The use of a smartwatch as a prompting device for people with acquired brain injury: A single case experimental design study

Jamieson, Matthew[1,2]; Monastra, Mattia[3]; Gillies, Graeme[3]; Manolov, Rumen[4]; Cullen, Breda[1]; McGee-Lennon, Marilyn[5]; Brewster, Stephen[2] and Evans, Jonathan[1]

[1]Institute of Health and Wellbeing, University of Glasgow, Glasgow, Scotland
[2]Human Computer Interaction, Department of Computing Science, University of Glasgow, Glasgow, Scotland
[3]Acquired Brain Injury Team, West Dunbartonshire, Scotland
[4]Faculty of Psychology, University of Barcelona, Spain
[5]Computer and Information Science, University of Strathclyde, Glasgow, Scotland

Funding Source

This study was funded by the Medical Research Council as part of a Doctoral Training Programme undertaken by author MJ. These funds paid for the smartwatch and smartphone hardware used in the study and the software was free to download. The funder had no role in the study design; collection, analysis, and interpretation of data; writing of the report; or decision to submit for publication.

Author correspondence: Matthew Jamieson (PhD student); m.jamieson.1@research.gla.ac.uk

The use of smartwatches as a prompting device for people with acquired brain injury: A single case experimental design study

Abstract

Prompting based memory compensation is a potential application for smartwatches. This study investigated the usability and efficacy of a Moto360 smartwatch as a memory aid. Four community dwelling adults with memory difficulties following ABI were included in an ABA single case experimental design study. Performance of everyday memory tasks was tested over six weeks with the smartwatch and software provided during weeks three and four. Participants were asked to use their usual memory aids and strategies during the control phases (weeks 1-2, 5-6). Three participants successfully used the smartwatch throughout the intervention weeks and gave positive usability ratings. A fourth participant experienced a seizure and subsequently left the study before the intervention phase. Three participants showed improved memory performance when using the smartwatch. NAP analysis showed a non-significant small increase in memory performance between baseline and intervention phases (mean NAP = 0.1, $p = 0.84$). There was a larger, significant decline between the intervention and return to baseline (mean NAP = 0.58, $p < 0.01$). The use of off-the-shelf smartwatch device and software was feasible for people with ABI in the community. It was effective compared to practice as usual, although this was only apparent on withdrawal of the device.

Keywords: Assistive technology, Acquired brain injury, Memory rehabilitation, Smartwatch, Prospective memory

Introduction

The term acquired brain injury (ABI) refers to injury to the brain arising from a head trauma (e.g. road traffic accidents and falls), cerebrovascular events (e.g. stroke), illnesses or diseases (e.g. brain tumour or encephalitis). Individuals who have suffered ABI or who have a neurodegenerative disease have a high prevalence of memory impairments. In particular in ABI, prospective memory is often impaired (Evans, 2003). People with ABI may also experience disorganized thinking, problems with planning, language impairment, poor self-monitoring and difficulty switching between or initiating tasks (Wilson, Gracey & Evans, 2009). These impairments make it difficult for people to perform everyday tasks such as shopping, personal care or cooking, or healthcare tasks such as remembering appointments, treatment plans and medication. Furthermore, health problems that directly or indirectly result from the ABI, such as physical disabilities, sensory/motor impairments and chronic illnesses can increase the number of health-based memory demands.

Prompting technology which can remind people about a task at a set time is effective in improving the frequency of remembering, and successfully completing, everyday activities compared to practice as usual or a non-technological equivalent (Jamieson, Cullen, McGee-Lennon, Brewster & Evans, 2014). Different hardware can be used to prompt; for example, Svoboda, Richards, Leach and Mertens (2012) used a smartphone app, McDonald, Haslam, Yates, Gurr, Leeder and Sayers (2011) used a calendar program on a computer which sent text messages to the participant’s phone, and Lemoncello, Sohlberg, Fickas and Prideaux (2011) used a television set to prompt participants with ABI about their exercise routines. Some researchers have tested wearable prompting devices, most notably Wilson, Emslie, Quirk and Evans (2001) in a randomised controlled trial testing NeuroPage, a wearable paging device. Although this aspect of its use has not been investigated explicitly, the wearability of NeuroPage may be an advantage compared to pencil or paper reminding strategies and other prompting devices that cannot be worn. This is because - as long as it is accepted by the user, and provided the user remembers to put it on - worn devices do not risk becoming ineffective due to being misplaced, or placed in clothes or bags in a way which would prevent the prompt being detected. Furthermore, wearable devices have the advantage of sending tactile, audio or visual alerts that can be highly noticeable because of the proximity of the device to the user, but which can also be subtle (e.g. tactile notifications) in social situations, for example during a meeting or a meal with friends. Previous research with older users has highlighted the importance of developing appropriate notification modalities for different types of reminders and in different social situations (McGee-Lennon, Smeaton & Brewster, 2012) as well as the differing impact of different types of notification modality (Warnock, McGee-Lennon & Brewster, 2013).

In the last few years, smartwatch technologies have grown in popularity and affordability. The current state of the art hardware can sync up to a smartphone, usually communicating using Bluetooth. Information and notifications which pop up on the phone can then be displayed on the watch and manipulated using voice or touch input. Reminding software which prompts on a phone can be made to be compatible with the smartwatch so that the notifications display on the watch. Many reminder and calendar apps, including those already provided as standard with a smartphone, have already been made compatible with watch hardware.

To the best of our knowledge, only one study has tested the efficacy of a watch as a reminder for people with memory impairments. Van Hulle and Hux (2006) used 'Watchminder' wristwatches with a vibrating alarm to prompt two participants about their medication. Participants had memory difficulties after traumatic brain injury and were receiving rehabilitation in a transitional living facility. A review by Jamieson and colleagues (2014) performed a Nonoverlap of all pairs (NAP; Parker & Vannest, 2009) analysis for both participants. Scores were between 0.5 and 0.66 indicating a small effect from the intervention. This was lower than the majority of NAP results for similar studies which used other types of devices to prompt participants with memory difficulties in single case experimental design studies (Jamieson et al., 2014). Therefore, it is unclear if the use of a watch with a vibration prompt is effective for people with ABI.

Watchminder pre-dates the development of smart technologies. Smartwatches can provide much more detail for a reminder than a vibration prompt alone (as was the case in the Watchminder study) because they have a display screen which can sync up to reminders set on a smartphone or computer-based calendar. However, it is possible that smartwatches may be unacceptable or unusable for participants with ABI because they are too complicated to use, especially for people living in the community who are unable to access daily help with the technology from clinicians or caregivers. Van Hulle and Hux (2006) did not report details about participants' use of the technology and included participants in a supported living environment. They also did not test use of watch-based prompting devices used by people within the community. To the authors' knowledge there is no study which has investigated the use of smartwatches as a prompting technology for people with memory impairment after ABI living in the community.

This paper reports a single case experimental design (SCED) study with four community dwelling participants with ABI. An ABA design was used to investigate the efficacy of a smartwatch reminder for prompting people with memory impairment after ABI about various events. SCED methodology was chosen because it allows a controlled trial to be performed to test efficacy when large scale recruitment is not possible (Barlow, Nock, & Hersen, 2009). A secondary aim of the study was to help understand whether or not a smartwatch using reminding software synced to a smartphone is a usable and acceptable off-the-shelf assistive technology to introduce within clinical practice. Reporting follows the guidelines detailed in the Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Checklist (Tate et al., 2016).

Methods

Participants

Participants were identified and recruited by staff in the Head Injury Care Team located within the West Dunbartonshire Community Health Care Partnership, Dumbarton, Scotland. This service assesses the client's neuropsychological profile and everyday functioning to establish their support needs and then helps to support clients in the community, working closely with other health and social services. Adults aged 18 or over who had experienced an ABI and who had been assessed as having memory impairment during clinical assessment by the recruiting service were considered for this study. National Health Service (Research Ethics Committee) ethical approval was granted for this study on 27.02.15 (reference number 15/WM/0079). Exclusion criteria were; i) the inability to provide informed consent for research participation, ii) inadequate writing or reading which would prevent them from completing the tasks required in the study; iii) severe verbal communication difficulties, severe physical impairment (which would prevent their use of a smartwatch or smartphone device). Participants did not currently use a smartwatch as a reminder, or any other technology based reminding device (such as a smartphone or personal computer based calendar) which successfully compensated for self-reported memory problems prior to the study. Four participants, who were adjudged by clinical staff within the care team to meet the study criterion, were initially recruited. One of these participants was recruited but only participated in the baseline phase of the study. The assistant psychologist at the recruiting service (MM) reported that this participant experienced a seizure during the A phase and failed to participate for several days after this. Prior to introducing the smartwatch intervention, MM and other clinical staff decided to remove this participant from the study for health reasons which were unrelated to the use of the smartwatch. Accordingly, their data was not reported in the results of this study. The cognitive profile of each participant whose data was analysed is reported in Table 2.

Neuropsychological Tests

The participants' scores from several neuropsychological tests and questionnaires were used to develop a cognitive profile. Many of the tests had already been completed prior to participation in the study (within the last three years) as part of their assessment by the neuropsychological team, although not all participants had completed the same tests. During the study, further tests were administered by the experimenters, in order to ensure that some information regarding intellectual functioning, memory and executive functioning was provided for each participant. Tests included the Test of Pre-morbid Functioning (TOPF) (Wechsler, 2011), the Wechsler Adult Intelligence Scale version 4 (WAIS-IV) – perceptual reasoning, verbal comprehension and processing speed sub-scales (Wechsler, 2008), Wechsler Memory Scale version 4 (WMS-IV) – auditory memory delayed and visual memory delayed sub-scores (Wechsler, 2009), the Behavioural Assessment of the Dysexecutive Syndrome (BADS) – including the key search and zoo map sub-tests (Wilson, Evans, Alderman, Burgess & Emslie, 1997), the Delis-Kaplan Executive Functioning System (DKEFS) – verbal fluency and letter number switching sub-tests (Delis, Kaplan & Kramer, 2001), Rey-Osterrieth Complex Figure (ROCF; Hubley & Jassal, 2006), Cambridge Test of Prospective Memory (CAMPROMPT; Wilson et al., 2005), and Rivermead Behavioural Memory Test (RBMT; Wilson, Cockburn & Baddeley, 1991). The assistant psychologist (MM) also noted demographic information, phone and technology use prior to the study, information about ABI and functional difficulties which the smartwatch could address.

TS

TS was a 45 year old man who suffered a brain haemorrhage 12 years previously, and had a stroke in 2007. He was reported to have a basal ganglia bleed and a colloid cyst in the lateral ventricle, with symptoms of hydrocephalus and damage to the corpus callosum. He reported experiencing memory loss, confusion, forgetfulness, executive difficulty and gait disturbances. His cognitive problems included language and communication difficulties and needing to rely on lists and calendars to aid prospective memory. Fatigue exacerbated these symptoms. He experienced decreased independence with cooking tasks after the haemorrhage due to difficulty remembering the steps involved. He lived by himself and requires prompts from the service and his parents to take care of his house and complete his weekly shopping. The service staff had previously encouraged him to utilise pencil and paper memory aids, such as calendars, and reported that he would forget how to use these or become confused when using them.

LA

LA was a 61 year old man who suffered spontaneous bleeding in his frontal lobe in 2004. This left some scarring which is now the focus for epileptic discharges. Scarring had a small impact on some but not all frontal lobe functions. The location of the damage is intra-axial and the symptoms he developed at that time were a poverty of conversation, reduced empathy and increased impulsivity. He had reduced independent initiation of activities including conversations, taking medication, cooking and chores. He also had poor insight into his difficulties as indicated by a discrepancy between the self-report scores on the DEX (Wilson et al., 1997) (score=4) and an independent observer (score=47). This confirmed that LA did not yet have full insight into the level of his difficulties. He lived alone and had family close-by who visited him regularly. A community service provider supported him for 12 hours a week to manage his tenancy and daily living tasks, prompt him to use his memory aids and monitor his health.

MA

MA was a 26 year old woman who was reported to have suffered reflex anoxic seizures as a child and head injury in 2013 resulting in damage in the region of the basal ganglia. Her difficulties included apathy, social anxiety, lack of insight and difficulty keeping track of goals. She also experienced fatigue and had poor attention, impaired learning and impaired executive functions. She had some memory difficulties and it was reported that she used a pencil and paper calendar to help with this. She lived alone with some support from family members. Prior to the study she required substantial prompting to be more active in order to help anxiety, and increase social interactions and confidence. The service also provided support for her to complete activities of everyday living, attend appointments and manage her budget.

Design and procedure

The design of this study was A-B-A which is a withdrawal / reversal design; for each participant, memory performance was assessed on various tasks during a baseline, intervention and return to baseline phase. Three participants were included to offer an indication of the external validity of the study effect. The study was designed so that at least five data points (the minimum recommended) would be collected for each phase of the study. Memory performance was the primary dependent variable and was used to calculate the efficacy of the reminding technology intervention. Memory tasks were decided upon during a meeting between MM and the participants during which the participant’s everyday memory difficulties were discussed. Memory tasks included leaving the house daily, eating at regular times and keeping a note of their memory tasks. To measure memory performance in this study participants were asked to send a message after a meal or going for a walk, text and email the experimenter at set times and fill out their memory log each night. This memory log, text and email data was recorded by the experimenter. Attendance during meetings with the experimenter was also recorded. See table 1 for details of memory tasks for each participant.

The independent variable was the study phase or condition. Phase A_1 lasted for 14 days and was a baseline control condition during which participants were instructed to use their usual memory strategies. Phase B lasted for 14 days and was the intervention condition during which the smartphone and smartwatch were given to the participants along with training (see the Training section below). Phase A_2 (return to baseline) lasted for 14 days and was the return to baseline condition during which the intervention was removed. The beginning of each phase of the study was not randomised and corresponded with bi-weekly meeting times scheduled by the service around participants' schedules. Due to the nature of the intervention, which required training with and use of a smartwatch and smartphone, neither participants nor researchers could be blinded to the study phases. No blinded assessors were used to assess performance during each phase. The majority of the memory performance data was collected using automatic measures such as phone or email logs (see discussion).

The study procedure was as follows:

- Week one: Meeting with MM to establish memory tasks that participants found difficult (1 hour). Participants were asked to perform these tasks using their usual memory strategies.
- Week two: Meeting with MM to gather demographic and neuropsychological data (1 hour).

- Week three: Smartwatch (and supplementary smartphone) given to participants along with 30 minute training from MM (see appendix).
- Week four: Meeting with MM to gather user experience data with the smartwatch and deal with any technological issues with the watch and phone (1 hour).
- Week five: Smartwatch (and supplementary smartphone) taken back. Participants were asked to continue performing the memory tasks using their usual memory techniques.
- Week six: Meeting with MM to give participants study debriefing (30 minutes).
- Week seven and after: MJ met participants to complete neuropsychological tests that had not been completed as part of their usual care within the service. He met

Participants were asked to hand in their completed memory logs for that week at each meeting. If people forgot to attend the meetings or could not attend then this was noted by the experimenters and another meeting was scheduled at the soonest possible date. The assistant psychologist (MM) was available by phone to answer queries about the technology. A manual with the same information given during the training session was given to participants to take away with them and refer to as required (see Appendix for details about training with the smartwatch). When carrying out the study MM followed the procedure reported above. There were no changes made to the procedure during the course of the investigation and there was no independent assessor available to evaluate procedural fidelity.

Materials

The hardware that was given to participants was a Moto 360 smartwatch and a Samsung Galaxy S3 or Google Nexus 5 smartphone. The purpose of the study was to assess smartwatch use but the smartphones were provided because reminding watch software which allows the setting of a weekly schedule does not exist. The reminding software was Google Calendar which was already available on both phones and was synced to the smartwatch by the assistant psychologist at the service (MM). Participants were instructed to keep the smartphone on charge and connected to Bluetooth and store it in the same place where they charged the smartwatch. This allowed the watch and phone to sync every night so that the watch notifications would update.

Participants were given memory log sheets and asked to fill these out each evening. Memory log sheets were written by MM and MJ and when filling them out participants were asked to enter the tasks they were supposed to do that day, the time they were supposed to do them and the time they actually did complete the task (see Appendix).

Measures

Memory performance, the primary outcome measure was calculated by recording the participant's memory log completion, recording meeting attendance and by assessing the texts and emails. A point was given if the memory was filled out for that day, if the meeting was attended at the right time, and if a text or email was sent at the right time. There were at least 3 data points combined to create a memory performance percentage each day, for each participant. A secondary dependent variable was user experience captured using the TLX and UTAUT measures. It was of interest whether or not the participants could learn to use the technology and whether or not they found it acceptable. The acceptability and user experience with the smartwatch and smartphone were assessed using the NASA Task Load Index (TLX; Hart & Staveland, 1988). Assessment was also performed on eight domains from the unified theory of acceptance and use of technology (UTAUT; Venkatesh, Morris, Davis & Davis 2003). Finally feedback was obtained from a recorded post-hoc interview in which participants were asked about their experience using the technology. TLX asks about mental demand, physical demand, temporal demand, evaluation of performance, evaluation of effort needed to achieve that performance and level of frustration. These scores (each on a scale of 1 to 20) were reported separately and aggregated together to create an overall task load score. The UTAUT includes groups of items concerning the following: performance expectancy (expectancy that the tech will be useful for its purpose), effort expectancy (perceived effort needed to use it), attitude towards the technology, social influence (the influence of others on the use of the technology), facilitating conditions (the extent to which their environment facilitates use of the tech), self-efficacy (estimations of their own ability to use the technology), anxiety (levels of anxiety felt when using the tech) and behavioural intention (an indication of whether the participant is intending to use the tech in the next 6 months). Scores for each item (on a scale of 1 to 6) within each domain were aggregated to give overall scores for each domain at each time point.

Statistical Analysis

Appropriate statistical analysis was decided upon based on visual inspection of the results, as suggested by Parker, Cryer, and Byrns (2006). It was decided that the *d*-statistic by Hedges, Pustejovsky, and Shadish (2013) was appropriate to give an overall summary of the results and NAP (Parker & Vannest, 2009) was used to give an indication of the level of change in memory performance between pairs of adjacent phases. The TLX and UTAUT scores were reported descriptively.

Training

Training with the technology consisted of a 5-10 minute demonstration followed by an assessment lasting up to 20 minutes. MM set the reminders on the smartphone during a meeting with the participants. Once the reminders had been set on the smartphone, the smartwatch software automatically notified participants as long as the phone and watch were synced. Therefore training was given as a back-up in case there were issues with the device after they were in the participants' homes. The 5 to 10 minute demonstration covered switching the watch on and off, using the watch touchscreen and button interactions, charging the watch and smartphone, making sure that the Bluetooth was switched on for the smartphone and clearing notifications on the watch. The training with the watch also covered receiving reminders, getting back to home screen and accessing agenda. Following this training there was an assessment of use which lasted up to 20 minutes. Participants were asked to turn the watch off and on again, switch on the Bluetooth on the phone, syncing the phone to the watch, put the smartwatch on the wireless charger, clear watch notifications, return to the watch home screen (clock-face), access the watch 'agenda' screen using the touchscreen. Training and testing continued until perfect performance was achieved.

INSERT TABLE 1 ABOUT HERE

Results

Cognitive profiles of participants

Table 2 summarises the cognitive profile for the participants.

INSERT TABLE 2 ABOUT HERE

Efficacy

Quantitative Summary of Results

Visual inspection focuses on the six data features mentioned in Kratochwill et al. (2010): level, trend, variability, immediacy of the effect, overlap, and consistency of data patterns across similar phases. Taking the percentage of tasks completed successfully as a dependent variable, the visual inspection of the data suggested that the data show a considerable variability both within and across cases. Moreover, there were no clear baseline trends. The visual inspection of the data suggested that there may be a general (i.e., consistent) upward shift in level, that is, on average task completion seems to have improved. Nevertheless, the effects of introducing the device are not immediate or visually evident, but the effect of its withdrawal is more pronounced, especially for participants LA and MA. Figures 1, 2 and 3 show the percentage memory performance for each participant over the three study phases.

In order to obtain an overall quantitative summary of the results, taking into account the variation within and between cases, we computed the *d*-statistic by Hedges et al. (2013) in its version for multiple-baseline designs. We obtained two separate quantifications expressed in standard deviations: an overall change from A_1 to B and an overall change from B to A_2. For obtaining the results we used the "scdhlm" package for R (https://github.com/jepusto/scdhlm). For the A_1 vs B comparisons, the value adjusted for small-sample bias was $d = 0.20$ (SE = 0.22; 95% confidence interval: [−0.22 to 0.62], thus, including zero as a plausible value) and the variability between participants was estimated to be 36% of the whole data variability and autocorrelation 0.18. For the B vs. A_2 comparisons, the value adjusted for small-sample bias was $d = -0.85$ (SE = 0.29; 95% confidence interval: [−1.42, −0.28], not including zero as a plausible value) and the variability between participants was estimated to be 38% of the whole data variability and autocorrelation 0.17. Given that this *d*-statistic was developed to be equivalent to the one applicable to between-group studies, Cohen's (1992) benchmarks could be used for interpreting the results, indicating a small not statistically significant increase in task completion with the introduction of the device (i.e., for the A_1 vs. B comparisons) and a large statistically significant decrease in task completion with the withdrawal of the device (i.e., for the B vs A_2 comparisons)

Nonoverlap of All Pairs Analysis

The amount of data variability also suggested that a nonoverlap index such as NAP (Parker & Vannest, 2009) may be useful, given that it uses all the measurements and is applicable to data that do not present improving baseline trend. Moreover, there is evidence that NAP performs well for single-case data (Manolov, Solanas, Sierra, & Evans, 2011). For obtaining the results, we used the R code mentioned in Brossart, Vannest, Davis, and Patience (2014) and available at (https://dl.dropboxusercontent.com/u/2842869/Tau_U.R). In contrast to the website by Vannest, Parker, and Gonen (2011), in the "A vs. B" comparison, which is part of the Tau index (Parker, Vannest, Davis, & Sauber, 2011) implemented in the R code ties (i.e., equal values) are counted as whole overlaps and not as half an overlap. We preferred the R code to the website, given that the latter provided confidence intervals that do not include the NAP value and sometimes exceeding the bounds of the index. For interpreting them we followed Parker and Vannest's (2009) suggestion for labelling NAP in the range 0 to .65 as small difference, .66 to .92 as medium one, and .93 to 1 as large difference. Additionally, we combine probabilities associated with the NAP values following Edgington's (1972) additive method, implemented in the "SCDA" plug-in for R (Bulté & Oghena, 2012). The additive method has been shown to control Type I error rates and to be sufficiently powerful when there are three cases and at least 20 measurements included in the conditions being compared, when the *p* value is obtained via a randomization test (Heyvaert et al., 2016). Once again, just as with the application of the *d*-statistic, we combined the *p* values for the A_1 vs B comparisons separately from the *p* values for the B vs A_2 comparisons as these share the phase B data and thus are not independent (a requirement for combining probabilities and effect sizes; Cheung & Chan, 2004; Jones & Fiske, 1953). The results are presented in Table 3.

INSERT TABLE 3 ABOUT HERE

TS

Figure 1 shows that TS's memory performance, while variable, was at a high level in phases A_1 and B. His memory performance then decreased between B and A_2.

INSERT FIGURE 1 ABOUT HERE

NAP analysis indicated that TS's memory performance did not change from the first A phase to the B phase (NAP = −0.03, p = 1). NAP score between phases A_2 and B was −0.81, indicating a significant medium effect of phase (p<0.01).

LA

Figure 2 shows that LA's memory performance was also highly variable. Overall his performance seemed to be highest during the intervention phase and lowest during the return to baseline phase.

INSERT FIGURE 2 ABOUT HERE

NAP analysis indicated that LA's memory performance improved between the first A phase to the B phase, though this was not significant (NAP = 0.28, p = 0.2 (small effect of phase)).

NAP score between phases A_2 and B was 0.58 indicating a significant, small effect of phase change (p<0.01).

MA

Figure 3 shows that MA's memory performance was quite poor throughout the trial. Overall her performance was highest during the intervention phase and was consistently at floor level at the end of A_1 prior to introduction of the intervention, and for the majority of A_2 after the intervention was taken away.

INSERT FIGURE 3 ABOUT HERE

NAP analysis indicated that MA's memory performance did not change a lot between the first A phase to the B phase (NAP = 0.05; p = 0.81). NAP score between phases A2 and B was −0.36 indicating a small, significant effect of phase change (p < 0.05).

Usability and User Experience

It was also of interest to know whether or not participants could be supported to use the smartwatch successfully and whether or not they found it acceptable. During the training sessions participants LA and MA, who had previous experience using smartphones and touch screen technology, found it easy to learn how to use the devices. TS found it more challenging but he was able to use the technology by the end of the 30 minute training session. The participants reported that they were able to use the technology without difficulty after the training session. MM reported that MA occasionally failed to charge the device properly at times during the study. No participants reported that the technology stopped working; however participant LA reported that the watch stopped prompting when it was taken too far from the phone (outside his house). As the participants were told to keep the smartphone in their home, next to the smartwatch charger, it is likely that the smartwatch only worked for LA while he was in his home. A usability problem reported by MA was that she could not feel the vibrations given by the watch and so would often miss the notification until she looked at the written prompt presented on the watch face.

Table 4 shows mean scores for each individual TLX and UTAUT category for each participant. TLX results show that MA reported the highest total task load when using the device and TS experienced relatively minimal task load with LA falling somewhere in between. It is clear that participants LA and MA viewed their own performance when using the technology as average or poor and that participant MA felt that a moderate amount of effort was required to achieve this level of performance (10/20 for the effort item in the TLX). MA also reported relatively high mental demand when using the devices (10/20). However, the majority of the task load scores were low, especially when asked about physical demand, temporal demand and frustration. Only one item, for one participant (MA) was over 10/20 (17/20 for performance item indicating she rated her performance with the device poorly). Therefore the overall scores indicated that participants did not experience a high amount of task load when using the technologies for two weeks. The UTAUT results show that TS had a slightly better experience using the technology than LA and MA but all three give quite high scores on the UTAUT. Encouragingly, all three participants scored maximum points on the self-efficacy questions confirming that they believed they could use the system without any help from either an on-screen tutorial or a carer or family member. TS and LA indicated that they would use the smartwatch again within the next six months if it was available to them and MA said she did not intend to use it in the next six months. In contrast to his results on the TLX effort and mental demand scales, LA reported low scores on the effort expectancy questions in the UTAUT, indicating that he felt like it would take a lot of effort for him to become skilful at using the system.

INSERT TABLE 4 ABOUT HERE

Discussion

Efficacy

The results of the efficacy analysis show that introduction of the smartwatch did not lead to any statistically significant change in memory performance for any of the participants, with MA and LA experiencing a small increase in memory performance, and TS experiencing no change. Memory performance of all participants declined when the smartwatch was removed. This effect was statistically significant and was small for MA and LA and medium for TS. There was, therefore, evidence of a relationship between the presence of the smartwatch and primary outcome measure. However there are different possible interpretations of the meaning of the results.

One interpretation is that, while people were able to remember to perform these tasks prior to the introduction of the memory aid, they became reliant on using the watches and so had reduced memory performance when the intervention was removed. The anxiety that they will become reliant on memory aid technology has been expressed by participants in studies canvassing the attitudes of end users towards prompting technology. For example Baldwin, Powell and Lorenc (2011) reported that some people with memory difficulties after brain injury believed that relying on memory aids would lead to their memory becoming 'lazy' and that remembering things by themselves was a step forward. McGee-Lennon et al. (2012) reported that some older users would prefer to be given a content-free prompt which allows them to remember for themselves what the task was that they needed to do. While these attitudes and opinions about assistive technology may affect people's willingness to use memory aids or memory aid technology, there is very little evidence in the literature of a decline in memory performance when the intervention is removed. In fact, many studies which have investigated the efficacy of prompts from technology to compensate for memory have found that task performance remains higher than it was at baseline, even after the intervention is removed. For example, Wilson, Evans, Emslie and Malinek (1997) reported a mean baseline percentage memory performance of 37.05% for 15 neurologically impaired participants. This increased to 85.46% with introduction of the wearable NeuroPage intervention and reduced only slightly to 74.46% when the NeuroPage was taken away. This indicates that the use of the NeuroPage facilitated habitual performance of the memory tasks. A similar result was found by van Hulle and Hux (2006) when they investigated the efficacy of a watch based prompt. The participant who responded well to this intervention then continued to have a good memory performance after the intervention was removed. While the return to baseline performance in prompting technology efficacy studies is not always higher than the baseline performance, it is rarely substantially lower. This kind of result, in which the return to baseline performance is better than or at least equivalent to baseline performance, is the most common among other studies investigating the efficacy of prompting devices, even amongst those in which the intervention failed to improve performance (e.g. Wilson et al., 2001; Lemoncello et al., 2011; Stapleton, Adams & Atterton, 2007). Therefore the findings in the current study are contrary to the majority of findings in the literature.

Another explanation may be that the participants' motivation was higher during the first phase of the study than it was during the return to baseline phase. This may have been because the study was new to the participants in phase A and study stimuli such as increased contact with the brain injury services and memory aid logs were novel. Motivation may also have increased with the prospect of receiving the smartwatch and smartphone technology, especially given that the participants lived in a very deprived area. A disparity in motivation between the first and final phases of the study was reported by members of the service and the assistant psychologist who ran the study. If this is the case then it may have had an effect on memory performance, particularly performance of memory tasks which were associated with the study. For example, participant LA stopped filling out his memory logs after the second day of the return to baseline phase and reported that he "didn't feel like doing it anymore." This highlights the importance of motivation in the success of neuropsychological rehabilitation interventions particularly when participants are apathetic; participant LA was reported to have particular difficulties with initiation of everyday activities. It can be difficult to separate behavioural difficulties with apathy and low motivation from memory because both apathy and poor memory can prevent the completion of activities of everyday living. If it was the case that LA's difficulties performing everyday tasks were due to low motivation then this may explain his pattern of results because the prospect of receiving the smartphone may have had given him a motivation boost that resulted in better memory performance until the device was taken away.

The efficacy results from this SCED study, when all study phases are taken into account, indicate that the introduction of the intervention did have an effect on memory performance, although this was only significant on the removal of the device. The results are the first to detail the impact of a smartwatch prompting system on everyday memory performance for people with ABI. However the reasons for the pattern of results found here are open to interpretation. The results are also limited by the fact that a stable baseline was not reached in the A phase for TS and LA. This makes it difficult to analyse the trends in the data which may have given insights into the reason for the substantial drop in performance between the B and A_2 phases. More research is needed to establish the clinical efficacy of a smartwatch intervention for people with ABI. For example a future study could give people the intervention for a longer amount of time, have a longer baseline phase in or order to achieve stable memory performance results or have a second intervention phase in an A B_1 A B_2 A single case experimental design. The use of smartwatches could also be assessed in a larger group study in which differences in the efficacy of the intervention could be quantitatively compared between participants' with different cognitive profiles.

Usability

The secondary aim of this paper was to investigate the user experience of participants when given the smartwatch and smartphone. The TLX and UTAUT scores are quite similar for all three participants and the measures were only given to participants once. It is therefore necessary to interpret the findings with caution. The results indicate that it would be feasible to provide this technology in practice to people with brain injury in the community, with minimal training and support from a clinician, without requiring a great deal of mental, physical or time demand from the end users. As the task load results were most favourable for TS and LA, one could argue that clients with a cognitive profile of impaired memory and executive functioning but preserved intellectual functioning may benefit the most from this kind of assistive technology intervention. However with only three participants there are not enough results to draw any conclusions about which service users would make the best use of this type of intervention. Future researchers could aim to further understand how the technology use differs between users with different cognitive profiles.

Methodological Limitations

The study followed the majority of the RoBiNT recommendations for external validity in SCED studies (Tate et al., 2013). Participants' baseline characteristics and the therapeutic setting were described, the independent and dependent variables were defined and operationalised, raw data was provided for each study phase for each participant, appropriate data analysis was used and the trial was replicated with three participants. There was no measure of generalisation of memory ability because it was not expected that this compensatory strategy would have any long term effects on memory ability after the completion of the study.

Fewer of the recommendations for internal validity were carried out. Tate et al. (2013) refer in the RoBiNT to the need for at least three repetitions of the treatment effect. This can be achieved in an ABAB design, or multiple baseline design with three Baselines or tiers. The ABA design used here can only demonstrate the treatment effect twice (AB and BA), though as we included three participants this allowed the treatment effect to be examined on six occasions.It was not possible to blind the therapist and participants to the study condition because the smartwatches had to be provided with training at the beginning of the intervention phase. Furthermore, there was no use of blind assessors so an inter-rater reliability measure was not conducted. The lack of blinding of the experimenter was unlikely to cause bias because only automatic measures such as text and email logs were used to calculate memory performance. Only one of the memory performance measures relied on judgement from the experimenters and that was whether or not the memory logs had been filled out each night. Future studies investigating a tech-based intervention may benefit from randomisation of the study phase. The study may also have benefited from an independent assessor of treatment fidelity, to assess how the intervention was delivered by the experimenter (e.g. training with the smartwatches).

MM reported that, during the initial meeting with participants, it was difficult to identify memory tasks which the participants carried out regularly, but that they also regularly forgot. One reason for this could be that, perhaps because of their difficulties with memory, participants did not lead lives that required many demands on their memory in a short period. Another reason could be that remembering the tasks you often forget is a challenging task. It may also be the case that events that occur regularly become habitual and so are less likely to be forgotten. Perhaps it is unexpected or unusual events that catch people out, and these are difficult to predict, measure and control for in a research study. This issue highlights a challenge with performing efficacy research investigating memory performance as an outcome variable. Ideally, we want to measure the impact of the intervention on memory difficulties that would occur in everyday life (e.g. forgetting an appointment). However, it is often necessary to develop experimental memory tasks as proxies for memory tasks that occur outside of an experiment, as they can be more easily controlled and measured (e.g. send a text at a certain time).

Conclusion

This study aimed to investigate whether a smartwatch memory aid intervention was effective for, and could feasibly be used by, people with ABI living in the community. The results of an ABA trial with three participants provided some evidence supporting the effectiveness of the intervention; however future work is required to understand the pattern of memory performance more fully. The user experience results show that, within the two week window in which they were given the device, participants were able to use it without a great amount of effort and reported positive user experiences with the technology. This indicates that it would be feasible to introduce smartwatch reminding technology off-the-shelf into clinical practice.

References

Baldwin, V. N., Powell, T., & Lorenc, L. (2011). Factors influencing the uptake of memory compensations: A qualitative analysis. *Neuropsychological Rehabilitation*, *21*, 484-501.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). Single case experimental designs: Strategies for studying behavior change. Boston, MA: Pearson.

Brossart, D. F., Vannest, K., Davis, J., & Patience, M. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation*, *24*, 464-491.

Bulté, I., & Onghena, P. (2012). When the truth hits you between the eyes: A software tool for the visual analysis of single-case experimental data. *Methodology, 8*, 104-114.

Cheung, S. F., & Chan, D. K.-S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology, 89*, 780-791.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. London, UK: Psychological Corporation.

Edgington, E. S. (1972). An additive method for combining probability values from independent experiments. *Journal of Psychology, 80*, 351−363.

Hart, S. G., & Staveland, L. E.(1988). Development of NASA-TLX: Results of empirical and theoretical research. In N. Hancock & P. A. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: Elsevier.

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*, 324-341.

Heyvaert, M., Moeyaert, M., Verkempynck, P., Van Den Noortgate, W., Vervloet, M., Ugille, M., & Onghena, P. (2016, March 2). Testing the intervention effect in single-case experiments: A Monte Carlo simulation study. *Journal of Experimental Education*. Advance online publication. doi:10.1080/00220973.2015.1123667

Hubley A. M., & Jassal, S. (2006). Comparability of the Rey-Osterrieth and Modified Taylor Complex Figures using total scores, completion times, and construct validation. *Journal of Clinical and Experimental Neuropsychology, 28*, 1482–1497.

Jamieson, M., Cullen, B., McGee-Lennon, M., Brewster, S., & Evans, J. J. (2014). The efficacy of cognitive prosthetic technology for people with memory impairments: A systematic review and meta-analysis. *Neuropsychological Rehabilitation, 24*, 419-444.

Jones, L. V., & Fiske, D. W. (1953). Models for testing the significance of combined results. *Psychological Bulletin, 50*, 375-382.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single case designs technical documentation. In *What Works Clearinghouse: Procedures and standards handbook* (Version 1.0). Available at http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_scd.pdf

Lemoncello, R., Sohlberg, M. M., Fickas, S., & Prideaux, J. (2011). A randomised controlled crossover trial evaluating Television Assisted Prompting (TAP) for adults with acquired brain injury. *Neuropsychological Rehabilitation*, *21*, 825-846.

Manolov, R., Solanas, A., Sierra, V., & Evans, J. J. (2011). Choosing among techniques for quantifying single-case intervention effectiveness. *Behavior Therapy, 42,* 533-545.

McDonald, A., Haslam, C., Yates, P., Gurr, B., Leeder, G., & Sayers, A. (2011). Google Calendar: A new memory aid to compensate for prospective memory deficits following acquired brain injury. *Neuropsychological Rehabilitation*, *21*, 784-807.

McGee-Lennon, M., Smeaton, A., & Brewster, S. (2012). Designing home care reminder systems: Lessons learned through co-design with older users. In *proceedings of the 6th international conference on pervasive computing technologies for healthcare*, 49–56.

Parker, R. I., Cryer, J., & Byrns, G. (2006). Controlling baseline trend in single-case research. School *Psychology Quarterly, 21*, 418-443.

Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357-367.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*, 284-299.

Stapleton, S., Adams, M., & Atterton, L. (2007). A mobile phone as a memory aid for individuals with traumatic brain injury: a preliminary investigation. *Brain Injury*, *21*, 401–411.

Svoboda, E., Richards, B., Leach, L., & Mertens, V. (2012). PDA and smartphone use by individuals with moderate-to-severe memory impairment: Application of a theory-driven training programme. *Neuropsychological Rehabilitation*, *22*, 408-427.

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and N-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation, 23*, 619-638.

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Brlow, H., Horner, R… and Wilson, B. (2016) The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Archives of Scientific Psychology, 4*, 1-9.

Van Hulle, A., & Hux, K. (2006). Improvement patterns among survivors of brain injury: Three case examples documenting the effectiveness of memory compensation strategies. *Brain Injury*, *20*, 101–109.

Vannest, K.J., Parker, R.I., & Gonen, O. (2011). Single Case Research: Web based calculators for SCR analysis. (Version 1.0) [Web-based application]. College Station, TX: Texas A&M University. Retrieved on Friday 25th September 2015 from http://www.singlecaseresearch.org/calculators/nap

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*, 425-478.

Warnock, D., McGee-Lennon, M., & Brewster, S. (2013). Multiple notification modalities and older users. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1091-1094). Association for Computing Machinery.

Wechsler, D. (2008). *Wechsler adult intelligence scale-fourth*. San Antonio, TX: Pearson

Wechsler, D. (2009). *Wechsler memory scale (WMS-IV).* New York, NY: The Psychological Corporation.

Wechsler, D. (2011). *Test of Premorbid Functioning. UK Version (TOPF UK)*. London, UK: Pearson Corporation.

Wilson, B. A. (1991). Long-term prognosis of patients with severe memory disorders. *Neuropsychological Rehabilitation, 1*, 117-134.

Wilson, B. A., Cockburn, J., & Baddeley, A. D. (1991). *The Rivermead behavioural memory test*. Bury St Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Emslie, H. C., Quirk, K., & Evans, J. J. (2001). Reducing everyday memory and planning problems by means of a paging system: A randomised control crossover study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *70*, 477–82

Wilson, B. A., Evans, J. J., Alderman, N., Burgess, P. W., & Emslie, H. (1997). Behavioural assessment of the dysexecutive syndrome. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 239-250). London, UK: Psychology Press.

Wilson, B. A., Evans, J. J., Emslie, H., & Malinek, V. (1997). Evaluation of NeuroPage: A new memory aid. *Journal of Neurology, Neurosurgery & Psychiatry*, *63*, 113-115.

Wilson, B. A., Shiel, A., Foley, J., Emslie, H., Groot, Y., Hawkins, K., & Watson, P. (2005). *Cambridge test of prospective memory (CAMPROMPT)*. San Antonio, TX: Pearson Assessment.

Appendix

Smartwatch training

The training will consist of a 5-10 minute demonstration followed by an assessment lasting up to 20 minutes. Once the reminders have been set on the smartphone, the smartwatch software will automatically notify participants as long as the phone and watch are synced. This means training is given as a back-up in case there are issues with the device.

5 – 10 minute demonstration

- Switching on and off, touchscreen and button interactions, charging, Bluetooth
- Clearing notifications, receiving reminders, getting back to home screen
- Accessing agenda

Assessment of use – up to 20 mins

- Turn on / off
- Switching on Bluetooth on phone, syncing phone to watch
- Put smartwatch on charger
- Clear notification
- Return to home screen
- Access agenda using touchscreen

The experimenter will be available by phone to answer queries about the technology. A manual with the same information given during the training session will be given to participants to take away with them and refer to as required.

Smartwatch Manual

Page 1

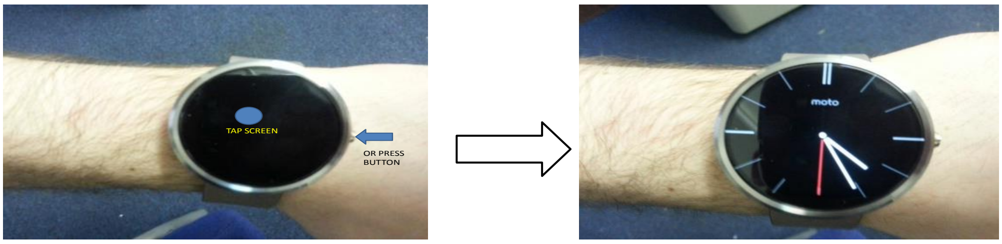
Turning smartwatch on and charging

The smartwatch will require charging every one or two nights depending on how much you use it. We would recommend that you charge the watch every night by placing it in the stand as shown.

Watch charging picture

To switch on the smartwatch press the button on the side once. When you are wearing the smartwatch it should also come on when you turn your wrist and look at the clock face. If it does not come on then try pressing the button or tapping the screen.

Blank -> clock face

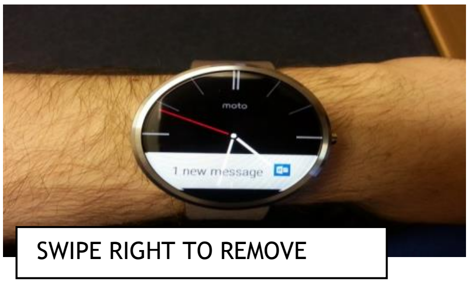


Page 2

Selecting and deleting notification

To select notifications just tap them on the screen. To remove a notification swipe it to the right as shown.

Notification picture

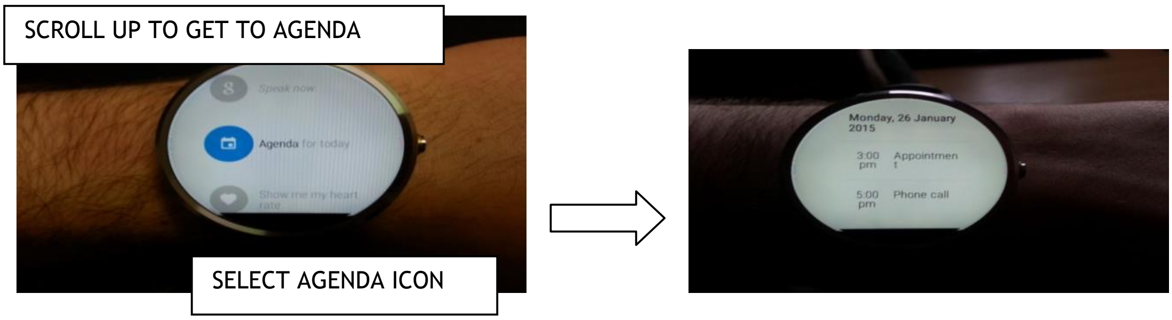


Sometimes heart monitoring, number of steps or email information comes up on the watch. If this happens please remove the notifications by swiping them to the right.

Accessing agenda: your reminders should appear on the watchface throughout the day. If you look at the watchface and cannot see any reminders then you can access them by viewing the 'agenda'. To access agenda simply tap the watchface and scroll down the list to agenda as shown.

Tap the agenda icon to see your events.

Picture – menu, arrows to scroll to agenda and tap icon. -> Agenda screen



Page 3

Setting an alarm

It may be helpful for you to set an alarm to remind you to do a task at a set time. To set an alarm tap the watchface until you get to the menu screen and scroll down to set an alarm option. Press the icon.

Scroll to the time you would like to set and select it – the watch will automatically set a one-off alarm for this time.

Scroll menu screen (selection) -> select time screen (selection) -> alarm setting screen

To remove the alarm scroll to show alarms, edit alarm and delete as shown.

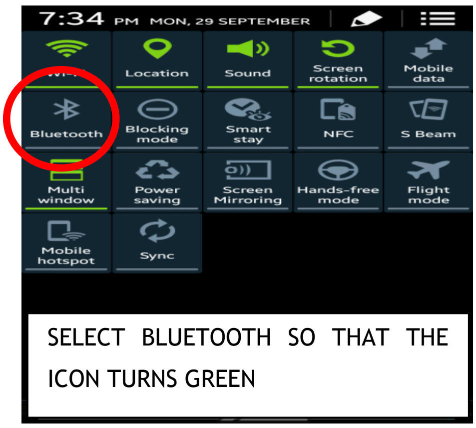
Scroll menu (select show alarms) -> edit alarm selection -> delete alarm selection

Page 4

Smartphone use

You may have been provided with a smartphone for this study. If you have been given a smartphone then please ensure that Bluetooth is activated (as shown)and that the phone is in near the smartwatch at some point every day.

Image of S3 Bluetooth selection



If possible we recommend keeping the phone on charge next to where you charge the watch every night.

Memory Log:

If you feel that memory difficulties might make it difficult to remember information such as whether or how often you use memory aids we would encourage you to ask a family member, friend or supporter to help.

Date_______

You have indicated that you would like to try remember the following events which you often forget. Please indicate whether or not you remembered to do these tasks today. If you cannot remember whether you did the tasks or not then please ask a family member, friend or supporter to help.

Memory tasks*	**At what time were you supposed to do this task?****	**What time did you do this task?**
Memory task 1		
Memory task 2		
Memory task 3		
Memory task 4		
Memory task 5		

***memory tasks will be decided during discussions with participants after they have given their consent to take part in the study.**

**** individual items on this table may be altered depending on the type of task selected by participants (e.g. some participants may need a prompt to help them stop a task rather than start one, for example watching T.V.)**

Other notes or comments:

Tables

Table 1. Details about the memory tasks on which each participant was assessed.

Initials	Daily tasks
TS	• Text experimenter once daily • Send email to experimenter once daily • Fill out memory log each evening • Come to an appointment once weekly
LA	• Send text after dinner daily • Send email after going for a walk daily • Come to an appointment once weekly • Fill out memory log each evening
MA	• Send a text after lunch daily • Send an email after going for a walk daily • Come to an appointment once weekly • Fill out memory log each evening

Table 2. Cognitive profile on tests of intelligence, memory and executive function for the study participants.

Test	TS (45 years old)	LA (61 years old)	MA (39 years old)
WAIS-IV perceptual reasoning score	110 (high average)	105 (average)	73 (borderline impaired)
WAIS-IV verbal comprehension score	-	89 (low average)	83 (low average)
WAIS-IV processing speed score	-	79 (borderline impaired)	59 (impaired)
TOPF estimated full-scale pre-morbid IQ	105 (average)	94 (average)	-
RBMT percentile rank (95% CI)	2 (0.4-10) (impaired)	2 (0.3-9) (impaired)	-
CAMPROMPT score (percentile rank)	12 (<5) (borderline impaired)	-	-
WMS-IV auditory memory delayed score	-	81 (low average)	87 (low average)
WMS-IV visual memory delayed score	-	81 (low average)	81 (low average)
ROCF score	27 (impaired)	-	28 (impaired)
DKEFS verbal fluency percentile rank (95% CI)	2.3 (0.2-14.9) (impaired)	15.9 (4.7-37.4) (low average)	1 (0-22.2) (impaired)

DKEFS letter number switching percentile rank (95% CI)	74.8 (18.4-98.7)(high average)	15.9 (1-62.4) (low average)	4.8 (0.1 to 46.1) (borderline impaired)
BADS key search profile score	4 (average)	4 (average)	-
BADS zoo map profile score	2 (low average)	0 (impaired)	-
BADS age corrected score	-	93 (average)	83 (low average)

Key:

WAIS-IV = Wechsler Adult Intelligence Scale version 4

TOPF = Test of Pre-morbid Functioning

RBMT = Rivermead Behavioural Memory Test

CAMPROMPT = Cambridge Test of Prospective Memory

WMS-IV = Wechsler Memory Scale version 4

ROCF = Rey-Osterrieth Complex Figure

DKEFS = Delis-Kaplan Executive Functioning System

BADS = Behavioural Assessment of the Dysexecutive Syndrome

Table 3. Nonoverlap of all pairs for each two-phase comparison for the three participants.

Labels are obtained according to the benchmarks provided by Parker and Vannest (2009). Statistical significance values were combined using Edgington's (1972) additive method.

Participant	**Comparison**	**NAP**	**Label**	***p* value**	**combined *p***
TS	A_1B	−.03	deterioration	1.000	
LA	A_1B	.28	small	.2040	.83847
MA	A_1B	.05	small	.8064	
TS	BA_2	−.81	medium	.0002	
LA	BA_2	−.58	small	.0062	.00001
MA	BA_2	−.36	small	.0340	

Table 4. TLX and UTAUT scores on each category for TS, LA and MA.

Lower scores in the TLX indicate lower task load and higher scores in the UTAUT indicates a better user experience. TLX items are out of 20, the total is out of 120. UTAUT items are out of 6, the total is out of 174.

TLX domain	TS	LA	MA	UTAUT domain	TS	LA	MA
Mental Demand	5	3	10	Performance Expectancy	5.75	4.38	4.75
Physical Demand	1	3	1	Effort Expectancy	5.75	1.63	5.5
Temporal demand	1	2	2	Attitude	5.38	4.5	3.5
Performance	4.5	10	17	Social Influence	4.33	3	2.66
Effort	1	3	10	Facilitating Conditions	4.33	2.33	4.33
Frustration	1	3	1	Self-Efficacy	6	6	6
Total score	**13.5**	**24**	**41**	Anxiety	6	5.25	4.75
				Behavioural Intention	4.33	5.67	1
				Total score	**154.5**	**120**	**122**

Figures

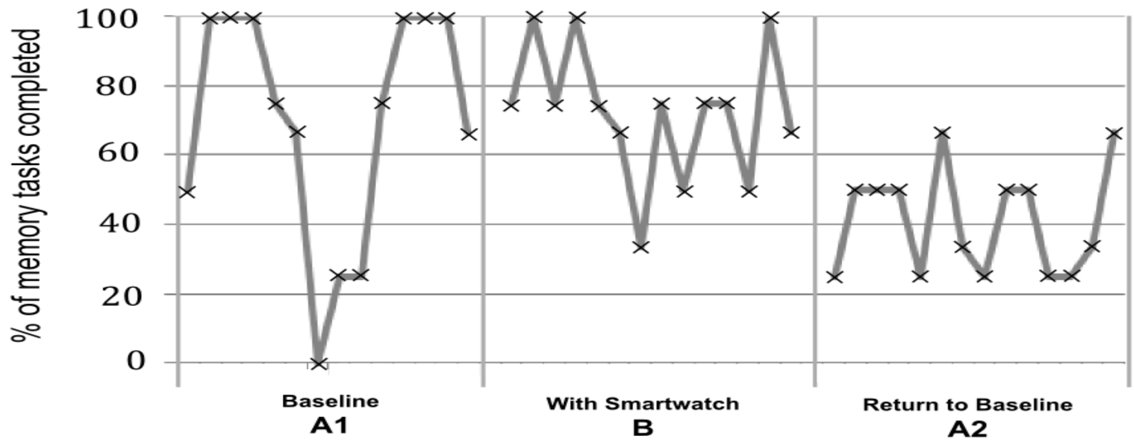


Figure 1. Percentage of memory tasks successfully completed in each study phase by participant TS.

The Y axis shows percent performance and X axis shows study day (each data point (x) in the figure represents one day in the study).

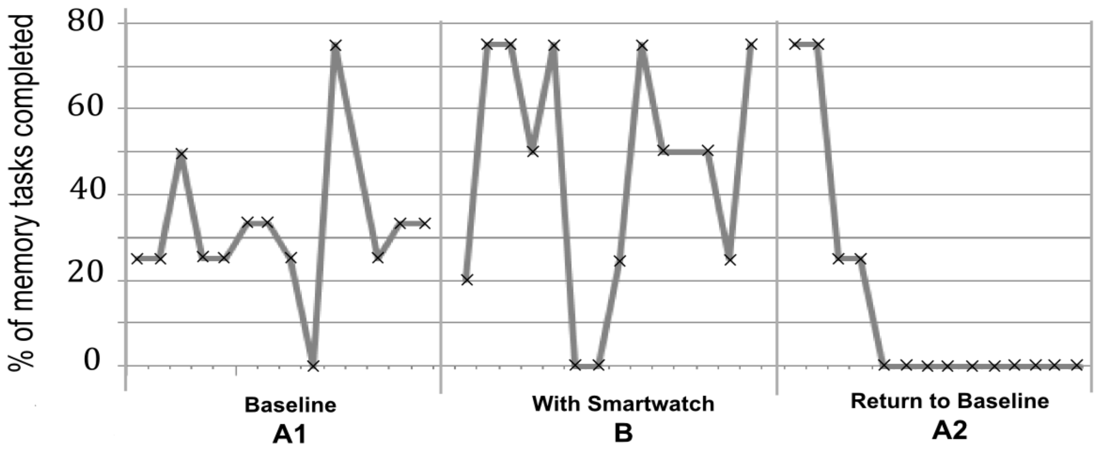


Figure 2. Percentage of memory tasks successfully completed in each study phase by participant LA.

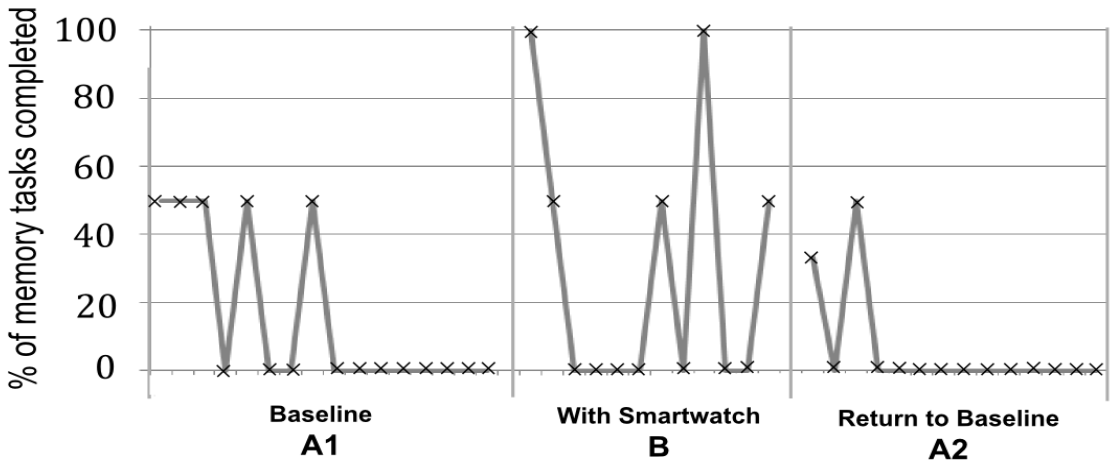


Figure 3. Percentage of memory tasks successfully completed in each study phase by participant MA.