Usability Evaluation

Group 3

Executive Summary

The aim of this usability evaluation was to test the application meets the specified usability requirements. Four people consisting of friends and family completed 4 tasks in a controlled useability study on a partially functioning PowerPoint prototype of the application. At the conclusion of the tasks, participants completed a SUS questionnaire. Several useability issues were found. The most severe of which was confusion in locating the person with dementia. The average SUS score was 82.5. Some limitations of the study were that none of the participants actually gave care to a person with dementia. The sample size was very small and some issues with Google Forms meant that some of the responses to the 1st post-study questionnaire have not been archived.

Context of Use

The application is intended to assist care givers of persons with dementia by providing tracking information that can be used to locate the person with dementia and a schedule that combines with the tracking to allow the caregivers to ensure that their person with dementia is not wandering or lost. The usability evaluation evaluated the tracking and location function of the app as well as adding, removing and editing plans from the schedule. The user group consisted of people between the ages of 18 and 64 most of whom are familiar and comfortable with smartphones. Unfortunately, all participants do not give care to a person with dementia so do not fall within the target demographic of the application.

Usability Test Tasks

Four tasks were tested.

Task 1: Find the current location of the person with dementia.

Reason for choosing the task: Locating the person with dementia is a core user task. The main way the app assists care givers support their PwD is via monitoring their location.

The usability requirements of this task specify that 80% of the participants should be able to successfully locate the PwD without assistance. The average number of errors should not exceed 3. The average difficulty rating should not exceed 3.

Task 2: Add to schedule a doctor appointment on Tuesday at 12:00 PM.

Reason for choosing the task: Keeping the person with dementias schedule is a core user task. Adding plans is a primary scheduling function of the application.

The usability requirements of this task specify that 80% of the participants should be able to successfully add a doctor's appointment to the schedule without assistance. The average unassisted completion time should not exceed 20 seconds. The average number of errors should not exceed 3. The average difficulty rating should not exceed 3.

Task 3: Remove the plan for going to the store on Wednesday at 10:30 AM.

Reason for choosing the task: Keeping the person with dementias schedule is a core user task. Removing cancelled plans from the schedule is a primary scheduling function of the application.

The usability requirements of this task specify that 80% of the participants should be able to remove the plan from the schedule without assistance. The average unassisted completion time should not exceed 10 seconds. The average number of errors should not exceed 3. The average difficulty rating should not exceed 3.

Task 4: Change the plan's time to 02:00 PM for visiting friend's house on Thursday.

Reason for choosing the task: Keeping the person with dementias schedule is a core user task. Being able to alter existing plans is an important scheduling function of the application. Technically not a required functionality since users could remove and then add the altered task but since it is expected that plans will be altered fairly regularly, this function is important and so should be tested.

The usability requirements of this task specify that 80% of the participants should be able to change the plan without assistance. The average unassisted completion time should not exceed 20 seconds. The average number of errors should not exceed 3. The average difficulty rating should not exceed 3.

A summary of the targets is provided in Table 1.

***Table 1**: Target Criteria Values*

	Unassisted Task Completion Rate	Unassisted Completion Time (sec)	Errors	Assists	Difficulty Rating	SUS Score
1. Locate PwD.	80%	10	1-3	1-3	1-3	
2. Add plan to schedule.	80%	20	1-3	1-3	1-3	
3. Remove plan from schedule.	80%	10	1-3	1-3	1-3	
4. Change plan to 2pm.	80%	20	1-3	1-3	1-3	
Overall	80%	60	1-3	1-3	1-3	68

**Note: SUS score for entire interface (not available for individual tasks).*

Method

The participants were tested in a usability laboratory equipped with logging software which recorded the participant's screen and a head shot from a webcam. 3 of the tests were conducted in person and 1 was conducted in a Zoom call.

After completing the informed consent procedure, they were administered with a screener questionnaire to determine their age, comfortability with smartphones, whether they gave care to a person with dementia and if so what stage of dementia, how many hours to the spend caring for them per day, how often do they wander off and how difficult is it to care for them. They were instructed to try to complete the tasks with assistance and that any problems they had were due to the interface and not them. They were asked to think aloud during the test. After finishing the test, participants were asked to complete the SUS questionnaire and respond to some general open-ended questions about what they did and did not like about the interface. The test tasks can be seen in the proceeding section.

Participant Characteristics

Four people consisting of family and friends participated the usability evaluation. See Table 2 for a summary of their characteristics. All participants did not, at the time of the evaluation, give care to someone with dementia. This means that none of the participants strictly fall within the target demographic of the application and because of the small number of participants, these results may not be representative of that demographic. All participants except for 1 reported being comfortable with smartphones which is a characteristic of the target demographic and so their results may still be somewhat representative. However, 2 of the participants were between the ages of 18 and 24, 1 between the ages of 35 and 49 and 1 between 50 and 64 which means most participants were younger than what would be ideal which might affect the test results as younger people tend to more easily operate software.

***Table 2**: Participants' characteristics*

	Age	Comfortability with Smartphones	Gives Care to a Person with Dementia	What Stage of Dementia?	How much time spent caring for PwD per day?	How often do they wander off?	How difficult is it to care for PwD?
P1	18-24	5	No	N/A	N/A	N/A	N/A
P2	18-24	5	No	N/A	N/A	N/A	N/A
P3	50-64	2	No	N/A	N/A	N/A	N/A
P4	35-49	4	No	N/A	N/A	N/A	N/A

**Note: Last 4 questions only apply if participant does give care to a person with dementia. Marked 'N/A' if they don't.*

Performance Data

Task 1 was to locate the person with dementia using the map function of the interface. The results can be seen in Table 3. All participants were able to complete the task without assistance giving the task a completion rate of 100%. None of the participants met the under 10 second requirement for Task completion time. This is most likely due to unfamiliarity with the prototype and evaluation process as well as an underestimation of the time required to complete the task in the first place. Encouragingly, all participants were able to complete the task without error or assistance which meets the usability requirements specified of between 1 and 3.

***Table 3**: Performance Data for Task 1 (Locate PwD)*

	Unassisted Task Completion	Errors	Assists	Task Completion Time (sec)
P1	100	0	0	15
P2	100	0	0	23
P3	100	0	0	30
P4	100	0	0	15

P4 said that he couldn't easily identify where the person with dementia was because it wasn't obvious to them that the map needed to be clicked to see the PwD's location.

P3 took very long to perform the task. However, this was due to unfamiliarity with the prototype and the evaluation process and once the participant was confident in what they had to do the prototype matched their expectations.

Task 2 was to add a doctor's appointment to the schedule. The results can be seen in Table 4. All participants were able to complete the task without assistance giving the task a completion rate of 100% which falls well within the 80% completion rate specified in the usability requirements. All participants were able to do this without making a mistake or requiring assistance which meets the usability requirements. All participants took longer than the 20 seconds specified in the usability requirements.

***Table 4**: Performance data for Task 2 (Add plan)*

	Unassisted Task Completion	Errors	Assists	Task Completion Time (sec)
P1	100	0	0	40
P2	100	0	0	54
P3	100	0	0	77
P4	100	0	0	40

P3 took the longest to complete the task because they had forgotten what the task was. However, once the they were reminded of the task, they required no assistance and completed it in a timely manner.

P1 took 40s not because of difficulty with the task but instead because they went into lots of detail of their thought process while performing the task.

The 20s usability requirement seems to have been misinformed and unrealistic. The data might be skewed due to the fact that being the second task perhaps participants were not yet comfortable with the prototype and hadn't gotten into the grove, so to speak, of the test.

Task 3 was to remove the plan to go to store from the schedule. The results can be seen in Table 5. All participants successfully completed the task without assistance which meets the 80% usability requirement and meets the usability requirement of between 1 and 3 errors and assists.

***Table 5**: Performance Data for Task 3 (Remove Plan)*

	Unassisted Task Completion	Errors	Assists	Task Completion Time (sec)
P1	100	0	0	15
P2	100	0	0	15
P3	100	0	0	14
P4	100	0	0	15

All participants were slightly above the desired 10 seconds for Unassisted Completion Time. However, since all participants were very similar and it isn't much over it can probably just be chalked up to unfamiliarity with the interface and that someone who regularly uses it would not have any trouble whatsoever in removing a plan from the schedule.

Task 4 was to change the plan to visit a friend's house in the schedule. The results can be seen in Table 6. All participants were able to complete the task without assistance giving the task a completion rating of 100% which meets the 80% specified in the usability requirements. None of the participants made an error while performing this task nor did they require assistance which meets the usability requirement of 1-3 errors and 1-3 assists. Only 1 participant met the 20 seconds completion time requirement, 2 weren't that far off but 1 took almost twice that amount of time.

***Table 6**: Performance Data for Task 4 (Change Plan)*

	Unassisted Task Completion	Errors	Assists	Task Completion Time (sec)
P1	100	0	0	20
P2	100	0	0	38
P3	100	0	0	24
P4	100	0	0	25

P2 took the longest with 38 seconds to complete the task. This was due to going back and forth between the prototype and the evaluation tasks questionnaire so this can probably be ignored especially since they gave the task a difficulty rating of 1.

Attitude Data

SUS scores were calculated for all 4 participants. The SUS score for P1 was 70, for P2 it was 90, for P3 it was 90 and for P4 it was 80. Overall, the participants found the interface to be easy to use but more specific conclusions are difficult to draw due to the small sample size. According to this data there seems to be very little correlation between time taken to complete a task and difficult attributed to that task. More data would need to be gathered to form more informative conclusions. The average SUS score was 82.5 which meets the useability requirement target specified of 68 or above.

An Average difficulty rating for each task was calculated. Task 1 was rated a 1 by P1, a 2 by P2, a 2 by P3 and a 4 by P4. This gives it an average difficulty rating of 2.25 which meets the usability requirement of 3 or lower. Task 2 was rated a 2 by all participants giving it an average rating of 2 which meets the usability requirement. Task 3 was rated a 2 by P1 and a 1 by all other participants giving it an average rating of 1.25 which meets the usability requirement. Task 4 was rated a 3 by P1, a 1 by P2, a 2 by P3 and a 1 by P4 giving it an average rating of 1.75 which meets the usability requirements.

All 4 participants were asked to give a rating of how likely it would be that they would recommend this prototype to a friend or colleague. However, due to technical difficulties with Google Forms, 2 of the responses have been lost and marked 'N/A'. Interestingly the 2 responses that have been recorded do not match with the SUS score that the participant gave the prototype. P1's SUS score was 70 and they rated the likelihood they would recommend to a friend was an 8 but P4 who gave a SUS score of 80 only rated it a 7. Clearly more data is needed to draw anything conclusive but perhaps it is because P4 rated task 1 as a 4 difficulty which perhaps colored their experience of the prototype leading to a lower recommendation.

***Table 7**: Satisfaction Data of the four participants*

	Task 1 Difficulty Rating	Task 2 Difficulty Rating	Task 3 Difficulty Rating	Task 4 Difficulty Rating	SUS Score	Recommend to a Friend
P1	1	2	2	3	70	8
P2	2	2	1	1	90	N/A
P3	2	2	1	2	90	N/A
P4	4	2	1	1	80	7

**Note: 1 = Very Easy and 5 = Very Hard for Difficulty Rating. 1 = Very Unlikely and 10 = Very Likely to recommend to a friend. Some records incomplete due to technical difficulties with Google Forms.*

Recommendations

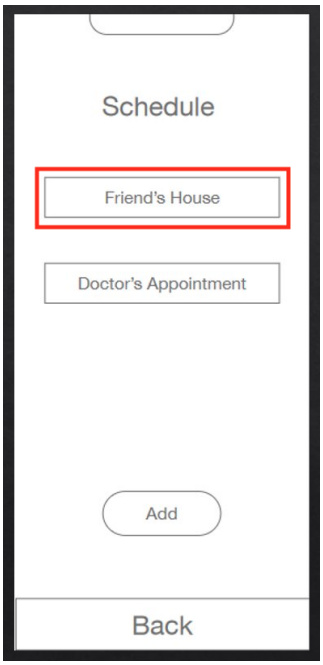
The results of the study revealed the following useability problems:

Problem 1: Map

Problem 1 Summary: Some users found it confusing that the map had to be clicked before locating the person with dementia.
Severity Assessment: Caused some users dissatisfaction and confusion with the map. If the user cannot overcome this problem, it will prevent them from locating the person with dementia entirely.
Severity Rating: Major, Locating the person with dementia is a core user task.
Recommendation: Remove the need to click the map to locate the person with dementia.

***Figure 1**: Home Page (Red square highlights problematic element).*

Problem 2: Editing Plans



Problem 2 Summary: Some users were confused on how to edit plans.
Severity Assessment: Caused some users dissatisfaction with the interface. If the user cannot overcome this problem, they could simply ignore the erroneous plans. However, this is very less than ideal.
Severity Rating: Minor, the user has workarounds for the problem.
Recommendation: Remove the need to click a plan to show the buttons to remove and edit the plan perhaps with a bin and wrench icon at the top of the Schedule page with radio button like icons next to each plan to signify if a task is selected or not.

***Figure 2**: Schedule Page (Red square highlights problematic element)*

Limitations

Participants for the evaluation didn't match the primary user group of the software as a result of limited options and time frame. On top of this the size of the sample group was incredibly small. These factors mean that the data collected in this evaluation is probably not completely representative of the primary user group. The tasks as they were expressed in the evaluation tasks questionnaire didn't provide much context for the participants which lead to some confusion while conducting the evaluation.

Our reliance on Google Forms to archive the participants' responses has resulted in some data being lost due to technical difficulties. In future the responses must be backed up in some way to reduce the likelihood of lost data.

Raw Data

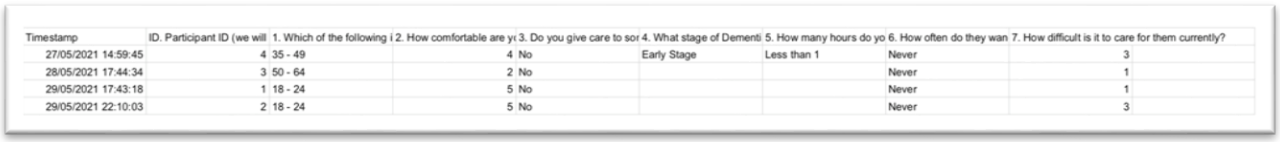
***Figure 3**: Screenshot of Post-Study Questionnaire 1 Responses Spreadsheet.*

Timestamp	1. Participant ID (we will h	2. What did you like most	3. What did you like least	4. What are your thoughts	5. What are your thoughts	6. Did you understand the	7. How likely is it that you would recommend this prototype to a friend or colleague?		
27/05/2021 15:57:04	4	The buttons are big so it i	The general design of the	These tasks are not too m	The menu is quite hard to	Yes, however the map sh	7		
27/05/2021 17:55:38	1	The simplicity when I was	The design, I think it can l	I found it easy to do, but I	The menus were very cle	Yes, The pin made the loc	8		

***Figure 4**: Screenshot of Post-Study Questionnaire 2 (SUS) Responses Spreadsheet.*

Timestamp	ID. Participant ID (we will	1. I think that I would like	2. I found the system unn	3. I thought the system wa	4. I think that I would nee	5. I found the various func	6. I thought there was too	7. I would imagine that mo	8. I found the system very	9. I felt very confident usin	10. I needed to learn a lot	Any other comments about your experience with the prototype:
27/05/2021 15:25:16	4	3	1	3	3	3	1	4	0	4	0	
27/05/2021 17:02:07	1	3	1	3	1	3	1	3	2	3	2	The prototype is well made but can be improved design wise.
28/05/2021 18:03:10	3	3	0	3	0	3	0	4	0	3	0	I can see that this system would be helpful for care givers.
29/05/2021 22:27:07	2	4	2	3	0	4	0	4	0	3	0	Fantastic

***Figure 5**: Screenshot of Demographic Questionnaire Responses Spreadsheet.*



Timestamp	ID. Participant ID (we will	1. Which of the following i	2. How comfortable are yo	3. Do you give care to sor	4. What stage of Dementi	5. How many hours do yo	6. How often do they wan	7. How difficult is it to care for them currently?
27/05/2021 14:59:45	4	35 - 49	4	No	Early Stage	Less than 1	Never	3
28/05/2021 17:44:34	3	50 - 64	2	No			Never	1
29/05/2021 17:43:18	1	18 - 24	5	No			Never	1
29/05/2021 22:10:03	2	18 - 24	5	No			Never	3

***Figure 6**: Screenshot of Evaluation Tasks Difficulty Rating Responses Spreadsheet.*

Timestamp	ID. Participant ID (we will	Task 1 - Please rate the d	Task 2 - Please rate the d	Task 3 - Please rate the d	Task 4 - Please rate the difficulty of this task:	
27/05/2021 15:15:50	4	4	2	1	1	
27/05/2021 16:59:03	1	1	2	2	3	
28/05/2021 17:51:29	3	2	2	1	2	
29/05/2021 22:16:41	2	2	1	1	1	

***Figure 7**: Screenshot of Informed Consent Form Responses Spreadsheet.*

Timestamp	Email address	I am over the age of 18 ar	2. In relation to this projec	Signature - Your name and date
27/05/2021 14:55:12	frankty123@gmail.com	Yes	I agree to be observed by	CHHAY 27/05/2021
27/05/2021 16:46:59	memari2255@gmail.com	Yes	I agree to be observed by	Mohammad Memari - 27th May 2021
28/05/2021 17:40:53	lacey.jd@bigpond.com	Yes	I agree to be observed by	John Lacey 28-05-2021
29/05/2021 22:03:46	103610596@student.swir	Yes	I agree to be observed by	Dooley Every 29/05/2021

***Figure 8**: Handwritten notes taken during the evaluation.*