Team bad HABT P4: Bus Buddy Prototype Evaluation

Thomas Gable, Blake Gruber, Harsahth JR, & Alan Zhang

Introduction

The following report is a discussion of project part IV for the Fall 2013 HCI class for team "bad HABT." The report details the process in designing and evaluating a bus and student tracking system for parents of elementary school children. This paper demonstrates the application of knowledge gathered in parts 1, 2, and 3 of this project as well as lectures and readings in an effort to evaluate the system. The summarization covers a description of the evaluation techniques, a discussion of why these techniques and the materials used for the evaluation were chosen, a report and discussion of the results, and then a discussion of the implications of these results for the design according to the usability specifications. Finally the paper discusses what should be improved on the prototype in response to the implications learned from the evaluation, as well as an overall conclusion of what was gained in this four-part project.

Bus Buddy Overview

Goal and users. The goal for this project is to transmit information to parents concerning the transportation of their elementary school children on buses from home to school and back. It should be noted that while all users involved in this ecosystem are not being included at this time, this prototype it is designed in such a way that the data and displays would be integrated within the larger ecosystem. The users of this particular system are parents of elementary school aged children who ride the bus so the age range of parents is widely variable, as are the user characteristics (e.g., job status, income, education level). The common denominator for these users is their concern about the whereabouts

and safety of their children, and in getting their children to the bus stop or picking them up on time.

Tasks. There are 3 main tasks and contexts in which the parents will be using the interface and should quickly be discussed in order to fully understand the design decisions made in the prototype and evaluation.

The first task is getting children to the bus stop, where a parent is attempting to get a young child up and ready for school on time. Their task is to keep track of the time and be aware of when the bus is arriving, and at some point making a decision if the child will make the bus. The context this takes place in is morning, in a household with children where there may also be more than one child who may or may not be getting on the same bus. Concurrently parents may also be trying to get ready for their own day and may be attempting to get to the workplace or run errands after dropping their children off.

Another task is picking up the child or children from the bus stop or ensuring that they arrived at the correct place safely. This means that the parent must reach the bus stop before the bus' arrival time, or if ensuring their safe arrival at a destination they must somehow confirm arrival without being there. While performing this task, parents may be driving home from their own workplace, at a store, or at home. Before parents leave to get their child they may be engrossed with work and in need of reminders. In addition parents may have other events to go to directly after picking up the child and delays or waiting could cause unwanted disruptions in their plans.

The final major task parents are involved in takes place outside of the times of dropping off or picking up their children. During this time parents may want to track if their child has arrived at school yet or make plans for where the child is going after school on certain days, like a friend's house, and convey this to bus drivers, teachers, and other parents. Parents may have to perform actions such as quickly changing plans, or checking plans for that day or viewing the status of bus and student. The context that this takes place is extremely variable as it could occur anytime in the day when parents are going about their daily schedule and may want information regarding current whereabouts of the child/bus in different environments.

Design implications. These tasks and the context led to the design implications on which the initial designs were based on and what was used to create the usability specifications for evaluation of this prototype. These implications include the ability to track the children and convey information to parents in real time, have a simple interface that is easily and quickly navigated, have an adaptable interface based on parent's preferences, allow for changes of schedules both current and in the future, and that any messages should be carefully designed and customizable (e.g., got on or off the bus, bus was delayed etc.). Additionally it should be considered that most parents will not want another device to keep track of, which lead to a mobile interface, and in order to be successful it must be cost-effective and scalable so school systems would be willing to implement, this meant a largely software based implementation.

Usability specifications. Based on these design implications a number of usability specifications were determined that would need to be considered during evaluation of the

prototype. The main design implications determined in Part 1 of the project were the ability to track the children and convey the information to parents in real time, have a simple interface that is easily and quickly navigated, have an adaptable interface based on parent's preferences, allow for changes of schedules both current and in the future, and that any messages should be carefully designed and customizable (e.g., got on or off the bus, bus was delayed etc.). From these specifications it was determined that a number of factors based on the novel user of the interface should be measured. Some of the factors considered are objective measures of goal completions, which can be captured through time to completion or simply ability to complete. The time to completion usability specifications were determined to be: if a parent can find out how long until the bus arrives within 5 seconds of turning on and logging into the application, if a user can find out what the day's plan is within 15 seconds, if a user can abort a current plan within 15 seconds, if a user can find the settings page, if a user can go to the live map and see where the bus is within 15 seconds, if a parent can change a future plan to a favorite within a minute, if a parent can cancel an aborted plan within 15 seconds. The second group of tasks are subjective measurements of users' opinions derived from a survey. These questions include: if a parent would actually use the device if it was available to them, if parents thought the interface was easy to use, if the tasks they would use the interface for are easy to complete on the prototype, if the users found the interface aesthetically pleasing, if the users thought they only needed short look times at the interface, and if they thought the interface was practical. The collection, analysis, and discussion of these two areas of information form the basic aim of this paper.

Prototype

Overview. The prototype designed for this project consists of a mobile phone application that allows parents to track their elementary school children while on their way to the bus before and after school, as well as during their child's ride. The system has countdowns for when a bus will arrive and 2 different types of location displays (a live map and tube-stop visualization) showing the user where the bus is. The system also allows for the setting and changing of daily plans at any time for that day or any day in the future. It includes favorites for where buses may take children, and the system can be customized to change alerts and the information displayed to the parents. The application is secure and has a login for each parent to ensure security of information.

The prototype includes nine screens that are used to interact with the system which all flow together as a cohesive user experience: login, home, details, planning, favorites, map, settings, admin and change plans. These screens can be seen in Figure 1. The user interacts with the application through a one dimensional flow as displayed in Figure 2, which was determined through flow testing and design iteration in part three of this project. Beginning at the home screen, parents are able to see an overview of their children's busses and how close they are to arriving. From this screen they can access a live map to see the exact location of the busses, click into settings, get a more detailed view of their child's plan, or set a new plan for another day. Upon clicking a child's name, users are taken one level deeper into a detail screen displaying further meta-data on the child's schedule and time until the bus arrives. The detail screen provides options for the parent to "abort" the current bus if the child won't make it on time, or change drop-off plans for the afternoon. Clicking on the plan button from the home screen gives users a series of inputs

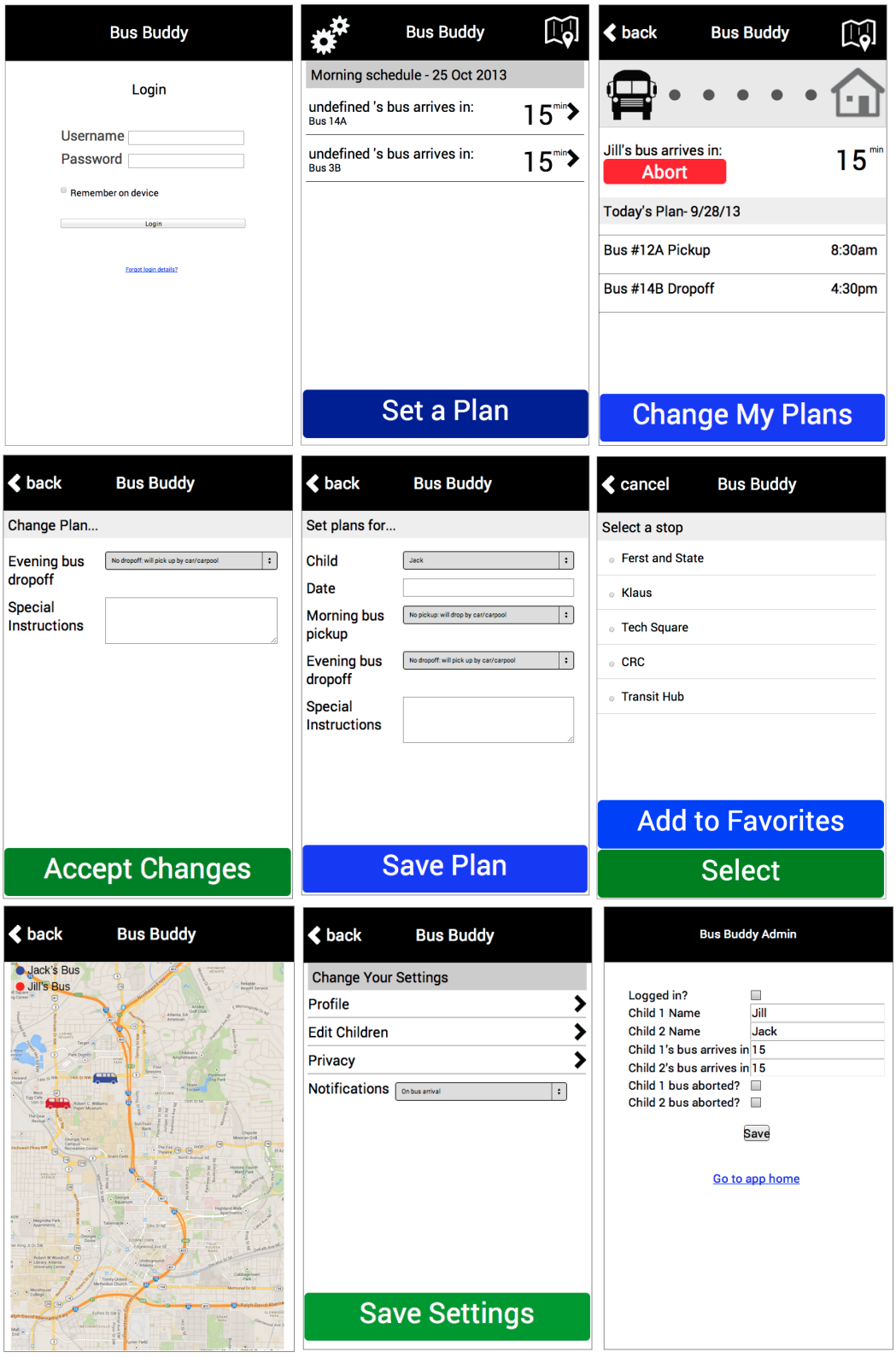

Figure 1. The nine prototype screens created for the interface.

to plan rides for their children throughout the week. These inputs include: child's name, date, morning pick-up location, evening drop-off location, and special instructions. It's these inputs that are communicated throughout the system to keep teachers, drivers, and administrators informed about the student's ride schedule. The planning screen inputs default to "favorite" options, but all stops can be accessed through the drop down selections and added to favorites, or they can be altered through the settings screen. The headline bar of the application conforms to the focused task of ride scheduling by only providing back and map buttons as a user dives deeper into application sections.

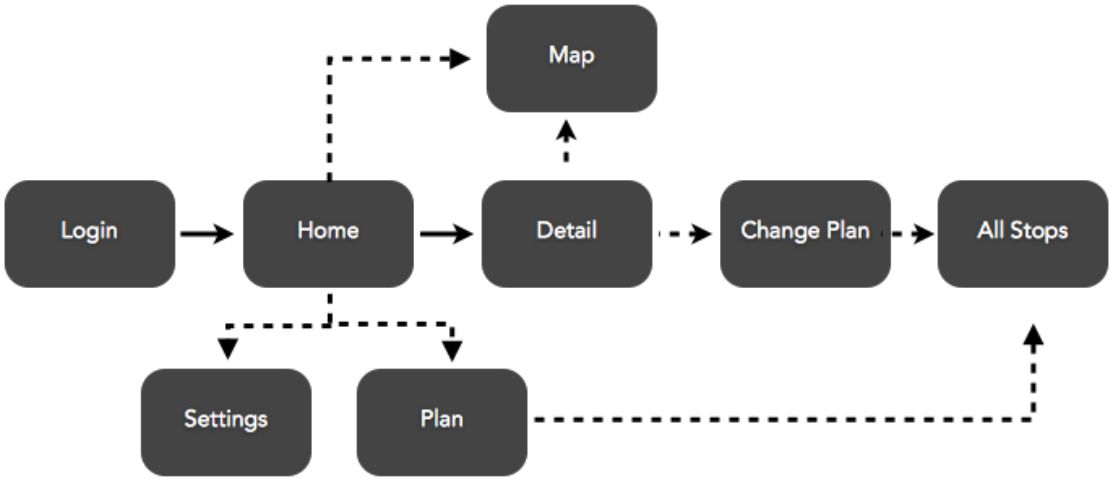

Figure 2. Flowchart of the flow through the prototype.

Rationale. The prototype was created as a web page because of the groups' familiarity with HTML, and the belief that it could implement a design that mimicked native application functionality. While there were some drawbacks to this method, it did give the group more control over interface components, allowed for a quicker build (which meant more screens), and enabled the faking components like logging entries (through cookies) without having to implement a full database. Additionally, building for the browser made

the system device agnostic, meaning it could be displayed on any mobile phone, or even a desktop browser if necessary. While building with HTML allowed for more flexibility and easier code handling, some of the components of phone browsers (most notably the elements of iOS 7) varied from a typical native application, causing some confusion for users unfamiliar with the platform. Regardless, the implemented design, with it's fixed screen size specifically for an iPhone, was realistically close to that of an actual application, with much less coding than a true application would have required.

Evaluation Methods

The following section aims to inform readers on the methods of evaluation used for this project. The goals of this evaluation were to measure the prototype on the usability specifications to determine what was done correctly and what could be improved. The type of evaluation can be classified as controlled settings involving users, as users are required to perform certain tasks in a laboratory for measurement (Sharp, Rogers, and Preece, 2007).

Participants. While the goal of the evaluation was to include only participants that fit the user group of the prototype, parents of elementary school aged children who rode the bus to school, it was determined that to get participants without funds some individuals who did not currently meet those specifications would have to be allowed. Therefore, the evaluation included a total of 9 participants, with a mean age of 45.78, 6 of whom had children. The participants also had an average of 4.33 years of smartphone experience.

Apparatus. The apparatus for this evaluation were three different 5-inch touch screen cell phones displaying the web site prototype discussed above, an iPhone 5, iPhone 5s, and a Nexus 4. The 5-inch screen phones were chosen since the prototype was built for

this size screen and if a smaller screen was used the whole interface could not be seen without scrolling on the page. Two phones were used because the aim here was to create a smartphone application, not an iPhone or Android application, so it was determined to use multiple phone types to ensure there were no device specific factors.

Measurements. The quantitative measurements during the test procedure included the time for each user to complete the set tasks and ratings about the interface on a Likert scale during the survey at the end of the study. These Likert scales were used since they are easy for users to understand (Sharp, Rogers, & Preece, 2007). The time was recorded to provide feedback on performance (Wixon & Wilson, 1997). Open-ended questions were also included on the survey regarding their experience. While users were interacting with the prototype an observer also took notes on user movements through the interface to try and understand any errors they made. Finally the users were asked to "think-aloud", i.e. to speak out what they were thinking when they were using the prototype, in order to help the experimenters understand the cognitive process (Sharp, Rogers, and Preece, 2007).

Procedure. Participants were recruited by word of mouth and no payment was given for participation. Upon arrival for testing, participants completed a consent form (Appendix A) and then read the instructions form (Appendix B). The participants were told what the interface did and what it was for so they understood what the application could do. They then had the role-play setting explained to them by reading the first paragraph of the evaluation script (Appendix C). The participants were then walked through the script by the experimenter and asked to perform each task while a timekeeper measured their time to completion, and a note taker watched their actions to take notes on what they were doing within the interface and to note any errors since the focus on the interface is on the

important factor of ease of learning and effective, error-free use (Courage, 2005). The role-play situation took them through three time periods: before the bus has picked up the kids from the bus stop and you are trying to get them ready, during the day at work and you need to make a change in schedules for the next day, and on your way to the bus stop to get the kids after school. After the completion of the tasks within the role-playing scenario participants were asked to complete four talk-aloud tasks similar to those that were done during the roleplaying in an effort to understand participants' reasoning behind the actions they made within the interface. Following the tasks participants were asked to fill out a survey regarding their thoughts on the application and some basic demographic information (Appendix D) and were then given the debrief form (Appendix E) before being released from testing.

Analysis. The data collected through measuring time until completion was averaged across participants and compared to the usability specifications to determine if the specifications were met. The subjective measures from the survey were also averaged and compared to the specifications. In order to help with visualization of the data bar charts with error-bar were created as well (Walker, 2013). The observational notes of the individual watching the user was investigated and taken into consideration along with any other feedback given by the participant and the information given through the talk aloud tasks. This information is used to direct the discussion portion of this paper and will be summarized and discussed in regards to any future redesign of the system.

Rationale. The rationale for the procedure was based entirely on the user specifications that were set prior to starting the evaluation. Since the desired information

was already known such as time to completion of certain tasks it was simple to design the study to cover those tasks. The participants involved in the study were chosen because a majority of them fit the users of the interface and would be able to relate to the ideal user of the finished application. The process of having participants role-play was to ensure they all had the same mind set going into each task and so they could try to consider the context they would be in when interacting with the interface. Outside of the timing of task completion it was determined that the subjective data would be most easily collected and analyzed in a quantitative way so a Likert scale was chosen. Other feedback can also be gained through qualitative data however, so the talk through protocol to understand users' thought process in interaction and written feedback on the surveys was used to gather this additional information. Finally, by watching participants and seeing what they were doing while performing the tasks, observers were able to get other data that the participants would not have thought about to mention. For the analysis the rationale was simply that the means and standard deviations were going to give the most informational data as the goal was to determine what the average user's time to interact with the device would be as well as how they felt about the device.

Results

The results of this study were broken into 3 sections. Firstly the timing result are given, compared to the desired times. Next the quantitative data from the survey results are given from the Likert scale ratings. Finally the qualitative data from the talk aloud, observations, and short answer questions on the survey are displayed in the best way possible.

Timing. The timing results of the study are displayed in Table 1 and the means and standard deviations are also shown in Figure 3. As can be seen the standard deviation in time to completion for all tasks is quite large, possibly due to sample size. Of the eight-desired time to completion goals, seven were obtained. The only goal not to be obtained may have been due to three participants having trouble finding the information and the median shows another way of looking at the averages for this and the other timed tasks and displays all goals being met. Some of the tasks were also completed much faster than the usability specifications, something discussed later in the paper.

Table 1. Time measured for each task

	Morning				**Midday**			**Afternoon**			
Participant	*Task 1*	*Task 2*	*Task 3*	*Task 4*	*Task 1*	*Task 2*	*Task 3*	*Task 1*	*Task 2*	*Task 3*	*Task 4*
1	1.4	6.2	15.9	10.6	10.5	7.7	61.2	34.5	8.5	3.3	10.2
2	5.7	6.7	15.5	15.2	14.1	15.9	57.1	33.2	3.2	6.8	26.8
3	4.3	7.9	14	7.5	2.2	9	35.6	9	5.2	7.3	11.8
4	22.9	20.3	58.9	25.7	9.9	58.3	91.6	11.1	24.5	7.3	13.2
5	3	10.4	34.4	6.3	8.3	18.1	56.4	14.5	10.2	10.9	14.3
6	1.5	9.1	18.2	13.3	9.1	11.2	42.3	21.6	7.4	8.3	11.4
7	46.3	1.3	2	40.3	24.3	1.3	2.3	40.1	52.3	24.4	45.7
8	3.1	7.1	25.8	6.6	18.7	25.7	26	10.4	16.5	7.4	16.6
9	21.2	32.8	32.2	7.3	8.6	16.4	40.1	11.3	4.5	6.5	14.1
Mean	12.2	11.3	24.1	14.8	11.7	18.2	45.8	20.6	14.7	9.1	18.2
Median	4.3	7.9	18.2	10.6	9.9	15.9	42.3	14.5	8.5	7.3	14.1
Stand Dev	15.3	9.5	16.4	11.4	6.5	16.6	22.3	12.2	15.6	6.1	11.4
Desired	5	15	60	15	15	60	60		60		
Goal?	No	Yes	Yes	Yes	Yes	Yes	Yes		Yes		

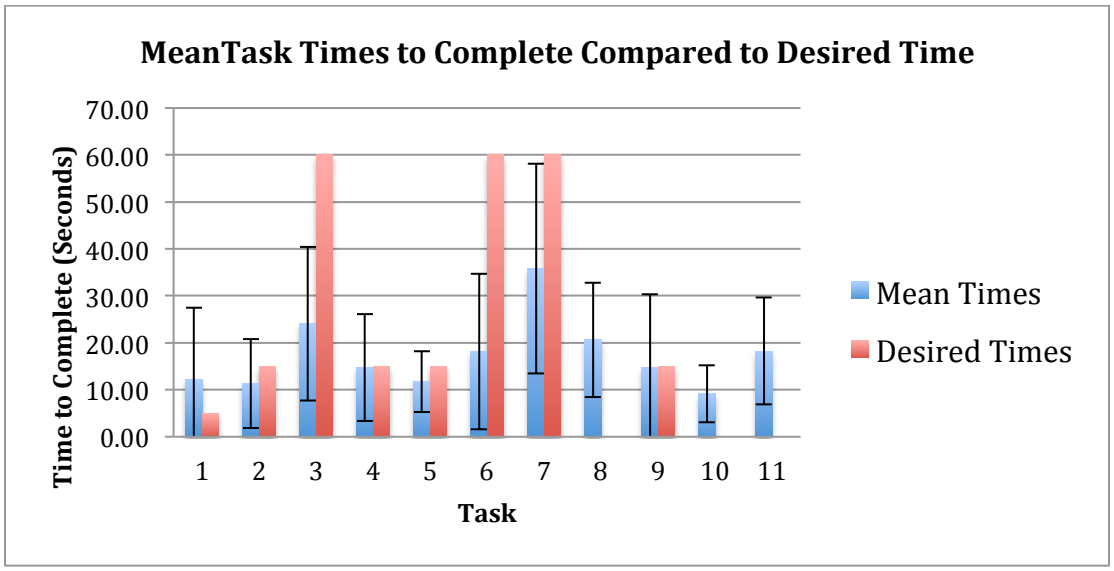

Figure 3. Average Time to Completion for Tasks

Survey Quantitative. The quantitative survey results for the study are displayed in Table 2 and parts of these data are also displayed in Figure 4. The data from the likert scale information display relatively low numbers, which means opinions leaning towards agree and strongly agree for the questions asked. These answers convey that participants liked the interface based on the questions asked and therefore the user specifications were met. Refer to Appendix D for a complete list of the questions.

Qualitative Data. While the timing and survey results were informative, the qualitative data from the evaluations was also interesting. Data received on the survey and information gathered during the talk aloud, as well as during the observation led to a few similar findings. Table 3 displays summaries of the written answers on the survey from participants. Additional qualitative data is examined in the discussion sections below.

Table 2. Participants' answers to quantitative questions

Participant	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	32	1	7	9	15	4	2	1	2	1	1	2	2
2	36	0		3	15	0	2	2	2	2	2	2	3
3	24	0		5	18		2	1	1	1	2	1	2
4	48	3	24,22,19	4	20	5	2	2	2	2	2	1	3
5	48	3	24,22,19	7	13	3	4	5	5	4	3	3	2
6	26	0		7	16		2	2	3	4	4	2	3
7	79	6	46-57	0	10	4	3	2	2	2	2	3	3
8	56	1	24	2	15	7	2	2	2	2	3	2	3
9	63	1	24	2	10		2	1	2	3	2	2	3
Mean	45.78	1.67	?	4.33	14.67	3.83	2.33	2	2.33	2.33	2.33	2	2.67
Median	48	1	?	4	15	4	2	2	2	2	2	2	3
Stand Dev	18.25	2	?	2.92	3.35	2.35	0.71	1.22	1.12	1.12	0.87	0.71	0.5

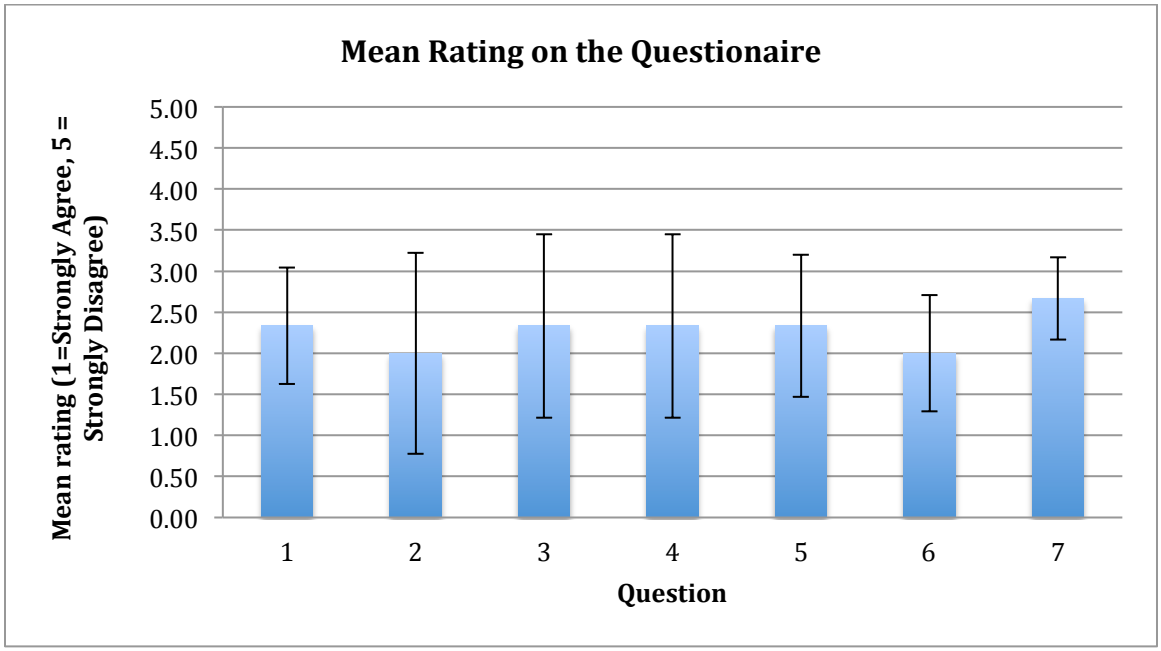

Figure 4. Average Likert Scale Ratings for Questionnaire

Table 3. Participants' answers to qualitative questions

Participant	14	15	16	17
1	Actual calendar	change bus info and track sons bus	some things looked clickable but were not	larger text on initial drop down
2	calendar icon	same as designed		
3	zoom on map	ability to place pins on map	date selector as calendar	
4				
5				
6				
7				
8		Knowing how long until bus arrives before and after school. Could cal neighbor to get kid after school	type too small	
9				

Discussion

Rationale for the choice of tasks. Part 1 of this project was spent discovering what is most important to users, and then developing a set of usability criteria for the application as discussed earlier. The benchmarks set to measure the usability criteria were also explained. Hence tasks were chosen for the evaluation study that would help to rate and compare our design with these benchmarks. The aim of these evaluation tasks was to obtain three distinct pieces of information:

1. How long it takes a user to perform these tasks. It was important that these tasks take as little time as possible, since the context of the application was primarily when the parent may be already otherwise occupied. Meaning, the usage of this application must not feel like a burden or a time sink to the parent.
2. How a user performs these tasks in different contexts. The intent behind the tasks may be different when done at different times or places. For example, checking how

long a bus will take to arrive at its stop has different purposes when performed in the morning versus the afternoon.

3. What their thought process is as they complete these tasks

   "Listening" to the users' thought process would help to gauge the overlap between users' conceptual model of the information presented in our application and its implementation in the design

Using these evaluation tasks we were able to validate the assumptions we made during the design of the app. Also, by comparing the results of the tests with the benchmarks we had set, it was possible to find out to what extent our design our usable as a mobile app.

Highlights and gained knowledge from prototype. The feedback gathered during the evaluation of the prototype was largely positive, but there was definite room for improvement in several aspects. The success can largely be attributed to our task centric design - most of the task times recorded during the evaluation were well under the set benchmarks. Some highlights of the successful aspects of the prototype:

- **Learnable.** The prototype to be very learnable. The participants did not need any special instructions as to how to use the application. Some users did show some hesitation when using a particular application flow for the first time, but were still successful at their task. By the time the participants were doing a similar task for the second time, they were very fluent at their usage. This was an indication that the information structure presented by the application to the users was very much in line with the mental models of that information of our participants. This was later validated by during the walkthroughs with users. They spoke sentences like "to change plan for Jack, I clicked on him first and then looked to see if there is

something to change plan" - perfectly aligned with the information structure presented by the application.

- **Quick response times for most tasks.** The design was found to be exceptionally successful for the most frequent tasks it had been designed for, like checking how many minutes away the bus is, and where the child is currently. When checking where their child is right now, the home screen displayed the current status of the child as a prominent piece of information. This validated the decision to put this kind of pertinent information on the home screen itself so that users needed only glance at the screen to view it. Some users did have a little hesitation the first time they saw the screen, but that can be attributed to learning effect. They were not sure whether the information they were seeing actually meant what they thought it meant.
- **Meaningful iconography.** The use of icons for the settings button and the map button was very successful. The icons were recognized instantly by all despite not having words to represent what they meant. Participants said words like "I just clicked on the map icon" when asked how they knew what to do to check where their children were. There was however, some lack in meaningful iconography within the planning portion of the application, which is described below.
- **Some broken task flows in the design.** Some tasks were not completed as successfully as others. One such plan was the task where the users had to set a future plan for their child. In this task, no participant followed the flow that we had planned. A "set plan" button was shown on the home screen, but participants did not realize on the first time that this button will help them set a future plan. Instead,

they randomly explored other sections of the application trying to find a function that would allow them to do this task, and used the “Set Plan” button after having exhausted all other options.

- **Flawed date selector.** The task given to users was to change the plan for “next Friday”. When attempting to complete this task, more than one user said “I don’t know which day next Friday is”, since the date selector provided by the prototype did not have day-of-the-week information.
- **Mistaking the “today’s plan” UI element to be touchable.** The details screen of our application has a section that lists the pickup and dropoff time for the child for that day. During the “change today’s plan” task, we observed that almost all participants tried to and trying to touch this section to change plans instead of using the big button at the bottom of the screen that proclaimed “Change today’s plans”. It seemed like people expected the plan element to be touchable. However once they realized that touching it did not work, the searched for a different action that could help them achieve their goal, and found the big bold button.
- **Some issues with the prototype.** The prototype was built as a web based application that ran inside a mobile browser on a phone, giving the illusion of a real mobile application. This technique of prototyping had some shortcomings that made it less usable than a real mobile application, and may have contributed to some increase in task completion times. Some of the repeatedly occurring problems are listed below:

  - Some participants reported that some buttons were not obvious enough as buttons, and hence they did not realize that it was an available action. This caused certain task delays and failures.
  - The text size on the prototype were not appropriate for a mobile screen. While testing it out on a desktop browser while development, it was not clear that this would become an issue when testing with a real phone. This was especially a problem with drop down menus, where the text was even smaller than normal text.
  - In certain cases, screen tapping was not working as well as would be expected from a native mobile application. Sometimes the touch gesture required to be held for just a little longer to activate a tap. This also caused a few task failures.

Avenues of prototype improvement. Some aspects of the design were not so successful, as listed in the above section. Below is a list of improvements that could be made to the prototype if more time was available to do a second iteration based on the results of the first round of testing:

- **Change broken task flows.** The "change future plan" task's flow was most notably broken. It could be fixed by either including another button immediately next the "change today's plan" button that would allow the user to change plans of a child for any arbitrary day in the future.
- **Better date selector.** The planning UI could be enhanced to include a more feature-complete date selector to allow easier identification and selection of dates based on day-of-the-week.

- **Touchable schedule listing.** The "today's plans" listing of the schedule for the day could be made touchable so that tapping it would allow the user to enter the "change today's plans" screen directly, and change that day's pickup or dropoff plans.
- **More realistic prototype.** The prototype had several cosmetic as well as responsiveness issues: text being too small, buttons not looking like buttons, or not being responsive enough. Based on how much time is available for a second iteration, the prototype could be quickly re-built using web based services like proto.io and hotgloo.com that can create more realistic prototypes that behave more like a native application. Alternatively, the existing prototype could be improved to have bigger and more realistic looking buttons, a larger and more comfortable font size etc.
- **Change task instruction terminology.** For the task where users were to cancel the bus pickup of their child, the instructions were phrased as "abort the bus". Many users repeated the word "abort" in phrases like "hm, abort the plan" or "abort the plan?". Then on the details screen they saw the button titled "abort" and clicked it. The same task could be evaluated with a change in the instruction terminology. It would be interesting to observe success rate for the task when using the phrase "cancel the pickup" instead of "abort the bus". In this evaluation the word "abort" in the instructions may have been a factor in the success rate of this task.

Knowledge gained from project. We would like to conclude this report by listing some items that we learned from the evaluation study:

- **Testing helps to validate assumptions.** A surprising find was the failure of an assumption made during the design. A link to the planning screen had been provided inside the “change today’s plan” screen thinking that if people come to that screen to change a different day’s plan they can find this link here. But the assumption failed - when looking for a way to change a future plan, most participants did not enter the “change today’s plan” screen.
- **Users rely heavily on convention and recognition.** A big bold high-contrast button that read “change today’s plans” had been placed on the details screen of the app. By all usual design principles, this was a good way of making sure it was obvious enough to the users. However, testing revealed that instead of going to this button directly, users attempted to click the “today’s plan” section of the screen expecting it to be clickable. Only after tapping a few times on the list of items in “today’s plan” did users realize that it did nothing, and then searched for the next best option, which was the big button at the bottom. This behavior also caused an increase in the task completion time for certain tasks by a second or two. It was interesting to note that some users tried clicking on the “today’s plan” items even when performing a second similar task.

  This revealed that users rely very much on existing conventions. The look and feel of the “today’s plan” section was similar to that of a clickable list, and users immediately recognized it as one and attempted to tap it to use it.

  Another instance, which revealed the reliance on convention, was when participants used the phone’s back button instead of the software back button provided in the prototype. The prototype had a method of navigating back to a previous screen

similar to android apps' "up" functionality, or to iPhone apps' back functionality. The test was conducted on an android device, and it had a hardware back button that served the same purpose as the prototype's software back functionality. It was observed that certain users used the back button on the phone instead of using the software back button in the prototype. This went to strengthen the case of how much users rely on convention to work.

- **Users think of certain tasks in certain ways.** Most of the participants had difficulty in completing the task where they had to change a plan for their child for "next Friday" - an arbitrary future date. All of them went down a different path than what had been designed, and backtracked once that path led them to a dead end. All participants recovered after backtracking and eventually found the right action to complete the task. However during the talk-aloud session for this task, several participants expressed the need for a better date selection mechanism - the mechanism shown by the prototype did not let the participants identify what day of the week a particular day was. Also, since this was a planning action, several participants said they looked around for a "calendar icon" first when they were given the task. To quote a particular participant - "I was looking for a calendar icon". This gives an interesting insight into how people subconsciously associate "planning" with a calendar, while the same people were able to perform the task of changing today's plan without looking for a calendar.

References

Walker, Bruce (2013). CS/PSY--6750 Georgia Tech Class Lectures,

Sharp, H., Preece, J., & Rogers, Y. (2007). Interaction design- beyond human computer interaction. Florence, Italy: John Wiley & Sons.

Courage, C. (2005). Understanding your users: A practical guide to user requirements methods, tools, and techniques (interactive technologies). Amsterdam: Elsevier.

Wixon, D., & Wilson, C. (1997). The usability engineering framework for product design and evaluation. In Helander, M. G., Landauer, T. K., & Prabhu, P. V. (Eds.), Handbook of human-computer interaction. 2nd ed. Amsterdam, The Netherlands: North-Holland.

Appendix A

Bus Stop Application Testing Consent Form

Study: Feedback on a Bus Tracker Application for Elementary School Aged Children
Principal Investigator: Dr. Bruce N. Walker (404-894-8265)
Location: School of Psychology, Coon (Psychology) Building, Georgia Institute of Technology
Duration of Each Session: 45 Minutes **Number of Sessions:** 1
Total Compensation: None
Approximate Number of Participants: 30
Participation limitations: Any parent who currently has or has had a child in elementary school who rode the bus.

General: You are being asked volunteer for a psychological experiment research project. The experiment is an effort to test the usability of an application built for parents to track a child's bus before and after school and to track information regarding their child being on it along with looking at the parent's interaction techniques. There will be a number of tasks for you to complete in the study followed by a survey regarding your experience with the application.

Study Purposes: This research is looking at how parents would interact with an application created to track their elementary school children while on their way to the bus before and after school as well as during their ride.

Procedures: If you decide to participate in this study we will introduce you to the interface that you will be using for this study. It is a phone application and/or a website that could be used to track a bus that children are on to go to school. During the study we will ask you to perform a number of tasks with the application such as "Find the time that Sally's bus should arrive" or "Change the plan for Jim next Wednesday to go to Jon's house" and see how you interact with the device. In Addition we may ask you to talk to us while you interact with the device. Following the tasks we ask you to perform, you will be asked to fill out a survey regarding your thoughts on the application and some basic demographic information.

Foreseeable Risks or Discomforts: This study is expected to involve no more than minimal risks associated with interacting with a phone or computer interface.

Confidentiality: The following procedures will be followed to keep your personal information confidential in this study: The data that is collected about you will be kept private to the extent allowed by law. To protect your privacy, your records will be kept under a code number rather than by name. Your records will be kept in locked files and only study staff will be allowed to look at them. Your name and any other fact that might point to you will not appear when results of this study are presented or published. To make sure that this research is being carried out in the proper way, the Georgia Institute of

Technology IRB will review study records. Again, your privacy will be protected to the extent allowed by law.

Injury/Adverse Reaction: Reports of injury or reaction should be made to Dr. Bruce Walker (404-894-8265). Neither the Georgia Institute of Technology nor the principal investigator has made provision for payment of costs associated with any injury resulting from participation in this study.

Contact Persons: If you have questions about this research, call or write Dr. Bruce Walker at 404-894-8265; School of Psychology, GA Tech, 654 Cherry Street, Atlanta, GA 30332-0170.

Statement of Rights: You have rights as a research volunteer. Taking part in this study is completely voluntary. If you do not take part, you will have no penalty. You may stop taking part in this study at any time with no penalty. If you have any questions about your rights as a research volunteer, call or write: The Institutional Review Board, Office of Research Compliance, 505 Tenth Street, Campus 0420. Phone: 404-894-6942; Fax: 404-385-2081.
Benefits: This study will provide no benefits to the participants other than knowing they assisted with the creation and testing of an application for tracking children on the bus.

Signatures: A copy of this form will be given to you. If you sign below, it means that you have read the information given in this consent form, and you would like to be a volunteer in this study.

Participant's Signature: _____

Date: _____

Person Obtaining Consent: _____

Date: _____

Appendix B

Instructions: Bus Stop Application Testing

Overall Project: This research is looking at how parents would interact with an application created to track their elementary school children while on their way to the bus before and after school as well as during their ride.
Purpose of Experiment: The purpose of this experiment is to see if the application we created fits well with the user needs that we had previously found through interviews and the way a typical user would want to interact with such a system.
Procedure
-In this study we will ask you to use an interface that we have created to track your hypothetical child's bus in different situations.
-We will first introduce you to the interface on the phone and/or website and allow you to become familiar with the interface through a short description of the application.
-You will then be asked to do a number of tasks within the application such as answering the question "how much time until Jill's bus arrives" and then record your actions to find that information and any issues you come across during that process.
-We may also request that you do a few talk aloud tasks where you will be asked to do a task similar to those that you had already done but describing what you are doing and why you are making those actions within the interface.
-Following the tasks we ask you to perform, you will be asked to fill out a survey regarding your thoughts on the application and some basic demographic information.

If you have any question about the task you are being asked to perform, or would like a copy of your consent form, please ask the experimenter now.

Before we begin, please turn off your cell phone (not to vibrate mode). If you need to take a break for any reason, please do so between steps.

Thank you.

Appendix C

Evaluation Script

Introduction

Hello and thank you for participating in our study. We will now move into the portion where we will ask you to roleplay that you are a parent of two kids, Jack and Jill who are both in elementary school. Jack is in 5th grade and Jill is in 3rd. We will ask that during the next few situations that we give you, you move through the device as best you can and do the tasks as quickly as possible. You can give us verbal "done" announcements or let us know you have the needed information and then we will ask you to let us know what the information we requested is.

Do you have any questions?

Use Observation (Role-play)

Morning

It is the morning on Wednesday December 4th and you are trying to get Jack and Jill ready for school this morning and out the door to the bus stop on time. Jill got up late. You need to check to see how long she has until the bus arrives. Task 1: Find out how long until **Jill's bus** arrives.

There might not be enough time for her to get dressed and ready for the bus to pick up. You decide to let her take time and will drive her to school. Task 2: **Abort** the plan for **Jill** to ride the bus that **morning**.

Also, you forgot you have to take Jill to the doctor after school today so Jack will go to Nana's house this afternoon. You are going to let the bus drop him at that stop. Task 3: Change the plan for **Jack** for that **afternoon** to go to **Nana's house**

Jill got ready very quickly and will make the bus.

Task 4: **Cancel** the aborted bus so she is riding this morning

Midday

It is midday and you are at work and want to make a few plans for later in the week as well as check plans for today. Task 1: Check on the plans for this **afternoon**.

You see that Jill is still scheduled to ride the bus today instead of being picked up to go to the doctor. Task 2: Change the plan for **Jill** to be **picked up today**.

You remember that Jack is planning to go to Keith's house on Friday next week (the 13th).

Task 3: Change the plan for **Jack** to ride the bus to **Keith's** house **next week on Friday**

December 13th.

Afternoon

You get to school to pick up Jill but she doesn't come out the regular door to your car. You need to check and make sure she didn't get on the bus. Task 1: Find out if **Jill got on the bus**.

Jill did get on the bus and is on the way to the regular stop and want to know where she is.

Task 2: Find where **Jill's bus** is on the **map**. *You have Jill in the car now but want to make sure that Jack got to Nana's house.*

Task 3: Check to see if **Jack arrived** at the **stop for Nana's house**.

After having all the issues today you realize you want more updates so before you want to change Task 4: Go to the settings screen and alarm me **everything** about my kids.

Talk Aloud

We will now ask you to do a repeat of some of the tasks from before and walk us through the actions that you made and why you did those actions to complete your goal.

Task 1: Find out how long until **Jill's bus** arrives.

Task 2: Find where **Jill's bus** is on the **map**.

Task 3: Change the plan for **Jack** for that **afternoon** to go to **Nana's house**

Task 4: Change the plan for **Jack** to ride the bus to **Keith's** house **next week on Friday December 13th**.

Appendix D

Survey: Bus Stop Application Feedback

1. Age: _____

2. Number of children: _____

3. Children's Age(s): _____

4. Years of smartphone experience: _____

5. Years of Internet experience: _____

6. On average how many times a month does your plans for picking or dropping off your child change? _____

For the following questions please circle one of the 5 options as it relates to your feeling about the statement

7. I enjoyed these interfaces

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

8. I would use these interfaces if they were available to me in my everyday life if I currently had a child

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

9. I found the tasks given to me easy to complete

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

10. I found the interfaces easy to use

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

11. I found I only needed a short glance to know what information was requested

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

12. I found the interfaces to be practical

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

13. I found the interfaces to be aesthetically pleasing

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	2	3	4	5

14. What additionally capabilities would you want to add to these interfaces?

_____

_____

_____

15. How would you use these interfaces? Would you use them the same way?

_____

_____

_____

15. What did you not like about the interfaces?

_____

_____

_____

16. Any additional comments?

_____

_____

Appendix E

Debriefing: Bus Stop Application Testing

First of all, thank you for your participation in this experiment.

Overall Project: This research is looking at how parents would interact with an application created to track their elementary school children while on their way to the bus before and after school as well as during their ride.

Purpose of Experiment: The purpose of this experiment is to see if the application we created fits well with the user needs that we had previously found through interviews and the way a typical user would want to interact with such a system.

Meaning of Expected Results: Based on the way you went about completing the tasks we asked you to do and any errors you encountered along the way, we can determine how usable our system was. The time that it took you to do certain tasks is also informative of how well we designed the system based on what information is important to be able to find quickly. The feedback from the questionnaire will also leads us further in our development and gives us your opinions of the application.

Confidentiality and Anonymity: The results of your experiment will be used for only psychological study and never used for any other purposes. The data that is collected from you will be kept private to the extent allowed by law. To protect your privacy, your records will be kept under a code number rather than by name. Your records will be kept in locked files and only research staffs will be allowed to look at them. Your name and any other fact that might point to you will not appear when results of this study are presented or published. To make sure that this research is being carried out in the proper way, the Georgia Institute of Technology IRB will review study records. Again, your privacy will be protected to the extent allowed by law.

Conclusion: All of the experiment procedures are finished. We very much appreciate your efforts again.

Contact Information
For further information of this research, contact:
Principal Investigator
Dr. Bruce Walker (bruce.walker@psych.gatech.edu)
Experimenters
Thomas Gable (thomas.gable@psych.gatech.edu), Blake Gruber (blakejgruber@gmail.com), Harshath JR (harshath.jr@gmail.com), & Alan Zhang (alandtzhang@gmail.com).