Using smartphone technology to improve prospective memory functioning: A randomized controlled trial

Michael K. Scullin PhD¹,² | Winston E. Jones PhD¹ | Richard Phenis PsyD² | Samantha Beevers BS² | Sabra Rosen PhD² | Kara Dinh PsyD² | Andrew Kiselica PhD³ | Francis J. Keefe PhD⁴ | Jared F. Benge PhD¹,²,⁵

¹Department of Psychology and Neuroscience, Baylor University, Waco, Texas, USA
²Department of Neurology and Plummer Movement Disorders Center, Baylor Scott and White Health, Temple, Texas, USA
³Department of Health Psychology, University of Missouri, Columbia, Missouri, USA
⁴Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina, USA
⁵Department of Neurology, University of Texas in Austin, Austin, Texas, USA

Correspondence
Michael K. Scullin and Jared F. Benge, Department of Psychology and Neuroscience, Baylor University, One Bear Place 97334, Waco, TX 76798, USA. Email: michael_scullin@baylor.edu and jared.benge@austin.utexas.edu

Funding information
National Institute on Aging, Grant/Award Number: AG053161; Microsoft Research, Grant/Award Number: N/A (donated Lumia phones)

Abstract

Background: A decline in the ability to perform daily intentions—known as prospective memory—is a key driver of everyday functional impairment in dementia. In the absence of effective pharmacological treatments, there is a need for developing, testing, and optimizing behavioral interventions that can bolster daily prospective memory functioning. We investigated the feasibility and efficacy of smartphone-based strategies for prospective memory in persons with cognitive impairment.

Methods: Fifty-two older adults (74.79 ± 7.20 years) meeting diagnostic criteria for mild cognitive impairment or mild dementia were enrolled in a 4-week randomized controlled trial. Participants were trained to use a digital voice recorder app or a reminder app to off-load prospective memory intentions. Prospective memory was assessed using experimenter-assigned tasks (e.g., call the laboratory on assigned days), standardized questionnaires, and structured interviews. Secondary dependent measures included days of phone and app usage, acceptability ratings, quality of life, and independent activities of daily living.

Results: Participant ratings indicated that the intervention was acceptable and feasible. Furthermore, after the four-week intervention, participants reported improvements in daily prospective memory functioning on standardized questionnaires ($p < 0.001$, $η_p^2 = 0.285$) and the structured interview ($p < 0.001$, $d = 1.75$). Participants performed relatively well on experimenter-assigned prospective memory tasks (51.7% ± 27.8%), with performance levels favoring the reminder app in Week 1, but reversing to favor the digital recorder app in Week 4 ($p = 0.010$, $η_p^2 = 0.079$). Correlational analyses indicated that greater usage of the digital recorder or reminder app was associated with better prospective memory performance and greater improvements in instrumental activities of daily living (completed by care partners), even when controlling for condition, age, baseline cognitive functioning, and baseline smartphone experience.

Michael K. Scullin, Winston E. Jones, and Jared F. Benge contributed equally and share authorship.
**Conclusions:** Older adults with cognitive disorders can learn smartphone-based memory strategies and doing so benefits prospective memory functioning and independence.

**KEYWORDS**
Alzheimer’s disease and related dementias, assistive technology, electronic memory aid, offloading, technology and dementia

**INTRODUCTION**

The ability to remember to perform daily intentions, or *prospective memory*, has been called a “signature” decline of Alzheimer’s disease and related dementias (ADRD). Everyday prospective memory tasks include remembering to attend appointments, deliver messages, and take medications as prescribed. In this way, prospective memory is critical to maintaining independent daily functioning.

With FDA-approved treatments for mild ADRD possibly extending duration of living with the disease, there is an urgent need for behavioral treatments that extend functionality and quality of life. Current nonpharmacological approaches to supporting prospective memory include cognitive “brain” training, mnemonic strategies, and electronic memory aids. With few exceptions, brain training paradigms have not produced clinically significant benefits to everyday prospective memory functioning. Training healthy adults to use mnemonic strategies has shown promise in laboratory settings, but the evidence is mixed as to whether persons with ADRD can effectively use mnemonic strategies in naturalistic settings.

Electronic memory aids allow intentions to be “off-loaded” onto digital devices so that they can be accessed anytime. A recent meta-analysis indicated that electronic memory aids may be the most effective intervention for prospective memory in persons with cognitive impairment. One example of an electronic memory aid is using a digital voice recorder to verbalize and later review one’s intentions. Another example would be recording the intention to an electronic device that can provide automated reminders to perform the intention at specific times or locations. Digital recorder and reminder apps are now widely available on smartphones, and smartphones are now owned by more than half of older adults (aged 65 and older), even those with geriatric cognitive disorders. Despite the prevalence and potential value of such devices, studies of electronic memory aids for prospective memory have typically lacked random assignment, rigorous blinding procedures, and adequate sample sizes (median $N = 5$).

We conducted a randomized controlled trial of persons with mild cognitive impairment (MCI) and mild dementia to investigate whether one could support prospective memory with a smartphone-based electronic memory aid: either a reminder app or a digital recorder app. We addressed whether such individuals could and would use smartphones (feasibility/acceptability) across 4 weeks (adherence), and whether doing so benefitted self-reported and objective prospective memory outcome measures (efficacy). We hypothesized that participants would use the smartphone on most days, but that usage would be lower in participants with worse cognitive functioning and less prior experience with smartphones. Additionally, we hypothesized that the reminder app would lead to larger improvements in prospective memory functioning than the digital recorder app because the former can provide time- and location-based reminders.

**METHODS**

**Participants**

The study commenced in February 2018 and concluded in February 2020 when the sample size goal was reached ($N = 52$). The sample size was determined a priori to...
detect large between- and within-group effects that would have clinical significance.\textsuperscript{5} Participants were older adults (74.79 ± 7.20 years; range: 55–92 years) who had received diagnoses of MCI or dementia (regardless of etiology) within 12 months of enrollment based on clinical neuropsychological or neurological evaluation at Baylor Scott & White Health in central Texas (Figure 1). Diagnoses followed the National Institute on Aging-Alzheimer’s Association work group core criteria,\textsuperscript{15,16} including a concern had been raised regarding cognition, there was objective impairment in cognition, and relative preservation of functional abilities (MCI); or there was report of cognitive symptoms that interfered with normal activities, objective impairment in cognition, and more than one domain of cognition was impaired (dementia).

To ensure the presence of no more than mild dementia, participant inclusion criteria required that informant-rated activities of daily living performance showed that participants maintained independence in basic physical self-maintenance tasks and that their telephone-administered cognitive screening instrument-modified (TICS-M) score was within $-1$ to $-2$ standard deviations of demographically corrected norms.\textsuperscript{17} The content of the TICS-M is similar to other screening instruments (e.g., mini mental status examination), but has the advantage of including a 10-word immediate and delayed recall test that increase sensitivity to memory deficits.\textsuperscript{18} The distribution of demographically corrected TICS-M scores across conditions is available in Figure S1; to summarize, overall cognitive functioning was slightly worse than in studies in which the sample was composed entirely of amnestic-MCI participants.\textsuperscript{19} Exclusion criteria included severe mental illness, disabling motor dysfunction, vision or hearing loss, aphasia, and limitations to English proficiency that prevented training.

Participants were either referred to the study by collaborating clinicians or received mailed invitations based on diagnoses from a prior clinical visit. Prior to
participation, a neuropsychologist reviewed the consent form with the patient and conducted a structured capacity to consent interview. Participants who could not demonstrate capacity to consent could still enroll via proxy consent from an authorized care partner along with continued expression of assent \((n = 3)\). The study was approved by the Baylor Scott & White IRB.

**Overview of study design**

The study was a parallel group trial design with participation lasting 4 weeks. Participants were randomly assigned to smartphone-based electronic memory aid conditions using a 1:1 allocation ratio, blocked randomization (sets of 2, 4, 6), and sealed envelopes. The PI generated the random sequences, the site PI enrolled the participants, and independent research staff trained the participants and conducted assessments.

**Smartphone devices**

Participants used a lab-provided phone (Microsoft Lumia 950 XL, Samsung Galaxy J7 Neo, or iPhone 6s) or their personal smartphone device. After the first six participants, we switched from requiring a lab-provided Lumia phone to allowing personal devices following difficulties with recruitment, user-experience, and logistics, such as problems transferring contact lists (Figure S2). This modification had minimal impact on the protocol (Supplementary Methods) and the study outcomes were similar across persons who used their own phone versus a lab-provided phone (Table S1).

**Procedure**

Following the consent process, participants completed baseline questionnaires, training in the general features of smartphones, and training in the intervention app (reminder app or digital recorder app). This baseline session generally lasted 2–4 h. A study partner attended the baseline session and provided a rating of instrumental activities of daily living (IADL) performance. Otherwise, participants completed all procedures themselves.

General training on phone functionality involved identifying and using: (1) on/off buttons, (2) volume buttons, (3) battery charging, (4) home-screen navigation, (5) phone calls, (6) adding contacts, (7) text messaging, (8) taking and viewing photos, and (9) browsing the internet. When possible, screen font size and volume were increased. Furthermore, the trainer monitored for, and corrected, tendencies to press or hold screen icons too hard or too long (which can lead to unexpected responses from the phones). App Usage (Android) or Moment (iOS) apps were installed to each device to allow for tracking of daily usage.

After structured training on the general features of smartphones, participants were trained to use a reminder app or digital recorder app. Participants were provided an app guide that included: (1) a screenshot of the app’s icon, (2) descriptions of the buttons necessary for using the app, and (3) description of how to phrase effective reminders for prospective memory tasks. In the reminder app condition, participants were trained to set, review, and receive reminders using the Cortana app, which was available on Windows, Android, and iOS devices. Cortana’s news and social media notifications were deactivated, and the reminder notifications were set to persist on-screen until actively dismissed. During the training session, the researcher also input addresses for locations that the participant might regularly visit or wish to receive reminders. In the digital recorder app condition, participants were trained to voice record and review their intentions using the Android Voice Recorder app, iPhone Voice Memos app, or an app that was native to their device that had identical functionality. We encouraged participants to record their intentions using the implementation intention structure (“When it is 7 pm, then I will call my brother”).

For 4 weeks following the baseline session, participants were assigned prospective memory tasks to perform on scheduled days and at specified locations. During this interval, a researcher also made three scheduled calls to the participant to troubleshoot difficulties with the smartphone (Supplementary Methods, Appendix S1). Call duration and references to memory strategies were tracked by an independent study team member to mitigate the potential for experimenter bias (Table S2). After 4 weeks, participants returned to the clinic to complete post-intervention questionnaires and a debriefing interview.

**Outcome measures**

A complete list of outcome measures and timing of assessments is provided in Table S3. The primary outcome measure was prospective memory performance. Twice each week, participants were instructed to remember to call an experimenter phone number on specific days and remember to take photos at specific locations. Locations cycled between the participant’s home, grocery store, church, and workplace (or gym/exercise location). To protect against experimenter bias, these prospective memory tasks were provided by an interactive voice response system (Plum Fuse+).
Self-reported prospective memory functioning was assessed using the Prospective-Retrospective Memory Questionnaire (PRMQ; Supplementary Methods, Appendix S1). In addition, the researcher administered a structured interview in which the participant and care partner identified 7–10 activities from the participant’s daily routine that required frequent use of prospective remembering (e.g., remembering to take medication). During the post-intervention session, participants rated whether performance on each identified activity had become much better (+2), better (+1), worse (−1), much worse (−2), or stayed the same (0), from which we calculated a composite score that was tested against the null hypothesis of 0.

Secondary outcome measures included prior experience with smartphones, evaluated using the 14-item Media and Technology Usage Smartphone Subscale (MTUSS), that has been adapted to include cognitive prosthetic features. Expected acceptability (baseline) and postintervention acceptability was assessed using the 26-item Smartphone Acceptability and Usability Scale (adapted from Reference 25). Quality of life was assessed at baseline and postintervention sessions with four Neuro-QoL short forms, which were also embedded into the biweekly interactive voice response system (Supplementary Methods, Appendix S1). Because this system had difficulty detecting participants’ voice responses, only Neuro-QoL data from the baseline and post-intervention sessions will be considered. Participants’ care partners completed IADL assessments at baseline and post-intervention sessions.

Statistical analysis

Two independent, blinded raters scored and entered data. Experimenter-assigned prospective memory performance was calculated as the proportion of tasks correctly completed each week (i.e., call on the scheduled day, photos of assigned locations on the week of assignment). Data were analyzed using analyses of variance (ANOVA), t tests, Pearson correlations, and linear regression. Alpha was set to 0.05. The study was registered at ClinicalTrials.gov (NCT03384043), and there were no interim analyses. Study materials and de-identified data are available at Open Science Framework (https://osf.io/v459f/).

RESULTS

Sample characteristics and safety

Figure 1 displays the CONSORT flow chart from screening (260 individuals) to enrollment (N = 52) to analysis (n = 48). The primary reason for screening failures was a TICS-M score that was out of the inclusion range (n = 86). Table 1 provides demographic information for the 48 participants with available data (n = 47 for some postintervention assessments). Though 73% of participants owned a smartphone, most had limited prior experience using it (Figure 2A). The study procedures were safe: the 11 adverse events recorded were deemed by the safety officer to be unrelated to the study protocol (e.g., fall in the shower).

Feasibility, adherence, and acceptability

Ninety percent of participants completed all study phases. At the baseline session, training in general smartphone features and apps was quicker in the digital recorder condition (M = 73.09 min, 95%CI: 60.00–85.63) than the reminder app condition (M = 95.41 min, 95%CI: 83.00–108.22), t(41) = 2.52, p = 0.016, d = 0.77 (Figure S3). These data reflect that digital recorder apps are simpler and more familiar than reminder apps.

After the baseline session, participants used their smartphones on 81.5% of days (meeting a priori estimates). Figure 2B illustrates that general usage of the smartphone was highest in Week 1 (85.8% of days) and
dropped only modestly over time (76.3% of days in Week 4), \( F(3,138) = 3.77, p = 0.012, \eta^2_p = 0.076 \). General smartphone usage was similar across conditions \( F(1,46) = 1.25, p = 0.27, \eta^2_p = 0.027 \); interaction: \( F(3,138) = 0.938, p = 0.424, \eta^2_p = 0.020 \).

We next examined records of digital recorder and reminder app usage. Across the 28 study days, participants showed a median of 20 app usages \( (M = 23.95, 95\% CI: 18.75–28.96) \). There were similar levels of app usage across conditions \( (F(1,46) = 1.25, p = 0.27, \eta^2_p = 0.027; \text{interaction: } F(3,138) = 0.938, p = 0.424, \eta^2_p = 0.020) \). Note that time stamps were not available for conducting analyses of app usage by study weeks.

Figure 2d demonstrates that ratings on the smartphone acceptability scale were similar across time points \( (F(1,45) = 0.74, p = 0.396, \eta^2_p = 0.016) \) and conditions \( (F(1,45) = 0.747, p = 0.39, \eta^2_p = 0.016; \text{interaction: } F(1,45) = 3.21, p = 0.080, \eta^2_p = 0.067) \). At postintervention, 76.6% of participants indicated agreement (rather than “neutral” or “disagree”) for the item “I would recommend the smartphone to a friend.” At baseline, 64.6% had agreed with this item.

Table S4 shows the correlation matrix for measures of feasibility/acceptability, demographics, global cognitive status, and previous smartphone experience. Participants with more prior experience with smartphones at baseline showed the highest levels of adherence and acceptability (Figure S4). Interestingly, older age and lower cognitive functioning were associated with lower acceptability at baseline, but these associations disappeared by the postintervention assessment (Figure S5). Furthermore, the frequency of digital recorder and reminder app usage was unrelated to age, gender, race/ethnicity, education, and cognitive functioning (all \( ps > 0.10; \text{Table S4} \)). Therefore, even though most participants began the study with limited prior smartphone experience, the majority of participants used their smartphone device and app regularly during the study, agreed that it was usable, and would recommend it to a friend.

### Prospective memory

Overall performance on experimenter-assigned prospective memory tasks was high (51.7% ± 27.8%) relative to typical performance levels reported in participants with geriatric cognitive disorders. \(^1\,^{5.6.29.30}\) Figure 3A demonstrates a significant main effect of time, \( F(1,45) = 21.45, p < 0.001, \eta^2_p = 0.318 \), that interacted with app condition, \( F(3,138) = 3.937, p = 0.010, \eta^2_p = 0.079 \) (condition main effect: \( F < 1 \)). The time by condition interaction occurred because the reminder app condition initially outperformed the digital recorder app condition, but this pattern reversed by Week 4 (largest posthoc comparison for Week 1: \( t(46) = 1.76, p = 0.09, d = 0.51 \)).
The overall high levels of performance (given an ADRD sample) suggest that both smartphone app interventions resulted in benefits to prospective memory performance. Consistent with this interpretation, Figure 3B illustrates a pre-to-post intervention improvement in self-reported everyday prospective memory (PRMQ), $F(1, 43) = 17.12, p < 0.001, \eta^2_p = 0.29$ (all other Fs < 1).

Furthermore, Figure 3c illustrates that participants reported an overall improvement in their personally relevant prospective memory tasks (structured interview), $t(46) = 5.94, p < 0.001, d = 1.75$. These improvements were similar across the reminder app and digital recorder app conditions ($t(46) = 1.61, p = 0.12, d = 0.47$). These benefits to prospective memory were observed in the absence of any pre-to-post changes in commission error rates, which refers to repeating an intention that has already been completed (Figure S6).31

**Quality of life**

The quality of life composite score significantly improved from pre-to-post intervention (Table S5), $F(1, 45) = 6.05, p = 0.018, \eta^2_p = 0.118$. Levels of improvement were
similar across conditions ($F$s < 1). IADL levels did not differ across conditions or time points ($F$s < 1; Table S5).

**Intervention adherence correlations with outcome measures**

Table S6 displays the correlation matrix for app usage (adherence) and the outcome measures. Two patterns are worth note. Figure 4A illustrates that participants who used their app more often performed better on experimenter-assigned prospective memory tasks, even after controlling for condition, age, baseline global cognitive status (TICS-M), and baseline smartphone experience (MTUSS), $r_p$ (38) = 0.335, $p = 0.035$. In addition, Figure 4B shows that when participants used their app more often, participants' care partners reported a positive improvement in the participant's IADL; again, statistically significant when controlling for baseline IADL, condition, baseline TICS-M, and baseline MTUSS, $r_p$ (35) = 0.326, $p = 0.049$.

**DISCUSSION**

A recent Cochrane review reported that there were zero randomized clinical trials on assistive technology for memory difficulties in dementia patients. Minimal progress has occurred since that review, but the current study demonstrated in a methodologically rigorous manner that training older adults with cognitive impairment to use smartphone memory aids was feasible, acceptable, and safe. After 4 weeks of using smartphone memory aids, 2/3 of participants reported an improvement in their prospective memory functioning, paralleling an improvement in quality of life. Though we predicted reminder apps to produce superior outcomes, digital voice recorder apps were similarly beneficial for this population. Given that the majority of prospective memory interventions have been tested over short time intervals in laboratory contexts with arbitrary assessments, it is notable that participants reported in the structured interview that their most personally relevant prospective memory difficulties had improved after the intervention.

**Efficacy for prospective memory**

The literature commonly reports 20% performance levels for experimenter-assigned prospective memory tasks in persons with MCI or mild dementia. By contrast, average performance levels with the two smartphone interventions were approximately 50%, and Week 1 performance approached 80% in the reminder app condition. Differences in demographic composition and/or types of prospective memory assessments may partially explain the levels of performance in our study relative to the broader literature. Nevertheless, the absolute performance levels in the current study highlight the value of interventions that involve “off-loading” intentions to assistive technologies.

We predicted that the reminder app would outperform the digital recorder app because of its ability to deliver reminders at specific times and locations. The results showed a more complex interaction: the reminder app outperformed the digital recorder app in Week 1, but this pattern reversed by Week 4. One explanation is that a reminder app is superior when it has been recently trained and can be used optimally (Week 1), but a digital recorder app remains simple to use over a longer time frame (Week 4). Another explanation is that the structured verbalization (implementation intention) in the digital recorder condition required a few weeks to be maximally beneficial. Both explanations are possible, but some evidence favors the former: participants who used their reminder app more showed better prospective memory performance and better improvement in IADLs. Thus, future research should focus on maximizing app usage via motivational interviewing for behavioral change, automated prompts, or by directly involving care partners. Furthermore, smartphone developers can improve the interface and usability of reminder apps. Such efforts might sustain higher levels of prospective memory functioning and support independence for longer periods in persons with mild ADRD.

**Feasibility and acceptability**

Some may be surprised that a smartphone intervention in persons with ADRD was even feasible or acceptable. Known as the “digital divide,” there is a pervasive view that most older adults dislike, and are unable to use, smartphone technology. Times change. Recent surveys and focus groups suggest that much of the digital divide reflects stereotyping, perceived ageism, and a cohort/generational effect (nonexposure to technology until too late). Given that smart technology can reduce prospective memory difficulties—as well as reduce social isolation, detect falls, monitor heart rhythms, and promote independence—it is time to discard stereotypes and train healthy older adults, persons with mild ADRD, and care partners how to optimally use smartphones to support functioning and enrich quality of life. In addition, there is a need for broadly improving internet access to ensure that advances in smart technology do not worsen health disparities across socioeconomic classes.
The digital divide may also reflect hesitancy to learn new technology on the part of the oldest individuals and those with relatively low global cognitive status.41 We observed these patterns at baseline, but after training, smartphone acceptability ratings no longer related to age or cognitive functioning, suggesting that extended practice overcomes these barriers. The greater determinant of smartphone usage and acceptability was prior familiarity with such devices. This result fits the technological reserve hypothesis,40,42 which predicts that establishing fluency in some technological skills prior to cognitive decline (e.g., texting) is a building block to learning new technological skills (e.g., using a reminder app) and preserving daily functioning.43–45 As the adult population aged 65 and older continues to show an increase in usage of smartphones, the feasibility and acceptability of smartphone-based interventions will increase, even in persons with ADRD.

Limitations and conclusions

Study limitations included a modest time interval (4 weeks), absence of baseline data for the experimenter-assigned tasks, and a relatively homogenous sample. In addition, this study did not have independent confirmation of whether the participant, a care partner, or both used the interactive voice response system each week, or to what degree care partners may have provided participants with assistance. However, at training, care partners were instructed to promote independent use of the strategy by the participant and never to take over the tasks for them.

The current work indicates that smartphone-based interventions for prospective memory are feasible and acceptable in persons with mild ADRD, and provides preliminary evidence that smartphone interventions can improve everyday prospective memory functioning. Importantly, the more participants used their apps the more care partners reported that the participant improved their IADL. Each of these outcomes is a promising signal that smart technology can reduce the negative impact of neurodegenerative conditions on prospective memory, quality of life, and independent functioning.

CONFLICT OF INTEREST

Microsoft Research, Inc., provided Lumia phones that were used for some participants.

AUTHOR CONTRIBUTIONS

Study conception: Scullin, Benge; Study design: Scullin, Benge, Keefe; Data acquisition: Jones, Phenis, Beevers, Rosen, Dinh, Kiselica; Data analysis: Jones, Beevers, Benge, Scullin; Data interpretation: Scullin, Benge, Jones; Manuscript preparation: Scullin, Benge, Jones; Reviewed the final manuscript: all authors. The authors are grateful to Madison Krueger, Michelle Dasse, PhD, Stacy Nguyen, Nicholas Rowley, Mary High, Nikita Chapagain, Taylor Terlizzese, Zachary Terlizzese, Taylor Luster, Daniel Zeter, Keyanna Gayden, Ivy Grossberg, Ariana Nail, and Erin Logue, PhD for contributions to data collection, data storage, data entry, and participant referral. We are additionally grateful to Gary Elkins, PhD for contributions to the project as safety officer.

SPONSOR’S ROLE

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

ACCESS TO DATA AND DATA ANALYSIS

Michael Scullin, Winston Jones, and Jared Benge had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

ORCID

Michael K. Scullin https://orcid.org/0000-0002-7578-7587

TWITTER

Michael K. Scullin @BaylorSleep

REFERENCES

7. Burkard C, Rochat L, Van der Linden ACJ, Gold G, Van der Linden M. Is working memory necessary for implementation...


**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.

**Appendix S1.** Supporting Information.