**A5\_ ppt**

[310\_A5\_PPT](https://docs.google.com/presentation/d/1b6qaDSIhAHiGCvmFFL-kp5HjZNmaq7tcyvsGMZkG8k4/edit#slide=id.ge107949c2f_1_256)

**A5 spreadsheet:**

[310-A5-problems](https://docs.google.com/spreadsheets/d/1rMGn1cSP9FI_uD8pj2KQDel5ibAN3yAirTBwTC9W31E/edit?usp=sharing)

## Study Description

### Before the Pilot Test

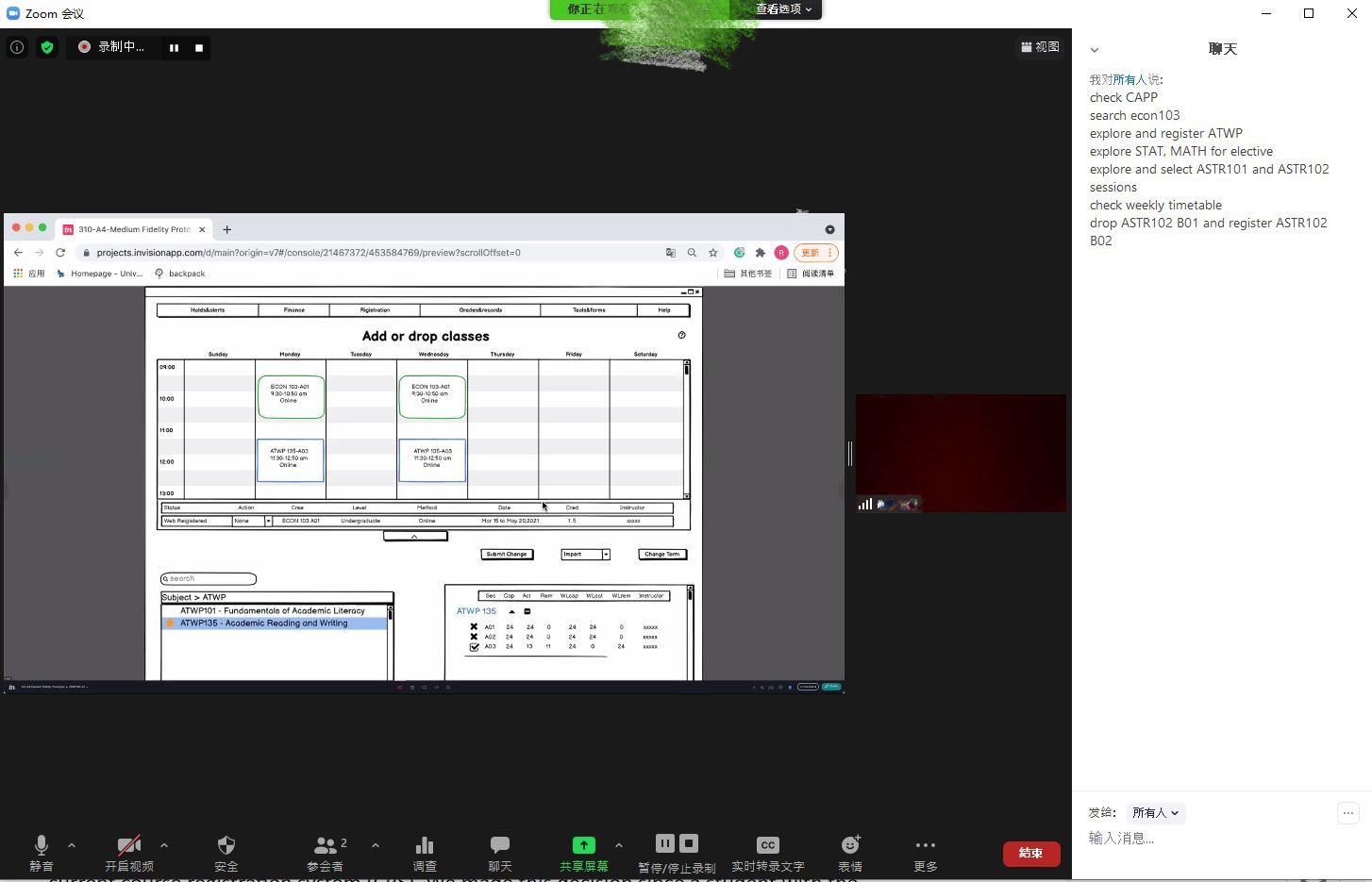
Before we ran a pilot test, we fixed a few spelling and placing mistakes that were found in our previous assignment. Then after some discussion, we decided to add a series of interactions that represent the processes of dropping a class, as dropping classes is also a crucial part of the CRS, and we did not want to overlook it. We also divided the prototype into 7 sections each representing a task, with task 3, 5 and 7 being our main focus. The tasks are:

1. check CAPP
2. search econ103
3. explore and register ATWP
4. explore STAT, MATH for elective
5. explore and select ASTR101 and ASTR102 sessions
6. check weekly timetable
7. drop ASTR102 B01 and register ASTR102 B02

### Pilot Test

We found a graduate student from SFU who has never interacted with our current course registration system (CRS) for the pilot test. We made this decision since a student with the knowledge of the current CRS might not spot the problems that exist in both our prototype and the current CRS. In addition, such a decision tests how a new user would interact with the system. We then let her conduct the pilot test on our prototype. During the process, she only found some minor problems, such as part of the text had consistent colouring and positioning, and a few tooltips were not showing up in the right position. We fixed most of them and also made some minor improvements regarding the wording of some pop-up messages.

### Participants Overview



We had a total of five participants(excluding the pilot tester), all students from the University of Victoria. There are two first-year students, one second-year student, one student from the third year and one student from the fourth year. The participants are from different faculties and have different year standings: one from psychology (3rd year), social science(1st year), business(2nd year), and two from engineering(2nd year and 4th year). For simplicity and anonymity, they will be referred to as Psyc, SS, Bus, Eng1 and Eng2, respectively. Three of our participants also participated in our previous research in assignment 2. To ensure users' consistency is helpful for us to complete the test, we have designed our prototype according to their feedback. The two new participants are also important as they do not know what our design aims to solve to represent the general population to some extent. Therefore, these people are highly relevant as testers to provide us with crucial modification ideas.

### Methods Used

The interview methods are kept consistent with the user evaluation plan in assignment 4. During the test, the users would use the “screen share” feature of the ZOOM meeting. The interviews were conducted in this way for easier interaction with each participant and to follow the social distance guidelines.

It is convenient for the interviewer to see the participants' operation during the test, record each task's completion time, pause, and make targeted guidance when there is a problem. Before we started the test, we obtained the participants' permission to record the video. We think that the video recording can let us look back and find some missing information. Like the time was taken to complete each task and some problems that we observed, the participants didn't give feedback.

### Data Sources Collected

Based on the prototype, we slightly changed the question from our original user evaluation plan. The questions are as follows:

Qualitative (answered by user):

1. Do the layout buttons of the interface enable you to easily complete tasks? Why?
2. Do you think the interface design of this new CRS is friendly? Why?
3. Compared with the current CRS, do you think this CRS layout meets your expectations (such as button and course search, page conversion)?
4. Do you think the interface displays the current operation status clearly(i.e. Do you know there are unsaved changes, do you know which button provides what function, etc.)
5. In general, how do you evaluate the experience of the interface and the tasks provided?

Quantitative (recorded by interviewer):

1. How long does it take the user to complete the main task?
2. How many times did the user request help?
3. In the process of task execution, when did the user deviate from the expected?

### Data Analysis Methods

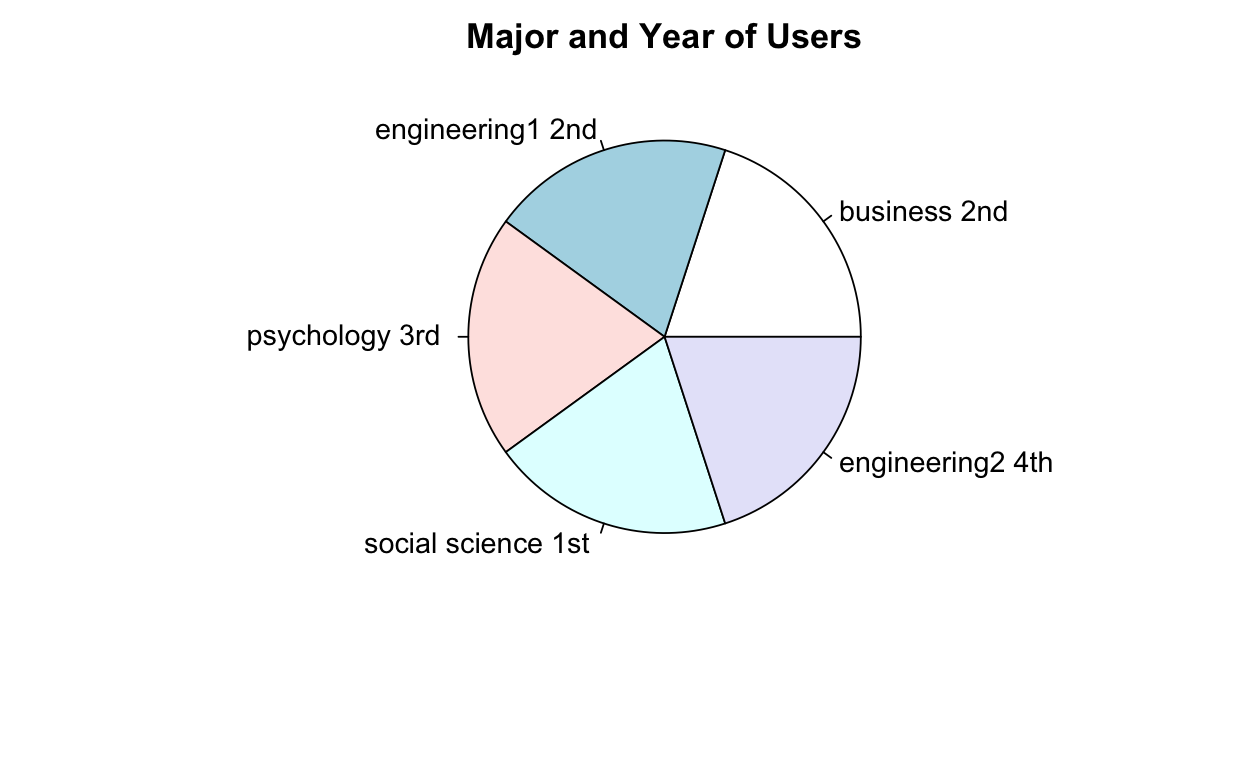
We split the data analysis phase into the qualitative and quantitative portions.

The results of the qualitative data were transcribed from the recordings and evaluated. The comments were then categorized according to the relevant prototype area. Finally, these findings were collected into a spreadsheet that displays the comments.

The data points for the quantitative responses were gathered and put into a spreadsheet, where they could be evaluated for trends. The mean, median, and mode of the data points were automatically generated with Google spreadsheet tools. Some graph tables were generated using the computed averages and the raw data to better visualize the data. For example, the stacked bar graphic could be beneficial as it showed how much time was spent on each activity.

## Result and discussion

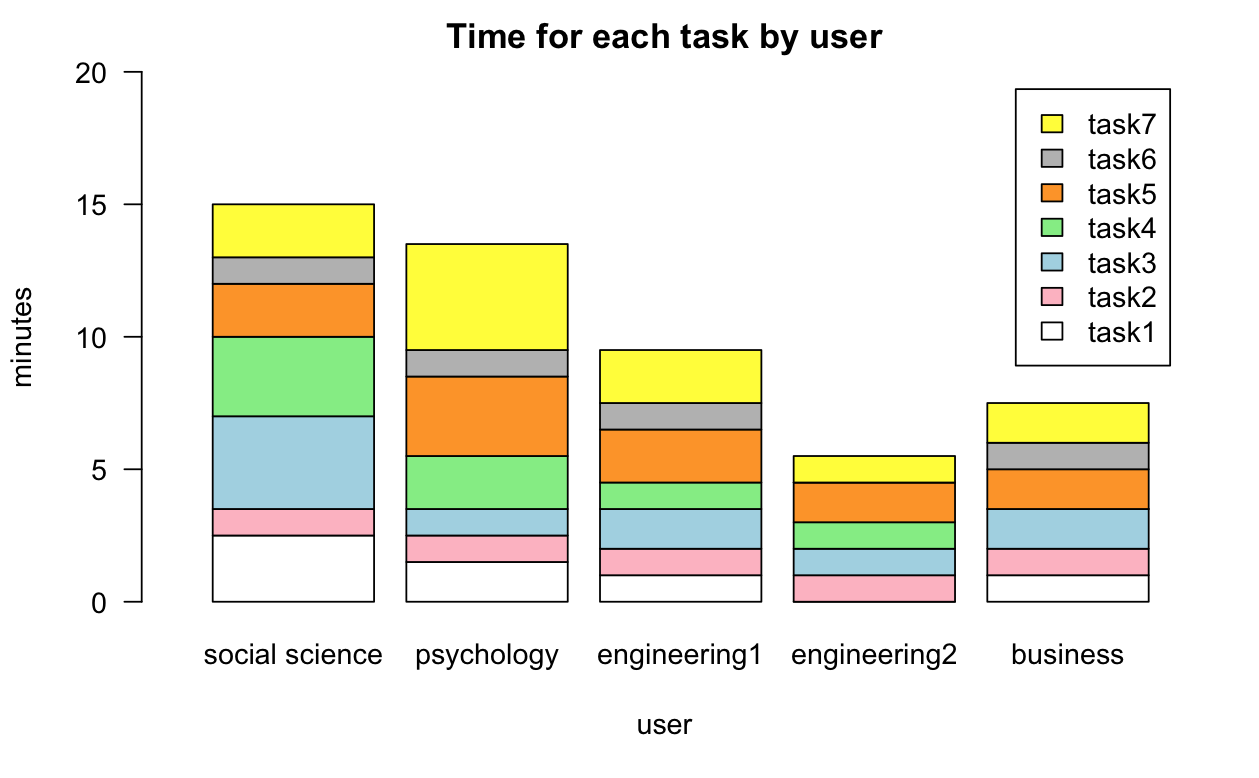
### Quantitative Results



Pie chart of participants’ faculties and year standings

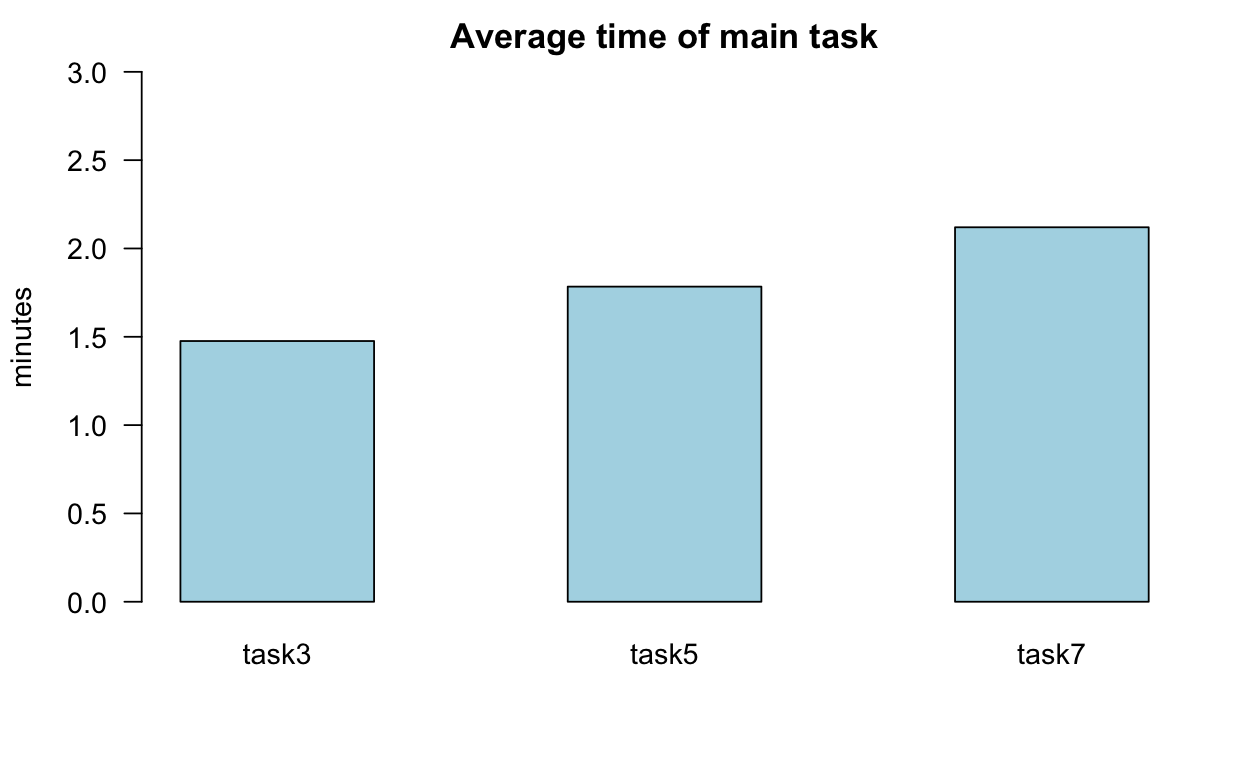
As mentioned above, we had 5 participants, mainly from different faculties and year standings.

#### 

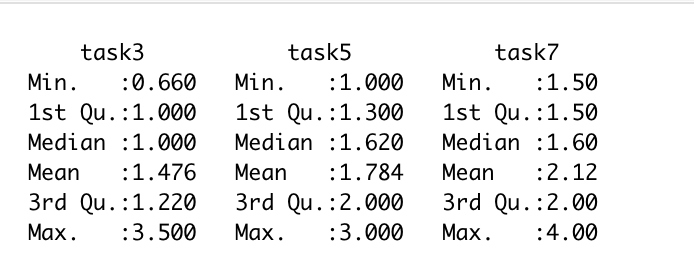


Barplot of times for each task by user

During the tests, we recorded each participant’s time spent on each task. Although the total time ranged from 5 to 15 minutes, the deviation was larger than expected, the outlier being Eng2. However, considering Eng2 skipped two minor tasks, the significantly shorter time can be tolerated.



bar plot of average time for users to complete their main tasks



Detailed analysis of time spent on main tasks

In the process of user testing, we asked the interviewer to record the quantitative data of three main tasks. The first is when the user completes each main task. The first main task is task3 (explore and register ATWP). The average time of users is 1.476 minutes. The second main task is task5 (explore and select astr101 and astr102 sessions). The average time of users is 1.784 minutes. The third main task is task7 (drop astr102 B01 and register astr102 B02). The average time of users is 2.12 minutes.

### 

In the test process, we let the interviewer record mistakes the user made during each task. Due to the users subconsciously clicking or pressing the “shift” key to reveal the interaction boxes, we shared and reviewed each recording, then decided to count some of them as mistakes by judging the users’ spoken thought process. The data that we recorded are: task 1: 1.2 times; task 2: 0 time; task 3: 0.2 times; task 4: 0 times; task 5: 0.8 times; task6: 0.6 times and task 7: 0.8 times.

During task 1(check CAPP), all the users were unaware of clicking the “year 1” on the CAPP page, this is caused by the task description being unclear. Some users have not yet interacted with the CAPP form, thus were confused. During task 2(search Econ103), one user was confused as they tried to use the subject list instead of the intended search bar. During task 5 (explore and select ASTR101 and ASTR102 sessions), most users were confused and unaware that multiple courses can be previewed simultaneously and tried to submit changes after selecting ASTR101 sessions. During task 7 (drop ASTR102 B01 and register ASTR102 B02), almost all users were unaware that there was a scroll bar on the list of registered courses, and were confused about it.

After the tasks were completed, the participants were asked to rate the experience with our prototype from 0 to 10, 0 being unusable, and 10 being perfect. The average we got was 8.92/10.

### Qualitative Results

## 

Thematic analysis from findings

During the usability test and the interview, we found some commonly asked questions and recurring answers, then categorized them into five themes: need to improve, Invision issue, suggestions, aesthetic and intuitive system. Most codes from "need to improve" are caught during the interaction, not brought up during the interview. We believe these complaints and confusions are also important feedback, as they could be forgotten as the interviewers did not intentionally mention the earlier parts of the prototype.

A lot of the complaints were that the participants did not know how to transition to another tool, for example, from “select term” to “CAPP/PPW”. These were resolved by the interviewers giving them hints. Some of the participants also failed to notice that there were tooltips on the symbols. These were also resolved by giving out hints. Both of these issues point out that some of the information representation was not intuitive enough. Moreover, some of them pointed out that the CAPP does not provide enough helpful information.

Although taking up the majority of our thematic analysis, these problems are only slight and minor inconveniences. Besides that, our overall design satisfies most of their expectations, and there were no major design flaws and usability issues.

In summary, the results show that our redesigned CRS is promising. Each participant completed (most of) their tasks, and most of them gave a positive response. Those who participated in our previous research responded that our design have indeed solved the problems they were unhappy with. The new participants liked our design and pointed out some minor details we overlooked in the designing phase.

## Study limitations and reflections

**Our study shortage and limitations:**

1. **Small sample size of only 5 people.** Because of time limitations, it is difficult for us to survey many people in a few days and get very accurate data. Although our 5 participants come from 4 different faculties, it is not enough to represent the whole population of students at UVic.
2. **Constraints on users’ freedom.** Due to the limitations of Invision and the rapid growth rate of pages, we could not implement a dynamically updated timetable and list of courses. Therefore we were forced to set the interactions in a fixed order, and we had to constraint the users’ freedom of choice by forcing them to complete the tasks in order, and thus it seems more like a scripted walkthrough.
3. **Only have a vertical medium-fidelity prototype.** Because of time and technology constraints, we did not do all the functions of CRS, and in terms of course selection, we only provided a limited number of courses for users to choose from. That means that it is difficult for users to understand our overall design of the Course Registration System. In addition, there are certain restrictions on the operation so that the interviewer may guide their behaviour.
4. **There exists bias.** As mentioned above, three participants are from our previous research. Since it was only a few weeks ago, they might still remember the topics and imply that we have designed the prototype, which would incentivize them to “sugarcoat” their feedback. In addition, our qualitative questions are also leading, which introduces bias.
5. **Test tool limitations.** Because of covid-19, the only way for us is to use an online tool to do the test. Furthermore, the big problem is that we can only interview those willing to share the screen, which makes the type of data we get constrained and may not represent everyone's idea of redesigning CRS.

In the process of our test, we found some difficulties and good things in this test. To make sure everything goes well next time, we summarize these points as follows.

**What went well:**

1. All participants finished all the main tasks successfully.
2. No critical technical issues occurred that stopped the interview.
3. Tasks were completed within a reasonable amount of time.
4. Although it changed slightly, we followed our evaluation plan.
5. Interviewers answered our qualitative questions and provided helpful feedback.

**What did not go well:**

1. Lack of indications to guide interviewer to finish all the tasks
2. Zoom overwrites existing recorded files and caused the loss of one interview footage. The interviewer had to re-record it with the participant.

**What could be improved**

The usability tests took between 5 minutes to 15 minutes to finish, which showed us that some of the participants might take longer to understand the user interface of the new CRS. As a result, we could improve it by implementing more help functions or a short tutorial to introduce the interface. Furthermore, during critical tasks, we can provide more indications to guide the user. If we were to conduct this research again, we would schedule more time to prepare our interface and implement a more mature prototype. Moreover, the strategies for recruiting interview participants and the number of participants will be reconsidered so that the participants can be distributed into more faculties and year standings, which would allow us to collect more data to evaluate our prototype better.

**--------------------------------------------------------------------------------------------------------------------------------------------------------**

**Ppt**

1. **预测试**
2. **找了五个人测试**
3. **测试结果**
4. **讲一下自己用invision 重新design crs， 用户可以自由选择我们设置的课程，然后找了5个人来测试，这五个人还是same person in rearch part 的老人，**

**--------------------------------------------------------------------------------------------------------------------------------------------------------**

Hello everyone, This is team six. I’m Rui. We have Zhanchi, Aaron, and Junhong.

Today, I will talk about our remote usability test results of our course registration system.

Participants side

We had a total of 5 participants, each from different faculties and have different year standings: one from psychology (3rd year), social science(1st year), business(2nd year), and two from engineering(2nd year and 4th year). 3 of them also participated in our previous research from assignment 2, and the other 2 are brand new. They are nowhere near to represent the whole population of UVic students, but that’s all we could do with the time given.

Task

We let the participants get hands-on experience on our prototype. And sent them a task list during the interview; It contains a total of 7 tasks which derived from our original prototype, 3 main tasks and 4 small tasks. Here are the 3 main tasks, and each of them represents the key function of the ‘Add/drop course‘ tool, involving registering courses, handling time conflicts, waitlisting, and dropping courses. The remaining 4 small tasks allow the users to gain the idea of the overall interface of the ‘Add/drzhanop course‘ tool and how the flow will be.

Quantitative questions

We recorded the process of the participants interacting with our protīotype. WE determined the length of time spent on each main task, recorded how many hints they asked from us and total mistakes they made.

Quantitative Result

This bar plot shows time for each main task used by participants. The mean time used for each task by participants range from 1.307 minutes to 2.367 minutes. The average is not the same because the number of steps required for each task is different, so the mean is not very meaningful for our result. Therefore, we mainly focus on the overall time spent to complete the task. And in the chart we could see time use for task1 is more than the other two tasks for new users. The time to complete the task showed a downward trend. So we can get the conclusion that after the participants are familiar with our redesigned interface, they can complete the tasks quickly.

In the mistakes/hints summary bar plot, participants make fewer mistakes in tasks 3 and 4 than in tasks 5, 6, and 7 because tasks 3 and 4 are less challenging. However, the trend for them to seek help is similar to the trend for the completion time of each task.

Finally, we can get such a result. Generally speaking, our design logic is clear, and the function is complete. The participants can still complete the task even though there are some page design problems, but some participants will still ask for help and make mistakes in each task, indicating that some details need to be improved.

Qualitative Result

For the qualitative result, we have asked several questions from participants, such as the interface design of this new CRS,and their experiences with the redesigned CRS interface.

The response we got was unanimously positive. For example, participants said that the navigation is simple and clear, and they could quickly find all the buttons and course information. We quote “compared to the current CRS, this one is more convenient and simple.”

Although the feedback is positive, we did observe some confusion and minor details that we could work on. For example, some participants did not realize that the icons have tooltips with them, and had to request hints. Some participants did not realize there is a scroll bar in the list of registered courses.

This is all our presentation. Thank you, are there any questions?

最后还是决定加了一页 。。 不然不好说hint 和mistake的趋势