

Measuring the Annoyance in Streaming Media Caused by Buffers and Interrupts

Joshua Allard
Andrew Roskuski

March 6, 2015

An Interactive Qualifying Project Report:
submitted to the Faculty of the
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements
for the Degree of Bachelor of Science

Approved by:
Professor Mark Claypool, Advisor

Abstract

With the increasing demand for streaming media, it is important to study the effects of quality defects in a video. We concentrated on how buffers and interrupts can impact a user's quality of experience by measuring how a user becomes annoyed with their presence. We determined annoyance by having the user watch a series of videos with different buffer times or interrupt count, asking for the perceived annoyance afterward. The results from 37 users show a polynomial relationship between buffer time and annoyance and a logarithmic relationship between interrupt count and annoyance. We were able to separate our videos into different levels of motion, and, while we found no relationship between annoyance caused by buffers and the motion level, we found an unexplained relationship between annoyance caused by interrupts and the motion level; the lowest motion level seemed to produce the highest annoyance rating. These results can be used by streaming services to understand how buffers and interrupts affect users, which can help find a better balance that minimizes annoyance.

Contents

1 Introduction.....	5
1.1 Hypotheses.....	7
1.2 Our Study.....	10
2 Related Work.....	12
2.1 Background Information.....	12
2.2 User Studies.....	13
2.3 Video Classification.....	16
2.4 Summary.....	16
3 Methodology.....	18
3.1 The Proposed Study.....	18
3.2 Preliminary Actions.....	19
3.2.1 Resources Needed.....	19
3.2.2 WPI IRB Approval.....	21
3.3 Video Preparation.....	21
3.3.1 Video Requirements.....	21
3.3.2 Video Categories.....	22
3.3.3 Process For Finding Videos.....	23
3.3.4 Encoding Videos.....	23
3.4 Pilot Studies.....	24
3.4.1 Process.....	24
3.4.2 Observations.....	25
3.5 Soliciting Participants.....	26
3.5.1 Promotion via Email.....	26
3.5.2 WPI Social Science Research Participation Program.....	26
3.6 Procedure.....	27
3.6.1 Lab Setup.....	27
3.6.2 Preliminary Participant Steps.....	29
3.6.3 Video Survey.....	29
3.6.4 Debriefing.....	32
4 Results & Analysis.....	33
4.1 Study Results.....	33
4.1.1 Demographics.....	33
4.1.2 Video Content.....	35
4.1.3 Annoyance Levels.....	37
4.2 Data Analysis.....	38
4.2.1 Buffers.....	38
4.2.2 Buffers and Motion.....	41
4.2.3 Interrupts.....	41
4.2.4 Interrupts and Motion.....	43
4.2.5 Buffers and Interrupts Combined.....	44
4.2.6 Effects of Video Content.....	46

4.3 Postmortem.....	47
4.3.1 User Feedback.....	48
4.3.2 Problems Encountered.....	48
5 Future Work.....	50
6 Conclusion.....	52
7 References.....	54
8 Appendix.....	56
8.1 Motion Levels.....	56
8.2 Script.....	57
8.3 Debriefing Form.....	58
8.4 Additional Documents.....	59

1 Introduction

There are two major ways of watching video content over a network. The first is to download the entire video and play it locally on a machine, but this requires delaying the playback of the video for a relatively long time. The second, streaming, is to watch the video as it is being downloaded. The fact that streaming allows the user to begin watching the video sooner is one of the reasons it is commonly used on the Internet.

The Quality of Experience (QoE) of streaming videos is the measure of a user's experience with the playback of the video. The QoE depends, among other things, on the network Quality of Service (QoS), which encompasses many technical factors, such as network delay and packet loss. Generally, as the QoS degrades, so does the QoE, but the exact relationship between QoS and QoE has not been thoroughly studied. Buffering and interrupts in the playback of a video are two common factors degrading QoE.

Buffering is the process of delaying the playback of the video until a portion of the video's data has been downloaded. For the sake of clarity, we will only use the word buffer to refer to the initial buffering time, and not the buffering that may occur during an interrupt. The downloaded data is stored in memory by the video player and once a certain threshold of data has been collected, playback of the video can begin while the rest of the video is still being downloaded, which helps prevent the occurrence of interrupts. Interrupts are the pauses in a video caused by the video player not receiving the data in time for it to be displayed. This is the result of the video playing at a higher rate than the video is being downloaded. Buffers and interrupts generally exist in an inverse relationship; the larger the buffer size, the fewer interrupts, and vice versa. On one extreme, when the buffer size equals the video size, then the video is guaranteed to have no interrupts, but this makes the process equivalent to downloading the entire video, defeating the purpose of streaming. On the other extreme, having no

buffer maximizes the chance for interrupts, since there is less protection against fluctuations in the network QoS, such as network congestion and low capacity Internet connections. The challenge is to select the proper buffer size to maximize the QoE given a particular network connection.

Buffering can affect the QoE by forcing the user to wait before the playback of the video can begin, which can cause the user to feel impatient. Interrupts can also make the user wait and feel impatient, but they occur after the video begins, which creates a discontinuous flow, similar to a phone call during dinner. Since buffering and playback interruptions are often unavoidable due to network limitations, determining how much these factors annoy the users and degrade their QoE can enable content providers to better serve their users as well as to inform video players about the proper buffer sizes.

When users attempt to utilize a streaming service, they expect to have a generally positive experience, with minimal buffers and interrupts. Therefore, it is important for a streaming service to try to minimize the annoyances received by the user, so that the service can remain competitive with providers of the same material. For example, if two companies both serve the same movie and their services both cost the same, then a user will generally pick the service that produces the best QoE. A company can modify the buffering policy in its video player, which can control the ratio of buffers to interrupts, given a certain network environment.

Video streaming players and providers could benefit if they know what users preferred; if users preferred buffers to interrupts, they can focus on that. The same can be said if users preferred interrupts. Assuming the preference is not binary, (i.e., not wholly one over the other), then a balance could be found that minimizes annoyance for users. With this balance, a video player could monitor a user's network status and create a buffering policy that would produce the balance of buffers and interrupts that would maximize QoE.

Companies may also benefit if they know a user's perceived annoyance received from buffering and interrupts varies based on the content of the video, such as its level of motion. The motion of a video is how the characters, background, and other objects in it change with respect to time. A user watching a video with higher motion may be more annoyed because the interrupt may be more noticeable, than if watching a video with lower motion. A streaming service could decide independently how to balance interrupts and buffers based on a video's motion in order to produce a better QoE for the user. In order to ascertain whether video motion impacts QoE, a video's level of motion must be measured to see if it affects the degradation of the QoE. The amount of motion of a video can be measured by how much of the image's pixels change from frame to frame on average.

1.1 Hypotheses

In our study, we use previous information and common inference to hypothesize the relationships between buffers, interrupts, and user annoyance. Our general hypotheses are discussed first, followed by specific hypotheses to illustrate the relationships.

We can determine the QoE for a video based upon the user's perceived annoyance caused by the QoS, namely buffers and interrupts; as the level of annoyance increases, the QoE decreases. We hypothesize that the annoyance caused by the buffer time of a streaming video starts low and rises logarithmically until a certain point is reached, where the level of annoyance increases exponentially. We make this assumption based upon Ebner's [1] depiction of a user's level of annoyance while waiting for a printer to respond. The level of motion in the video does not affect the annoyance produced by buffers, because the motion does not begin until after the initial buffer. For interrupts, the user's annoyance rises logarithmically as the number of interrupts grows, increasing rapidly at first, with the annoyance leveling off after a certain point, approaching a maximum amount. With respect to interrupts, high levels of motion amplify the rate of growth of annoyance, while low levels of motion

reduce it. This is theorized because users watching videos with more movement have more to focus on, which is disturbed by the occurrence of interrupts. Combining these hypotheses, we produce a relationship between the effect of buffers and interrupts on the QoE with three different datasets, based on the level of motion.

We summarize our hypotheses into the following statements:

1. As the buffer size increases, the annoyance level increases slowly at first, but grows exponentially after a point as shown in Figure 1.
2. As the buffer size increases, the level of motion does not affect the annoyance level.

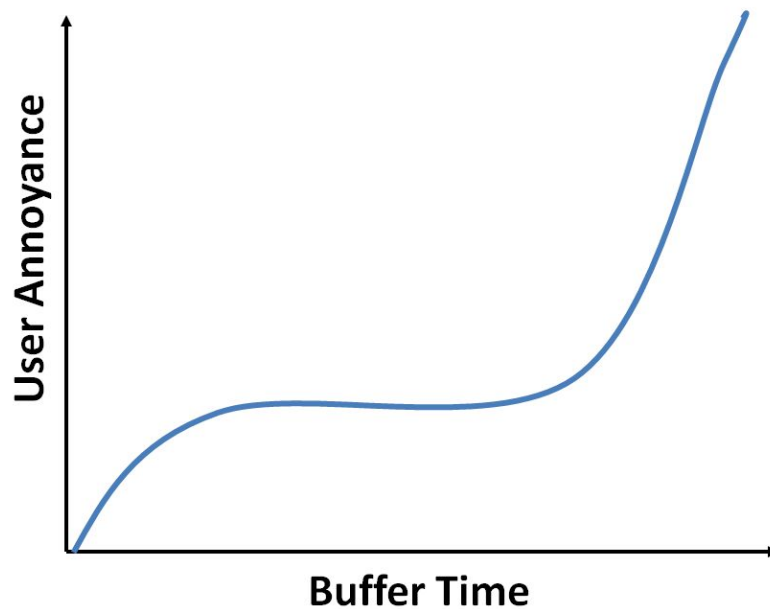


Figure 1: Hypothesis 1. Annoyance Increases Similar to Degree 3 Polynomial as Buffer Length Increases

3. As the number of interrupts increases, the user's annoyance increases logarithmically as shown in Figure 2.
4. The rate that the user's annoyance from interrupts is amplified by the level of motion in the

video as shown in Figure 2.

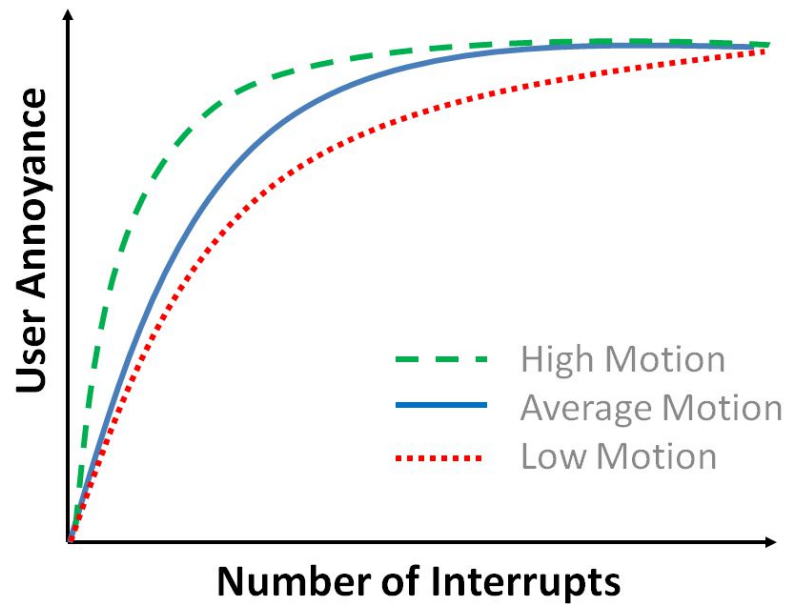


Figure 2: Hypotheses 3 and 4. Annoyance Rises Logarithmically with Number of Interrupts and is Amplified by Motion.

Combining the effects on QoE of buffering and interrupts and normalizing the levels of annoyance, we get the graph in Figure 3. Given a buffer size and a number of interrupts the annoyance level can be determined. The darker shades on the graph represent the user being more annoyed. The bands on the graph represent the annoyance on a discrete scale from 1 to 5. Given either buffers or interrupts, the other factor needed to remain at the desired level of annoyance can be determined.

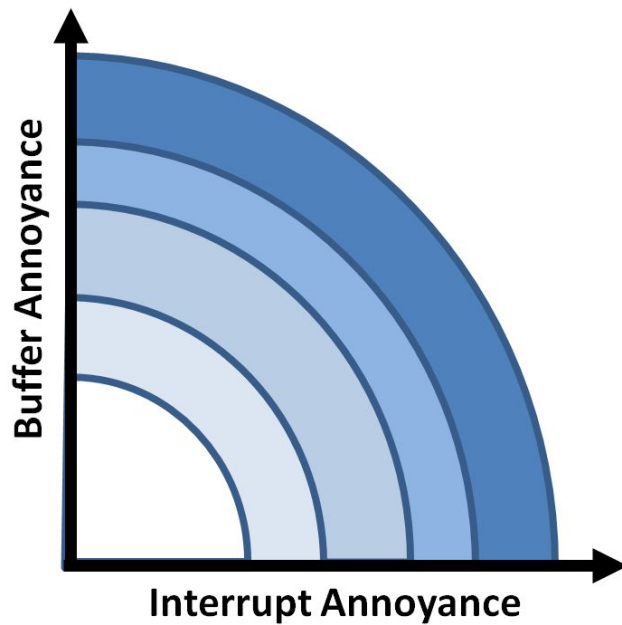


Figure 3: Theoretical Application of Results. Given Interrupt and Buffer Annoyance, Overall Annoyance Can be Determined

1.2 Our Study

In order to test the hypotheses in Section 1.3, we designed and executed a user study to determine annoyance cause by buffers and interrupts, independently. Each user was asked to watch a series of videos containing only buffers or only interrupts and rate annoyance with the defects on a scale from 1 to 5. The videos watched by the users were collected from YouTube according to guidelines we developed. Each video had varying amounts of buffers or interrupts, which allowed us to analyze annoyance over a range of buffers or interrupts. Each video had a varying level of motion so that the effect of motion on annoyance in the presence of buffers or interrupts could be observed.

We ran our study for four weeks; for the first two weeks we exclusively recruited participants via email advertising and referrals and for the remaining two weeks we solicited participants solely from WPI's Social Science Participant Pool. By the end of the study, we had a total of 37 participants with about 600 data points. 19 of these participants watched only videos with interrupts, while the

remaining 18 watched only videos with buffers.

After completing our study, we analyzed the results to test our hypotheses. Regarding Hypothesis 1, we found that the annoyance caused by buffers followed a degree 2 polynomial, instead of the degree 3 polynomial we were expecting. We also found that there was no correlation between motion and annoyance caused by buffers, confirming Hypothesis 2. For Hypothesis 3, the annoyance cause by interrupts followed a logarithmic curve, as predicted. However, the annoyance values were consistently highest in videos with low motion, leaving Hypothesis 4 unsupported.

Section 2, Related Work, discusses previous research and studies that we modeled our study after, including background information on streaming media, similar user studies, and ways to classify videos. Section 3, Methodology, describes our process for preparing and executing our study. Section 4, Results & Analysis, discusses our findings from the study, divided into three parts: the raw data, the relationships found, and the reflection on the study. Section 5, Future Work, briefly outlines future studies that can build upon the relationships we found. Section 6, Conclusion, summarizes our findings.

2 Related Work

There has been research that can aid a study of a user's preference for buffer lengths against playback interrupts. In order to understand our study, knowing prior work can help inform the reader about the current situation in streaming media, such as the necessity of buffers and interrupts, and general concepts of user studies. The research has been divided into three sections: background information, consisting of information pertaining to buffers, interrupts and general ideas; user studies, consisting of previous studies that test streaming media concepts, from which we drew our methodology; and classification information, containing details that helped determine how to classify the videos in our experiment.

2.1 Background Information

Kumar et al. [2] investigated maximizing the QoS of mobile video. The QoS was measured using various metrics such as starting delay for buffering, frames dropped, and rebuffering percentage. The methodology focused on purely improving QoS using simulations to determine optimal conditions, while our study focuses more on the QoE.

The Online Reporter [3] gave background information into the current situation with streaming media. It detailed how too many people are faced with buffering and video interrupt problems, such as how 20% of the video streams they monitored contained at least one interrupt and 20% had long buffer times. It also discussed how video services with longer buffer sizes cause people to abort the video much more than videos without. This article shows how a quantitative user study is needed to show streaming services the relationship between buffers and interrupts for the user's experience.

Entertainment Newsweekly [4] described in its article a new invention that is intended to reduce

the buffer delay at the beginning of streaming media to a minimum while still avoiding interrupts. The article provides a great technical overview of the current situation with streaming media, explaining why buffers and interrupts existed in the first place, despite the fact that the solution it discusses is not relevant to our study.

Eufemia et al. [5] investigated the relationship between buffer size and interrupts for YouTube videos. They first captured actual playback data, and then created simulations based on that data of how the playback would have gone with various initial buffer sizes. The data that they collected was useful for determining realistic values for buffer time and interrupt count.

Liang and Liang [6] numerically analyzed methods of compensating for interrupts in the video stream. They did so by means of running simulations based on existing captured network data. While the findings in the paper will be useful for drawing conclusions, such as the odds of interrupts occurring based upon certain buffer sizes, the methodology does not translate well to what we tried to do.

2.2 User Studies

Qiu and Cui [7] mined 2 years worth of usage data from MSN video in order to find general patterns and relationships. They used the data to build a model to predict user satisfaction based on factors such as initial buffer time, number of interrupts, and video popularity. The methodology in this paper focused on judging user satisfaction by how long the video was watched, while we focused on asking the user directly about the experience.

Balachandran et al. [8] examined various metrics in an effort to find what actually matters to the quality of the user experience. While they admit that their findings are limited by simply using play time as the measure of the experience, the paper made important observations about the content of the videos used for testing, and raised certain challenges, such as difficulty keeping the users' interest

levels constant, in running user studies related to streaming video. Similarly to Qiu and Cui, they tested the users' annoyance with the video by monitoring when they aborted it, while we directly asked the user about their annoyance.

Wu [9] investigated methods for automatically adjusting the stream for optimal experience. The work contains useful information related to subjective assessments of video, such as for determining video motion levels. This helped to inspire how we classified videos based on motion.

Tavakoli et al. [10] studied the effect of video resolution changes on a user perceived quality. They measured different conditions to evaluate the user's level of annoyance: increasing, decreasing, and oscillating quality. They broke each video clip into segments and had users evaluate each part and the clip as a whole. This article gave us ideas on how to organize a user study further and even provided an efficient evaluation scale similar to the one we used in our pilot studies, but later dropped.

Claypool and Tanner [11] detailed that there is not much information on how jitter affects a user perception of a video's quality, and that there is little information comparing the difference between jitter and packet loss for the user's experience. They also present another methodology for a user study; they presented 25 videos from 5 categories of temporal movement (the amount of motion in a clip), and asked each user to rate the video using a slider, which was mapped to values from 1 to 1000. Besides varying the amount of jitter and packet loss for each clip, all other factors were held constant. They divided the 41 users into groups that assured an equal amount of users saw each clip at the varying levels of jitter/packet loss. This user test is extremely similar to our study.

Chen et al. [12] performed a massive user study in China. They attached themselves to a major content provider in China and monitored how its users were watching streaming videos. This represents a more passive approach to user testing, where the user is not fully informed, and the results are drawn from their non-manipulated actions. The results of their paper indicate common behaviors that are beneficial to consider when performing user studies with streaming video, such as how 80% of the

time, a user did not watch a full video. While they were investigating user behaviors without QoE issues, concepts from this study were applied to our study, such as maintaining minimal manipulation over the results.

Mu et al. [13] created user tests to determine the affect of QoS on user experience. Their process included showing a user an unaltered version of the film first, showing them the same video again with a little lower QoS, and then asking the user how annoying the lowered quality was on a scale of 1 – 5. We based our scale off of the one used in this study. They also indicate that it is important to get background information on the participants. This experiment designates one way to perform a user experiment on video media.

Finamore et al. [14] designed experiments to measure YouTube's network traffic, varying by two different media: mobile devices or traditional PCs. Their results show that users' behaviors do not vary too much amongst the different devices, with users leaving a video normally within the first 40 seconds - this indicated that we should use short videos in our testing, because we did not want this behavior to be a factor in our study. The paper also discussed YouTube's aggressive buffering policy, indicating the need for a more optimized process. This showed that a study comparing the balance of buffer size and interrupt occurrences could help companies change their policy to fit what the user desires.

Tripathi and Claypool [15] investigated user response of various techniques of scaling the quality of video. The videos were classified based on having high or low levels of motion. The amount of motion was determined by the amount of interpolation in the encoded file, found using a tool named mpeg_stat. They prepared videos to use in their user study by pre-processesing the video clips to have a version of the clip for each level of scaling. Users were shown 2 clips from each motion category at each of the scaling levels used, including not scaled at all. The users were asked to rate each version of the clips on a scale from 1 to 100, and mean ratings were used to analyze results. This methodology for

this study is similar to what we needed so we adopted it for our use.

2.3 Video Classification

Polacheck [16] investigated technical means of separating the traffic of live and on demand video. This was accomplished by analyzing the length of the packets traveling through the network, as well as their flow. The goal was to optimize the network for the playback of each one. The information in the paper was useful to help characterize videos.

Claypool [17] discussed the classification of games and videos as components of scene complexity and temporal movement. He discussed how he and others were able to determine these measurements from video captures of video games, as well as from the clips he was comparing the video games to. He mentioned separating a video frame into 16 equal rectangles, and estimating complexity and motion for each one through time, as well as using a tool to do the same analysis; we used the tool, `mpeg_stat`, that was referenced in this study to determine the level of motion. This resource presented the means for us to do our experiment setup.

Song et al. [18] described a method to organize videos. While we did not utilize a lot of the measures they use to classify videos, and the 1000+ categories they introduce, the study is a resource on using mechanics to distinguish video types, which was similar to how we decided to classify videos by motion. They helped us focus on measures of a video, as opposed to classifying video by genre.

2.4 Summary

Using the resources we gathered, we were able to develop our study. We found background information about buffers and interrupts, such as ways to measure QoS and how buffer size affects the occurrence of interrupts, which helped us throughout this study. We were also able to research many different user studies, so that we could develop our own. There were some complications that we

learned to avoid, such as simulating live content, and indirect user feedback. There were some concepts that we learned to apply, such as trying to hold as many factors constant as possible, and how motion affects the results. From several of the different models we found, we were able to create a methodology that allowed us to gather the data we desire, such as the ideas presented by Tripathi and Claypool and Tavakoli et al., which included methods for measuring the level of motion in videos and for surveying the users for their perception. The resources on classification turned us away from some methods, but also reaffirmed the idea to sort our videos by motion by providing several more methods for measuring it.

3 Methodology

In this section, we discuss the methodology for our study. First, we outline the resources needed for our experiment. Then, we discuss our process for gathering videos. Finally, we explain our procedure for the user study.

3.1 The Proposed Study

During A-Term of the 2014-2015 year, we planned a study for how to measure the level of annoyance associated with the presence of buffers and interrupts. We developed the idea to institute a study to analyze annoyance using a video survey. We ran the video survey with the following requirements/components:

- The videos we selected for the survey are modified with defects manually, with either an initial buffering time or a series of equally spaced interrupts.
- A user completes the following process when completing our survey:
 - The user fills out demographic information, (e.g. gender, age, major, and experience with streaming media),
 - The user watches a control video to understand what we define as “perfect conditions” (i.e., a video with no initial buffering time and no interrupts),
 - The user then watches a series of videos with only one factor, buffers or interrupts, and these videos contain varying levels of motions,
 - For each video, the user rates the video in terms of annoyance,

- After all of the videos, the user then answers feedback questions about how they felt about their factor and what annoyed them the most.
- In order to avoid real streaming issues due to network issues, the videos will need to be stored on the user's computer prior to their viewing.
 - We accomplished this requirement by having the study done in a physical lab environment with all of the videos stored on each lab computer prior to each run of a user study. This guarantees that there will be no network delays when watching the videos.
- We need an application to record users' feedback as they watch our videos.
 - To accomplish this, we used a HTML page that could host the local video files and create an Iframe (i.e., a webpage embedded within a webpage) for displaying an online survey.

3.2 Preliminary Actions

Before we could proceed further with our study, there were some activities we needed to complete beforehand. First, it was crucial that we gathered all of the resources necessary to complete our study, such as the videos and tools used. It was also important that we received the proper authorization to run our study, through approval from WPI's Institutional Review Board (IRB).

3.2.1 Resources Needed

For the purpose of this study, we needed:

- Videos – We needed a set of videos to manipulate and show to test subjects. Exact details about

the videos will be discussed in Section 3.3.

- `mpeg_stat` – We used this tool for analyzing the level of motion for each video. This tool takes in a video file, analyzes its properties, and reports statistics. From this tool, we looked at the skipped macroblocks statistic to determine the motion of a video. Macroblocks are the processing unit used for compressing the video. The skipped macroblocks measurement is derived from analyzing the average compression rate over the entire video; the use of this tool is covered in Section 3.3.2.
- `ffmpeg` – We used this tool alongside `mpeg_stat`. This allowed us to transcode the mp4 videos into the mpeg format needed for `mpeg_stat`. All transcoding was done using the default settings for `ffmpeg`. This tool can be found online at ffmpeg.org.
- Video Editing Tool – We used a video editor to add artificial buffering and interrupts to videos. Since the main task was inserting a video clip representing buffering at various points in the video, we kept things simple and used Windows Movie Maker, because it met our requirements and did not require further expenditures to acquire.
- Survey Tool – In order to collect the responses from our participants, we needed a tool that would handle our survey. We used Qualtrics, which is free for WPI students, as a tool for creating a survey design and for storing our data during the study's duration.
- Video Player Application - In order to present the videos we modified, we needed a video player application. As a result of trying to be compatible with our survey tool, we decided to display our videos using the Internet browser, Google Chrome.
- Data Analysis and Graphing Tool – We needed to analyze and produce visual representations of the data that we collect. We used LibreOffice Calc for our data analysis and presentation.

- Lab Space – We needed somewhere to conduct the tests where computers are provided. We needed to be able to install our test suite beforehand, and have the users be able to log on to the computers to take the test. We needed all of the computers to be running the same OS, and for them to be capable of running our selected video player. In order to meet these condition, we selected the Zoo Lab, which is in the sub-basement of WPI's Fuller Laboratories. The computers in this lab all were 64-bit, Windows 7 machines, with an Intel Core i7-3770 processor and 12 gigabytes of RAM.

3.2.2 WPI IRB Approval

In order to run our study using human participants, WPI required that we first seek approval from its own IRB. This process was to verify that we knew how to perform studies ethically, protecting the rights of our human subjects. We had to undergo the “Protecting Human Research Participants” training program, hosted by National Institute of Health. After that, we needed to submit an IRB expedited-review application, which consisted of listing what we would do with the participant's information and study data, as well as a consent form that we completed and needed each user to sign. Our IRB application and approval forms are included in Appendix 8.4.

3.3 Video Preparation

To perform the anticipated study, it was important to carefully select videos that will enable us to have reliable and consistent data. To do so, first we had to define the requirements and categories, then find videos that met the requirements, then finally modify the videos to our needs.

3.3.1 Video Requirements

In order to keep the image quality and level of detail in the videos fairly consistent, we set some standards for the videos we would use:

- All of the videos that we chose had a resolution of 1280x720 with 30 frames per second; this was chosen because this combination is very common, while still maintaining a satisfactory level of quality.
- All of the videos were in mp4 format, because of its widespread use as a video format.
- As for video content, all of the videos were primarily live action (i.e. real humans and objects), since recordings of realistic scenes have a relatively consistent level of detail compared to animated content, since animated content has many different artists and styles.
- For length, we initially chose to use 1 minute clips, since that seemed like a reasonable length to get proper feedback, but not take too much time. As explained in Section 3.4, this was eventually reduced to 30 seconds.
- All of the videos had to be a single scene with no cuts. This limitation was necessary due to the fact that the presence of scene changes would skew our measurements of the level of motion in the video.

3.3.2 Video Categories

In order to determine if motion made a difference in people's annoyance, we needed a way to classify the videos based on their level of motion. To this end, we used the skipped macroblocks statistic, as measured by `mpeg_stat`. This statistic represents how many macroblocks stayed the same between frames and thus were skipped during the encoding process for compression purposes. A high skipped macroblocks percentage indicates low levels of motion, and vice versa. We began by measuring a video with no motion, and a video where everything was moving to find the range of skipped macroblocks values we could expect for videos that we found. Once we had this range, we split it into 4 even chunks to form our 4 categories: High Motion, Medium High Motion, Medium Low

Motion, and Low Motion. For the values we used from mpeg_stat for these categories, see Appendix 8.1.

3.3.3 Process For Finding Videos

Once we determined the ranges for our motion categories, we searched YouTube to locate videos to fill them. All of the videos were downloaded via the site keepvid.com in mp4 format, and were tested in mpeg_stat after being transcoded to the mpeg format using ffmpeg. Four videos were selected for each category, in order to avoid bias resulting from video content. Of particular note is the fact that videos in the Medium High Motion category which met our restrictions were particularly difficult to find. For the actual content of the videos, we attempted to get as wide a variety as possible. As much as possible, we tried to cover a wide variety of tastes and interests, and ensured that there was minimal repetition of types of videos; we used content types such as music videos, talk shows, speeches, cat clips, dances, and sport videos. We also made an effort to keep the content of the videos fairly inoffensive. All this was to avoid bias in our results due to video content. To see a list of the videos used, view Appendix 8.4.

3.3.4 Encoding Videos

Since inducing actual buffers and interrupts during playback was determined to be too unreliable, we inserted fake buffers and interrupts into the videos. Each video was modified four different ways, two of which had buffers inserted into the video and the other two had interrupts inserted. Buffers and interrupts were kept separate in order to only vary one factor at once. The buffers and interrupts were represented using a video clip of the YouTube buffering animation on a black background. For buffer lengths, 2 seconds, 4 seconds, 8 seconds, and 16 seconds were chosen, as these values worked well in our pilot studies. Similarly, versions of the videos with 2, 4, 8, and 16 evenly spaced one-second interrupts were produced. Using the exponential scale allowed us to see more data

points with low buffer times and interrupt counts, while still having data points in the higher ranges; we believed the data would be more interesting at these lower levels, containing more drastic changes, but that we would still need the higher values to study our hypotheses.

3.4 Pilot Studies

Before moving onto the official user study, it was essential to run some pilot studies in order to fix as many possible problems and confusions that we could have overlooked. Our pilot studies were not extremely extensive, but were thorough enough to catch some major errors early on. We used three of our friends as subjects during this phase and asked them to be critical of our design. We followed a set procedure during these pilot studies and kept notes of the observations made.

3.4.1 Process

When performing these pilot studies, we adhered to the following process:

- We stored all of our videos and our HTML Survey page onto a flash-drive.
- We had them install our flash drive, and open the HTML page.
- We described what the study was for - determining a user's annoyance with wait-times and interruptions in a video's playback, (For the actual study, we wrote a script that we would read from).
- We had them fill out the demographic information.
- Then we explained how the video watching and rating process would work.
- We then sat near them, and observed them as they watched all 16 videos, plus a control.
- They would fill out the questions at the end, and then tell us what they thought.

Note: We were with the participant at all times during the pilot study, so that if any immediate problems arose, we were there to take note of it and so that they could make comments on any minor observations as well.

3.4.2 Observations

During the course of the user studies, the following observations were made:

- One user found that videos with a lot of sound made him more annoyed than videos with lesser amounts of sound. This was something we considered might be a factor in a user's annoyance, and is talked about in our Section 5. Another problem with sound, in respect to speech, was discovered during the official user study, which is discussed in the Section 4.3.
- One user found that when the slider was placed in the middle of the range for video quality, he felt much more comfortable moving the slider up, then down. We modified the value to start at the lowest value, 0, on a 0 – 10 scale, and later 1, when we moved to the 1 – 5 scale.
- One user found a mistake in our survey, where we wrote a "100" instead of a "10" when we were using a 0 - 10 scale.
- Some noticed that some of the video's content may be affecting their annoyance score, outlining the need for a future study that focuses on a video's content. They said they were more annoyed when a video they enjoyed was interrupted, as opposed to videos they cared less for. In order to account for this possible factor, we included a question about the video's content in order to gauge if they appreciated it or not.
- One user couldn't tell that the levels of interrupts were changing, but his scores did vary appropriately, (i.e., his scores said he was more annoyed by videos with 16 interrupts as

opposed to 8). This showed us that even though a user might not pay attention to the exact amount of defects, they can still become more annoyed at their more frequent occurrence.

3.5 Soliciting Participants

During the formation of our study's goals, we discovered that in order to get a proper participant population, that we would need between thirty and fifty people. In order to get these participants, we utilized multiple tools to broadcast our study, allowing us to achieve our goal, by having 37 total participants.

3.5.1 Promotion via Email

For the first two weeks of our study, we solely advertised our study via email and word of mouth. The template for the email we sent out was included as part of our IRB application (see Appendix 8.4). We used a raffle for a \$25 dollar Visa gift card as an incentive for people to partake in our study, which we distributed at the end of the two week period. While this method originally seemed to be fruitful, with 5 participants on the first day, we soon discovered that only one or two people would show up per day for the remainder of the two weeks. During the duration of this emailing advertising period, we received 12 participants; it became apparent that another measure was needed in order to solicit more participants.

3.5.2 WPI Social Science Research Participation Program

During the promotion via email phase, we knew we had to do more to attract participants to our study; from the advice of our adviser, we looked into WPI's Social Science Research Participation Program, which is commonly referred to as the Participant Pool. This program is used by WPI's Social Science Department as a tool to engage WPI's students in the area of Social Science. By using this

participant pool, the incentive was already supplied, because Psychology classes often require that their students participate in multiple studies.

We applied to run our study for the remainder of B-Term, which we did by signing up for 30-minute sessions, that allowed 5 participants each. We were remarkably surprised by the success of using this program. During the first full week that we ran the study using the Participant Pool, (we ran two sessions the week before, but no one signed up, possibly because of short-notice), we received 23 more participants for our study, allowing us to meet the minimum goal of 30 we had set earlier. After this success, we saw that having more participants would allow us to have a better data set, so we ran 8 more sessions during the first week of C-Term, but we only received 2 more participants. We are not entirely sure of why the difference in the number of participants occurred; it may have been because there is more of a rush toward the end of the term for people to complete these studies. In total, we received 25 participants from the Participant Pool.

3.6 Procedure

This section details the steps we took during our study. We outline the setup of the lab, what we required the participants to do before the survey, the actual process for the survey, and the debriefing section afterward.

3.6.1 Lab Setup

Before the entire procedure began, it was necessary for us to configure the Zoo Lab for our use. The first action taken was the choice of what computers in the lab to use, for we had reserved the whole computer lab, but we did not want to stop others from using the lab; as a result, we chose only a limited selection of computers. The computers that we chose were arranged in such a way that no one at one of our selected computers could see the screen of another participant, as Figure 4 shows;

Figure 4 shows where our utilized computers, marked with an “X”, were placed in respect to each other and the other objects in the room. It is important to note that the original selection of computers was dependent upon whether or not the computer was available during our selection process.

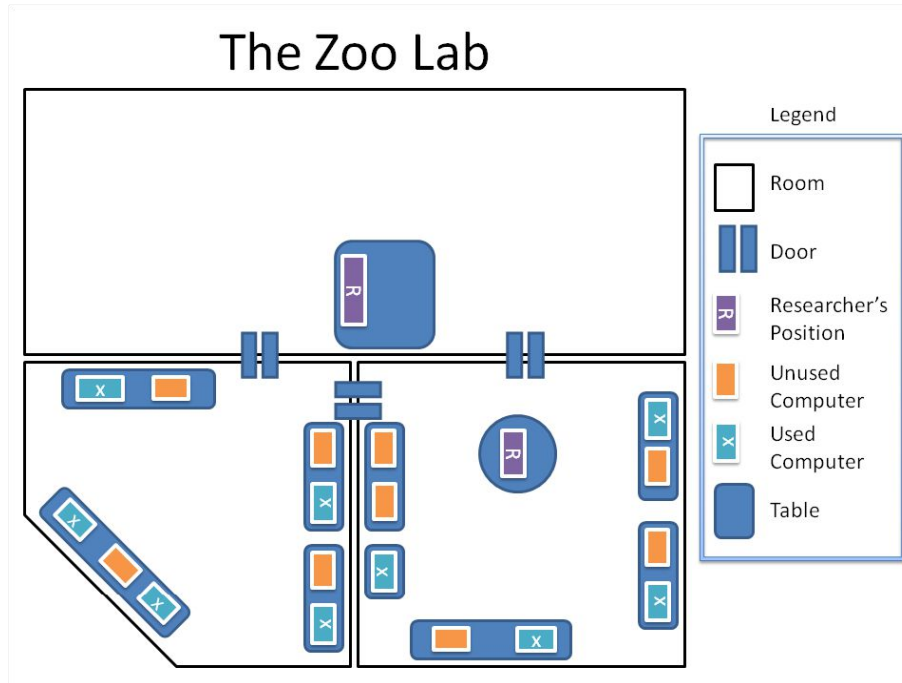


Figure 4: The layout of the Zoo Lab for our study

After choosing which computers to use, we stored our videos and HTML survey page in a local shared folder, allowing our users to log into their account, and have access to the files without downloading or streaming them. We tested each computer to make sure the videos would play in the survey, using Chrome as the HTML page viewer and video player. We also confirmed that each computer had the ability to play sounds through the headphones we borrowed for the study.

Before every study testing period, we also prepared the study room. The first thing we had to do was reserve the lab a few days early, 24 hours at minimum; doing this required that we contacted WPI staff with the reservation request. After the reservation was complete, we guaranteed that we would have enough available computers for our participants. From Figure 4, one can see that we had 9 lab computers configured, but it was only necessary for us to reserve at most five computers at a time,

because we never had more than five participants during one half-hour session. At least twenty minutes before each study period, we left paper reservations at the computers we wanted to keep available, stating the date and time period the computers would be reserved, which we would remove at the end of the day's testing period.

Once the computers were set up for the study, we created a station where we would be located. This involved finding a table in the room, which would allow us to grant Participant Pool credit, display a raffle sign-up sheet, store our study's consent forms and other tools, such as headphones. We positioned ourselves on this table, so that the reserved computer screens were not directly in our line of sight, while maintaining our view of the door to see incoming participants. Once everything was positioned, we would await our participants.

3.6.2 Preliminary Participant Steps

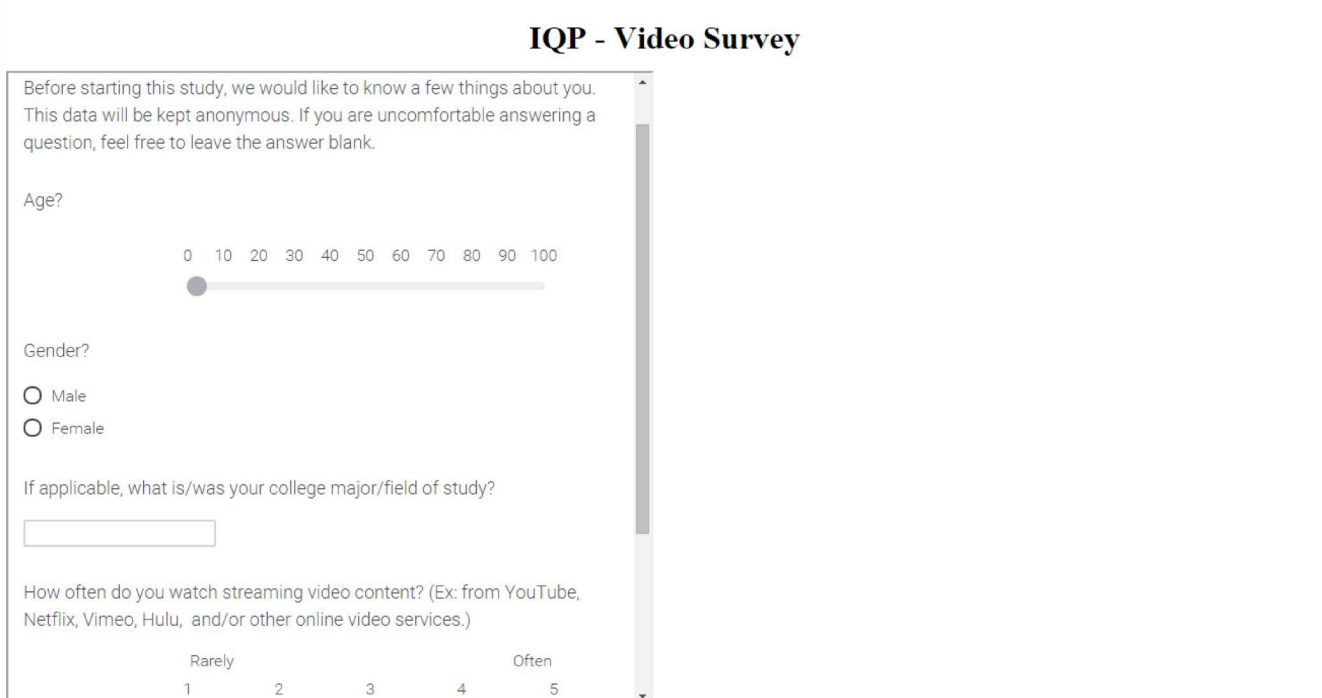
Once a participant arrived, there was a specific process we followed. First, we welcomed the participant and thanked them for participating. After this we followed a script (see Appendix 8.2), which led the participant through filling out our consent form, and detailed what the study would require of them. We also asked them if they would like to enter our raffle, if they were someone who responded to our email advertising, or we would enter their credit into the respective program, SONA, if they were from the Participant Pool. If the user needed to borrow headphones, we supplied them with the ones we borrowed from WPI's Academic Technology Center.

3.6.3 Video Survey

While the participant was completing the consent form, we selected one of the four groups we created for our study; we assigned them to either Buffer 1, Buffer 2, Interrupt 1, or Interrupt 2. The group they were in did not change the procedure, but it did change what they experienced. Everyone would watch the same content videos, but would experience different quality defects; those in the

buffer groups would watch the videos with varying buffer lengths, and those in the interrupt groups would watch the videos with varying interrupt counts. Buffer 1 and Buffer 2 all watched the same videos, but, in order to work against content bias, we had each equivalent video encoded with different buffer times; the differences between Interrupt 1 and Interrupt 2 were the same, but with interrupt counts instead of buffer times.

After we picked the group and all the steps in the previous section were complete, we lead the user to an open reserved computer. We had them log into their WPI Account on the machine and plug in their headphones, because it would occasionally take a little while for them to be configured for the system. In the meantime, we told them how to fill out the demographic information, which was the first page of the survey, (shown in Figure 5). This demographic information asked for age, gender, major, and streaming media experience, but the users could skip any of these questions.



The image shows a web browser window titled "IQP - Video Survey". The page content includes a consent statement: "Before starting this study, we would like to know a few things about you. This data will be kept anonymous. If you are uncomfortable answering a question, feel free to leave the answer blank." Below this are four questions: 1. "Age?" with a horizontal slider ranging from 0 to 100 in increments of 10, with a black dot at approximately 10. 2. "Gender?" with two radio button options: "Male" and "Female". 3. "If applicable, what is/was your college major/field of study?" with an empty text input box. 4. "How often do you watch streaming video content? (Ex: from YouTube, Netflix, Vimeo, Hulu, and/or other online video services.)" with a Likert scale from 1 to 5. The scale is labeled "Rarely" at 1 and "Often" at 5, with intermediate labels 2, 3, and 4.

Figure 5: The Demographic Survey Page

After the demographic information, the user watched a control video page; this video contained

no quality defects, (i.e. no buffers or interrupts). The user pressed a “Play Video” button to start the video, which would display the video in full-screen mode, without any user playback controls, such as pause or seek. Once they were done watching the video, the video exited, and the user could rate their annoyance with the video's quality defects, which we told them before hand only referred to buffers or interrupts, from one to five. They were told that a value of one meant they were minimally annoyed and a value of five meant they were maximally annoyed. After that they were able to rate how much they liked the video's content, in a similar scale, with one being disliked and five being liked. After they chose the values for those two questions, they could move on to the next page. This process repeated itself sixteen more times, for a total of 17 videos, but in these sixteen new videos, the user experienced the factor that applies to their group. As described in Section 3.3.4, the videos would be encoded with one of the four different levels for their factor, (i.e. buffer or interrupts). Figure 6 shows the video survey page for the control video.

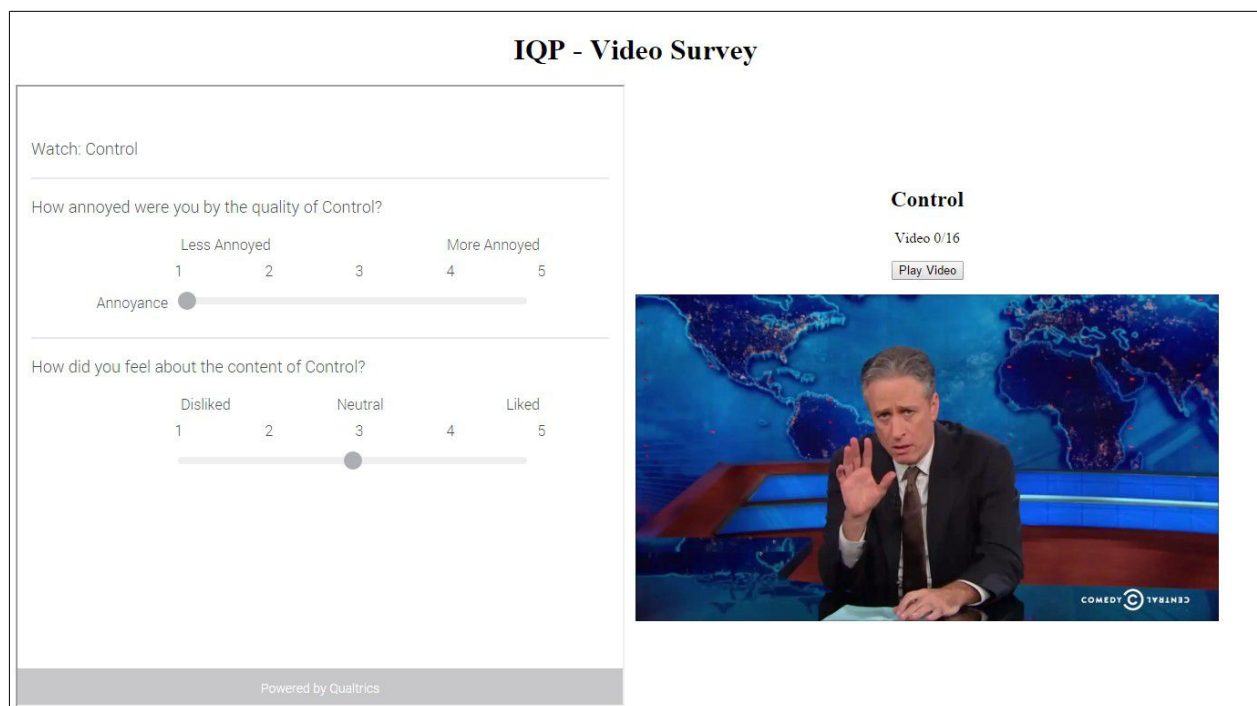


Figure 6: The Video Survey Page

Once all 17 videos were watched, additional questions were presented to the user so we could gain information on their opinion on the survey, as well as miscellaneous data. In order to determine how people learned of our experiment, the survey asked them how they found out about our study. We also asked them what they found the most annoying about the videos as well as any general comments on our study. These additional questions were optional. After answering these questions, we moved on to the debriefing stage.

3.6.4 Debriefing

In order to gather participants from WPI's Participant Pool, we were required to provide a debriefing after the survey was done. Prior to our use of the participant pool, we simply thanked the user and asked if they had any final questions or suggestions for us. The debriefing form which we created for the Participant Pool, (see Appendix 8.3), was more detailed, and covered what our hypotheses were, explained what factors were being tested, and offered the user the ability to see our results via this paper. For the new debriefing section, we read the form, asked them if they had any last comments, and then finally thanked them for their participation.

4 Results & Analysis

In the following sections, the results of our study are displayed and analyzed. We broke up this information into separate sections; the “Study Results” section covers data that was not thoroughly analyzed in respect to our hypotheses, such as demographic information and raw information, while the “Data Analysis” section explains how the data can be applied to our hypotheses and mentions some other relationships we found.

4.1 Study Results

This sections displays the raw data gathered while performing the study. It discusses the demographic background of our study and explores how participants rated our videos with respect to content and annoyance.

4.1.1 Demographics

Before exploring the results of our study, it is important to understand who the participants were. In total we received 37 participants for our study, with varied backgrounds. The following graphs display the makeup of our participants:

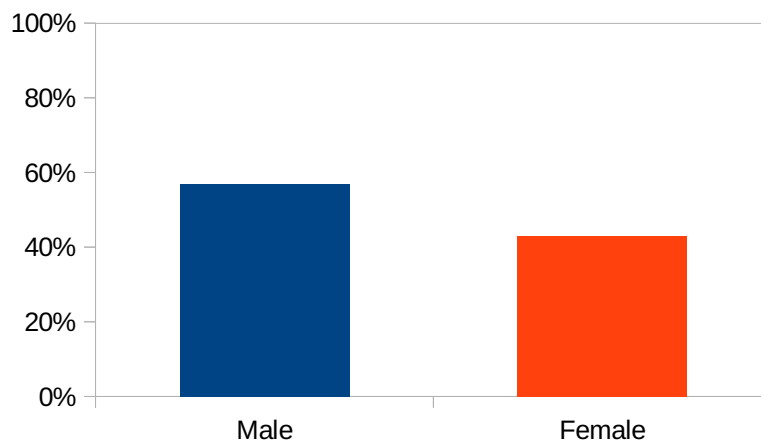


Figure 7: Gender Distribution

Figure 7 shows that the distribution of males and females, while not perfectly equal, is reasonably split. In fact, the ratio is closer to a state of equivalence than the overall gender ratio of WPI undergraduates, which is 68% male according to WPI's 2014 Fact Book.

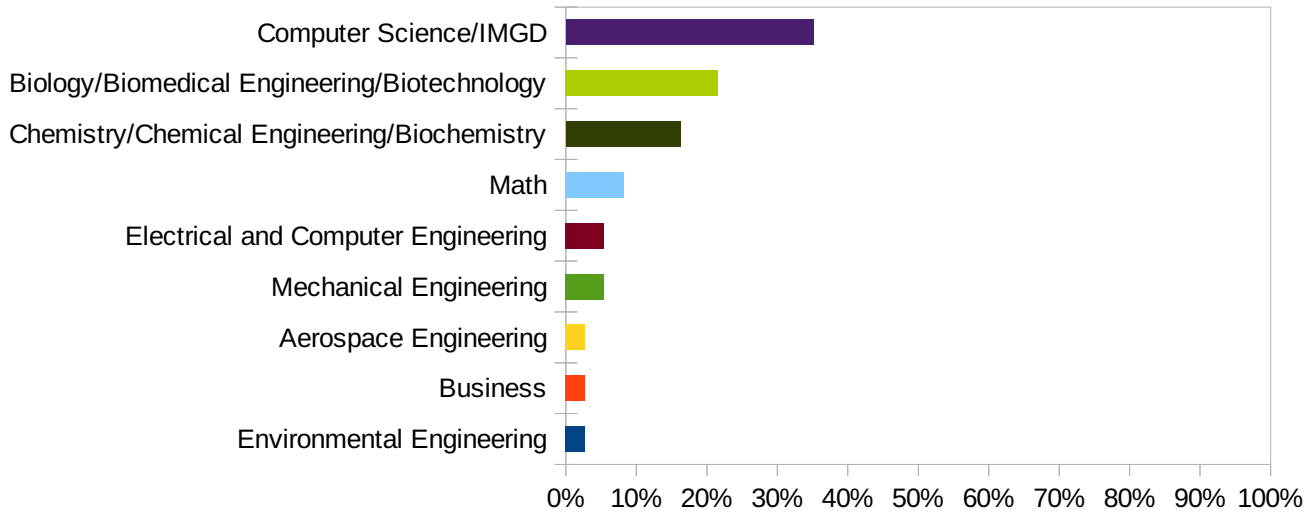


Figure 8: Major Distribution

Figure 8 shows the how the participants are broken up by their major at WPI. This shows that we received a broad distribution of students at WPI, with people representing most of the prominent fields of study at WPI, with WPI's 2014 Fact Book listing the top 3 undergraduate majors as Mechanical Engineering, Computer Science, and Biomedical Engineering. Note that Computer Science/IMGD is the largest group, because the emails that we sent out were only distributed to the CS and IMGD email aliases; the people from the Participant Pool were much more diverse in regard to major.

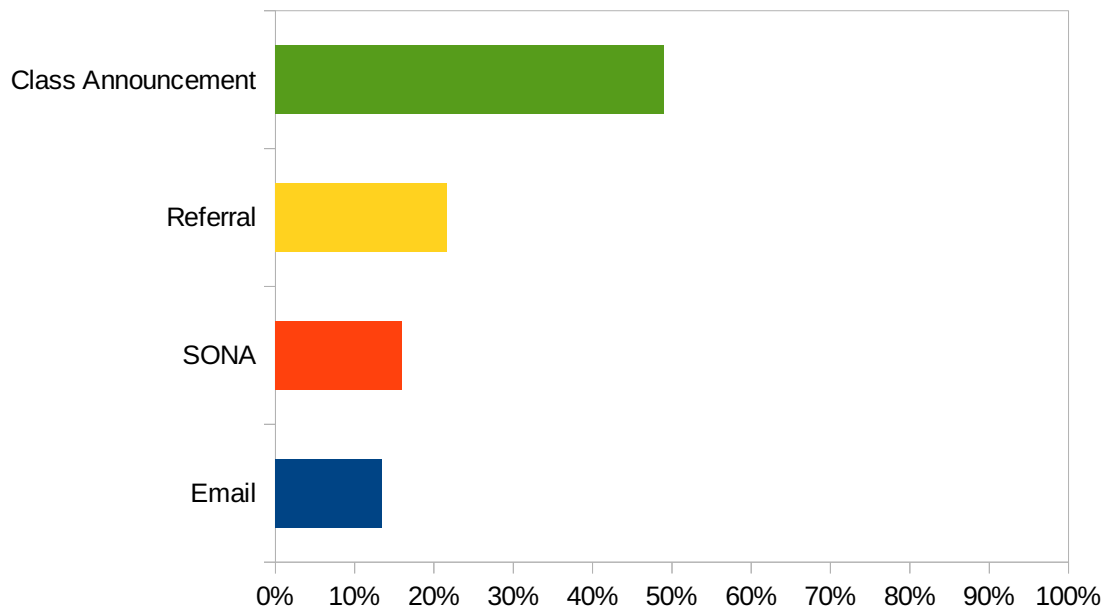


Figure 9: Participant Solicitation Distribution

Figure 9 shows how the participants discovered our study. The Referral and Email categories generally cover the participants from the pre-Participant Pool phase, where users were solicited via email and word-of-mouth. The Class Announcement and SONA, a cloud based participant management tool provided by Sona Systems, categories refer mostly to the users gained through the Participant Pool; for the participant pool, there were one or two teachers who actively promoted our study, so that their students would receive the respective sociology credit, and SONA was the tool where users could sign up for our study.

4.1.2 Video Content

As part of our study, we asked each user to rate every video's content between 1 and 5, with 1 meaning dislike and 5 meaning like. The following graph, Figure 10, shows how the participants rated each video; the graph displays the mean content rating for all 16 testing videos and the control video, with the standard error represented by the black lines extending from the ends of the bars.

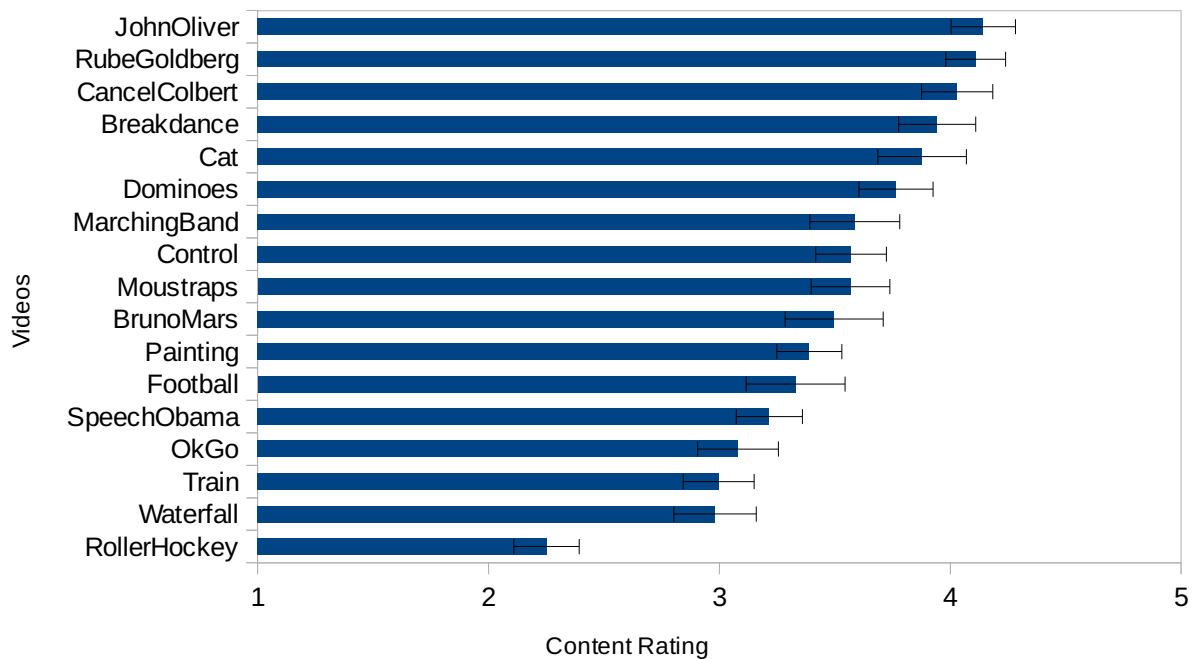


Figure 10: Video Content Rating

As the graph shows, most of the video were above the value of 3, with 3 being the neutral response between disliked and liked. This means that for most videos, people generally liked the video more than disliked. The video with the highest mean score was “JohnOliver.” However, there are a few videos that received an average score of about 3, which indicates that it was liked as much as disliked. There was only one video with a rating that was less than 3; this video, which we called “RollerHockey”, was disliked a lot more than liked.

Overall, the videos we chose for study were generally liked. Only a few videos were not liked nor disliked on average and only one was generally disliked. In Section 4.2.6, we discuss the relationship between a user's video content rating and their annoyance due to buffering.

4.1.3 Annoyance Levels

This section is meant as a broad overview of what the annoyance scores were for our study.

However, in the subsections of 4.2, we discuss the annoyance scores received in respect to buffer times, interrupt counts, motion, and content.

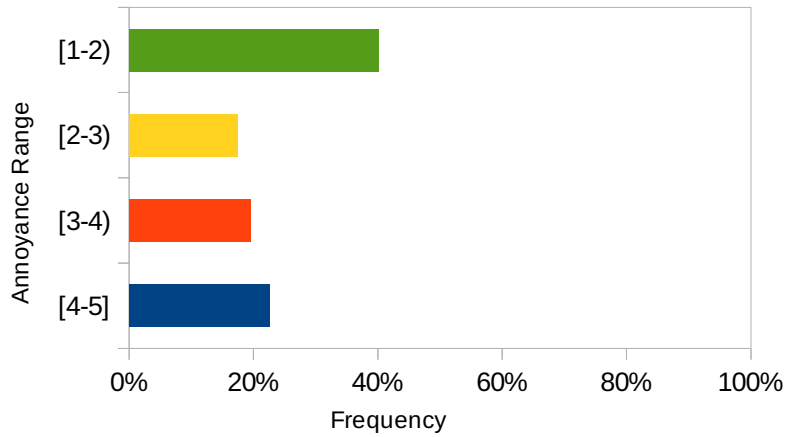


Figure 11: Annoyance Range Frequency

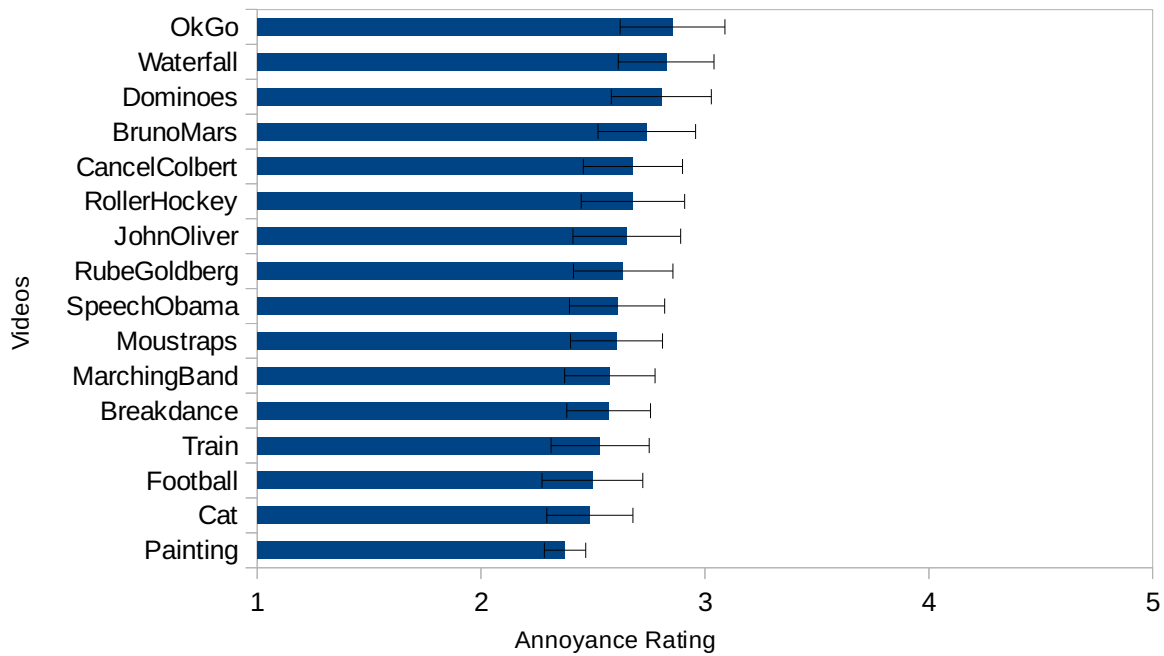


Figure 12: Average Annoyance Ratings Per Video

Figure 11 shows the percentage of time a certain range of annoyance values was selected; the

main thing this graph shows is that the likelihood of an annoyance value being between 1 and 2 is almost twice as likely as any other category. Figure 12 shows the average annoyance rating and standard error bars for each video, without any other factors considered. The graphs shows that on average the participants were not generally annoyed by any video in particular.

4.2 Data Analysis

This section shows how buffers and interrupts are related to annoyance, as well as how these findings compare to our hypotheses. It also describes other relationships that were discovered during data analysis.

4.2.1 Buffers

For buffers, we found that annoyance values were low overall. Figure 13 shows the cumulative distribution function, (CDF), of annoyance with respect to buffer time, with the x-axis representing the annoyance level and the y-axis representing the percent of user annoyance scores with a rating of x or less. The different trend lines represent the CDF for each buffer time. From the graphs discovered that the distribution is clearly skewed toward the lower end for each buffer time. Even with the longest buffer time that we tested, nearly three-quarters of the results fell on the lower half of the range for annoyance values.

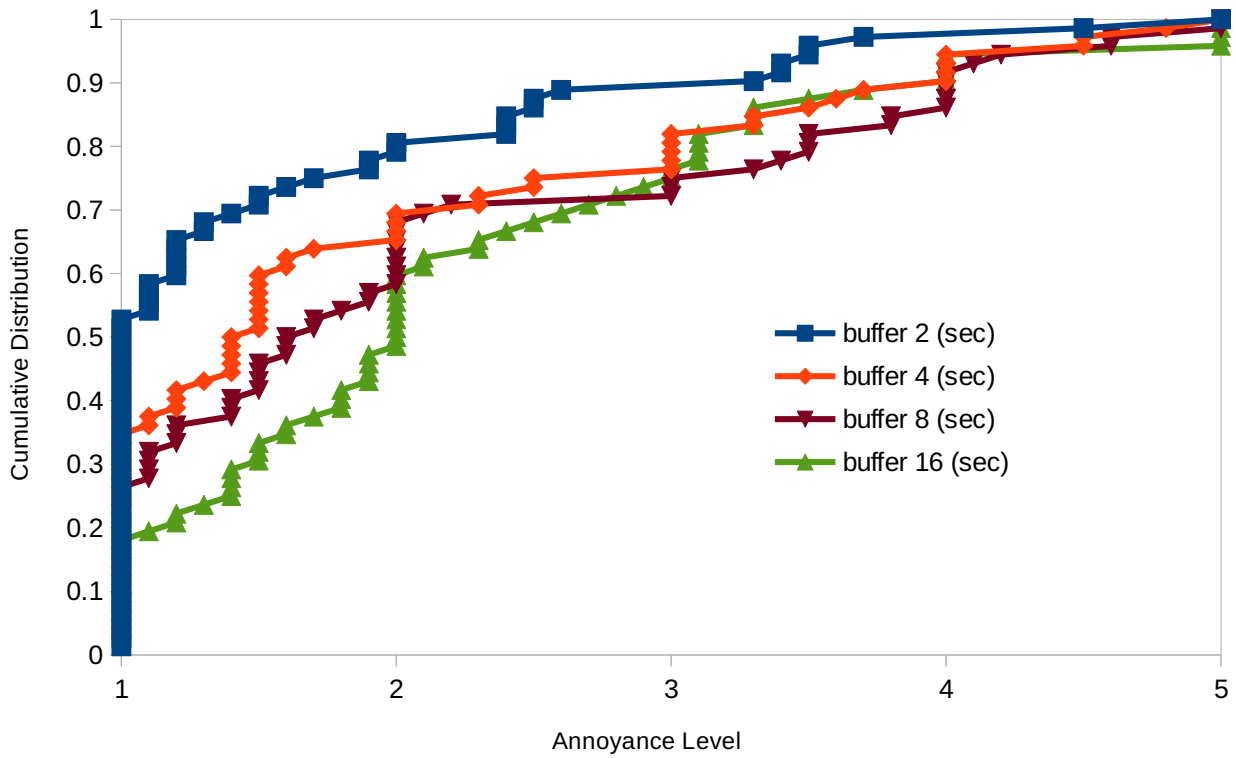


Figure 13: Cumulative Distribution Function for Annoyance Level,
With Respect to Buffer Time for Video

When we graphed the averages of the annoyance values for each buffer length, we got the graph in Figure 14, with the buffer time as the x-axis and the mean annoyance rating as the y-axis. The black lines on the graph represent the standard error for a given mean value. We fit the graph to a degree 2 polynomial, since we were expecting a degree 3 polynomial from Hypothesis 1, but the data that we got did not show the second turning point (i.e. the point where the annoyance level increases rapidly on the right side of the graph for Hypothesis 1) that we expected. The R^2 value for this trend line, 0.95, matches our data well. The trend line can be specified with the following formula:

$$f_b(x) = -0.01x^2 + 0.16x + 1.32$$

Equation 1: Annoyance Function with Respect to Buffer x in Seconds

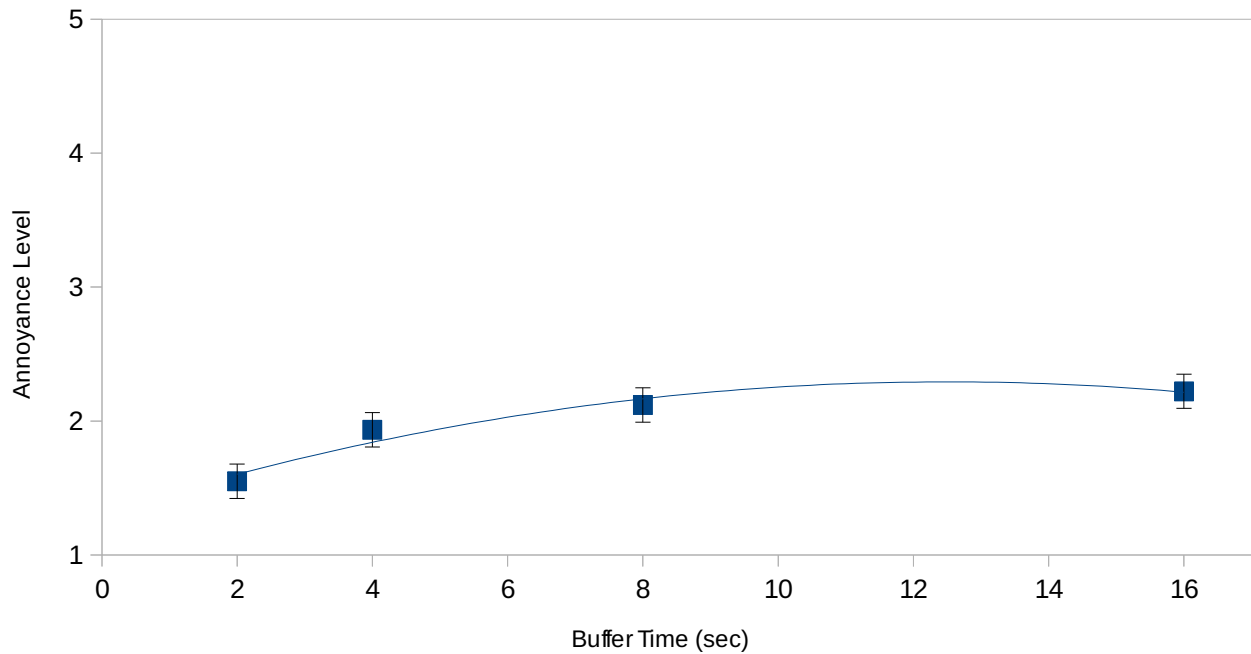


Figure 14: Mean Values of Annoyance Levels for Each Buffer Time, with 2nd Degree Polynomial Best Fit Line

Since the distribution of annoyance is skewed and our data does not contain the second turning point that was expected, our maximum buffer length was too small to show more annoyance. Seeing as the data that we did collect forms the first half of the curve that we expected in Hypothesis 2, using buffer times greater than 16 seconds, the data may start looking like the second half; it is important to note that, if the second turning point was present, then the downward trend on the right side of the graph would disappear. The low annoyance ratings that we collected also support this, since they indicated that there is more potential for the user to get annoyed. It is also possible that there may be other confounding variables that may be affecting our results, such as the amount of time between experiencing the buffers and filling out the annoyance ratings, or maybe even the content of the videos themselves.

4.2.2 Buffers and Motion

Figure 15 shows the relationship between buffer time and annoyance for each motion level. The x-axis represents the buffer time and the y-axis represents the mean annoyance level. When we separated the annoyance ratings from the four motion categories, there was no clear separation in the trend lines. All of them overlapped and converged, as shown by Figure 15. This suggests that there is no correlation between the level of motion and the annoyance caused by buffers; this fully supports Hypothesis 2.

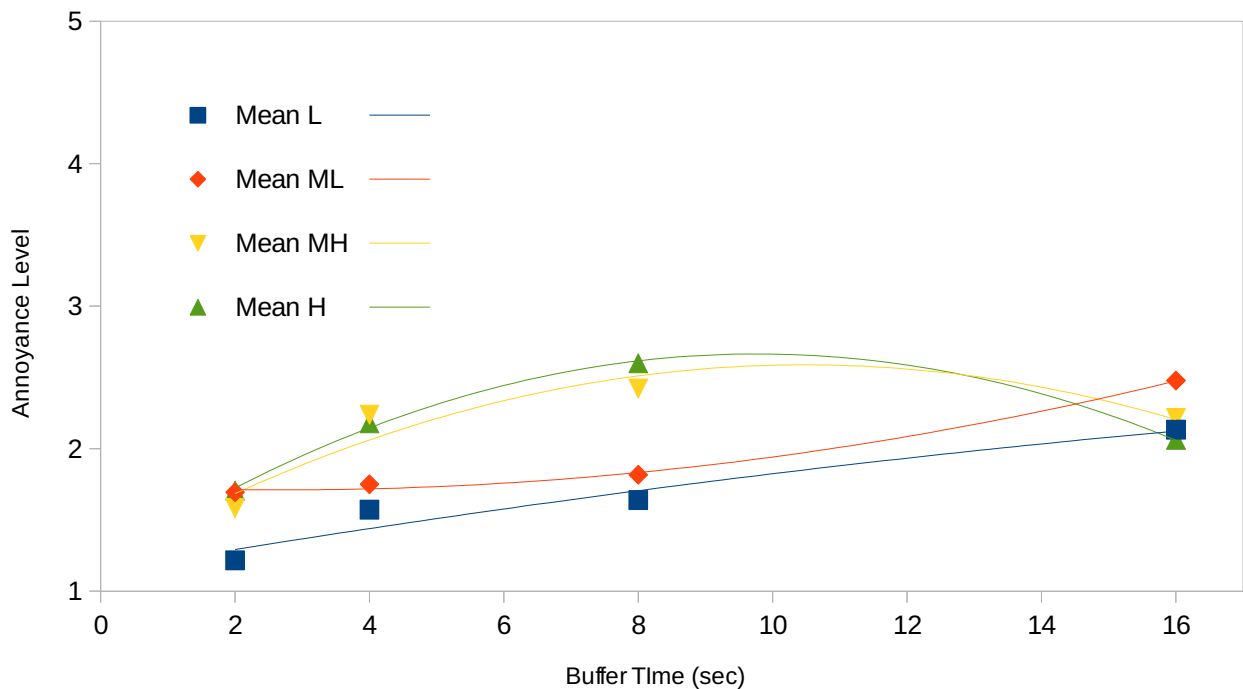


Figure 15: Mean Values for Buffers vs Annoyance, Separated by Motion Level

4.2.3 Interrupts

Figure 16 shows the cumulative distribution function, (CDF), of annoyance with respect to interrupt count, with the x-axis representing the annoyance level and the y-axis representing the percent of user annoyance scores with a rating of x or less. While the buffer length factor tended to produce

lower annoyance values, the interrupt count factor varied more greatly. As Figure 16 shows, the distribution of values varied between the different interrupt count categories, but the values were also much greater than with buffers; with all the categories, more than half of the ratings were above an annoyance level of 2.5. From the distribution in Figure 16, it is clear that people are generally more annoyed as the interrupt count increases.

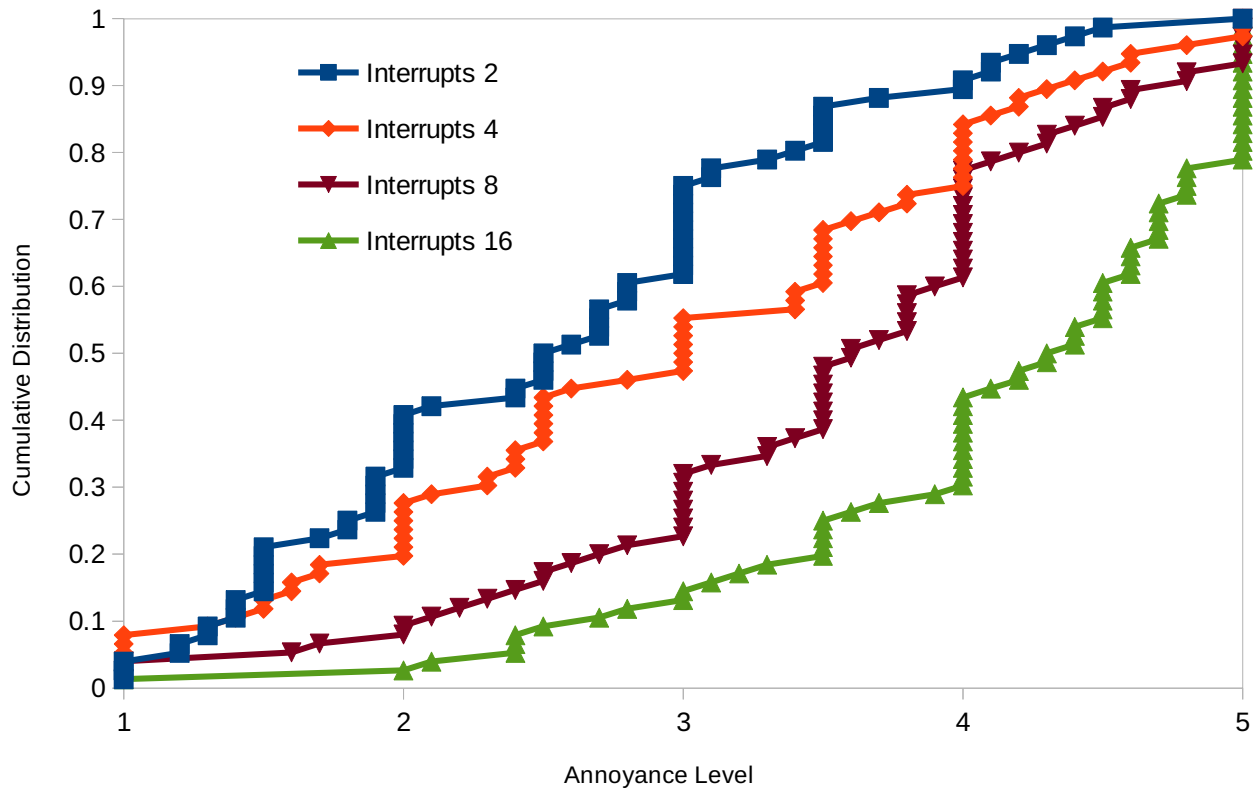


Figure 16: Cumulative Distribution Function for Annoyance Level, with Respect to Interrupt Count for Video

We graphed the mean values for the varying interrupt counts, as shown in Figure 17, with the interrupt count as the x-axis and the mean annoyance rating as the y-axis. The black lines on the graph represent the standard error for a given mean value. We used a logarithmic trend line to represent the relationship amongst the varying interrupt counts, because we hypothesized that this was the proper

relationship before gathering the data. The logarithmic relationship can be specified using the following function, with a 0.99 R² value:

$$f_i(y) = 0.75 \ln(y) + 1.98$$

Equation 2: Annoyance Function with Respect to Interrupt Count y

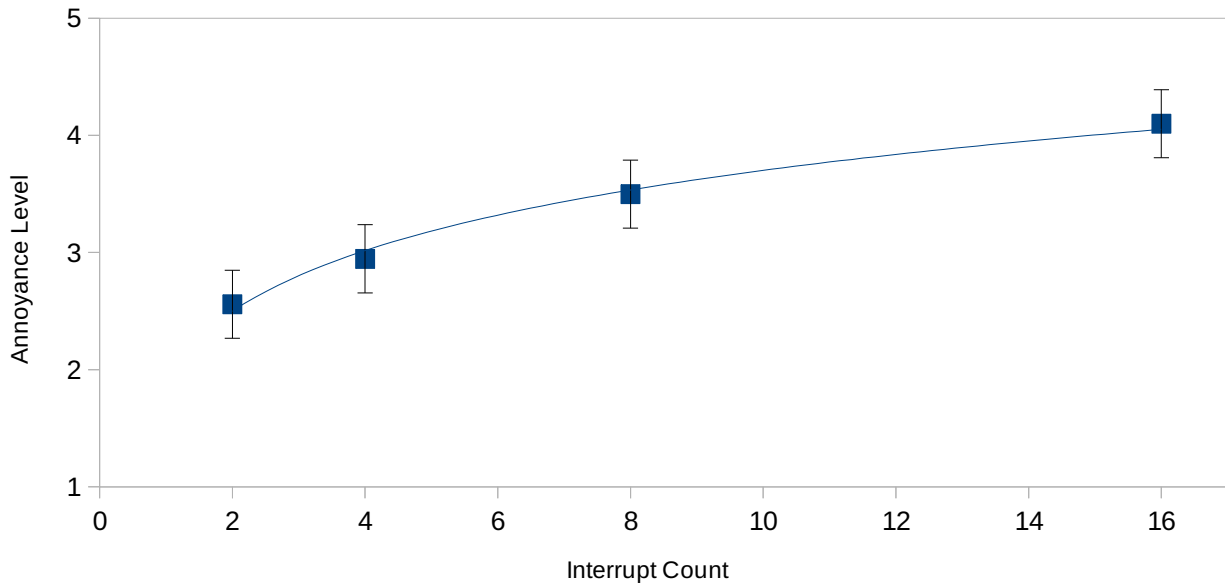


Figure 17: Mean Values of Annoyance Levels for Each Interrupt with Logarithmic Best Fit Line

As the distribution in Figure 16 and mean values in Figure 17 show, the interrupt counts we chose cover a larger range of annoyance values than buffers did. Interrupts have a much larger effect on user annoyance than buffer times do, at least for the buffer times and interrupt counts we observed; the lowest interrupt mean value, which occurs at the interrupt count of 2, is higher than any of the other mean values for the buffer times.

4.2.4 Interrupts and Motion

Figure 18 shows the relationship between interrupts and annoyance for each motion level. The

x-axis represents the interrupt count and the y-axis represents the mean annoyance level. While we were unable to find a relationship with annoyance due to buffers and motion, we found an unexplained anomaly when examining how the annoyance level varied with interrupts and motion. While most of the motion levels seemed to be following the same pattern with overlap, as seen in Figure 18, the low motion level is consistently higher than the others. It is difficult to understand why this might be; there could be some other factor that caused the low motion videos to be more annoying or this relationship might actually be true. We would need to have more users and more videos in each level to test this relationship more thoroughly. Since we were unable to find the relationship predicted, Hypothesis 4, regarding interrupts and motion levels, is not supported by the data.

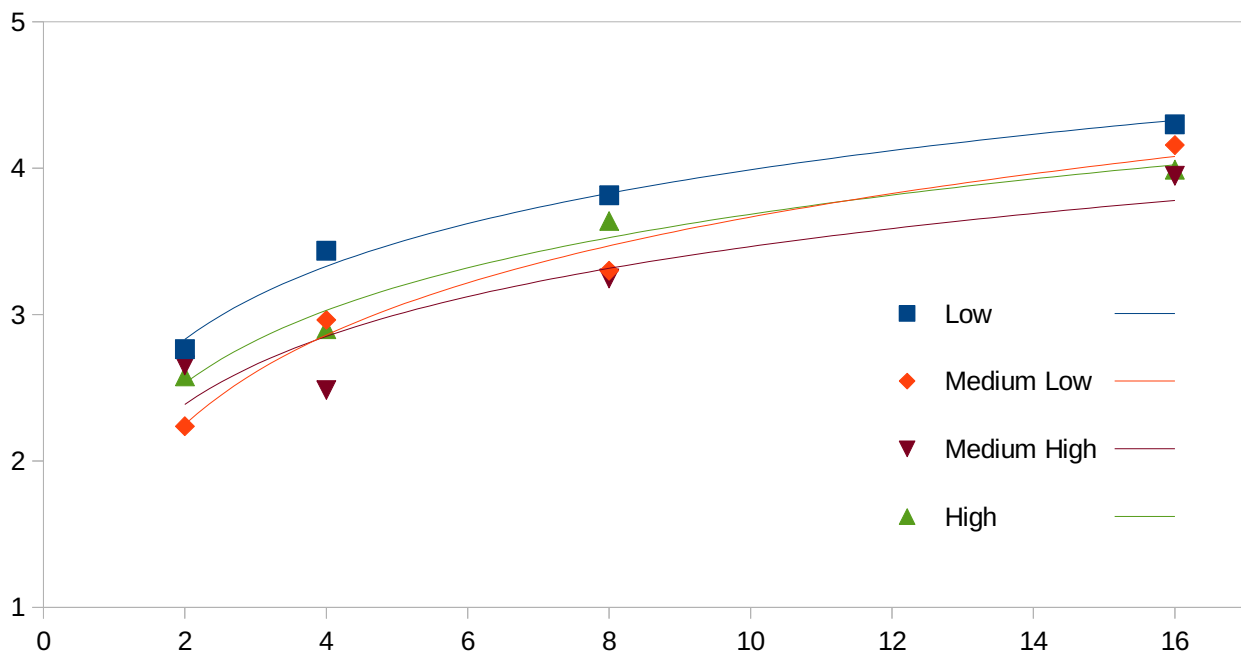


Figure 18: Mean Values for Interrupts vs Annoyance, Separated by Motion Level

4.2.5 Buffers and Interrupts Combined

In the previous sections, we were able to develop functions that dictate the annoyance level for

a given buffer time or number of interrupts, with the variables falling in the range of 2 and 16:

$$\text{Equation 1: } f_b(x) = -0.01x^2 + 0.16x + 1.32$$

$$\text{Equation 2: } f_i(y) = 0.75 \ln(y) + 1.98$$

It would be useful to combine the individual annoyance data in terms of buffer times and interrupt counts. The simplest solution would be to add the annoyance levels together from the buffer-annoyance function and the interrupt-annoyance function, as demonstrated by the following function:

$$f_t(x, y) = \text{Min}(f_b(x) + f_i(y), 5)$$

Equation 3: Estimated Total Annoyance with Respect to
Buffer x in Seconds and Interrupt Count y

The function takes the minimum of the total and 5, because 5 was the maximum annoyance level, and the sum of the two functions has the ability to be a number more than this max annoyance. This relationship could be an accurate description of what would happen in the presence of buffers and interrupts. However, the accuracy depends upon the two functions being completely independent of each other in terms of user annoyance; it may be possible that when presented with a buffer, the user might become even more annoyed when interrupts occur, which our functions would not predict. More research would have to be done to understand the specific relationship between the user annoyance caused in the presence of buffers and interrupts, but we propose the following function to estimate the minimum annoyance:

$$f_m(x, y) = \text{Max}(f_b(x), f_i(y))$$

Equation 4: Minimum Annoyance with Respect to
Buffer x in Seconds and Interrupt Count y

This function takes the maximum of the buffer-annoyance and interrupt-annoyance values for a given buffer time and interrupt count; this means that the annoyance for the two variables will be the

larger of the two respective function calls.

4.2.6 Effects of Video Content

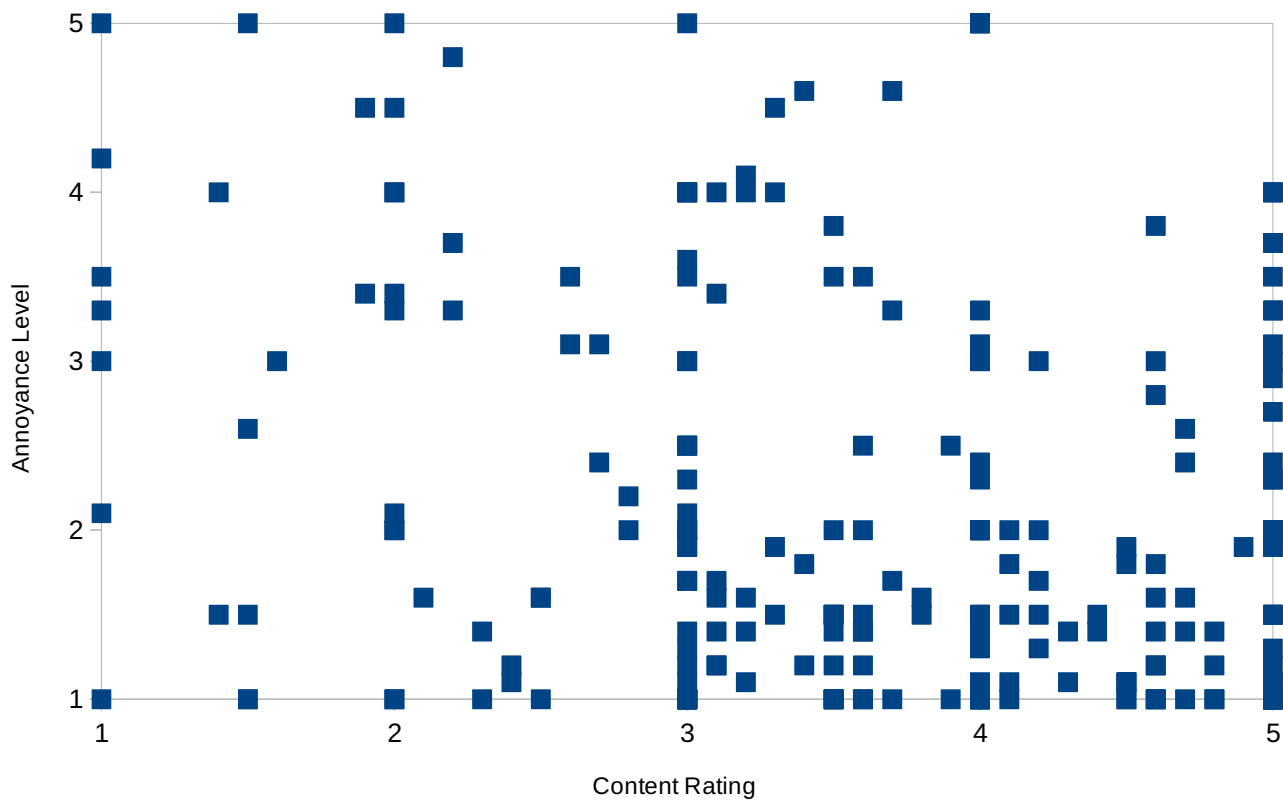


Figure 19: Scatter plot of Annoyance Level vs Content Rating for Tests Involving Buffers

Since we collected information content ratings from users as well as the annoyance ratings, we analyzed the correlation between those values. Figure 19 depicts the scatter plot between content rating and annoyance involving buffers; the x-axis represents the content rating value, with the y-axis depicting the annoyance level value. Each dot represents one data point, with the matching content rating and annoyance level. We found a weak correlation between content rating and annoyance level for our tests involving buffers. Looking at the cluster in the graph, the videos that were better liked tended to get lesser annoyance ratings. The correlation is -0.28 . This suggests that there is likely a relation between video content and annoyance.

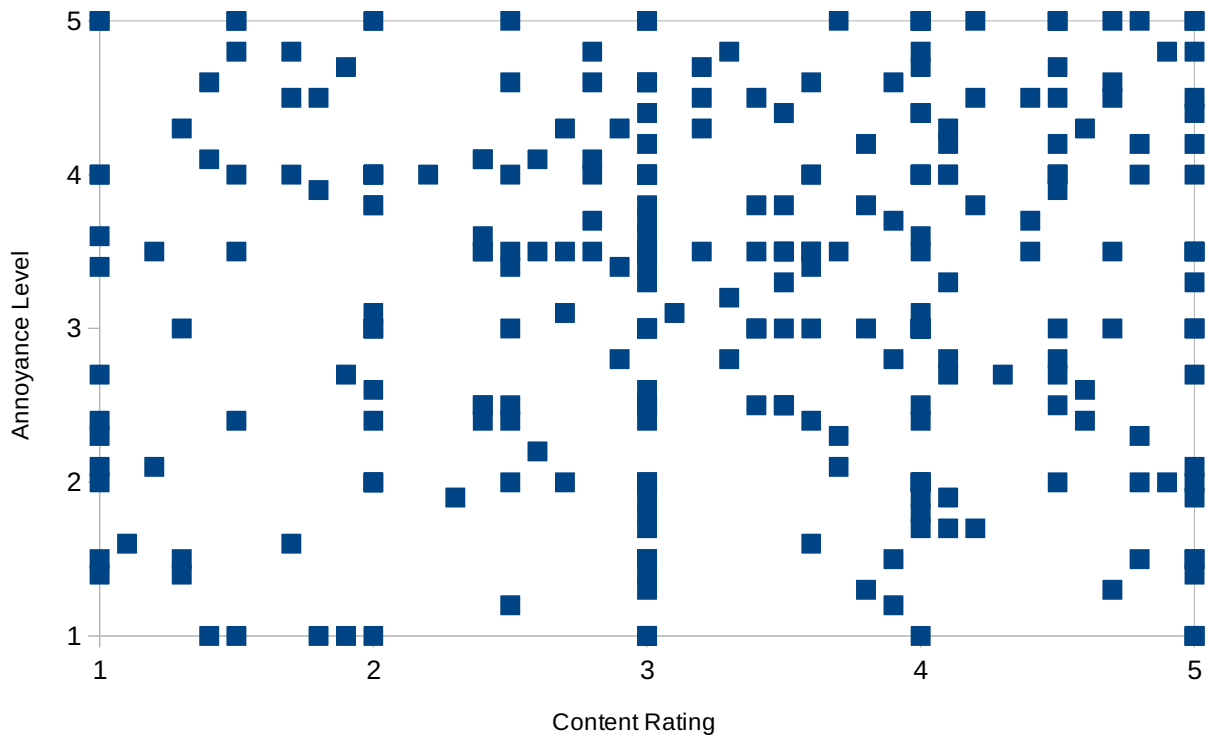


Figure 20: Scatter plot of Annoyance Level vs Content Rating for Tests Involving Interrupts

Figure 20 depicts the scatter plot between content rating and annoyance involving interrupts; the x-axis represents the content rating value, with the y-axis depicting the annoyance level value. Each dot represents one data point, with the matching content rating and annoyance level. When we did the same analysis for videos with interrupts, as shown in Figure 20, the correlation was only 0.05, and the only noticeable pattern on the graph is that people tended to leave content ratings of 3, which was the default value. There seems to be no apparent relations between content rating and annoyance caused by interrupts.

4.3 Postmortem

The Postmortem reflects upon what we could have done better with our study. It covers the user feedback we received as well as some of the problems we encountered during our study.

4.3.1 User Feedback

As part of our survey, we asked users how they felt about our study and if there was anything we should have changed. In general, the responses to this were mostly positive, with only a few recommendations. While we appreciated the comments, it was too late to apply them after the study, but they are something to take note of for future studies. The following is a summarized list of significant comments:

- The videos were good, but more variety in the content would have been better.
- A more concrete scale could have been used; instead of labeling the values for 1 and 5, all of the integers could have been labeled.
- The videos varied in volume, causing an extra annoyance; (this concern is addressed in Section 4.3.2).
- We could have made the videos play automatically when the users gets to the next page.
- We could have allowed a user to move at their own pace, by allowing the user to control the video.
- We could have clarified the directions a bit more, such as what we meant by quality; (we had a statement we thought was clear in our script, but it is possible that the user did not understand us or that it was not as clear as we imagined).

4.3.2 Problems Encountered

We encountered a few problems that affected the participants or end results. In these situations, we tried to rectify the situation, if it could be done without invalidating the study.

The first major problem we observed was with the variation of the default volume in the video clips. Some clips had a low volume, but the next might be much louder, causing an extra annoyance for

the participants. We became aware of this after a participant commented on it during the first day. The best solution would have been to modify the videos to all have a similar volume level, but since the study had already begun, we found another solution. In order to account for the change in volume, we added to our script, as a verbal addition, a warning about the changes and told the participants how they could modify the volume of the sounds outputted to the headphones via the keyboard, allowing them to find a suitable volume level.

Another major problem that we observed was a rounding issue on one of the videos. While all of the other videos rounded to the nearest tenth for both sliders, (annoyance and content rating), one video rounded to the nearest integer on both sliders. Each slider had a displayed value to the side of it, but it is not known whether all of the participants realized that their range of values was extremely limited. We did not know of this problem until the last session of our study during B-Term; a participant brought it to our attention after noticing the change in the displayed value. We fixed this rounding issue before the sessions we hosted at the start of C-Term. In respect to the data we had already gathered for this video, we decided to allow the data to be part of our final results. Since the slider rounded to the nearest integer, (i.e. 1.4 to 1, and 1.5 to 2), we assume that the data for the video was acceptable because the loss of precision was minimal; the numbers may not have been exactly what the users wanted, but it should be a good estimation of how they felt.

The last major problem we encountered was the lack of data validation on one of the videos. On the video “OkGO”, a user could move on to the next video without entering a value for annoyance, which would leave a blank value in the corresponding .csv file we would download. However, this problem was minimal, because there was only one user who did not enter a value for the video. To resolve this, we simply ignored that user's entry for that video.

5 Future Work

The first major study we suggest is a look into the combined annoyance relationship between buffers and interrupts. As we mentioned in Section 4.2.5, we do not know how to effectively combine the relationships we gathered in order to estimate the total annoyance for videos with buffers and interrupts. This might be done by completing a study that incorporates buffers and interrupts into videos together and determining how the annoyance ratings differ from our proposed annoyance equations. A study that was able to understand how these two factors work together, would be helpful for being able to apply our study to real world situations, allowing it to be used by streaming media providers.

Since the annoyance values that we collected for buffers were low overall, it would be helpful to repeat our study with higher values for buffer time. A study doing this will show if Hypothesis 1 was accurate, which would modify Equation 1, the buffer-annoyance function, to be a degree 3 polynomial.

During the course of study, we observed other factors that need to be studied in depth. As seen in Section 4.2.6, the content of a video might have a correlation with the level of annoyance a user perceives when dealing with buffers alone. this could be measured by doing a similar study to the one we completed, but with more videos and more users; since this relationship was with buffers, interrupts could be ignored in this new study. A study that thoroughly examines this correlation might reveal some important relationships between content rating, buffers, and annoyance.

Another factor worth studying is the presence of sound; during our pilot studies we gained some feedback that people were more annoyed when interrupts occurred during talking in the video than during quieter moments. This might warrant a future study to see if it is possible to place interrupts

during periods lacking talking or other sounds and if that will lower user annoyance with interrupts.

Another factor to question is regarding why the low motion category for interrupts was consistently more annoying to users. We were not able to state why this occurred; it might have been a flaw with our study, some hidden factors, or an actual relationship. This factor could be studied by repeating our study for videos with only interrupts, while also introducing more videos per motion category. This would reveal whether the data we found was misleading or whether there is a relation between low motion, interrupts, and annoyance. A study into this matter might be helpful, but it seems less fruitful than the other ideas suggested in this section, because we believe that most likely there was some other factor affecting the annoyance.

6 Conclusion

Streaming services and video players benefit from increased QoE for users. There are several aspects that degrade the QoE, such as the annoyance caused by buffers and interrupts in the video. With knowledge of how buffers and interrupts affect the annoyance of users, it is possible to tune the experience to minimize annoyance. We aimed to discover the relationships that would enable this tuning to be done. We also searched for any effect motion might have on the relationships.

In order to explore the effect of interrupts and buffers on annoyance, we developed a user study, in the form of a survey, that compared user annoyance with each factor independently. We had 37 diverse participants watch 16 modified videos and a control video each and rate each video on the annoyance caused by buffers or interrupts, depending upon which group they were placed in. We also measured the motion level of the videos, as well as the participants' opinions on the content of the videos.

Using the data we collected, we were able to come to some conclusions. For buffers, we hypothesized that as buffer times increases, the users' annoyance increases as a degree 3 polynomial (Hypothesis 1). Additionally we predicted in Hypothesis 2 that motion had no affect on these annoyance ratings, as buffers occur before the motion of the video begins. While the data supported our hypothesis about the effect of motion, the annoyance curve, independent of motion, we found followed a degree 2 polynomial, instead of a degree 3 one. Due to how low the annoyance values that we collected for buffers, we suspect that if longer buffer lengths were tested, the annoyance curve for buffers might look more like the degree 3 polynomial that we expected.

For interrupts, Hypothesis 3 stated that as the number of interrupts increases, the annoyance increases in a logarithmic manner. Additionally, Hypothesis 4 stated that higher levels of motion

increases the level of annoyance, and that lower levels of motion similarly decreases it. Our data is consistent with Hypothesis 3, with the annoyance rising as we suspected, however, motion had a completely unexpected effect, failing to prove Hypothesis 4; instead of the relationship we expected, the low motion category consistently scored highest for the annoyance ratings, with the other 3 categories having roughly equivalent annoyance ratings. This suggests that there is some other factor at play that is causing the results that we observed.

The results that we found are important toward understanding how users are annoyed by defects in streaming video. Furthermore, we proposed two equations that can estimate the annoyance caused by buffers and interrupts together. With these equations, streaming services and video players can pick the buffer size, which indirectly picks the interrupt count with a given network condition, that will cause the least annoyance for each user. Additionally, our findings can be used as a basis for future studies concerning QoE in streaming media.

7 References

- [1] F. F. Ebner, Y. Hoppenot, and J. A. Wright, “Resource saving while avoiding customer wait annoyance,” US20130083339 A1 (US Patent), 04-Apr-2013.
- [2] U. Kumar, O. Oyman, and A. Papathanassiou, “QoE Evaluation for Video Streaming over eMBMS,” *Journal of Communications*, vol. 8, no. 6, pp. 352–358, 2013.
- [3] “Streaming Video Still Full of Low Bitrates and Buffering,” *The Online Reporter*, p. 9+, 10-May-2013.
- [4] Editing Staff, “WAG Acquisition, LLC; Patent Issued for Streaming Media Buffering System,” *Entertainment Newsweekly*, Dec. 2013.
- [5] D. Eufemia, X. Li, W. Parker, and H. P. Phyu, “Finding Optimal YouTube Buffer Size For Mobile Devices On a Campus Network,” WPI IQP, 2014.
- [6] G. Liang and B. Liang, “Balancing Interruption Frequency and Buffering Penalties in VBR Video Streaming,” in *IEEE INFOCOM. 26th IEEE International Conference on Computer Communications*, Anchorage, AK, pp. 1406–1414. 2007.
- [7] F. Qiu and Y. Cui, “A Quantitative Study of User Satisfaction in Online Video Streaming,” in *IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, NV, pp. 410–414. 2011.
- [8] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, “A Quest for an Internet Video Quality-of-experience Metric,” in *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, New York, NY, USA, pp. 97–102. 2012.
- [9] H. Wu, “ARMOR-Adjusting Repair and Media Scaling with Operations Research for Streaming Video,” PhD Thesis, Worcester Polytechnic Institute, 2006.
- [10] S. Tavakoli, J. Gutierrez, and N. Garcia, “Subjective Quality Study of Adaptive Streaming of Monoscopic and Stereoscopic Video,” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 4, pp. 684–692, Apr. 2014.
- [11] M. Claypool and J. Tanner, “The Effects of Jitter on the Peceptual Quality of Video,” in *Proceedings of the Seventh ACM International Conference on Multimedia (Part 2)*, New York, NY, USA, pp. 115–118. 1999.
- [12] L. Chen, Y. Zhou, and D. M. Chiu, “Video Browsing - A Study of User Behavior in Online VoD Services,” in *22nd International Conference on Computer Communications and Networks (ICCCN)*, Nassau, Bahamas, pp. 1–7. 2013.
- [13] M. Mu, R. Gostner, A. Mauthe, G. Tyson, and F. Garcia, “Visibility of Individual Packet Loss on H.264 Encoded Video Stream: A User Study on the Impact of Packet Loss on Perceived Video Quality,” *The Multimedia Computing and Networking*, San Jose, CA, vol. 7253, pp. 725302–725302–12. 2009.
- [14] A. Finamore, M. Mellia, M. M. Munafò, R. Torres, and S. G. Rao, “YouTube Everywhere: Impact of Device and Infrastructure Synergies on User Experience,” in *Proceedings of the ACM SIGCOMM Conference on Internet Measurement Conference*, New York, NY, USA, pp. 345–360. 2011.
- [15] A. Tripathi and M. Claypool, “Improving Multimedia Streaming with Content-Aware Video Scaling,” In *Proceedings of the Second International Workshop on Intelligent Multimedia Computing and Networking (IMMCN)*, Durham, North Carolina, USA, March 8-12, 2002.
- [16] S. Polacheck, “Online Classification of VoD and Live Video Streaming Applications.” M.Sc. Dissertation, Efi Arazi School of Computer Science, Jun-2003.
- [17] M. Claypool, “Motion and scene complexity for streaming video games,” in *Proceedings of the*

4th International Conference on Foundations of Digital Games, Port Canaveral, FL, USA, pp. 34–41. 2009.

- [18] Y. Song, M. Zhao, J. Yagnik, and X. Wu, “Taxonomic Classification for Web-based Videos,” in *IEEE Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA, USA, pp. 871–878. 2010.

8 Appendix

8.1 Motion Levels

Motion Levels	Lower Bound	Upper Bound
L	0.00%	48.58%
ML	48.59%	62.23%
MH	62.24%	76.01%
H	76.02%	100.00%

Table 1: Motion Levels

Interrupts:	Count	(1 second each)
i1	2	
i2	4	
i3	8	
i4	16	

Buffers:	Seconds
b1	2
b2	4
b3	8
b4	16

Table 2 and 3: Study Factor Levels

Title	Skipped Macroblocks (%)	Motion Level	Buffering 1	Buffering 2	Interrupts 1	Interrupts 2
Mousetraps	87.78	L	b4	b1	i4	i1
Waterfall	64.42	ML	b4	b1	i4	i1
FootBall	57.37	MH	b4	b1	i4	i1
Rube Goldberg	38.44	H	b4	b1	i4	i1
Obama	83.93	L	b3	b2	i3	i2
Marching Band	72.38	ML	b3	b2	i3	i2
RollerHockey	59.7	MH	b3	b2	i3	i2
OkGo	46.5	H	b3	b2	i3	i2
Colbert	83.77	L	b2	b3	i2	i3
Painting	75.81	ML	b2	b3	i2	i3
Train	58.62	MH	b2	b3	i2	i3
Dominoes	36.44	H	b2	b3	i2	i3
JohnOliver	79.59	L	b1	b4	i1	i4
Cat	72.71	ML	b1	b4	i1	i4
BrunoMars	55.93	MH	b1	b4	i1	i4
BreakDance	36.73	H	b1	b4	i1	i4

Table 4: Video Categorization by Motion and Factor

8.2 Script

Introduction:

Hello! Thank you for participating in our study. This study is designed for measuring how much quality defects in videos annoy you. The process will take about 15-20 minutes. First, please read and sign this consent form. If you have any questions feel free to ask.

After signing the form:

Thank you. Are you here for the psychology pool? If so, please write down your information here so we can give you credit. (If yes: skip the rest of this section) Would you like to participate in our optional raffle to win a \$25 Visa gift card. All we need is your name and email. (If yes: Please put your name and email on this form).

For seating:

Please follow me, I will take you to a lab computer to start the survey. <load up the survey for their group> Do you have any headphones on you? <if not hand them a pair and say “Please return these as you leave the study”>. Plug in your headphones: change the volume with your mouse to see if they are functional. Next, this page is for collecting demographic information. I will not watch you fill out these questions, and they are completely optional. Once you are done filling out as much information as you are willing to give, you will click the next button.

After that there will be a control video which has no defects in its quality. When we talk about quality in this study, we are only referring to wait-times or interruptions. You will start the video using the “Play Video” button. After watching the video, you will use the respective slider to indicate how annoyed you were with the video's wait-times or interruptions. 1 means you were minimally annoyed, 5 means you couldn't be more annoyed. You will also rate the video's content using the second slider. Just tell us how you felt about the video disregarding its quality. A thing to note is that you need to move both sliders for the survey to recognize that you entered a response, so if you want to use the default value, just slide it off and then back. Do you have any questions? <answer questions> There are 16 more videos besides the control that you will be rating. Please follow this process for each one. At the end, there are a few questions to gather general feedback about the survey. Please continue until you are on the page that says “Your response has been recorded.” During the course of the survey, it is important that you do not refresh the page, because it will restart the process. If you have any questions, feel free to ask, and please tell us when you are done.

8.3 Debriefing Form

Debrief Form

Study: B1 - Buffering Versus Interrupts

Thank you for participating in the study Buffering Versus Interrupts! The purpose of this study is to record how defects in a video's quality may annoy a user – we check to see how annoyances change as the buffer time and the number of interruptions in a video fluctuates. We also wanted to know if the level of motion in a video affects how annoyed you were. We believe that videos with higher motion will cause a user to feel more annoyed when defects occur. This study is important because the information gained from this can help streaming services find a better ratio for interrupts and buffers, so as to minimize a user's annoyance.

In this study we asked participants to watch 16 videos, experiencing either buffers at the start of a video or interrupts during, and to rate the video from 1 to 5 in terms of annoyance; we also asked them to rate their opinion of the video, so as to verify that the video's content wasn't changing their annoyance level. The only deception you may have experienced in this study is the fact that the defects in the video were not real, but encoded into the video by us. We expect to find that there are two different relationships, one between buffer time and annoyance that resembles a smooth incline, but becomes exponentially related at some transition point and one between interrupts and annoyance, that resembles a logarithmic function, where annoyance rises quickly, but relatively slows down as it approaches a higher annoyance value.

Previous research has shown that our predicted relationship for buffer times does resemble data other organizations have found. If you would like to learn more about this topic we can give you our references. Do you have any questions about this study? When you were doing the study what did you think the study was about? Was there any part of the study that was difficult? What would you change about the study?

Again, thank you for your participation in our research. If you have any questions you can ask me now or you can contact Prof. Mark Claypool at claypool@cs.wpi.edu / (508) 831-5409, Joshua Allard at jnallard@wpi.edu, or Andrew Roskuski at ajroskuski@wpi.edu at a later date. If you would like to see our final analysis of our results, via our IQP paper, we can email them to you at the end of the study. We will update your credit for taking the survey momentarily.

8.4 Additional Documents

The following documents are included in this section of the Appendix:

LIST OF VIDEOS

IRB – APPLICATION

IRB – APPROVAL

2D Photography Rube Goldberg Machine

Type Video Recording
Contributor 2D House
URL http://www.youtube.com/watch?v=qKpxd8hzOcQ&feature=youtube_gdata_player
Date 2011-07-10
Accessed 11/1/2014, 8:50:02 PM
Library Catalog YouTube
Running Time 253 seconds
Abstract <http://facebook.com/2dhouse/> <http://2dhouse.com/blog/2011/07/rube-goldberg/> - This is our Photography-themed Rube Goldberg Machine. We hope you all enjoy the clip! Watch our video explaining some of the process of getting this video done here: <http://youtu.be/qfrmTN0Ly94> You can watch the teasers we put out in the months leading up to our Rube Goldberg Machine here: Teaser #1: <http://youtu.be/1Z-yvoslbn4> Teaser #2: <http://youtu.be/M-4qPNObTcI> Special thanks to everyone who helped make this video possible!! Check out the blog post above for more information.
Date Added 11/1/2014, 8:50:02 PM
Modified 11/1/2014, 8:50:02 PM

3D Waterfall-Tranquil Mountain River-Relax-Calming Nature Sounds-Bird Song-Dripping Water

Type Video Recording
Contributor johnnielawson
URL http://www.youtube.com/watch?v=2lnxIdOTxog&feature=youtube_gdata_player
Date 2011-03-08
Accessed 11/3/2014, 6:56:45 AM
Library Catalog YouTube
Running Time 301 seconds
Abstract 3D Waterfall - Relax with the meditative sound of Nature and this Tranquil Waterfall. Take 5 minutes for yourself, ENJOY! Use this video and the other videos on my channel as part of a Stress Management plan, they are very suitable for a Stress Relief programme. You can freely use this and my other videos as part of your Meditation Relaxation plan. These pure crystal clear waters flowing off the mountain energise and heal the body and mind. Lose yourself in the tranquil beauty of this peaceful waterfall scene. I filmed this

video whilst standing in the cool waters of Glencar waterfall County Leitrim in the Northwest of Ireland. Please share this video with anyone you think may find it useful to relax with. Thank you very much for watching, commenting, rating and supporting this and my other relaxing videos, This video is filmed and produced by myself, Johnnie Lawson. I am a visual Artist living and working in the North West of Ireland. Thanks to melarancida Benboncan and Mugurm who have generously granted permission to use their respective sound recordings under creative commons attribution, www.freesound.org
CHANNEL:- <http://www.youtube.com/user/johnnielawson> VIDEO:-
<http://youtu.be/2lnxIdOTxog>

Date Added 11/3/2014, 6:56:45 AM

Modified 11/3/2014, 6:56:45 AM

A Kitten A Cat & A Bag...

Type Video Recording

Contributor Cole and Marmalade

URL http://www.youtube.com/watch?v=xY4MUwoBj2s&feature=youtube_gdata_player

Date 2013-09-02

Accessed 11/15/2014, 11:59:42 PM

Library Catalog YouTube

Running Time 94 seconds

Abstract Follow COLE & MARMALADE on Facebook! <https://www.facebook.com/ColeTheBlackCat> Crazy Kitten (Marmalade) trying to entice Cole the Black Cat to play too! Follow COLE & MARMALADE on Facebook! <https://www.facebook.com/ColeTheBlackCat> 10 Reasons Why Cats Are Better Than Men! <http://www.youtube.com/watch?v=ovjhnMugWmc> Tiny Kitten - BIG MEOW! <http://www.youtube.com/watch?v=Dnx0z1cC8u8> NEW iPhone made for CATS! <http://www.youtube.com/watch?v=2nRKU8KG6Z0> Black Cat Stick 'em Up! <http://www.youtube.com/watch?v=-Q5QR82Fu4E> Cat Sees Owner for First Time in 12 Months! <http://www.youtube.com/watch?v=H6j7UnpUMy4>

Date Added 11/15/2014, 11:59:42 PM

Modified 11/15/2014, 11:59:42 PM

Breakdance Battle - Chelles Battle Pro 2014 Final

Type Video Recording

Contributor CANAL STREET Dance

URL http://www.youtube.com/watch?v=9tG-xwv0kw0&feature=youtube_gdata_player

Date 2014-03-07

Accessed 11/7/2014, 8:10:08 PM

Library Catalog YouTube

Running Time 699 seconds

Abstract Breakdance Battle Chelles Battle Pro 2014 8 vs 8 final, the 1st of March between Team Rhône Alpes and Gamblerz ! Don't forget to suscribe to get a new video once a week : http://www.youtube.com/user/canalstreetdance?sub_confirmation=1 Battles, reports, coulisses des évènements majeurs de street dance. Une fois par semaine, découvrez une vidéo exclusive de CANAL STREET, le site des cultures urbaines par CANAL+ : - Voir toutes les battles : http://www.youtube.com/playlist?list=PLphHRzzaPKox0_UCSTbSsq0Gd5FIVxPZ8 - Voir tous les backstages : <http://www.youtube.com/playlist?list=PLphHRzzaPKoyOKh8OBSKubvR28LR15wMC> Abonnez-vous pour recevoir une nouvelle vidéo chaque semaine ! http://www.youtube.com/user/canalstreetdance?sub_confirmation=1 Follow us: Site : <http://canalstreet.canalplus.fr/danse> Facebook : <http://www.facebook.com/canalstreet> Twitter : <http://twitter.com/canalstreet> Google+ : <http://plus.google.com/+canalstreet> Instagram : <http://instagram.com/canalstreet>

Date Added 11/7/2014, 8:10:08 PM

Modified 11/7/2014, 8:10:08 PM

Bruno Mars - The Lazy Song [OFFICIAL VIDEO]

Type Video Recording

Contributor Bruno Mars

URL https://www.youtube.com/watch?v=fLexgOxsZu0&feature=youtube_gdata_player

Date 2011-04-15

Accessed 11/7/2014, 8:52:52 PM

Library Catalog YouTube

Running Time 209 seconds

Abstract Moonshine Jungle Tour 2014 tickets and more info: <http://www.brunomars.com/moonshinejungletour> Available now on iTunes! <http://smarturl.it/Doo-Wops> © 2011 WMG. The official music video for 'The Lazy Song' by Bruno Mars from doo-wops and hooligans - available now on Elektra Records! Directed by Bruno Mars and Cameron Duddy Choreographed and Performed by Poreotics: <http://poreotics.com> Connect with Bruno: <http://www.brunomars.com> <http://www.twitter.com/brunomars> <http://www.facebook.com/brunomars> <http://www.instagram.com/brunomars> Get 'Unorthodox Jukebox' on iTunes

<http://smarturl.it/unorthodoxjukebox> Follow Bruno on Spotify <http://bit.ly/1gDnkxr> LYRICS Today I don't feel like doing anything I just wanna lay in my bed Don't feel like picking up my phone So leave a message at the tone 'Cause today I swear I'm not doing anything. Uh! I'm gonna kick my feet up Then stare at the fan Turn the TV on, throw my hand in my pants Nobody's gonna tell me I can't I'll be lounging on the couch, Just chillin' in my snuggie Click to MTV, so they can teach me how to dougie 'Cause in my castle I'm the freaking man Oh, yes I said it I said it I said it 'cause I can Today I don't feel like doing anything I just wanna lay in my bed Don't feel like picking up my phone So leave a message at the tone 'Cause today I swear I'm not doing anything Nothing at all! Ooh, hoo, ooh, hoo, ooh, ooh-ooh Nothing at all Ooh, hoo, ooh, hoo, ooh, ooh-ooh Tomorrow I'll wake up, do some P90X Meet a really nice girl, have some really nice sex And she's gonna scream out: 'This is Great' (Oh my God, this is great) Yeah I might mess around, get my college degree I bet my old man will be so proud of me But sorry pops, you'll just have to wait Haha Oh, yes I said it I said it I said it 'cause I can Today I don't feel like doing anything I just wanna lay in my bed Don't feel like picking up my phone So leave a message at the tone 'Cause today I swear I'm not doing anything No, I ain't gonna comb my hair 'Cause I ain't going anywhere No, no, no, no, no, no, no, no, no, no I'll just strut in my birthday suit And let everything hang loose Yeah, yeah, yeah, yeah, yeah, yeah, yeah, yeah, yeah, yeah, yeah Ooh Today I don't feel like doing anything I just wanna lay in my bed Don't feel like picking up my phone So leave a message at the tone 'Cause today I swear I'm not doing anything Nothing at all Nothing at all Nothing at all Comic Chimp Mask used by permission of Easter Unlimited, Inc. / FunWorld Div. All Rights Reserved. Copyright 2009.

Date Added 11/7/2014, 8:52:52 PM

Modified 11/7/2014, 8:52:52 PM

Concert Band - Coldplay on Stage

Type Video Recording

Contributor Hattie Matthews

URL https://www.youtube.com/watch?v=79Y7h1OnXo8&feature=youtube_gdata_player

Date 2013-04-08

Accessed 11/1/2014, 9:35:18 PM

Library Catalog YouTube

Running Time 342 seconds

Date Added 11/1/2014, 9:35:18 PM

Modified 11/1/2014, 9:35:18 PM

Guinness World Record - Longest / Biggest Domino Line Ever

Type Video Recording
Contributor ShanesDominoez
URL http://www.youtube.com/watch?v=vDy2xWpZWVc&feature=youtube_gdata_player
Date 2010-05-15
Accessed 11/1/2014, 8:31:22 PM
Library Catalog YouTube
Running Time 229 seconds
Abstract Like me on Facebook! <http://www.facebook.com/pages/ShanesDominoez/114082981938704> This is the longest domino line ever. My subscriber special for 1,500 subscribers is coming. So if you want to see an awesome domino line, subscribe. Just sayin'. A special thanks to my friend sabs who helped me clean up the dominoes. We cleaned up this domino line in exactly 12:43 seconds. Which is a lot shorter than most people would expect. This is the longest domino line, nothing else. Just a line, no fields, no walls, just a line. So I don't want any dumb comments. Music by : Blarsa
www.newgrounds.com/audiportal Search : Garden Party Dominoes : 4,400 Breaks : 0!!!! ENJOY!!!!!!
Date Added 11/1/2014, 8:31:22 PM
Modified 11/1/2014, 8:31:22 PM

HD Train scenery from Tirano, Italy to St. Moritz Christmas day part 6

Type Video Recording
Contributor EdWaldrup
URL http://www.youtube.com/watch?v=AmpKX8tEGs&feature=youtube_gdata_player
Date 2010-07-04
Accessed 11/1/2014, 8:32:13 PM
Library Catalog YouTube
Running Time 590 seconds
Abstract Return from Tirano, Italy to St. Moritz Christmas day 2004. Shot from the rear window of the Bernina Express. Some of the most spectacular scenery in the Alps.
Date Added 11/1/2014, 8:32:13 PM
Modified 11/1/2014, 8:32:13 PM

Landscape oil painting in just 18 minutes wet on wet

Type Video Recording
Contributor KevinOilPainting
URL http://www.youtube.com/watch?v=Coih_-y4tVw&feature=youtube_gdata_player
Date 2013-03-16
Accessed 11/1/2014, 8:33:32 PM
Library Catalog YouTube
Running Time 1101 seconds
Abstract My website <http://paintwithkevin.com/index.html> Facebook <http://www.facebook.com/paintwithkevin>
Date Added 11/1/2014, 8:33:32 PM
Modified 11/1/2014, 8:33:32 PM

Last Week Tonight with John Oliver: Pumpkin Spice (Web Exclusive)

Type Video Recording
Contributor LastWeekTonight
URL https://www.youtube.com/watch?v=DeQqe0oj5Ls&feature=youtube_gdata_player
Date 2014-10-13
Accessed 11/1/2014, 8:35:13 PM
Library Catalog YouTube
Running Time 192 seconds
Abstract Pumpkin spice. Why? John Oliver investigates. Well, he doesn't really investigate. He says things about it, though! Connect with Last Week Tonight online... Subscribe to the Last Week Tonight YouTube channel for more almost news as it almost happens: www.youtube.com/user/LastWeekTonight Find Last Week Tonight on Facebook like your mom would: <http://Facebook.com/LastWeekTonight> Follow us on Twitter for news about jokes and jokes about news: <http://Twitter.com/LastWeekTonight> Visit our official site for all that other stuff at once: <http://www.hbo.com/lastweektonight>
Short Title Last Week Tonight with John Oliver
Date Added 11/1/2014, 8:35:13 PM
Modified 11/1/2014, 8:35:13 PM

Moanalua High School Football Varsity Entrance 2011 vs Farrington High School

Type Video Recording
Contributor daniyoh
URL http://www.youtube.com/watch?v=wePuwI2pSdg&feature=youtube_gdata_player
Date 2011-12-14
Accessed 11/8/2014, 11:12:23 PM
Library Catalog YouTube
Running Time 124 seconds
Abstract Moanalua High School vs Farrington High School 2011 (last game of the regular season) Some athletes wore and/or dyed their hair pink and ran out with a pink flag at the end in-tribute to breast cancer awareness. From a spectators stand point, by far the best game of the season. Shout out to #44 of MoHs. \nm/ \nm/ \nm/ \nm/
Date Added 11/8/2014, 11:12:23 PM
Modified 11/8/2014, 11:12:23 PM

Mousetrap Chain Reaction in Slow Motion - The Slow Mo Guys

Type Video Recording
Contributor The Slow Mo Guys
URL http://www.youtube.com/watch?v=tO7LIRhGbfo&feature=youtube_gdata_player
Date 2014-03-27
Accessed 11/1/2014, 9:20:28 PM
Library Catalog YouTube
Running Time 187 seconds
Abstract Gav sets up a chain of 150 mousetraps. Dan sets them off bravely. Follow Gav on Twitter - <https://twitter.com/GavinFree> Follow Dan on Twitter - <https://twitter.com/DanielGruchy> Filmed at 2500fps with a Phantom Flex
Mousetrap Chain Reaction in Slow Motion - The Slow Mo Guys
Date Added 11/1/2014, 9:20:28 PM
Modified 11/1/2014, 9:20:28 PM

OK Go - The Writing's On the Wall - Official Video

Type Video Recording
Contributor OK Go
URL https://www.youtube.com/watch?v=m86ae_e_ptU&feature=youtube_gdata_player
Date 2014-06-17
Accessed 11/1/2014, 8:35:30 PM
Library Catalog YouTube
Running Time 257 seconds
Abstract Buy Hungry Ghosts on iTunes: <http://smarturl.it/HungryGhostsiTunes> For dozens of exclusive behind-the-scenes video clips from the set, visit <http://okgo.net/wotw/> Buy the EP now on iTunes: <http://smarturl.it/OKGOEP>
LYRICS: Listen, I know it's been hard, you know it's no different for me. We're less than a zero-sum game now, and baby we both know that's not how it's supposed to be. The writing's on the wall — it seems like forever since we had a good day. The writing's on the wall, and I... I just want to get you high tonight. I just wanna see some pleasure in your eyes. And I go too hot, and you go too cold and we both fall apart. and you bring your mind to rest against mine, but the mind has no say in affairs of the heart. The writing's on the wall — it seems like forever since we had a good day. The writing's on the wall, and I... I just want to get you high tonight. I just wanna see some pleasure in your eyes. Even if it's the last thing that we do together. **CREDITS:** Created By SpecialGuest.tv for OK Go Directed by Aaron Duffy, Damian Kulash, Jr. & Bob Partington Production Co - 1stAveMachine Exec. Producer - Cheri Anderson at SpecialGuest Exec. Producer - Sam Penfield at 1stAveMachine Producer - Garrett Fennelly Creative Consultant - Mary Fagot Director of Photography - William Rexer Production Designer - Ethan Tobman Wardrobe Designer - Ciera Wells Color - Ricart + Co -- Seth Ricart Audio - Dig It Audio -- Jeffery Seelye
Date Added 11/1/2014, 8:35:30 PM
Modified 11/1/2014, 8:35:30 PM

President Obama Makes a Statement on Hurricane Sandy

Type Video Recording
Contributor The White House
URL https://www.youtube.com/watch?v=oyZltYdr8Ao&feature=youtube_gdata_player
Date 2012-10-29
Accessed 11/1/2014, 8:36:27 PM
Library Catalog YouTube
Running Time 390 seconds

Abstract President Obama discusses efforts underway to prepare for Hurricane Sandy and urges those who may be affected by the storm to take all needed precautions. October 29, 2012.

Date Added 11/1/2014, 8:36:27 PM

Modified 11/1/2014, 8:36:27 PM

The Colbert Report: Who's Attacking Me Now? - #CancelColbert

Type Video Recording

Contributor Comedy Central

URL http://www.youtube.com/watch?v=MBPgXjkfBXM&feature=youtube_gdata_player

Date 2014-04-01

Accessed 11/1/2014, 8:33:08 PM

Library Catalog YouTube

Running Time 454 seconds

Abstract The Interwebs nearly swallow Stephen whole after the dark forces of Twitter descend on The Report.

Short Title The Colbert Report

Date Added 11/1/2014, 8:33:08 PM

Modified 11/1/2014, 8:33:08 PM

The Daily Show: Consequence-Free Speech

Type Video Recording

Contributor Comedy Central

URL https://www.youtube.com/watch?v=D3Lv6tINW8I&feature=youtube_gdata_player

Date 2014-05-01

Accessed 11/9/2014, 12:04:19 AM

Library Catalog YouTube

Running Time 219 seconds

Abstract Sarah Palin gives a speech conflating a sacrament of her faith with torture.

Short Title The Daily Show

Date Added 11/9/2014, 12:04:19 AM

Modified 11/9/2014, 12:04:19 AM

The Ohio State University Marching Band Performs their Hollywood Blockbuster Show

Type Video Recording
Contributor osumbvideo
URL https://www.youtube.com/watch?v=DNe0ZUD19EE&feature=youtube_gdata_player
Date 2013-10-27
Accessed 11/1/2014, 8:54:17 PM
Library Catalog YouTube
Running Time 530 seconds
Abstract The "Hollywood Blockbuster Show" was performed by members of the Ohio State University Marching Band on October 26th, 2013 at the Penn State Game. Coming off of their Michael Jackson Tribute show, students had a week to learn the drill associated with this show and a little over a week to learn the music. For more information on The Best Damn Band In The Land Visit our website at: <http://tbdbitl.osu.edu/>
Date Added 11/1/2014, 8:54:17 PM
Modified 11/1/2014, 8:54:17 PM



WORCESTER POLYTECHNIC INSTITUTE
Institutional Review Board
Application for Approval to
Use Human Subjects in Research

WPI IRB use only
IRB #
Date:

This application is for: (Please check one) <input checked="" type="checkbox"/> Expedited Review <input type="checkbox"/> Full Review	WPI IRB use only
Principal Investigator (PI) or Project Faculty Advisor: (NOT a student or fellow; must be a WPI employee)	
Name: <u>Mark Claypool</u> Tel No: <u>508-831-5409</u> E-Mail Address: <u>claypool@cs.wpi.edu</u>	<input type="checkbox"/>
Department: <u>IMGD/CS Departments</u>	
Co-Investigator(s): (Co-PI(s)/non students)	
Name: _____ Tel No: _____ E-Mail Address: _____	<input type="checkbox"/>
Name: _____ Tel No: _____ E-Mail Address: _____	<input type="checkbox"/>
Student Investigator(s):	
Name: <u>Joshua Allard</u> Tel No: <u>603-657-4258</u> E-Mail Address: <u>jnallard@wpi.edu</u>	<input type="checkbox"/>
Name: <u>Andrew Roskuski</u> Tel No: <u>413-887-9162</u> E-Mail Address: <u>ajroskuski@wpi.edu</u>	<input type="checkbox"/>

Check if: Undergraduate project (MQP, IQP, Suff., other) IQP
 Graduate project (M.S. Ph.D., other) _____

Has an IRB ever suspended or terminated a study of any investigator listed above?
No Yes (Attach a summary of the event and resolution.)

Vulnerable Populations: The proposed research will involve the following (Check all that apply):
pregnant women human fetuses neonates minors/children prisoners
students individuals with mental disabilities individuals with physical disabilities

Collaborating Institutions: (Please list all collaborating Institutions.)

Locations of Research: (If at WPI, please indicate where on campus. If off campus, please give details of locations.)
Campus - Computer Lab (We will be using the Zoo lab in Fuller Laboratories)

Project Title: Streaming-Buffering vs Interrupts

Funding: (If the research is funded, please enclose one copy of the research proposal or most recent draft with your application.)

Funding Agency: (none) WPI Fund: _____

Human Subjects Research: (All study personnel having direct contact with subjects **must** take and pass a training course on human subjects research. There are links to web-based training courses that can be accessed under the Training link on the IRB web site <http://www.wpi.edu/offices/irb/training.html>. The IRB requires a copy of the completion certificate from the course or proof of an equivalent program.)

Anticipated Dates of Research:

Start Date: 11/24/2014 Completion Date: 12/18/14



WORCESTER POLYTECHNIC INSTITUTE
Institutional Review Board
Application for Approval to
Use Human Subjects in Research

WPI IRB use only
IRB #
Date:

Instructions: Answer all questions. If you are asked to provide an explanation, please do so with adequate details. If needed, attach itemized replies. Any incomplete application will be returned.

1.) Purpose of Study: *(Please provide a concise statement of the background, nature and reasons for the proposed study. Insert below using non-technical language that can be understood by non-scientist members of the IRB.)*

In this experiment, we will investigate the effects of buffering and interrupts on a user's video watching experience. We will also investigate whether a video's motion has an effect. This information will be used to improve buffering algorithms in streaming media players, in order to minimize the level of annoyance experienced by users.

2.) Study Protocol: *(Please attach sufficient information for effective review by non-scientist members of the IRB. Define all abbreviations and use simple words. Unless justification is provided this part of the application must not exceed 5 pages. Attaching sections of a grant application is not an acceptable substitute.)*

A.) For **biomedical, engineering and related research**, please provide an outline of the actual experiments to be performed. Where applicable, provide a detailed description of the experimental devices or procedures to be used, detailed information on the exact dosages of drugs or chemicals to be used, total quantity of blood samples to be used, and descriptions of special diets.

B.) For applications in the **social sciences, management and other non-biomedical disciplines** please provide a detailed description of your proposed study. Where applicable, include copies of any questionnaires or standardized tests you plan to incorporate into your study. If your study involves interviews please submit an outline indicating the types of questions you will include.

C.) If the study involves **investigational drugs or investigational medical devices**, and the PI is obtaining an Investigational New Drug (IND) number or Investigational Device Exemption (IDE) number from the FDA, please provide details.

D.) Please note if any **hazardous materials** are being used in this study.

E.) Please note if any **special diets** are being used in this study.

3.) Subject Information:

A.) Please provide the exact number of subjects you plan to enroll in this study and describe your subject population. *(eg. WPI students, WPI staff, UMASS Medical patient, other)*

Males: 25 Females: 25 Description: Mostly WPI students, maybe staff.

B.) Will subjects who do not understand English be enrolled?

No Yes *(Please insert below the language(s) that will be translated on the consent form.)*

C.) Are there any circumstances under which your study population may feel coerced into participating in this study? No Yes *(Please insert below a description of how you will assure your subjects do not feel coerced.)*

We will advertise and invite people to attend - We will provide an incentive, but there will be no negative effects for not attending.



WORCESTER POLYTECHNIC INSTITUTE
Institutional Review Board
Application for Approval to
Use Human Subjects in Research

WPI IRB use only
IRB #
Date:

D.) Are the subjects at risk of harm if their participation in the study becomes known?
No Yes (Please insert below a description of possible effects on your subjects.)

We are not matching names to our data. If names are stored, it will only be for use in an incentive raffle.

E.) Are there reasons for excluding possible subjects from this research?
No Yes (If yes, please explain.)

F.) How will subjects be recruited for participation? (Check all that apply.)

- | | |
|--|---|
| <input checked="" type="checkbox"/> Referral: (By whom) <u>The investigators</u> | <input checked="" type="checkbox"/> Direct subject advertising, including: (Please provide a copy of the proposed ad. All direct subject advertising must be approved by the WPI IRB prior to use.) |
| <input type="checkbox"/> Other: (Identify) _____ | <input type="checkbox"/> Newspaper |
| <input type="checkbox"/> Database: (Describe how database populated) _____ | <input type="checkbox"/> Radio |
| | <input type="checkbox"/> Television |
| | <input type="checkbox"/> Internet |
| | <input checked="" type="checkbox"/> Bulletin board |
| | <input checked="" type="checkbox"/> Flyers |
| | <input type="checkbox"/> Letters |
| | <input checked="" type="checkbox"/> E-mail |

F.) Have the subjects in the database agreed to be contacted for research projects? No Yes N/A

G.) Are the subjects being paid for participating? (Consider all types of reimbursement, ex. stipend, parking, travel.)

- No Yes (Check all that apply.) Cash Check Gift certificate Other: Raffle for a 25 dollar gift certificate.
- Amount of compensation _____

4.) Informed Consent:

A.) Who will discuss the study with and obtain consent of prospective subjects? (Check all that apply.)

- Principal Investigator Co-Investigator(s) Student Investigator(s)

B.) Are you aware that subjects must read and sign an Informed Consent Form prior to conducting any study-related procedures and agree that all subjects will be consented prior to initiating study related procedures?

No Yes

C.) Are you aware that you must consent subjects using only the IRB-approved Informed Consent Form?

No Yes

D.) Will subjects be consented in a private room, not in a public space?

No Yes

E.) Do you agree to spend as much time as needed to thoroughly explain and respond to any subject's questions about the study, and allow them as much time as needed to consider their decision prior to enrolling them as subjects?

No Yes

F.) Do you agree that the person obtaining consent will explain the risks of the study, the subject's right to decide not to participate, and the subject's right to withdraw from the study at any time?

No Yes

G.) Do you agree to either 1.) retain signed copies of all informed consent agreements in a secure location for at least three years or 2.) supply copies of all signed informed consent agreements in .pdf format for retention by the IRB in electronic form?

No Yes

(If you answer No to any of the questions above, please provide an explanation.)



WORCESTER POLYTECHNIC INSTITUTE
Institutional Review Board
Application for Approval to
Use Human Subjects in Research

WPI IRB use only
IRB #
Date:

Participants will have the ability to sign the consent form in a private room, prior to the study, but walk-ins will also be able to sign the form in the public lab.

5.) Potential Risks: *(A risk is a potential harm that a reasonable person would consider important in deciding whether to participate in research. Risks can be categorized as physical, psychological, sociological, economic and legal, and include pain, stress, invasion of privacy, embarrassment or exposure of sensitive or confidential data. All potential risks and discomforts must be minimized to the greatest extent possible by using e.g. appropriate monitoring, safety devices and withdrawal of a subject if there is evidence of a specific adverse event.)*

A.) What are the risks / discomforts associated with each intervention or procedure in the study?
Eye-strain from screen viewing, but this will be minimal because the survey will be less than 30 minutes. Participants may feel a sense of annoyance, but that is the feeling we are measuring.

B.) What procedures will be in place to prevent / minimize potential risks or discomfort?
Users may take breaks from the survey if needed, and they can leave it at any time.

6.) Potential Benefits:

A.) What potential benefits other than payment may subjects receive from participating in the study?
None

B.) What potential benefits can society expect from the study?
This data can be used by streaming services to improve the user experience for streaming videos, hopefully limiting the level of annoyance a user may encounter while viewing streaming videos.

7.) Data Collection, Storage, and Confidentiality:

A.) How will data be collected?
Data will be collected using an online Qualtrics survey.

B.) Will a subject's voice, face or identifiable body features (eg. tattoo, scar) be recorded by audio or videotaping?
No Yes (Explain the recording procedures you plan to follow.)

C.) Will personal identifying information be recorded? No Yes (If yes, explain how the identifying information will be protected. How will personal identifying information be coded and how will the code key be kept confidential?)

D.) Where will the data be stored and how will it be secured?
Data will be stored on Qualtrics, and will be secured by their own security measures.

E.) What will happen to the data when the study is completed?
Data will be archived in a secure space; it will not be available online.

F.) Can data acquired in the study adversely affect a subject's relationship with other individuals? (i.e. employee-



WORCESTER POLYTECHNIC INSTITUTE
Institutional Review Board
Application for Approval to
Use Human Subjects in Research

WPI IRB use only
IRB #
Date:

supervisor, student-teacher, family relationships)

No. _____

G.) Do you plan to use or disclose identifiable information outside of the investigation personnel?
No Yes (Please explain.)

H.) Do you plan to use or disclose identifiable information outside of WPI including non-WPI investigators?
No Yes (Please explain.)

8.) Incidental findings: *In the conduct of information gathering, is it possible that the investigator will encounter any incidental findings? If so, how will these be handled? (An incidental finding is information discovered about a subject which should be of concern to the subject but is not the focus of the research. For example, a researcher monitoring heart rates during exercise could discover that a subject has an irregular heartbeat.)*

No. _____

9.) Deception: *(Investigators must not exclude information from a subject that a reasonable person would want to know in deciding whether to participate in a study.)*

Will the information about the research purpose and design be withheld from the subjects?
No Yes (Please explain.)

10.) Adverse effects: *(Serious or unexpected adverse reactions or injuries must be reported to the WPI IRB within 48 hours using the IRB Adverse Event Form found out at <http://www.wpi.edu/offices/irb/forms.html>. Other adverse events should be reported within 10 working days.)*

What follow-up efforts will be made to detect any harm to subjects and how will the WPI IRB be kept informed?

There are no adverse effects in consideration that we need to be prepared for.

11.) Conflict of Interest: *(A conflict of interest occurs when an investigator or other key personnel in a study may enjoy material benefits based on study results. Relationships that give rise to a conflict of interest or the appearance of a conflict of interest must be disclosed in the informed consent statement provided to study subjects. More information, including examples of relationships that require disclosure and those that do not, can be found [here](#).)*

A.) Do any of the investigators listed on this application have a potential or actual conflict of interest with regard to this study?

- a. Investigator (name) Mark Claypool No Yes
- b. Investigator (name) Joshua Allard No Yes
- c. Investigator (name) Andrew Roskuski No Yes

B.) If any of the answers to 11A. are "Yes," please attach an explanation of the nature of the conflict to this



WORCESTER POLYTECHNIC INSTITUTE
Institutional Review Board
Application for Approval to
Use Human Subjects in Research

WPI IRB use only
IRB #
Date:

application and identify appropriate language for use in the consent form. Examples of consent language are found on the IRB website, [here](#).

C.) Does each investigator named above have a current WPI conflict of interest disclosure form on file with the appropriate supervisor/department head? No Yes

D.) Do any of the investigators' COI forms on file with WPI contain information regarding this research?
No Yes
a. If "Yes," identify the investigator(s) _____

12.) Informed consent: (*Documented informed consent must be obtained from all participants in studies that involve human subjects. You must use the templates available at <http://www.wpi.edu/offices/irb/forms.html> to prepare these forms. **Informed consent forms must be included with this application.** Under certain circumstances the WPI IRB may waive the requirement for informed consent.*)

Investigator's Assurance:

I certify the information provided in this application is complete and correct.

I understand that I have ultimate responsibility for the conduct of the study, the ethical performance of the project, the protection of the rights and welfare of human subjects, and strict adherence to any stipulations imposed by the WPI IRB.

I agree to comply with all WPI policies, as well all federal, state and local laws on the protection of human subjects in research, including:

- ensuring the satisfactory completion of human subjects training.
- performing the study in accordance with the WPI IRB approved protocol.
- implementing study changes only after WPI IRB approval.
- obtaining informed consent from subjects using only the WPI IRB approved consent form.
- promptly reporting significant adverse effects to the WPI IRB.

Signature of Principal Investigator _____

Date

11-19-19

Print Full Name and Title _____

Mark Claypool

Professor of CS

*Please return a signed hard copy of this application to the WPI IRB c/o Ruth McKeogh 2nd Floor Project Center
Or email an electronic copy to irb@wpi.edu
If you have any questions, please call (508) 831-6699.*

Methodology

1 Resources Needed

For the purpose of this study, we will need:

- Videos – We need a set of videos to manipulate and show to test subjects. Exact details about the videos will be discussed in section 2.
- mpeg_stat – We will use this tool for analyzing the level of motion for each video. This tool takes in a video file, analyzes its properties, and reports statistics. From this tool, we will be looking at the skipped macroblocks statistic to determine the motion of a video. Macroblocks are the processing unit used for compressing the video. The skipped macroblocks measurement is derived from analyzing the average compression rate over the entire video.
- Video Editing Tool – We will use a video editor to add artificial buffering and interrupts to videos. Since the main task we need to do is inserting a video clip representing buffering at various points in the video, programs such as Windows Movie Maker will suffice as long as we use the same one for everything, because the insertion operation is relatively basic.
- Data Analysis and Graphing Tool – We will need something to help us analyze and produce visual representations of the data that we collect. Excel or similar programs would work for this task.
- Lab Space – We need somewhere to conduct the tests where lab computers are provided. The Zoo Lab in Fuller Laboratories would be a possible candidate, as well as the computer labs in Salisbury Laboratories and Higgins Laboratories. We will need to be able to install our test suite beforehand, and have the users be able to log on to the computers to take the test. We will need all of the computers to be running the same OS, and for them to be capable of running our selected video player.

2 Gathering Videos

We want to find videos with specific levels of motion. We will set ranges based on the percentage of interpolated macroblocks as determined by mpeg_stat. This statistic represents the number of blocks that did not change between frames of the video on average, which is determined as dictated in section 1. We will use live action videos, which are videos that capture the movement of real people, to keep complexity fairly constant. We will attempt to select videos with inoffensive content so as to not influence the results. We will also be asking users their opinion on the video

content during the survey to determine if the content is affecting their responses. The videos we are looking for will be about 30 seconds in length, not including the time added by buffers and interrupts, because this allows for a period of time long enough for interrupts to be dispersed throughout and for a longer reasonable buffering time. These videos will only contain one scene, because the switch between scenes would cause a miscalculation in our level of motion. For an example of videos to be chosen, our low motion clips could consist of news anchors talking for a minute, while those with high levels of motion could be videos following soccer players, as long as they contain no perspective changes.

3 Procedure

For this experiment it is important to classify the videos we collect into four levels of motion.

We suggest the following categories, which will be determined by ranges of the percent skipped macroblocks:

L – Low Motion

ML – Medium Low Motion

MH – Medium High Motion

H – High Motion

We need to categorize the users into two groups, one consisting of users experiencing various buffer sizes, and one of users experiencing various numbers of interrupts. Each group will be split into two subgroups, each observing the entire range of their test factor, but seeing it distributed differently across the set of videos. For example, both subgroups watch a talking head while experiencing interrupts, but one would see few interrupts, but the other would see many. Every group will watch the same set of videos, but they will be manipulated differently according to their test factor and subgroup.

Experimental groups:

I – Buffering Group 1

II – Buffering Group 2

III – Interrupts Group 1

IV – Interrupts Group 2

The experiment will begin by each user watching a control clip which contains no interrupts or buffering time. They will be instructed to use this video as a reference for perfect QoE, meaning no annoyance. They will then watch their first clip, experiencing their factor in question. After watching the video, they will be asked to rate, using a paper survey provided at the beginning of the test, on an 11 point scale, ranging from 0 to 10, their level of annoyance with the QoS provided. They will then repeat this process for each video in their group's set, which will consist of 16 videos, with 4 from each motion category. At the end, they will be asked for general feedback about the test and their opinion on the factor that they experienced. During the study we will collect demographic information, such as age, sex, and major, but will not store any identifying information. The entire test should take no more than 30 minutes total.

4 Analysis/Expected Results

From the data received, we can analyze each factor separately, creating statistics and graphs which we can compare to our hypotheses. We will take note of any relationships that may develop, as well as look for any outliers in the dataset. If possible we will attempt to describe the relationships formulaically. After we have analyzed each factor separately, we can combine the data, to create a generic graph which can predict the level of user annoyance for a given combination of buffer size and interrupt amount. If adequate relationships exist, we can create an equation to estimate user annoyance in any situation.



The University of
Science and Technology.
And Life..

Informed Consent Agreement for Participation in a Research Study

Investigators: Mark Claypool, Andrew Roskuski, Joshua Allard

Contact Information: CS Department
WPI
100 Institute Road
Worcester, MA 01609
Tel. 508-831-5409, Email: claypool@cs.wpi.edu

Title of Research Study: Streaming-Buffering vs Interrupts

Introduction:

You are being asked to participate in a research study. Before you agree, however, you must be fully informed about the purpose of the study, the procedures to be followed, and any benefits, risks or discomfort that you may experience as a result of your participation. This form presents information about the study so that you may make a fully informed decision regarding your participation.

Purpose of the study:

In this experiment, we will investigate the effects of buffering and interrupts on your video watching experience. This information will be used to improve buffering algorithms in streaming media players, in order to minimize the level of annoyance experienced by users.

Procedures to be followed:

You will be seated at a computer station. One of the investigators will open an online survey and tell you how to proceed. You should plug in your headphones (either personal or supplied) to the computer and wear them. First you will be asked to answer some demographic questions, but your name will not be recorded. After that, you will watch a control video to use as a baseline. Following that you will watch 16 short videos, answering a few questions after each. At the end, you will be asked a few general questions about your overall experience. After the survey is complete, you are free to leave.

Risks to study participants:

You may experience eye strain from looking at the screen. You also will likely experience a slight sense of annoyance.

Benefits to research participants and others:

There is no direct benefit to you.

Record keeping and confidentiality:

Records of your participation in this study will be held confidential so far as permitted by law. However, the study investigators, the sponsor or its designee and, under certain circumstances, the Worcester Polytechnic Institute Institutional Review Board (WPI IRB) will be able to inspect and have

access to confidential data that does not identify you by name. Any publication or presentation of the data will not identify you.

Compensation or treatment in the event of injury:

In the unlikely event of physical injury resulting from participation in the research, you understand that medical treatment may be available from WPI, including first aid emergency care, and that your insurance carrier may be billed for the cost of such treatment. No compensation for medical care can be provided by WPI. You further understand that making such medical care available, or providing it, does not imply that such injury is the fault of the investigators. You do not give up any of your legal rights by signing this statement.

Cost/Payment:

You will be given the option to enter a raffle for a \$25 Visa prepaid gift card.

For more information about this research or about the rights of research participants, or in case of research-related injury, contact:

Prof. Mark Claypool, CS Department, WPI, 100 Institute Road, Worcester, MA (Tel. 508-831-5409). You may also contact the chair of the WPI Institutional Review Board (Prof. Kent Rissmiller, Tel. 508-831-5019, Email: kjr@wpi.edu) or WPI's University Compliance Officer (Michael J. Curley, Tel. 508-831-6919).

Your participation in this research is voluntary. Your refusal to participate will not result in any penalty to you or any loss of benefits to which you may otherwise be entitled. You may decide to stop participating in the research at any time without penalty or loss of other benefits. The project investigators retain the right to cancel or postpone the experimental procedures at any time they see fit. Data obtained in this experiment will become the property of the investigators and WPI. If you withdraw from the study, data already collected from you will remain in the study.

By signing below, you acknowledge that you have been informed about and consent to be a participant in the study described above. Make sure that your questions are answered to your satisfaction before signing. You are entitled to retain a copy of this consent agreement.

Study Participant Signature

Date: _____

Study Participant Name (Please print)

Signature of Person who explained this study

Date: _____

Proposed Advertising Text:

Dear Members of WPI,

We are trying to research the relationship linking a video's quality of service to your level of annoyance. All you need to do is watch 10 minutes of videos and tell us how annoyed you felt! It's simple!

Participate and you will have the option be entered into a raffle to win a \$25 Visa gift card!

Where: The Zoo Lab - In the sub-basement of Fuller Laboratories, A21

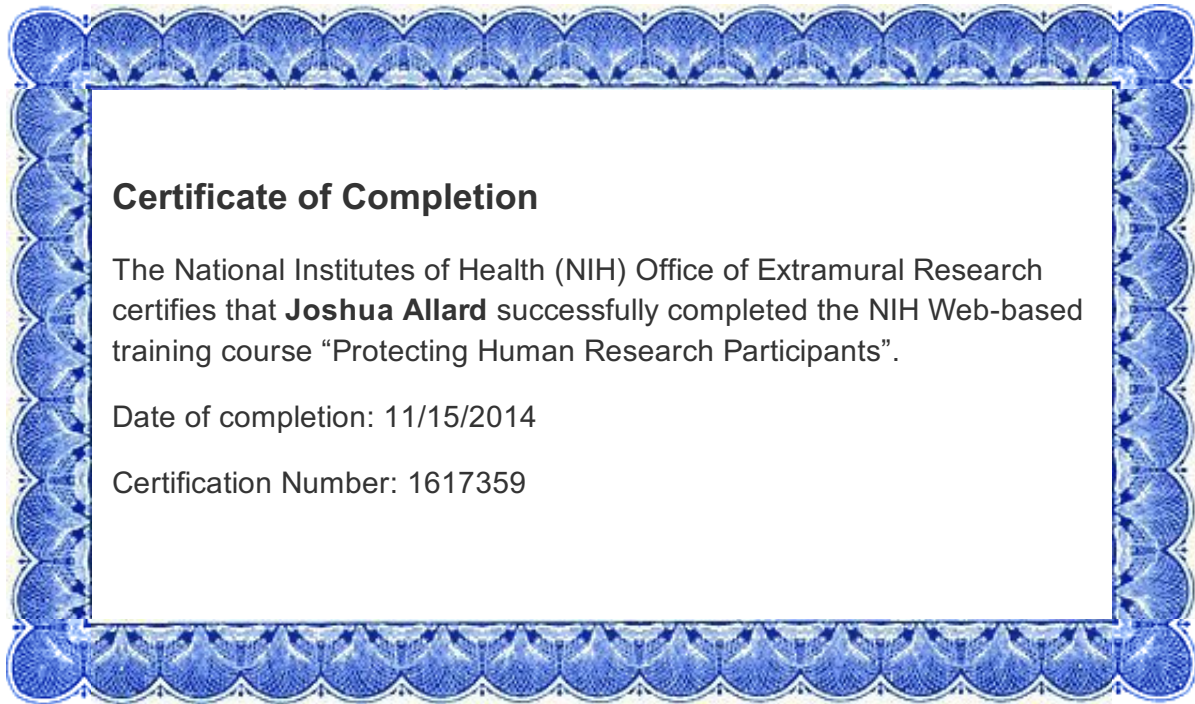
When: (Show up for any 15-20 minute section of the following dates and times)

Day 1: [TBD]

Day 2: [TBD]

Note: If you can, please bring your own headphones. If you cannot bring your own, public-use ones will be available.





WORCESTER POLYTECHNIC INSTITUTE

Worcester Polytechnic Institute IRB# 1
HHS IRB # 00007374

24 November 2014
File: 14-243

Re: IRB Expedited Review Approval: File 14-243 "Study 2:14-242 Claypool"

Dear Prof. Claypool,

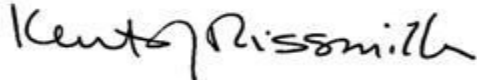
The WPI Institutional Review Committee (IRB) approves the above-referenced research activity, having conducted an expedited review according to the Code of Federal Regulations 45 (CFR46).

Consistent with 45 CFR 46.116 regarding the general requirements for informed consent, we remind you to only use the **attached stamped approved consent form** and to give a copy of the signed consent form to your subjects. You are also required to store the signed consent forms in a secure location and retain them for a period of at least three years following the conclusion of your study. You may also convert the completed consent forms into electronic documents (.pdf format) and forward them to the IRB Secretary for electronic storage.

The period covered by this approval is 24 November 2014 until 23 November 2015, unless terminated sooner (in writing) by yourself or the WPI IRB. Amendments or changes to the research that might alter this specific approval must be submitted to the WPI IRB for review and may require a full IRB application in order for the research to continue.

Please contact the undersigned if you have any questions about the terms of this approval.

Sincerely,



Kent Rissmiller
WPI IRB Chair