMed: 28462758]

Higginson CI, Lanni K, Sigvardt KA, & Disbrow EA (2013). The contribution of trail making to the prediction of performance-based instrumental activities of daily living in Parkinson's disease without dementia. Journal of Clinical and Experimental Neuropsychology, 35(5), 530–539. 10.1080/13803395.2013.798397 [PubMed: 23663116]

Hirsch MA, van Wegen EEH, Newman MA, & Heyn PC (2018). Exercise-induced increase in brain-derived neurotrophic factor in human Parkinson's disease: A systematic review and meta-analysis. Translational Neurodegeneration, 7(1), 7. 10.1186/s40035-018-0112-1 [PubMed: 29568518]

Hogan D, Bailey P, Black S, Carswell A, Chertkow H, Clarke B, … Thorpe L (2008). Diagnosis and treatment of dementia: 5. Nonpharmacologic and pharmacologic therapy for mild to moderate dementia. Canadian Medical Association Journal, 179(10), 1019–1026. 10.1503/cmaj.081103 [PubMed: 18981443]

Hughes AJ, Daniel SE, Kilford L, & Lees AJ (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. Journal of Neurology, Neurosurgery, and Psychiatry, 55(3), 181–184. [PubMed: 1564476]

Hurd MD, Martorell P, Delavande A, Mullen KJ, & Langa KM (2013). Monetary costs of dementia in the United States. The New England Journal of Medicine, 368(14), 1326–1334. 10.1056/NEJMsa1204629 [PubMed: 23550670]

Jefferies LN, Smilek D, Eich E, & Enns JT (2008). Emotional valence and arousal interact in attentional control. Psychological Science, 19(3), 290–295. 10.1111/j.1467-9280.2008.02082.x [PubMed: 18315803]

Johns MW (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep, 14(6), 540–545. 10.1093/sleep/14.6.540 [PubMed: 1798888]

Kudlicka A, Hindle JV, Spencer LE, & Clare L (2018). Everyday functioning of people with Parkinson's disease and impairments in executive function: a qualitative investigation. Disability and Rehabilitation, 40(20), 2351–2363. 10.1080/09638288.2017.1334240 [PubMed: 28597694]

Lange EB (2005). Disruption of attention by irrelevant stimuli in serial recall. Journal of Memory and Language, 53(4), 513–531. 10.1016/j.jml.2005.07.002

Lipsmeier F, Taylor KI, Kilchenmann T, Wolf D, Scotland A, Schjodt-Eriksen J, … Lindemann M (2018). Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial: remote PD testing with smartphones. Movement Disorders, 33(8), 1287–1297. 10.1002/mds.27376 [PubMed: 29701258]

Llerena K, Park SG, McCarthy JM, Couture SM, Bennett ME, & Blanchard JJ (2013). The Motivation and Pleasure Scale–Self-Report (MAP-SR): reliability and validity of a self-report measure of negative symptoms. Comprehensive Psychiatry, 54(5), 568–574. 10.1016/j.comppsych.2012.12.001 [PubMed: 23351831]

Lo C, Arora S, Baig F, Lawton MA, El Mouden C, Barber TR, … Hu MT (2019). Predicting motor, cognitive & functional impairment in Parkinson's. Annals of Clinical and Translational Neurology, acn3.50853. 10.1002/acn3.50853

Manly T, Lewis GH, Robertson IH, Watson PC, & Datta AK (2002). Coffee in the Cornflakes: Time-of-Day as a Modulator of Executive Response Control (*Vol.* 40). 10.1016/S0028-3932(01)00086-0

Miller IN, Neargarder S, Risi MM, & Cronin-Golomb A (2013). Frontal and posterior subtypes of neuropsychological deficit in Parkinson's disease. Behavioral Neuroscience, 127(2), 175–183. 10.1037/a0031357 [PubMed: 23398433]

Moore RC, Campbell LM, Delgadillo JD, Paolillo EW, Sundermann EE, Holden J, … Swendsen J(2020). Smartphone-based measurement of executive function in older adults with and without HIV. Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 35(4), 347–357. 10.1093/arclin/acz084 [PubMed: 31942632]

Moore RC, Swendsen J, & Depp CA (2017). Applications for self-administered mobile cognitive assessments in clinical research: a systematic review. International Journal of Methods in Psychiatric Research, 26(4), e1562. 10.1002/mpr.1562 [PubMed: 28370881]

Morris R, Martini DN, Smulders K, Kelly VE, Zabetian CP, Poston K, … Horak F(2019). Cognitive associations with comprehensive gait and static balance measures in Parkinson's disease.Parkinsonism & Related Disorders, 69, 104–110. 10.1016/j.parkreldis.2019.06.014 [PubMed: 31731260]

Murray DK, Sacheli MA, Eng JJ, & Stoessl AJ (2014). The effects of exercise on cognition in Parkinson's disease: a systematic review. Translational Neurodegeneration, 3(1), 5. 10.1186/2047-9158-3-5 [PubMed: 24559472]

Overdorp EJ, Kessels RPC, Claassen JA, & Oosterman JM (2016). The combined effect of neuropsychological and neuropathological deficits on instrumental activities of daily living in older adults: a systematic review. Neuropsychology Review, 26(1), 92–106. 10.1007/s11065-015-9312-y [PubMed: 26732392]

Pal R, Mendelson J, Clavier O, Baggott MJ, Coyle J, & Galloway GP (2016). Development and testing of a smart-phone-based cognitive/neuropsychological evaluation system for substance abusers. Journal of Psychoactive Drugs, 48(4), 288–294. 10.1080/02791072.2016.1191093 [PubMed: 27260123]

Parsey CM & Schmitter-Edgecombe M (2013). Applications of technology in neuropsychological assessment. The Clinical Neuropsychologist, 27(8), 1328–1361. 10.1080/13854046.2013.834971 [PubMed: 24041037]

Rentz DM, Dekhtyar M, Sherman J, Burnham S, Blacker D, Aghjayan SL, … Sperling RA (2016). The feasibility of at-home iPad cognitive testing for use in clinical trials. The Journal of Prevention of Alzheimer's Disease, 3(1), 8–12. 10.14283/jpad.2015.78

Reynolds GO, Hanna KK, Neargarder S, & Cronin-Golomb A (2017). The relation of anxiety and cognition in Parkinson's disease. Neuropsychology, 31(6), 596–604. 10.1037/neu0000353 [PubMed: 28425730]

Roth R, Gioia G, & Isquith P (2005). BRIEF®-A – Behavior Rating Inventory of Executive Function®—Adult Version.

Salazar RD, Moon K, Neargarder S, & Cronin-Golomb A (2019). Spatial judgment in Parkinson's disease: contributions of attentional and executive dysfunction. Behavioral Neuroscience, 133(4), 350–360. [PubMed: 31294590]

Schuster RM, Mermelstein RJ, & Hedeker D (2016). Ecological momentary assessment of working memory under conditions of simultaneous marijuana and tobacco use. Addiction, 111(8), 1466–1476. 10.1111/add.13342 [PubMed: 26857917]

Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, & Lipton RB (2018). Reliability and validity of ambulatory cognitive assessments. Assessment, 25(1), 14–30. 10.1177/1073191116643164 [PubMed: 27084835]

Sliwinski MJ, Smyth JM, Hofer SM, & Stawski RS (2006). Intraindividual coupling of daily stress and cognition. Psychology and Aging, 21(3), 545–557. (2006-11398-009). 10.1037/0882-7974.21.3.545 [PubMed: 16953716]

Timmers C, Maeghs A, Vestjens M, Bonnemayer C, Hamers H, & Blokland A (2014). Ambulant cognitive assessment using a smartphone.Applied Neuropsychology: Adult, 21(2), 136–142. 10.1080/09084282.2013.778261 [PubMed: 24826507]

Tombaugh TN (2004). Trail making test A and B: normative data stratified by age and education. Archives of Clinical Neuropsychology, 19(2), 203–214. 10.1016/S0887-6177(03)00039-8 [PubMed: 15010086]

Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, & Clarke CE (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Movement Disorders: Official Journal of the Movement Disorder Society, 25(15), 2649–2653. 10.1002/mds.23429 [PubMed: 21069833]

Torous J, Wisniewski H, Bird B, Carpenter E, David G, Eduardo E, … 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, 73–85. 10.1007/s41347-019-00095-w

van der Velden RMJ, Mulders AEP, Drukker M, Kuijf ML, & Leentjens AFG (2018). Network analysis of symptoms in a Parkinson patient using experience sampling data: An n = 1 study: symptom network analysis in Parkinson's disease. Movement Disorders, 33(12), 1938–1944. 10.1002/mds.93 [PubMed: 30288791]

Wardenaar KJ, van Veen T, Giltay EJ, de Beurs E, Penninx BWJH, & Zitman FG (2010). Development and validation of a 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (MASQ). Psychiatry Research, 179(1), 101–106. 10.1016/j.psychres.2009.03.005 [PubMed: 20472297]

Watson D, Clark LA, & Tellegen A (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of Personality and Social Psychology, 54(6), 1063. [PubMed: 3397865]

Wechsler D (1997). Wechsler Memory Scale–Third Edition. San Antonio, TX: The Psychological Corporation.

Weizenbaum E, Torous J, & Fuford D (2020). Cognition in context: understanding the everyday predictors of cognitive performance in a new era of measurement. JMIR Mhealth Uhealth, 8(7), e14328. 10.2196/14328 [PubMed: 32706680]

West R, Murphy KJ, Armilio ML, Craik FIM, & Stuss DT (2002). Effects of time of day on age differences in working memory. The Journals of Gerontology: Series B, 57(1), P10. 10.1093/geronb/57.1.P3

Wu JQ & Cronin-Golomb A (2020). Temporal associations between sleep and daytime functioning in Parkinson's disease: a smartphone-based ecological momentary assessment. Behavioral Sleep Medicine, 18(4), 560–569. 10.1080/15402002.2019.1629445 [PubMed: 31199163]

Zhan A, Mohan S, Tarolli C, Schneider RB, Adams JL, Sharma S, … Saria S (2018). Using smartphones and machine learning to quantify Parkinson disease severity: the mobile Parkinson disease score. JAMA Neurology, 75(7), 876. 10.1001/jamaneurol.2018.0809 [PubMed: 29582075]

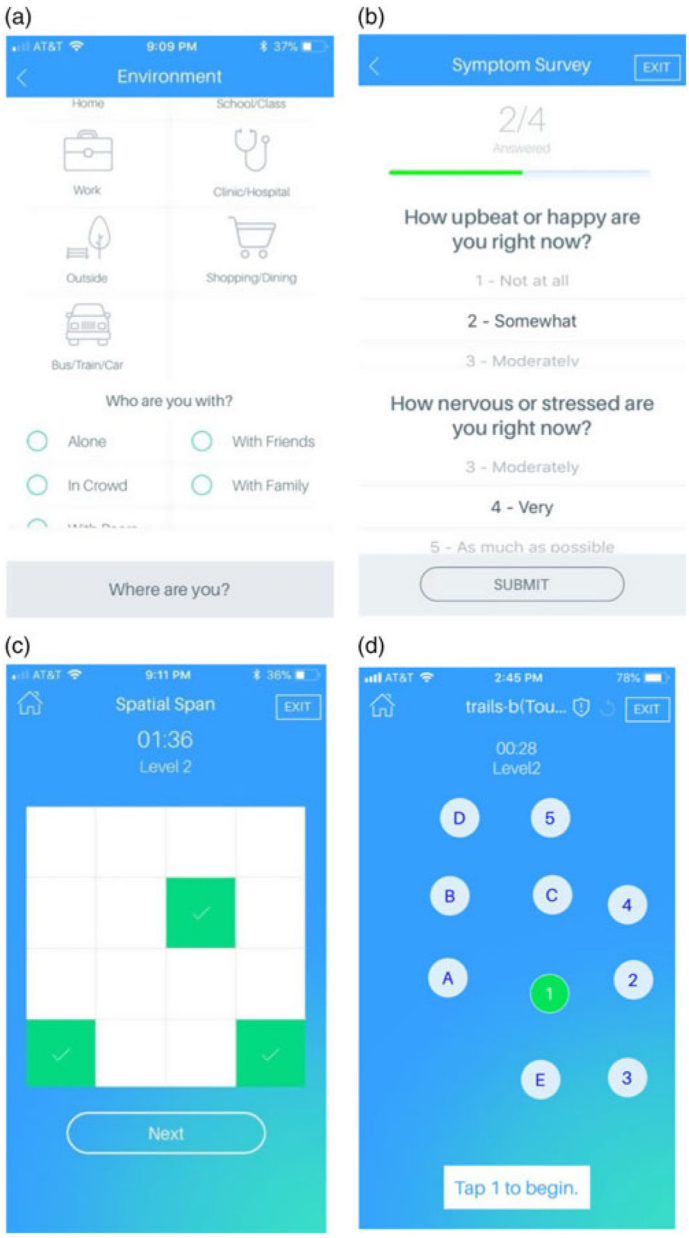

Fig. 1.

mindLAMP App Assessments. (a) Location & Social Context Survey: Indication of type of current location and social context. (b) Mood & Alertness Survey: five questions including rating happiness, sadness, anxiety, alertness, and level of motivation to try one's hardest on subsequent cognitive games. (c) Backwards Spatial Span Test: four increasingly challenging trials of a task requiring one to tap the sequences of boxes that appeared on the screen during that trial but in reverse order. (d) Trails-B Test: four increasingly lengthy trials of a task requiring one to tap numbers and letters on the screen in a sequential and alternating manner.

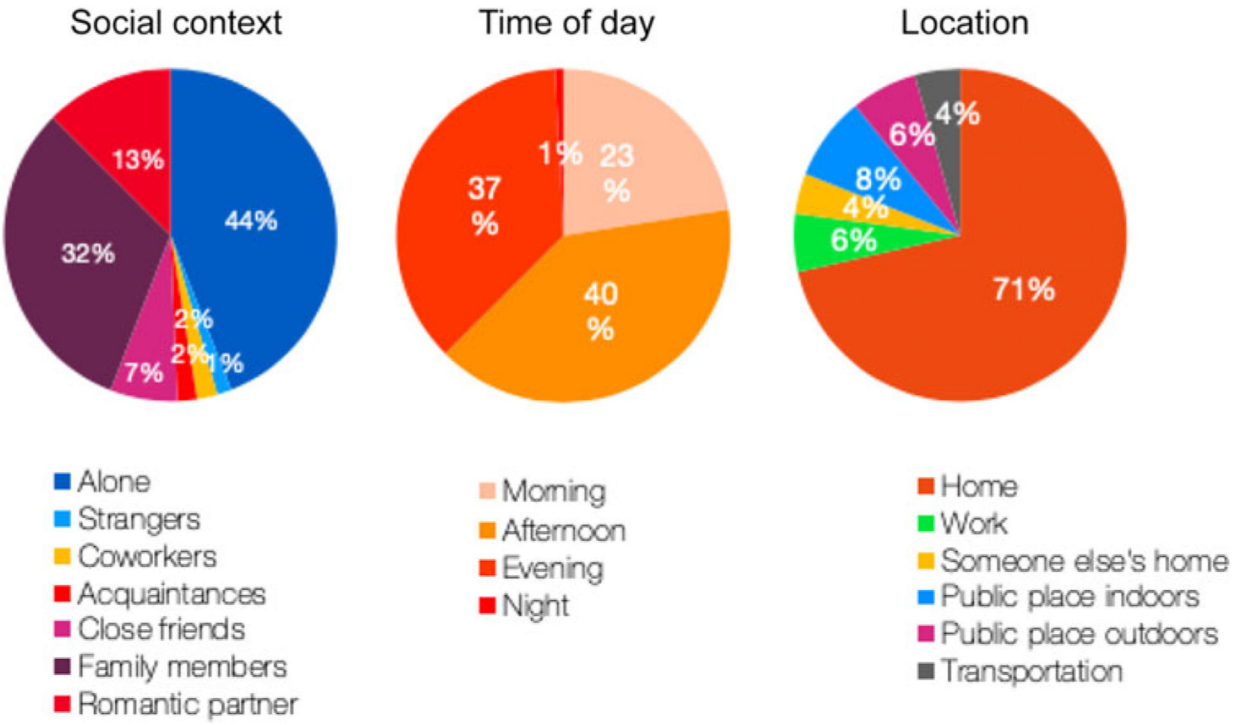

Fig. 2.
Summary of Smartphone Survey Responses.

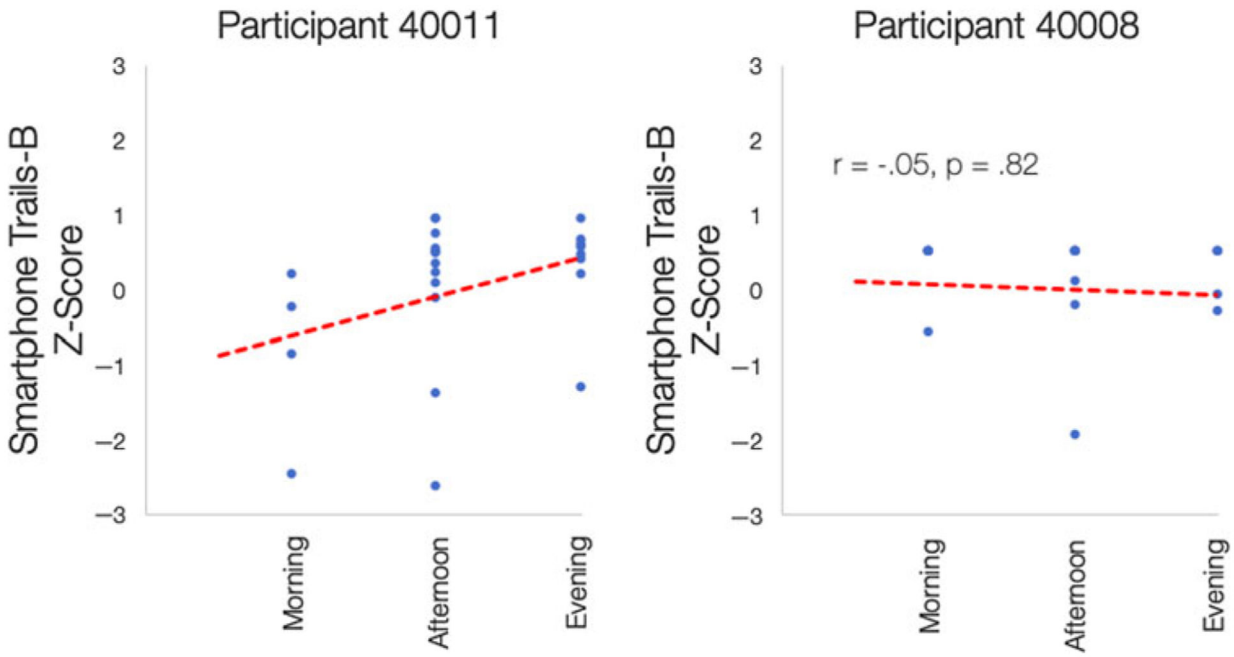

Fig. 3.

Examples of Between-Person Variability in the Association Between Time of Day and Smartphone Trails-B Accuracy.

Table 1.

Participant characteristics (n = 27)

Age (years) Mean (*SD*)	Men: Women	Education (years) Mean (*SD*)	Race & Ethnicity	MoCA Score Mean (*SD*)	PD Stage Hoehn & Yahr Median (Range)	UPDRS Total Score (*SD*)	LEDD Mean mg/day (*SD*)
63.2 (8.7)	14 Men 13 Women	17.7 (3.2)	26 White 1 Middle Eastern 1 Native American	27.7 (2.5)	2.0 (1.0–3.0)	27.0 (11.3)	420 (244)

MoCA, Montreal Cognitive Assessment; UPDRS, Unified Parkinson's Disease Rating Scale (disease severity); LEDD, levodopa equivalent daily dose.

J Int Neuropsychol Soc. Author manuscript; available in PMC 2023 September 02.

Table 2.

Study design

Study Component	Tasks	Timeline	Time Required
Individual difference questionnaires	Online self-report questionnaires related to trait mood, sleep, and executive function	Following eligibility phone screening, before in-lab assessment	30 min
In-lab assessment	UPDRS, MoCA, Trail Making Test, WMS-III Spatial Span; download study app & practice smartphone tasks	Study Day 0	90 min
Remote smartphone assessments	Brief survey of context, mood, alertness, motivation, caffeine, recent exercise, and medication ON-OFF state; Trails-B task (2 min) and Backwards Spatial Span task (2 min)	Study Day 1–10; prompted every day at 9:00a, 12:00p, 3:00p, 6:00p, 9:00p	5 min, 5x/day, 10 days

Table 3.

Smartphone cognitive test psychometrics

	Within-Person Mean Accuracy	Correlation between Within-Person Mean and *SD*	Between-Person Reliability Value
Backwards Spatial Span	85.9 % (*SD* = 5.2)	r = −.70 (*p* < .001)	.885
Trails-B	96.2% (*SD* = 3.3)	r = −.93 (*p* < .001)	.874

Table 4.

Convergent validity between smartphone and traditional analog cognitive tests

	Smartphone Backwards Spatial Span			**Smartphone Trails-B**		
	Unstd. Beta Coefficient	**SE**	***p*-value**	**Unstd. Beta Coefficient**	**SE**	***p*-value**
MoCA Total	***B* = .757**	**.361**	**<.001**	*B* = .128	.272	.637
Trail Making Test B time	***B* = −.146**	**.027**	**<.001**	*B* = −.035	.027	.186
WMS-III Spatial Span Total	*B* = 1.147	.266	<.001	*B* = .410	.223	.066
BRIEF-A Total	*B* = −1.281	3.465	.712	*B* = −3.411	2.359	.148

Table 5.

Multilevel model results of PD-specific individual differences and smartphone cognitive test scores

	Smartphone Trails-B			Smartphone Backwards Spatial Span		
	Unstd. Beta Coefficient	**SE**	***p*-value**	**Unstd. Beta Coefficient**	**SE**	***p*-value**
H&Y Score	−.333	1.298	.797	2.41	1.81	.183
UPDRS Total	**−.132**	**.051**	**.009**	−.025	.084	.767
UPDRS Tremor	**−.768**	**.271**	**.005**	.122	.456	.789
UPDRS MMB	.225	.312	.471	.126	.457	.784
UPDRS ADL	−.28	.149	.06	−.301	.228	.186
UPDRS Motor	**−.246**	**.074**	**.001**	.026	.129	.841
Age	−.096	.075	.202	**−.349**	**.089**	**<.001**

J Int Neuropsychol Soc. Author manuscript; available in PMC 2023 September 02.

Table 6.

Contextual predictors of smartphone score as individual multilevel models

Backwards Spatial Span					
	Direct Effect			BP Residual	
Level 1 Predictor	**Unstd. Beta**	**SE**	***p*-value**	**Unstd. Beta (SE)**	***p*-value**
Day	**.329**	**.129**	**.011**	.080(.151)	.597
Time point	**.042**	**.016**	**.008**	.001(.002)	.595
Time of Day	.683	.43	.112	.176(.061)	.951
Home	**7.188**	**3.012**	**.017**	3.421(5.951)	.565
Alone	.778	2.069	.715	.368(8.413)	.965
Upbeat	−.367	2.804	.896	.391(1.391)	.779
Nervous	−2.017	2.435	.408	.950(.782)	.224
Sad	1.651	9.24	.858	.629(7.473)	.933
Alert	1.066	1.045	.308	.262(1.023)	.798
Motivation	−8.075	13.71	.556	1.522(11.482)	.895
ON-Med Period	6.538	6.046	.286	8.836(18.118)	.626
Caffeine	.857	7.806	.940	1.544(7.056)	.827
Exercise	**2.157**	**.778**	**.006**	**1.262(<.001)**	**<.001**
Trails-B					
Day	.125	.08	.117	.005(.038)	.893
Time point	.016	.01	.121	<.000 (.001)	.607
Time of Day	.279	.474	.556	**3.753 (1.646)**	**.023**
Home	1.259	2.114	.552	**57.761 (23.494)**	**.014**
Alone	.946	1.274	.458	.178(1.057)	.866
Upbeat	−1.294	2.862	.651	4.762(2.636)	.071
Nervous	.121	2.303	.958	3.863 (3.144)	.219
Sad	2.333	2.887	.419	.147(.366)	.688
Alert	.568	.756	.452	.756(.401)	.059
Motivation	−2.274	9.747	.816	1.630(2.300)	.554
ON-Med Period	1.21	3.789	.749	31.02 (22.265)	.164
Caffeine	.587	7.806	.943	1.544(7.056)	.827

Backwards Spatial Span						
	Direct Effect			BP Residual		
Level 1 Predictor	Unstd. Beta	SE	*p*-value	Unstd. Beta (SE)	*p*-value	
Exercise	.195	.966	.842	1.048(2.668)	.694	

Table 7.

Multilevel model results of between-person predictors of context-smartphone performance association

Smartphone Backwards Spatial Span				
Level 1 Direct Effect	**Level 2 Predictor**	**Unstd. Beta**	**SE**	***p*-value**
Exercise	UPDRS Total	.062	.052	.467
	Age	−.037	.140	.792
	MoCA	.135	.386	.726
Smartphone Trails-B Level 1 Direct Effect	Level 2 Predictor	Unstd. Beta	SE	*p*-value
Time of Day	UPDRS Total	−.022	.016	.163
	UPDRS Motor	−.364	.023	.218
	Time since diagnosis	−.243	.037	.586
	H&Y score	.124	.298	.678
	Age	−.012	−.618	.536
	ESS	−.666	.249	.008
Home	UPDRS Total	.027	.038	.478
	Time since diagnosis	−.104	.093	.261
	Age	.027	.055	.627
ON-OFF Meds	MoCA	.054	.149	.739
	UPDRS Total	−.096	.075	.199
	Time since diagnosis	−.091	.252	.723

J Int Neuropsychol Soc. Author manuscript; available in PMC 2023 September 02.