User Evaluation

# Users

There were five (5) users who tested the application from both of the user groups defined. The skills ranged from novice computer usage to intermediate computer usage. The five users were Mom (Novice), Rick (Novice), Ashley (Intermediate), Natasha (Intermediate) and Julie-Anne H (Intermediate). The time it took them to complete the task list averaged about 10 minutes. The users then completed a survey based on the tasks they completed. This survey asked about how they felt after doing the list and gave opinions to the best and worst task.

# Task List

#1 - Enter a new movie into the application  
#2 - Edit a movie already in the application  
#3 - Delete a movie from the application  
#4 - Bring up the movie details of a movie in the library  
#5 - Add a movie to a user list (My Wishlist/Gift Movies/Watch Next)  
#6 - Remove a movie from a user list (My Wishlist/Gift Movies/Watch Next)  
#7 - Get a movie suggestion based on Rating  
#8 - Perform a tagged search  
#9 - Add a new tag to search by  
#10 - Use the search bar in the library and search for a movie name  
#11 - Using the criterion on the left bar, pick one or two and then search based on those

# Variables

Time – I recorded the time it took each of the users to complete the list (10 min is dependant)  
Percent stuck in task – I estimated the time they spent figuring out a task (15% is dependant)  
Hard Tasks – Number of tasks not easily figured out/required assistance(2 task is dependant)  
Questionnaire (rated 1-5; 5 is very, 1 is not very)  
Toughness of task list – How tough users found the task list to be (3 rating is dependant)  
Frustration level – How frustrated users became at any point (1 rating is dependant)  
Fatigue level – How fatigued a user was after completing the tasks (1 rating is dependant)  
Feel of complexity – How complicated did the user feel the application is (2 rating is dependant)  
Opinion questions – There is no dependants here. Users gave opinions on hardest, easiest, most favorite and least favorite tasks. This way we can see if there is any part of the application that can/needs to be improved.

# Evaluation Techniques

Each user was required to fill out a survey/questionnaire after they completed the task list assigned. This questionnaire had questions pertaining to the variables defined in the above section. During the users experimenting with the task list we sat observing the users as they interacted with the application. This way we could see if they had recurring problems or where a certain user would get stuck. After the list of tasks was completed a few questions were asked interview style as well as a few questions thought of during the observation process. These questions were:

*“Would you use this application if you wanted to keep track of movies?”  
“What would suggest changing in the application to make it easier to use?”  
“Do you have any other suggestions or comments?”*

# Evaluation Criteria

Please see the PDF file included for the results of the users completing the task list.

# Analysis

It was found that only 2 of the users came close to the estimated time to complete the list of tasks. The majority of users took longer than expected (3 out of 5). Although if we look at it from a different perspective, half the users came close or under the expected time while 2 users took quite a while longer than the rest.

It was also found that the percent of time spent stuck on tasks during the completion of the list was greater than anticipated. This was not expected because it means that the interface was not clear at some points. It was also observed that only 1 user figured out how to show the movie details and that was by accident. So this in specific needs to be redesigned to be better for a user.

The percent of time is directly correlated with the amount of tasks that were hard to complete. It was found the higher percentage of time stuck on a task meant that there were more hard tasks to complete. Our initial estimation of 2 tasks giving the user trouble was a good estimation. Two users got stuck on 2 tasks with one user getting stuck on 1, and one user getting stuck on 3. Out of 11 tasks this is a good outcome.

It was found that the users thought the task list was relatively easy. The lowest ranking was 4 where 5 was the easiest rating. Even though there were a few times they had trouble they found it easy to do what they were assigned to complete. It was expected that the average between all the users would be about a 3 rating, but it turns out the rating users gave as better than the expected one.

The level of frustration was very low (1) for everyone except for one user who was frustrated after using the application (4). This shows that the design did not make the users want to get up and walk away but that there were some people who would not use the application; which is expected for every application. Our estimated rating for this area was confirmed.

The complexity feeling of the application was recorded to be on average 2 with three users giving it the estimated rating. This was expected as with no feedback until the end of development it was hard to gauge exactly how complex a user would feel with the initial design, even though they were asked what they wanted. There will always be the trade-off for the number of features versus the complexity of the program.

The users were asked to rate their confusion level at any point during the tasks. This helped us then figure out if there were any certain points that confused the user and would need to be redesigned. This question was a write off though because it seems that the confusion came from not fully understanding the task assigned.

Another opinion question asked was if the user ever felt like not doing a task. Though the intended results of this question were to figure out where in the task list the user wanted to not keep going, but instead this answer was more dependent on the mood of the user at the time. The user ratings for this area were all over the map.

As a bit of an evaluation of how the tasks themselves were designed the users were asked to rate them based on easiest, hardest, favorite, and least favorite. The more complicated steps (further down the task list) seemed to be the harder tasks. This is in part due to the more complex tasks being at the bottom. This makes sense when we look at the results. It can also mean that some tasks were not designed easy enough to follow so even though they are complex tasks there is a way to make them easier for the user. These are the parts that would need more extensive testing and redesign. The easiest tasks were either task 1 or task 2. This is expected as those are the easy tasks and they are the easiest concepts to understand. When we asked for a favorite task half of the results were what we expected, they users listed a task within one or two of the task they picked to be the easiest. Although this was the majority case, we also found some users liked some of the complex tasks. This we believe attributes to the users personality as they may have liked the challenges. It was found that the mean of users listed their least favorite task as was expected. But in other cases we saw that users disliked tasks that they did not find the hardest. This is interesting because we think that the answers vary because it is up to the user to decide why they like, or why they do not like a task for whatever criterion they reason out.

When asked if users would like to use the application and would find it useful the answer was a unanimous yes, although some wanted to see some improvements. Such suggestions for improving the application consisted of showing a star rating or other information about a movie. This could mean all the time or based on certain criteria that would be searched for. It was also said that the search bar was not labeled so it was not clear that you could type in it to filter the search results. Every user said that they could not get into the movie details right away and it was not obvious so this would have to be improved as well. One user suggested that there be a help section in the application that someone could read to educate them on how to use it properly. Something like a crash course guide would be helpful, although with better redesign the need for this will dissipate. There was also trouble observed with not only viewing movie details but also with deleting a movie and adding a movie to a user list. To do this you had to open a window first and it was suggested you do it with drag and drop to move movies into lists. Besides that users commented most everything was fairly straightforward.

# Experiences

We found that the observation strategy was very useful when getting a user to complete tasks from the list. This really showed us where a user got stuck and what they didn’t understand how to do. We could also see where they moved to mouse and the things they tried, so based on some of that we could figure out how to fix some of the difficulties they were having. We also saw the tasks that the users did not have problems with. This was useful because we realized that these are the tasks we need to pay attention to as they are the ones that have been designed the best for the user. A structured questionnaire is useful be we learned that the questions need to be posed in such a way that if we want a closed answer such as a rating, we can’t let things like their mood or personal judgement skew their answers.

Next time we would try to pose a few better questions. As mentioned earlier some were not the best so they would have to be posed differently the next time. User testing was done in a somewhat cramped space and we feel that this may have attributed to some of the lower ratings on the questionnaire. Next time we would want to change the environment and see how a user would respond compared to this time with everything the same on the task list. It would also be beneficial next time to do the user testing in intermediate stages so that it would be clear to see what needs work before moving on to the next thing. This would require more time with the user but in the end we feel it would bear a better end product.

User testing is very good for seeing what a user would want and does not like about an application. In future projects we can incorporate the techniques used in this project (albeit refined and a little shorted maybe) to evaluate how we as developers are aligning with user needs. It would also be good to try out other techniques with different projects as the functions of different projects differ and require other forms of evaluation. Any techniques used would be beneficial if we were in a work term for coop because this would really show they employer that we can do a few important things well. These things would include working with clients, redesigned to align with client needs, ensuring client relation, ensuring client happiness, and as well honing our evaluation skills so we can become more efficient in doing all of these tasks.