Investigating usability of the smartwatch as an everyday prospective memory aid for older adults with and without cognitive impairment

E. Sanders, W. Boot

Florida State University (Tallahassee)

INTRODUCTION

Prospective memory (PM), the ability to remember and carry out an intention in the future, is crucial for maintaining health, independence, and social connections. Normative age-related changes, changes that most people can expect to experience as they age, can result in meaningful PM failures, and studies have shown a reliable link between PM and activity of daily living performance and quality of life among older adults (OAs) [1]. Thus, PM decline represents an obstacle for functional independence in older adulthood. This issue can be exacerbated for individuals with cognitive impairment (CI) such as mild cognitive impairment (MCI) [2], CI due to traumatic brain injury (TBI) [3], and post-stroke CI (PSCI) [4].

Wearable technologies, such as smartwatches, have the potential to serve as effective PM aids. A wearable device can move with the person; it can be used as an aid inside and outside of the home [5-6]. Additionally, it can send visual, audio, and tactile alerts that can be highly noticeable to the user but subtle enough to not be disruptive in social settings. However, the user must continually wear the device and keep it charged for it to be effective [6]. Potential barriers to the effective use of a wearable device include challenges due to small and hard-to-see buttons, difficult-to-navigate menus, confusing terminology, and other usability problems [7]. These types of challenges are likely to be exacerbated for OAs experiencing CI [5, 8].

Few studies of mobile device efficacy include a diverse population of OAs, such as individuals with disabilities and CIs [5]. Little is known about how OAs, particularly those with CIs, might interact with smartwatches. Thus, further research following an evidence-based, user-centered design process is needed before an intervention can be developed for this population. This study aimed to 1) assess the usability of smartwatches for OAs with a diverse range of cognitive functioning, including those with CIs, and 2) gain an initial understanding of the efficacy of smartwatches as PM aids for this population.

METHODS

Study procedure

A telephone screening was conducted to identify potential participants. Participants were excluded if they scored below a 22 on the Telephone Interview for Cognitive Status-Modified (TICS-M), as these individuals would be unlikely to provide informed consent. Once enrolled in the study, participants first completed telephone interviews in which demographic, technology acceptance [9], and neuropsychological information were collected. Initial cognitive status was measured by the Montreal Cognitive Assessment (MoCA)- BLIND. Subjective measures of memory included the Memory Functioning Questionnaire (MFQ) [10] and the Prospective and Retrospective Memory Questionnaire (PRMQ) [11]. Participants were then randomly assigned to the Intervention Group or the WL Group.

Participants in the Intervention Group received in-person training on how to use the smartwatch. Participants with Android phones were given a Samsung Galaxy Watch4 Classic (46mm; Bluetooth). Participants with iPhones, as well as one participant whose Android phone was not compatible with the Galaxy Watch4, were given a Samsung Galaxy Watch Active2 (44mm; Bluetooth). The smartwatch screens were customized to display the same digital face. Participants were asked to wear the smartwatch and use its reminder notifications to remind them to complete an automated telephone survey once each day for 10 days. Reminders to complete the survey were programmed to appear on the smartwatch each day at a time chosen by the participant. The daily survey consisted of one question in which participants were asked to rate from 1 to 5 how difficult it was for them to remember to complete the survey that day. The completion of this survey served as an objective memory test. A reminder to charge the watch was also programmed to appear on the smartwatch each night at a time chosen by the participant. At the end of the 10-day period, participants completed a final phone interview in which they completed several usability measures including the System Usability Scale (SUS), NASA Task Load Index (TLX), and the Perceived Usefulness and Ease of Use Scale (PUEU) [12].

Participants in the WL Group were first asked to use their usual memory strategies to remember to complete the daily survey for 10 days. These participants then followed the same procedure as those in the intervention group for an additional 10 days.

RESULTS

Participants were 27 community-dwelling OAs with either normal cognitive functioning or a self-reported CI due to MCI, TBI, or PSCI. Demographic information is summarized in Table 1.

Table 1. Demographic information for participants within and across study conditions

Group	n	% Female	Age (yrs)	% White/Caucasian	% Non-impaired
Intervention Group	12	16.7	70.25 (6.03)	66.7	50.0
Wait-List Control Group	15	80.0	72.27 (5.91)	93.3	60.0
All Participants	27	51.9	71.37 (5.94)	81.5	55.6

Usability

Descriptive statistics for usability measures are reported in Table 2. SUS ratings of the smartwatches were average to below average. Participants who used the Samsung Galaxy Watch4 gave the watch an average rating of 68.18%, corresponding to a score of "C," or average. Participants who used the Samsung Galaxy Watch Active2 gave the watch an average rating of 58.44%, corresponding to a score of "D," or below average. There was not a significant difference in the SUS rating between the two smartwatches.

Ratings on the NASA TLX relate to the task of responding to (i.e., dismissing or snoozing) the reminder notifications on the smartwatch screen. For both smartwatches, participants gave relatively low ratings for all scales, representing low task demands (i.e., mental, physical, temporal), little total effort required for the task, low levels of frustration involved in the task, and good performance on the task. Compared to the Galaxy Active2, the Galaxy Watch4 generally had lower ratings of task demands, lower effort levels required, lower frustration levels, and better ratings of success, though these differences were not significant. Ratings of mental demand were marginally lower for the Galaxy Watch4 compared to the Galaxy Active2, $t(25) = -1.86$, $p = .075$, $d = -0.73$.

For both smartwatches, participants gave mean perceived usefulness ratings of approximately 50% of the possible scale maximum, suggesting that there was not an overall positive or negative view of usefulness. For the Galaxy Watch4, participants gave an average mean ease of use rating at 73.81% of the scale maximum, indicating a slightly above average view of ease of use. For the Galaxy Watch2, ease of use ratings were slightly lower, though not significantly, at an average of 63.54% of the possible scale maximum.

Table 2. Means and standard deviations for usability measures for both smartwatches

Device	SUS	NASA TLX						PUEU	
		Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Usefulness	Ease of Use
Galaxy Watch4 (n = 11)	68.18% (11.73)	10.00% (8.66)	5.91% (6.25)	8.64% (14.51)	13.18% (17.65)	10.45% (12.14)	15.91% (17.58)	3.33 (1.99)	5.17 (1.04)
Galaxy Active2 (n = 16)	58.44% (20.27)	25.94% (27.40)	15.00% (18.44)	14.06% (10.83)	20.63% (24.55)	18.13% (22.05)	34.06% (34.41)	3.65 (1.43)	4.45 (1.04)

Memory performance

Memory performance (Figure 1) was measured by the percentage of daily surveys completed during each study period. A between-subject ANCOVA was conducted to compare memory performance between the Intervention Group and WL Group while the Intervention Group had smartwatches and the WL Group did not, controlling for initial cognitive status (MoCA). There was not a significant effect of condition on the proportion of daily surveys completed after controlling for initial cognitive status, $F(1, 24) = .01$, $p = .928$, $\eta_p^2 < .01$. In other words, participants with the smartwatches did not complete a higher or lower proportion of daily surveys ($M = .93$, $SD = .15$) compared to those without the smartwatches ($M = 0.91$, $SD = 0.26$).

A paired-samples *t*-test compared performance within the WL Group before and after the introduction of the smartwatches. There was not a significant difference in the proportion of daily surveys that participants completed without the smartwatch ($M = 0.91$, $SD = 0.26$) compared to the proportion completed with the smartwatch ($M = 0.93$, $SD = 0.12$), $t(14) = -.185$, $p = .856$.

Ratings provided in the daily surveys were used as a self-reported measure of how difficult it was for the participants to remember to complete the daily survey. There were no significant differences in survey ratings between groups or within the WL Group before and after the introduction of the smartwatches.

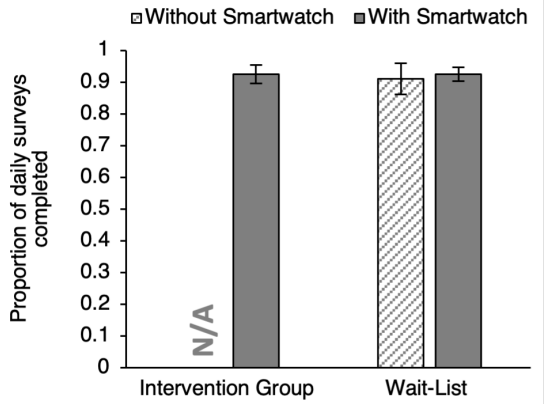

Figure 1. Memory performance with error bars (one standard error above and below the mean)

Relating memory and technology acceptance to perceived usefulness

Correlations were run to examine relationships between perceived usefulness and scores from the MoCA, MFQ subscales, PRMQ subscales, and technology acceptance subscales. Significant results are reported.

There were significant negative correlations between perceived usefulness scores and two subscales of the MFQ: the retrospective functioning scale, $r(25) = -.41$, $p = .035$, and the frequency of forgetting scale, $r(25) = -.50$, $p = .008$. Better perceived memory was associated with lower ratings of perceived usefulness of the smartwatches.

There were significant positive correlations between perceived usefulness scores and four subscales of the PRMQ: prospective long-term environmentally-cued, $r(25) = .48$, $p = .011$, retrospective short-term self-cued, $r(25) = .45$, $p = .018$, retrospective long-term self-cued, $r(25) = .56$, $p = .002$, and retrospective long-term environmentally-cued, $r(25) = .50$, $p = .008$. Reports of more frequent memory failures on these subscales were associated with higher ratings of perceived usefulness of the smartwatches.

DISCUSSION

Usability

Participants gave relatively low usability ratings for the smartwatches overall. Participants gave SUS scores corresponding to “C” and “D” grades for the Galaxy Watch4 and Galaxy Active2, respectively, indicating that the overall usability of the smartwatches was average to slightly below average. This trend was consistent with the perceived ease of use subscale of the PUEU, where participants reported perceived ease of use scores at approximately 70% of the scale maximum for the Galaxy Watch4 and approximately 60% for the Galaxy Active2.

Despite giving poor usability ratings for the smartwatches overall, participants seemed to respond more positively when it came to their main task with the smartwatch, which was to respond to the reminder notifications that appeared daily on the watch screen. Participants reported in the NASA TLX that the task demands were relatively low and required little effort, and they reported good performance on the task. This could suggest that poor ratings on the SUS may have originated from experiences with the smartwatches that participants had outside of responding to the reminder notifications. Notably, 11 out of the 16 participants who used the Galaxy Active2 reported that at some point during their 10-day period with the smartwatch, the reminder notifications stopped appearing on the watch. No participants who used the Galaxy Watch4 reported this issue. It is possible that for participants who used the Galaxy Active2, responses were influenced by negative perceptions of the reliability and consistency of the reminder notifications.

Examining relationships between participants’ perceptions of their memory and perceived usefulness of the smartwatches revealed some intriguing findings. Overall, as participants rated their memory more positively and memory failures as less frequent, the perceived usefulness of the smartwatches decreased. This is theoretically consistent with popular models of technology acceptance and adoption in which perceived usefulness is a major determinant. These results have important implications for smartwatch adoption; it is possible that individuals with more memory problems may be more likely to choose to use a smartwatch, as they may perceive them as being more potentially useful compared to individuals with fewer memory problems. However, it would be important to determine whether peoples’ perceptions of their memory problems are accurate. If an individual has memory impairments but doesn’t perceive it as a problem, they may not be interested in adopting a smartwatch.

Smartwatches as a prospective memory aid

Overall, the smartwatches did not appear to have an effect on participants' abilities to remember to complete the daily surveys. Comparisons within the WL Group revealed that there was not a significant difference in either the proportion of daily surveys completed or the ratings provided by participants before and after the introduction of the smartwatches. A comparison between participants in the WL Group without the smartwatches and participants in the Intervention Group also revealed no significant differences. Thus, the smartwatches did not serve as an effective PM aid above and beyond participants' typical memory strategies. However, results might have been different if the sample included OAs with greater impairments experiencing more severe disability.

CONCLUSIONS

Overall, participants gave relatively low usability ratings for the smartwatches and reported a lack of consistency and reliability in the Galaxy Active2. Further, the smartwatches did not appear to make a difference in participants' abilities to remember to complete the daily PM task. Thus, the smartwatches used in the present study, particularly the Galaxy Watch2, may not be the best fit for providing PM support for OAs with a diverse range of cognitive abilities. Findings provide important implications for the potential adoption of smartwatches by OAs with and without CIs. In particular, OAs' perceptions of their memory may predict their perceived usefulness of smartwatches, an important motivator for adoption. Future work should examine the reliability and efficacy of a variety of smartwatches as potential PM aids for OAs with and without CIs and should investigate predictors of the adoption of smartwatches within this population.

REFERENCES

[1] Woods SP, Weinborn M, Li YR, Hodgson E, Ng AR, Bucks RS. Does prospective memory influence quality of life in community-dwelling older adults? Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2015;22(6):679-92.

[2] Lajeunesse A, Potvin MJ, Labelle V, Joubert S, Rouleau I. Characterization of prospective memory in mild cognitive impairment by using the Ecological test of Prospective Memory. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2021 May;28(3):367-391.

[3] Schmitter-Edgecombe M, Wright MJ. Event-based prospective memory following severe closed-head injury. Neuropsychology. 2004 Apr;18(2):353-61.

[4] Hogan C, Cornwell P, Fleming J, Man DWK, Shum D. Self-reported prospective memory after stroke. Neuropsychol Rehabil. 2021 Sep;31(8):1190-1206.

[5] Charness N, Boot WR, Gray N. Mobile monitoring and intervention (MMI) technology for adaptive aging: Proceedings of a workshop. National Academies of Sciences, Engineering, and Medicine. Washington (DC): National Academies Press (US); 2020 Sep 25.

[6] Ehrler F, Lovis C. Supporting elderly homecare with smartwatches: advantages and drawbacks. Stud Health Technol Inform. 2014;205:667-71.

[7] Khakurel J, Knutas A, Melkas H, Penzenstadler B, Fu B, Porras J. Categorization framework for usability issues of smartwatches and pedometers for the older adults. Univers Access in Hum Comput Interact. Methods, Technol, and Users. 2018;91-106.

[8] Malinowsky C, Almkvist O, Kottorp A, Nygård L. Ability to manage everyday technology: a comparison of persons with dementia or mild cognitive impairment and older adults without cognitive impairment. Disabil Rehabil Assist Technol. 2010;5(6):462-9.

[9] Chen K, Lou VWQ. Measuring senior technology acceptance: Development of a brief, 14-item scale. Innov Aging. 2020 Jun 27;4(3):igaa016.

[10] Gilewski MJ, Zelinski EM. Memory Functioning Questionnaire (MFQ). Psychopharmacol Bull. 1988;24(4):665-70.

[11] Smith G, Della Sala S, Logie RH, Maylor EA. Prospective and retrospective memory in normal aging and dementia: A questionnaire study. Memory. 2000;8: 311-321.

[12] Mitzner TL, Savla J, Boot WR, Sharit J, Charness N, Czaja SJ, Rogers WA. Technology Adoption by Older Adults: Findings From the PRISM Trial. Gerontologist. 2019 Jan 9;59(1):34-44.