Methods Assignments: M5

Redesigning Youtube Recommendation System

Gaurav Pande
gpande@gatech.edu

Abstract. We all want to design things that touch our lives somewhere so that we can interact with them better. I want to design a better intuitive recommendation system for the Youtube mobile application. Specifically, I want to focus on the task of playing with the recommendations of videos in the mobile application like it's a child's play to remove recommendations, to add recommendations, to mark/block videos from recommendations.

Qualitative Evaluation

Report on evaluation as a whole

As part of my qualitative evaluation, I chose to conduct a survey for my wireframe prototype build in M3 exercise. I have attached all the **wireframes** in the **appendix** for the reference. The survey had total of **11 questions**, most of which are to decide the qualitative analysis of the prototype, like for example "How useful do you think wirefram1 will be in managing your recommendation?". The questions were designed to get qualitative opinions of the students on efficiency, quality, learnability of the prototype. I had **6 questions** based on YES/NO answers, and the rest of **5 questions** were based on the filling the textual opinions about the prototype(like if you do not agree with wireframe, then why do you think so?). I limited the number of responses to be 25, and I managed to get all the **25 feedbacks quickly**. I find most of the survey response genuine because students who have selected "No" for any wireframe, have explained why do they think so as well. In some questions maybe it is possible to get inaccurate opinions because students might fill the survey very quickly to earn extra credit. I should have opened the survey for more than 25 people to get a better approximation of the qualitative evaluation. Overall the survey and its results very worthwhile to carry forward my analysis.

Report Raw Results

The raw results of the survey have been exported to a CSV file and I have attached the link in the **appendix**. I have also attached the analysis or graphs for most of the questions which ask the opinion from the students on whether they think a particular prototype will be useful or not. Here are the main points which I analyze from the survey result:

There were a total of **4 wireframes** involved(see **appendix**). You can clearly see from the bar graph in figure 7 in the appendix that **96 percent** of the students agree that the **wireframe 1**(figure 1) will be useful in marking a video a **"Not interested"** will be discoverable and comfortable to execute. Similarly, for **wireframe 2**(figure 2), **88 percent** of the responses strongly agree that swiping right a video one can mark a video as **"Interested".** Some confusion I saw that some people suggest why not swipe up to mark a video as interesting, however, it does not seem an appropriate gesture because swiping up and down resembles more towards the browsing the videos, so to avoid the confusion I chose to make swipe right for marking Interested. **92 percent** of people strongly agree that **UNDO** option is a good idea to undo an action done in wireframe 1 and wireframe 2, which provides a good tolerance in the design. In response to the question asked from the expert users of the youtube mobile application about whether they think it is a good idea to have gestures drawn in the screen using fingers in the screen to mark a video "Interested/Not Interested", mostly everyone disagrees complaining that it will be uncomfortable and will involve using both the hands, which is not a good idea. **80 percent** of the responses believe that marking the video or **tagging** a video will help them in managing their recommendations better. Overall **88 percent** responses **strongly agree** that the prototype will be useful in managing the recommendation in the youtube mobile application.

Analyze the survey responses

I think I have received overall positive feedback on the prototype build using wireframes. From all the wireframes, **wireframe 4**(figure 4 in the appendix) is not considered as a good option to have because respondents of the survey feels it will have a larger **Gulf of execution**, but I only considered this option for the **expert** users as they should be comfortable with even drawing the gesture quickly on the

screen in between the video and once they complete the gesture the video either will be added to the recommended list or removed from list. Overall, mostly everyone agrees with the idea that this prototype will be useful in managing their recommendation.

Changes that feedback would already suggest in the prototype

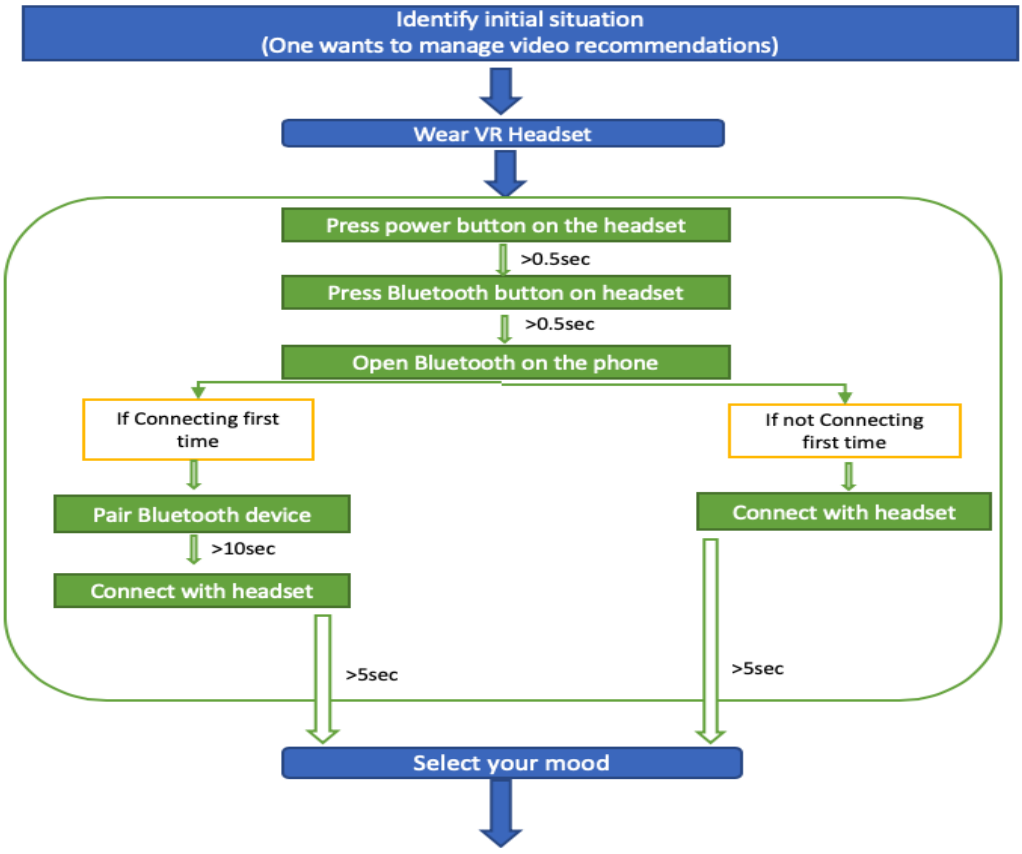
Some of the changes I would like to make based on the survey feedback is, one I will not use the gesture-based wireframe or will try to do more needfinding on how expert users would like to control their recommendations. In the **wireframe 3**(figure 3 in the appendix), I realize that I should have described it a bit more on the wireframe how it can control the recommendations.

Predictive Evaluation of textual Prototype

To perform the predictive evaluation of the **textual prototype** mentioned in the **appendix**, I have used a **GOMS** model to analyze the Goals, Operators, Methods and Selection rule associated with my textual prototype.

The GOMS model in **figure 10** below 2 sections or flows to complete the task of managing the recommendations. The first is to install or equip the VR headset and that installation itself involves sub methods and operators and selection rules. If the headset is connected for the first time then the device needs to be connected or paired with the mobile application for the first time, else the device can be directly connected once it is powered on. The first operation would take longer(>15 sec) than the second one(>5 sec). Once the the device is connected, the next phase is to **select** the **mood manually** from the youtube mobile application which involves **operators** like opening the application and choosing a mood that user is feeling. Once this selection is done then the role of VR headset comes which **auto detects** the mood of the user and does **actionable intelligence** based on the mood selected by the user and the mood detected by the VR headset(EEG). The GOMS model here gives an example but the application or usability of this application can be very wide as there are many emotions other than mentioned in the GOMS model here. Let's take an example, if the user is feeling happy than the VR can simple based on the watch history recommend cheerful videos and detects the mood. If the mood of the user changes or deviates, then it's time to recommend a new video which can bring

back the mood of the user to its original. Now another selection here is that what if the user is having violent thoughts or aggressive thoughts than the VR application should be smart enough to suggest videos which can pacify the mood of the user and make him cheerful or calm or peaceful. The whole idea of the VR headset here is to balance the emotions or thoughts of the user using the youtube application.



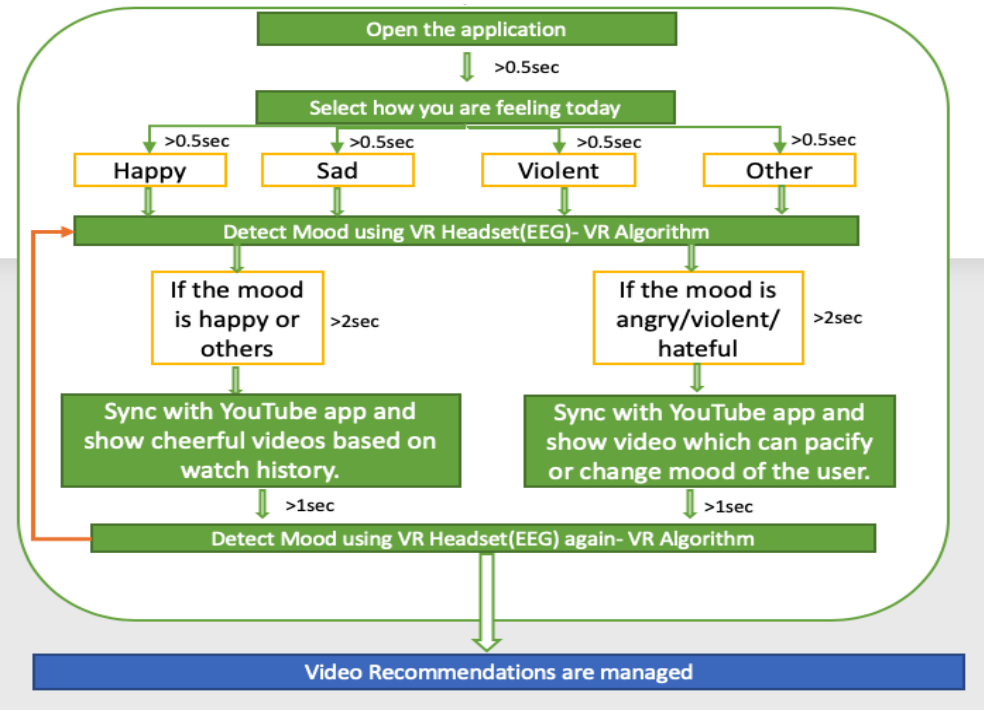


Figure 10: GOMS model for Textual prototype

Potential Improvements

As previously mentioned, there can be many many emotions or thoughts that a user might be feeling before starting the application. The VR application should be able to auto detect the emotions using EEG and balance the emotions by recommending the videos. I should have included the manual approach in the GOMS model so that the user can always fall back to the default behavior whenever he wants. One more improvement here which i did not consider before is the headset physical size because it will be quite clumsy to wear a big device and watch the video. But this would involve a completely different design and needfinding approach.

Evaluation Summary

Additional Needfinding

It is clear from the survey that users who use the application frequently find the use of the gesture-based marking a video as Interested/Not interested inappropriate, however, it was unclear what other alternatives we have to provide better options to the expert users.Some better questions needs to be asked in the needfinding exercise to know the correct behavior of the expert user. From the feedback, it was also evident that the discoverability might be an issue as swiping left or right would not be self-evident at least to the novice users, so what other options we have. These options we have to find out using another exercise of needfinding. There can be solutions like adding a tutorial when the user opens the application for the first time or provide an option/button for the tutorial. But as explained before we can only say about the better solution after doing another needfinding exercise.

Future surveys should be focused to ask users to walk through specific tasks and rate how easy it was to find an option or task.

Prototype Revision and Design alternatives

As it was apparent from the survey that I need to come up with design alternatives which can make discoverability as one of the prime focus. The alternate design should include a tutorial to walk through the users the first time or provide a button/icon to turn the tutorial on or off. The target would be the direct manipulation of all the task involve managing the recommendations.

From GOMS model I need to design the headset physical parameters and there needs to be a new needfinding exercise to design the headset itself because the whole VR application depends upon how easy or comfortable it is to wear the Headset and watch the videos. The success of the GOMS model also depends on the algorithm or design of the VR headset, through which it will be able to sync the thoughts/emotions of the user and recommends the video. In the textual prototype, there needs to be another interface which will be an algorithm or software acting as an intermediary whose design we did not discuss.

Next Steps

Based on the valuable feedback received from the survey I can say that the wireframe prototype is ready for making some positive impact on the user experience for controlling the recommendations. Probably one more iteration of needfinding will include the discoverability and direct manipulation to make the design better.Based on the GOMS model I can say that there are significant areas of improvement which I think are very critical for the success of the GOMS model. We will need a couple of more needfinding exercises to suggest the revised prototype which will result in a fully blown VR application for managing the recommendations.

Appendix

Wireframes

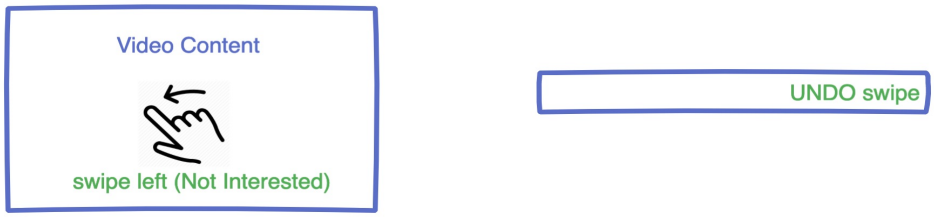


Figure1: Wireframe 1 for marking a video as "Not interested" by swiping to the left and the Corresponding UNDO option

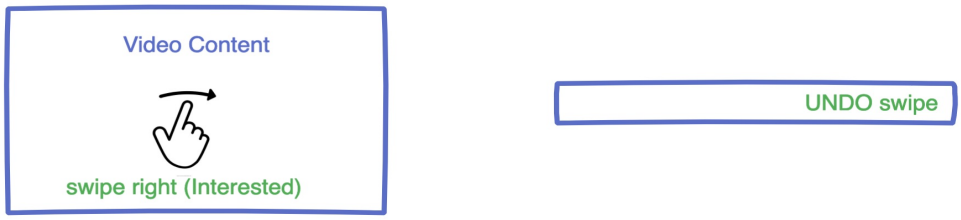


Figure 2: Wireframe 2 for marking a video as "interested" by swiping to the left and corresponding UNDO option

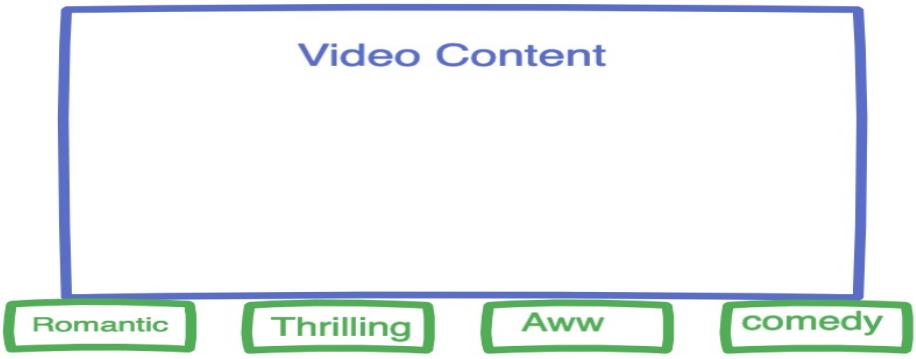


Figure3: Wireframe 3 for categorizing/tagging a video for better management of recommendations

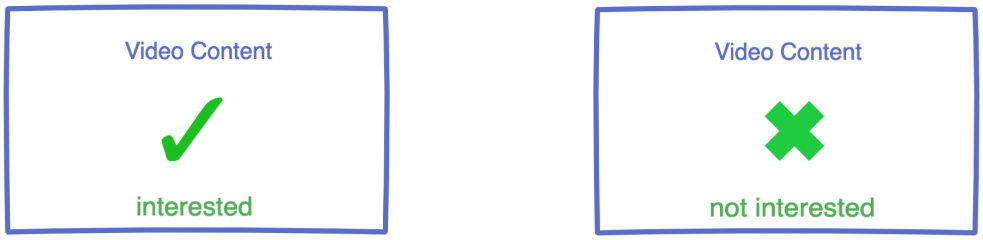


Figure4: Wireframe 4 for gesture-based marking a video as Interested/Not Interested by drawing the gestures in the screen.

Survey results

Raw results in the form of a CSV file: link

Graphs on survey results:

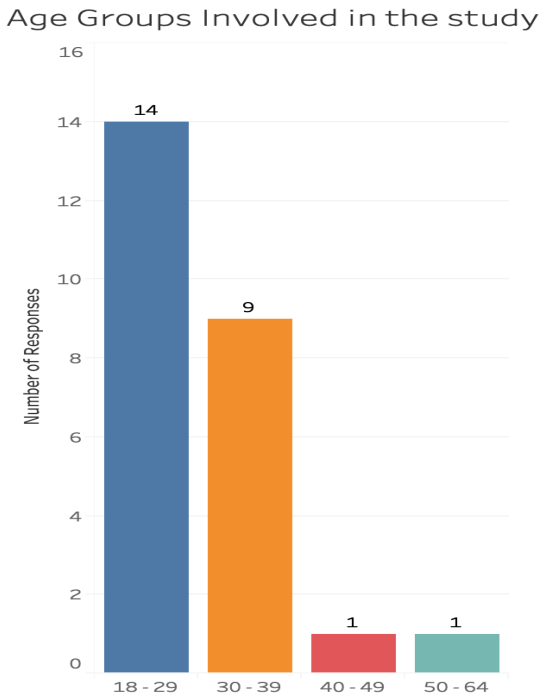


Figure 5: Age group distribution

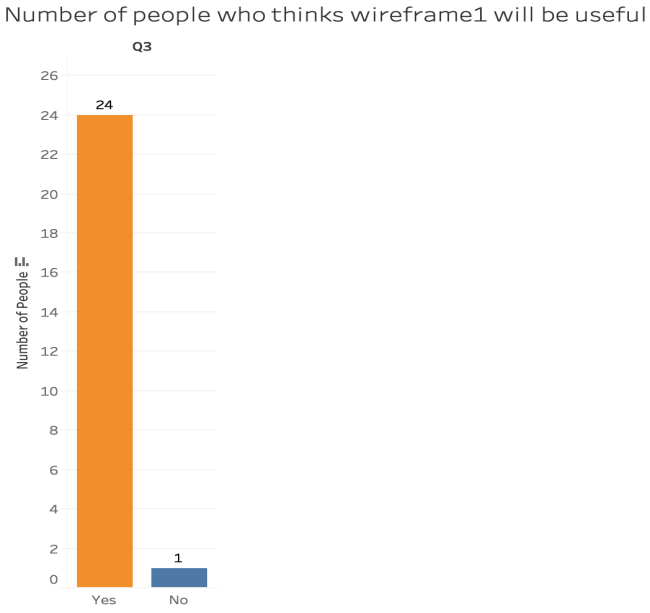


Figure 6: Wireframe 1 responses

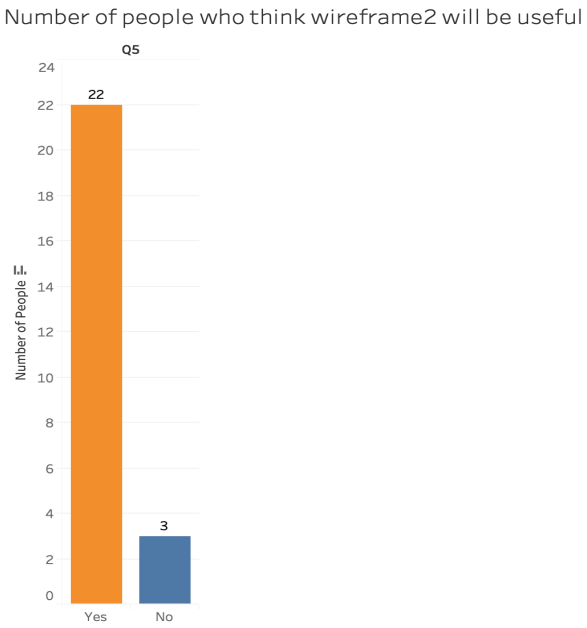
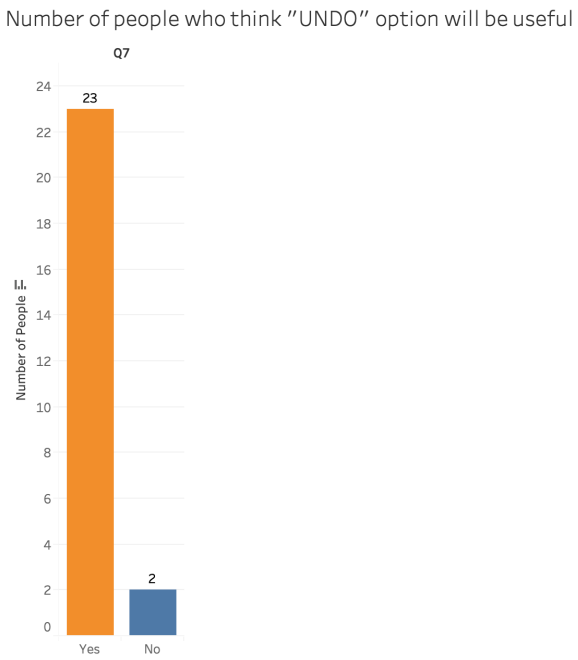


Figure 7: Wireframe 2 responses



Firgure 8: Responses for an UNDO option

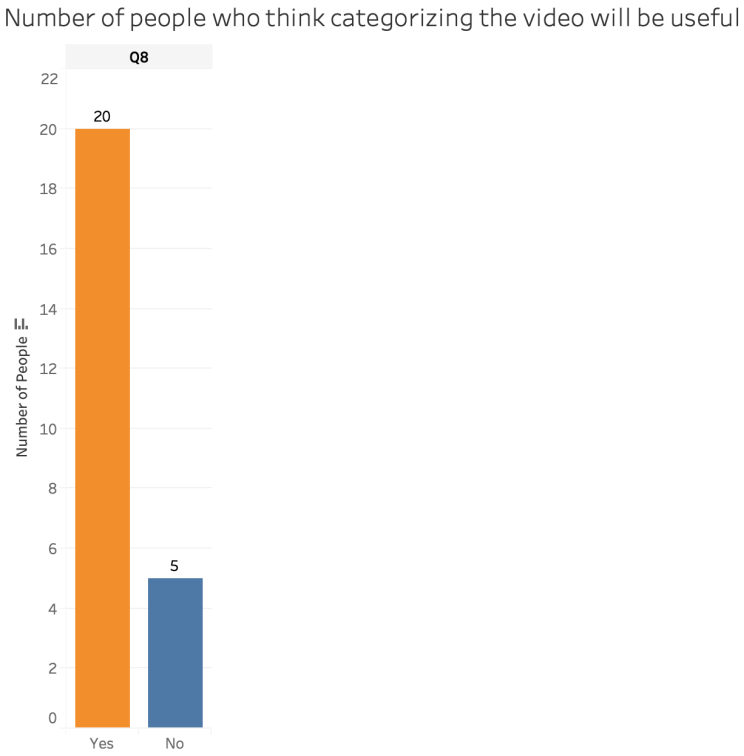


Figure 9: Responses for tagging a video

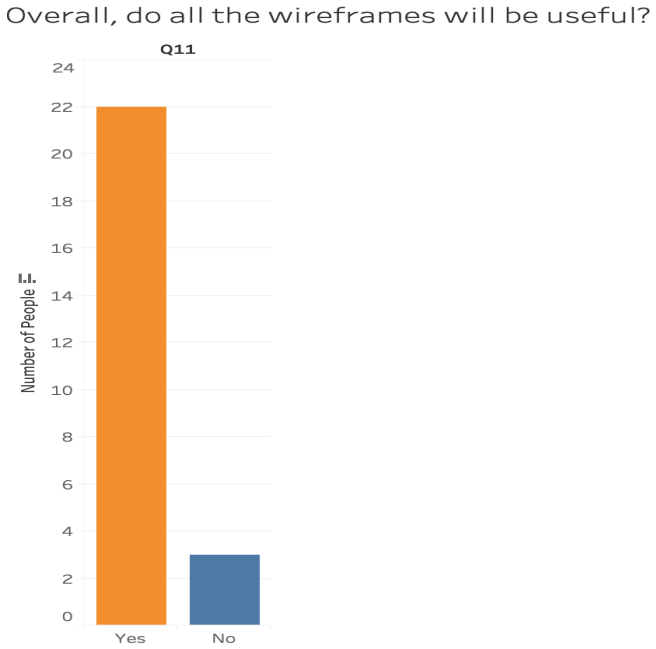


Figure 10: Overall responses on all the wireframes

Textual Prototype:

This prototype will include the virtual reality and I would like to add it as a textual prototype to better describe my intentions while designing the 3rd design alternative:

*"Integration of the **virtual reality** to control the recommendation with **zero cognitive loads**"*

How can I achieve zero cognitive load

Imagine a wearable device like neurable[1] which is made of electrodes to measure your brain activity. The purpose of this device will be to measure your brain neuro-signal to determine a user's intention or mood while watching a video, like for example if while watching a video I am feeling good that the device should be able to capture the brain signals and accordingly mark that video or channel as "Interested" or "Not Interested". The device will use the dry electrodes to record the brain activity via EEG(electroencephalography[2]) and then a software can be used to determine the activity or mood of the person and according to the software will add or remove the video from the recommendation list. The user has to literally do nothing except watching a video thereby reducing the cognitive load to zero as the user will not have to remember his interest or whether he like or not liked the video the device will start improving quickly as it will consider the brain activity.

[1] http://www.neurable.com/
[2] https://en.wikipedia.org/wiki/Electroencephalography