Project Part 5: Improving When2Meet

An even better way to schedule meetings & events

Human-Computer Interaction
December 6th, 2017

Youbeen Shim (ys3ep@virginia.edu)

The goal of the evaluation and prototype testing was to collect enough quantitative and qualitative data to determine the next steps of the project as well as evaluate the progress of the project. The questions of interest were both broad and specific for this reason. By the end of the evaluation phase, we would have a concrete idea on the best visual layout (and for what reason), clear idea of how a user may respond to the product, how accessible and intuitive the features are, and, if all goes well, the confidence that we progressing perfectly on schedule.

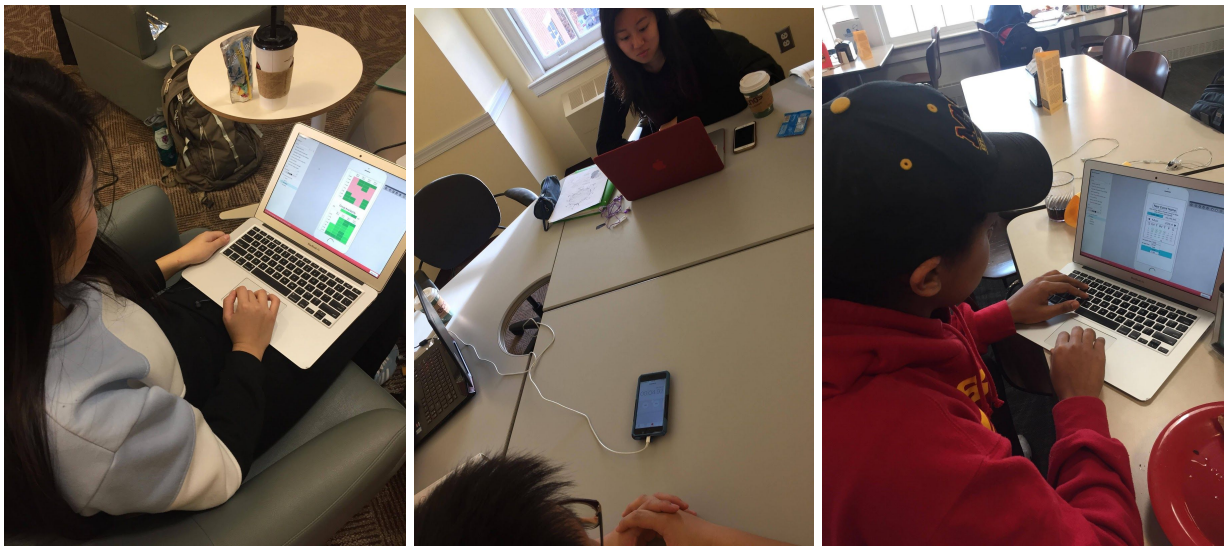
The users were asked to interact with the prototype on the software designed to simulate a mobile interface, Mockplus. Each prototype was designed with interaction capabilities such as progressing from one page to next, scrolling, selecting times and dates, or typing on an empty text box. The users were not given any specific instruction outside of instructions meant to help to interact with the simulation software. The tasks that the testers were asked to do were derived from task analysis and are as follows:

- Create an Event
    - Event name (type)
    - Choose date mode (button)
    - Select date (click and drag)
    - Select time (scroll box)
    - Create event (button)
- Sign in
    - Name (type)
    - Password (type; recognize as optional)
    - Log-in (button)

- Fill out schedule
    - (click and drag; repetitive)
- Determine ideal time
    - Look at the completed schedule and identify the time where everyone is available

During the testing, testers would first be asked to provide a background information about themselves. Basic profiles with name, age, year of study, major, and frequency of non-consistent group meetings were created. Then, the users were given a brief introduction on Mockplus and were timed while they were asked to complete the tasks listed above. Testers were asked to think out aloud as they progressed through the tasks. This was followed up by an interview asking basic questions including but not limited to what they enjoyed/disliked about the design, confusing elements that they faced (alternatively they were asked to elaborate on confusing elements they commented on during the tasks), orientation of their preference (horizontal vs vertical), and layout of their preference followed up with reasons.

This mix of broad ranged questions, specific reasons, timed data, and the commentary during the completion of the tasks helped to draw a more complete picture of the interaction process. The testers helped to evaluate decisions on qualities that were desirable but with consequences, such as adding more features and risking to lose the intuitive nature of some of the interactions. Important questions, such as how likely it is that users will use certain features, were also addressed.



During the evaluation phase, a total of twenty-four individuals was asked to test and assess the prototypes. Out of the twenty-four, ten people both interacted with the system and stayed to provide feedback. Each of them were asked to try two horizontal prototypes or two vertical prototypes to remove preference for orientation as a deciding factor. The preference for orientation was asked as an interview question after completing the requested tasks.

Each tester took an average of eight to twelve minutes for the interview. Interacting with the system took around 2-3 minutes per prototype and 2-6 minutes were spent providing feedback and answering different questions. Each individual was given a role -Planner (create the event and decide on a time + role as a Participant) or Participant (log in and fill out the time). The time to complete a specific task, such as scanning through the completed schedule table and choosing the ideal time, were measured. The overall time was defined as the summation of the time it took to complete all the specific tasks. Testers were also asked to rate the system from one to ten with one decimal point increments to determine user satisfaction and hopefully

identify the main source of confusion. Directed questions, especially those asking for a positive feedback (i.e. “What did you like about this prototype”) were asked to identify reasons for a certain preference but also followed up with questions that had a contrasting sentiment (i.e. “What did you dislike about this prototype”).

Very few of the testers were currently taking a Human Computer Interaction or a related course, or have previously taken the class. While this may skew the feedback and these peers from class is likely an inaccurate representation of the complete user base due to their knowledge in the human-computer interaction and in the designing process, they were able to provide a feedback with more depth and was able to be more critical. Overall, the testers were peers from around the university with no specific criteria. The ratio for gender were even. Age ranged from eighteen to twenty-two with no exceptions. Taking into account the setting (University of Virginia), the demographic was also largely educated and literate and also largely democratic.

Though the demographics of the testers do not match the demographics of all the potential users for the system, the American population, it did capture the demographics of the users most likely to interact with the system. In order to address the largest user base, the purpose of the product -When2Meet- and its resulting user base must be addressed. When2Meet is a website that will create and save a schedule table so individuals with complicated schedules can scope out an ideal time of meeting. This makes When2Meet ideal for individuals with following qualities: 1) technologically apt; 2) has a complex or non-traditional schedule (i.e. not on a nine to five schedule); and 3) has to plan multiple meetings with different people. Thus, college students, with multiple courses, assignments, and other time commitments as well as

numerous reasons to meet up with different people, were a incomplete but overall an accurate selection.

Results:

Name	Type	Time	Like	Dislike	Rating
Ahra	Prototype 3	40 s	B	1	7
	Prototype 4	50	A	2	8
Vicky	Prototype 4	44	A, C	1	5
	Prototype 3	30	F	3	7
Judy	Prototype 4	44	n/a	4	7
	Prototype 3	44	E	6, 7	8.5
Aisha	Prototype 3	39	D	4	8
	Prototype 4	50	C	2, 3	7.5
Yaman	Prototype 4	57	B	n/a	7.5
	Prototype 3	52	B,C	1	8
Nikkita	Prototype 1	42	E	7	9
	Prototype 2	39	C	4	7
Catherine	Prototype 2	51	F	n/a	8
	Prototype 1	48	A	1	6
Graham	Prototype 1	30	A	7	8.5
	Prototype 2	38	B	n/a	8
Angie	Prototype 1	46	F	n/a	7.5
	Prototype 2	41	D	5	9

Vasudha	Prototype 2	39	C, F	7	8
	Prototype 1	40	E	2, 3	7.5

Like: A. Layout; B. Option; C. Intuitive; D. Orientation; E. Scroll; F. Better than original

Dislike: 1. Creating an account is bothersome; 2. Lack of feedback; 3. Deselection/error correction unclear; 4. Scroll; 5. Inconsistent; 6. Box; 7. Unintuitive

It is critical to first note that the interview questions, commentary, and the data collected during the tasks were helpful, but not conclusive. It was possible to answer questions such as overall preference for orientation or how much time each task consumed. We were also able to rate/rank certain features over others, better visualize how a user may interact with the system, and identify areas that require more thought. However, this most of the result is gained by analyzing a mix of responses and data, not using a formal statistical analysis. While this was expected from the starting phase of the evaluation process, when the goals for the evaluations were set, it is still important to remember this as we move on.

Moving forward, it appears that Prototype 2: Vertical Mobile When2Meet (No Scroll) is the design best suited for the project. Prototype 2 wasn't just the most popular design with the most positive feedback. What was interesting was that the question that we did not realize that we were asking, but was at the core of varied prototypes, was answered. "Do we create a design that stays true to When2Meet's original design or do we deviate towards a more mobile-friendly design at the risk of being unfamiliar?". Prototype 2 deviated from When2Meet and the other three prototypes in that it cut out the dual-screen interaction feature and only presented the screen that was relevant to the particular task one at a time. It strayed the furthest from the

original design but at the same time was the most intuitive when it came to being on a mobile interface.

From the tester evaluations, it appears that while we were initially concerned about deviating from the original design, the users themselves did not mind much that the designs were different. As long as the core concept of creating a new event each time and uploading schedule remained the same, the features provided should make the most sense to the device being used.

However, Prototype 2 was not perfect either. Most importantly, the design for Prototype 2 was too isolated. There was less consistency from task to task and each screen did not point towards the progression of the tasks. While most users were able to move on from one page to the next without a significant issue, more than a few users struggled for at least a few seconds to figure out how to progress. This was especially true when the testers were given the Participant role and not the Planner role. This suggests that there should be a better guidance towards the progression of the tasks -a goal for the next portion of the process.

We also cannot deny the considerable advantage that Prototype 2 had over other prototypes in regards to orientation and scrolling. The four prototypes each had either a vertical or horizontal orientation and either a scrolling or no-scroll layout. After the full evaluation, it became clear that both the vertical orientation and no-scroll layout were preferred. Perhaps the only reason that Prototype 2 was preferred was because of these elements and not the concerns addressed above. The divisiveness of scrolling versus no scrolling may also not be as significant of an issue than the overall evaluation suggests. The simulation software ran on a laptop instead of a mobile device, meaning scrolling feature was only possible by clicking, holding, and

dragging, which is considerably easier on a laptop compared to a mobile device. Testing higher fidelity prototypes on actual mobile devices remains a future goal that our team aims to reach.