RICE UNIVERSITY

# Why Does Content Desirability Impact Subjective Video Quality Ratings and What Can Be Done About It?

by

**Rochelle E. Evans**

A THESIS SUBMITTED

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

**Doctor of Philosophy**

APPROVED, THESIS COMMITTEE:

Philip Kortum, Professor in the Practice, Chair
Psychology

Michael Byrne, Associate Professor,
Psychology

David M. Lane, Associate Professor,
Psychology, Statistics, and Management

Lin Zhong, Assistant Professor,
Electrical and Computer Engineering

HOUSTON, TEXAS

MAY 2011

# ABSTRACT

Why Does Content Desirability Impact Subjective Video Quality Ratings and What Can Be Done About It?

by

Rochelle E. Evans

This study attempted to determine why users who like a product (in this case, film clips) rate the product higher in quality (in this case, video quality). In this study, film clips were located that were high or neutral in enjoyment. Experiment 1 determined that participants liked the enjoyable film clips more than the neutral clips. In Experiment 2, liking, affect, and content immersion were positively correlated with video quality. Additionally, liking partially mediated the relationship between affect and video quality. A halo effect was also found whereby all items assessing each film clip were rated highly due to heuristic reliance on affect. In Experiment 3, training participants on video quality moderated the relationship between affect and video quality; therefore, training was able to remove the halo effect. Training also increased participants' accuracy at rating video quality. Experiment 4 demonstrated that participants' video quality ratings could also be improved by instructing participants to focus on quality. Focusing on quality did not moderate the relationship between content immersion and video quality, as participants maintained high levels of content immersion even when focusing on video quality. All experiments demonstrated that enjoyable clips were rated higher in quality than neutral clips. Both training and focusing attention were helpful in increasing the accuracy of subjective video quality ratings; therefore, both approaches could be utilized in usability testing to improve the quality of feedback received.

TABLE OF CONTENTS

List of Tables.......... vii

List of Figures.......... viii

List of Appendices.......... ix

Introduction.......... 1

Video Quality.......... 3

Content.......... 4

The Current Study.......... 5

Media Enjoyment.......... 6

Affect.......... 8

Reliance on Affect.......... 9

Heuristic Assessment.......... 9

Halo Effect.......... 10

No Reliance on Affect.......... 11

Immersion.......... 12

Attention and Immersion.......... 13

Inattentional Blindness.......... 15

Affect and Immersion.......... 17

Experiment 1.......... 17

Method.......... 18

Participants.......... 18

Design.......... 18

Materials.......... 18

Procedure.......... 20

Results.......... 21

Experiment 2.......... 21

Method.......... 22

Participants.......... 22

Design.......... 22

Materials........ 23
Procedure........ 24
Results........ 24
Analysis........ 25
Familiarity with Film Clips........ 25
Gender........ 26
Liking........ 26
Affect........ 27
Immersion........ 31
Media Enjoyment........ 33
HLM........ 37
Discussion........ 37
Experiment 3........ 39
Method........ 39
Participants........ 39
Design........ 40
Materials........ 40
Procedure........ 41
Results........ 43
Manipulation Check........ 43
Affect........ 44
Quality........ 49
HLM........ 52
Discussion........ 53
Experiment 4........ 55
Method........ 56
Participants........ 56
Design........ 56
Materials........ 56

Procedure........................................................................................ 57
Results.............................................................................................. 58
Manipulation Check.................................................................... 58
Immersion........................................................................................ 59
Quality............................................................................................ 59
HLM................................................................................................ 63
Discussion........................................................................................ 64
General Discussion................................................................................ 65
Experiment 1................................................................................... 65
Experiment 2................................................................................... 66
Experiment 3................................................................................... 66
Experiment 4................................................................................... 69
Video Quality Discrimination.......................................................... 72
Limitations and Future Directions.................................................. 73
Conclusions..................................................................................... 76
References............................................................................................ 78

LIST OF TABLES

Table 1: Difference in ratings for film clips not seen before/seen before .................. 26

Table 2: Dependent variable ratings by gender.......................................... 26

Table 3: Intercorrelation matrix for Experiment 2...................................................... 30

Table 4: Exploratory factor analysis ............................................................... 31

Table 5: Full model from HLM for Experiment 2...................................................... 37

Table 6: Difference in ratings for film clips not seen before/seen before..................... 43

Table 7: Intercorrelation matrix for untrained participants in Experiment 3................ 46

Table 8: Intercorrelation matrix for trained participants in Experiment 3.................... 47

Table 9: Exploratory factor analysis for Experiment 3................................................ 48

Table 10: ANOVA for fixed effects when film clips and participants are random....... 52

Table 11: Full model from HLM for Experiment 3..................................................... 53

Table 12: Difference in ratings for film clips not seen before/seen before................... 58

Table 13: ANOVA for fixed effects when film clips and participants are random...... 63

Table 14: Full model from HLM for Experiment 4..................................................... 63

LIST OF FIGURES

Figure 1: Model........................................................................................................ 12

Figure 2: Correlation between liking and quality (MOS) in Experiment 2.................. 27

Figure 3: Correlation between affect and quality (MOS) in Experiment 2................... 28

Figure 4: Testing liking as a mediator between affect and quality (MOS)................... 29

Figure 5: Correlation between immersion and quality (MOS) in Experiment 2........... 32

Figure 6: Testing liking as a mediator between immersion and quality (MOS)........... 33

Figure 7: Rating of liking for each film clip in Experiment 2..................................... 34

Figure 8a: Quality (MOS) rating for both film clip types (neutral and enjoyable)........ 35

Figure 8b: Liking rating for both film clip types (neutral and enjoyable).................... 35

Figure 8c: Affect rating for both film clip types (neutral and enjoyable)..................... 36

Figure 8d: Immersion rating for both film clip types (neutral and enjoyable).............. 36

Figure 9: Interaction of training between affect and quality (MOS)............................ 45

Figure 10: Quality x condition interaction.................................................................. 50

Figure 11: Main effect of enjoyment.......................................................................... 51

Figure 12: Quality x enjoyment interaction................................................................ 52

Figure 13: Lack of interaction for direction of focus between immersion and quality (MOS)......................................................................................................... 60

Figure 14: Quality x condition interaction.................................................................. 61

Figure 15: Quality x enjoyment interaction................................................................ 62

# LIST OF APPENDICES


Appendix A: Film Clips Utilized in Experiment 1 ..................... 92

Appendix B: Experiment 1 Film Clip Questionnaire ..................... 93

Appendix C: Experiment 1 & 2 End of Experiment Questionnaire ..................... 94

Appendix D: Affect Grid ..................... 95

Appendix E: Immersion Scale ..................... 96

Appendix F: Experiment 2 Film Clip Questionnaire ..................... 97

Appendix G: Video Quality Assessment Overview ..................... 99

Appendix H: Experiment 3 Film Clip Questionnaire ..................... 102

Appendix I: Experiment 3 End of Experiment Questionnaire ..................... 104

Appendix J: Experiment 4 Film Clip Questionnaire ..................... 106

Appendix K: Experiment 4 End of Experiment Questionnaire ..................... 108

Oxenfeldt, as "all attributes of a product which yield consumer satisfaction" (p. 300). A positive correlation between quality and satisfaction exists for a variety of products and experiences, from mobile phones (Zhang, Rau, & Salvendy, 2010), food (Ha & Jang, 2010; Verbeke, Wezemael, Barcellos, Kugler, Hocquette, et al., 2010), vehicle navigation systems (Lin & Chien, 2010), athletic shoes (Tsiotsou, 2006), patient care (Pascoe, 1984), and video quality (Kortum & Sullivan, 2010). The subjective nature of satisfaction lends to a variety of antecedents influencing how users rate a product on satisfaction. For example, Walker and Dubitsky (1994) evaluated a commercial, measuring satisfaction (how much did users like the commercial?). When asked why users were satisfied with, or liked, the commercial, many aspects of the commercial were referenced, such as its creativity, liveliness, cuteness, realism, and how convincing it was.

One powerful antecedent to both satisfaction and quality is affect. In regards to satisfaction, users assess their mood state when making satisfaction ratings. Specifically, the satisfaction response can bear emotional foundations. For example, 64% of users asked to describe their satisfaction with a product in an interview setting were very likely to use affective terminology such as "like love", "excited", and "pleasantly surprised" in lieu of the word "satisfied" (Giese & Cote, 2002). In a study of users who had recently purchased a new car, both emotion and satisfaction was measured (Westbrook & Oliver, 1991). The positive dimensions of affect measured (happiness/contentment; delight) were positively correlated with satisfaction. The negative dimensions (angry/upset; unpleasant surprise) were more likely to display a lack of dissatisfaction, as scores fell around the midpoint (users did not demonstrate low levels of satisfaction despite rating low emotions). Compeau, Grewal, and Monroe (1998) explained that users retrieve decreasing buffering time. As bandwidth necessary to accommodate analog signals is only in the range of a few megahertz (Mhz), the difference between analog and the pure digital signal is very large. Although HDTV signals are encoded at different bit rates than digital broadcasting (Riley & Richardson, 1997), many HDTV owners were not sure if they did or did not receive HD programming (Oxman & Pollack, 2007). Therefore, users are not necessarily sensitive to decrements in bit rate. Subjective measurements of video quality can assess whether or not users notice degradation. Users are ideal for video quality testing, as objective measurements cannot perfectly model human assessment (Corriveau, 2006). For example, users examined multiple film clips across five different encodings (Kortum and Sullivan, 2004). There was a linear trend from the lowest encoding, 490 kilobits per second (kbps), which was deemed “somewhat good” to the highest encoding, DVD quality, which fell between a “good” and “very good” rating. However, VHS quality ratings were not significantly different from DVD quality ratings. While users noticed quality differences at low bit rates, users were not sensitive to quality differences at slightly higher rates, indicating that not all bit rate reductions are noticeable.

*Content*

Many factors during the TV-viewing experience may affect subjective ratings of video quality. For example, increased attention to the auditory modality decreased the degree to which users noticed degradation in video quality (Rimmell, Mansfield, & Hands, 2008). A decrease in audio quality can also reduce video quality ratings (Hands, 2004). Additionally, content can impact video quality requirements. Video redundancy varies based on the video’s content, where an increase in movement is tied to less redundancy (Todorivic, 2006). A very static shot can contain over 90% redundancy. Corriveau (2006) noted that if there is little movement in the standard test sequence being used, users may not notice the major artifacts that occur with a drop in bit rate. Content does not only affect standard test sequences, but film clips from popular media as well. On mobile devices, users preferred an increase in video quality for high motion content (such as ice hockey and TV series), and required less video quality and more audio quality for content such as music videos and news (Jumisco-Pyykko & Hakkinen, 2005). However, a user's bias can also enter this equation, making it more complicated to accurately assess video quality. Kortum and Sullivan (2004) found that users who disliked a film clip rated its quality as "somewhat good" whereas those who liked a film clip rated its quality between "good" and "very good". This same trend was found across encoding levels, such that film clips were rated lower in video quality when users did not want to see content, and rated higher when users did want to see the content, this pattern existing across all encoding levels from 550 kbps to DVD quality (Kortum & Sullivan, 2010).

## The Current Study

The current study aimed to understand why content affects video quality ratings. Two explanations were explored. First, viewers may be rating certain clips higher in video quality because the content of the clips increases their affective state. This explanation could have a secondary impact such that individuals in a good mood who lack a well-formed mental model of video quality assessment may succumb to a halo effect. In this scenario, video quality would be rated highly because other items being evaluated were also given a high rating. Second, users may attend more to the content of film clips they like, indicating that they have less attentional resources that can be allocated to the detection of image quality.

Participants viewed film clips that differently affected enjoyment. It was believed that enjoyable film clips would increase mood, immersion, and film clip liking, whereas neutral film clips would decrease ratings of mood, immersion, and film clip liking. By locating the source that contributed to the inaccuracy in video quality ratings and removing rating inaccuracies in an experimental setting, it was hoped that this source could be removed or controlled for in future studies in the field where accurate subjective reports of video quality are paramount.

## Media Enjoyment

Similar to satisfaction, media enjoyment is highly subjective. Users will not like every movie, television show, or video game they are exposed to. However, there are key features that are correlated to the media enjoyment experience. Nabi and Krcmar (2004) indicated that a component of media enjoyment is affective, where positive and negative mood play a role in determining a user's level of enjoyment. This is related to the aforementioned correlations between satisfaction and emotion. A second behavioral component was composed of immersion, or attention, such as leaning forward and engaging in the film versus looking at a clock frequently. This attentional component was not discussed in the previous research as it is unique to the media enjoyment experience as defined in this paper.

Supporting the affective side, Zillmann (1988) indicated that enjoyable stimuli derived from media entertainment can alter moods. In fact, the aim of media entertainment is to alter a user's mood (Zillmann, 2003). Good moods can be maintained or enhanced, whereas bad moods can be altered depending on the entertainment selected. In order for a drama to create enjoyment, the creation of both positive and negative affect towards media characters was necessary (Zillmann & Bryant, 1994). By empathizing with characters, characters became liked or disliked, creating anticipation for certain outcomes to occur (Vorderer, Klimmt, & Ritterfeld, 2004). Beyond characters, users reacted to scenes or experiences relatable to personal experiences, which induced a variety of emotions such as happiness, sadness, fear, or embarrassment (Vorderer, 2001). Specifically, media entertainment required emotional enjoyment. Users listening to music preferred songs which they felt stronger positive emotions for, and did not prefer music where strong negative emotions were present (Schubert, 2010). Viewers watching Super Bowl XL reported higher levels of entertainment when their level of affect averaged over the course of the entire game was positive–average negative affect (walking away from the game in a negative mood) was not related to entertainment (David, Horton, & German, 2008). While situational positive (i.e. a touchdown for one's team) or negative (i.e. a fumble for one's team) affect induced entertainment, a viewer's mean level of affect gave a more accurate report of overall entertainment. Users reported that movies can be enjoyed when they contain fun, suspense, a lasting impression, or a moving/thought provoking experience (Oliver & Bartsch, 2010). These concepts encompass a variety of genres from comedies to actions to dramas, and emotions conveyed while watching movies can indeed vary. The above research is in agreement with Eliashberg and Sawhney (1994), whose model for movie enjoyment included users' temporary moods, variability of these moods, the movie's emotional content, and individual differences related to sensation-seeking behavior. Indeed, affect plays a crucial role in media enjoyment, and enjoyment can be experienced independent of genre.

Supporting the behavioral side, a film must maintain the audience's attention in order to create media enjoyment (Anderson & Kirkorian, 2006). Specifically, Green et al. (2004) indicated that a visually-rich story was necessary during television or film viewing in order to capture attention. Documentaries or news reports also induced immersion if a narrative was present. Across three movies, presence, flow, and transportation (concepts related to immersion) were positively correlated with each other and with media enjoyment (Wissmath, Wiebel, & Groner, 2009). An increase in attention to a soap opera was positively correlated with program satisfaction (Kim & Rubin, 1997). This concept of attention, or immersion, was highest surrounding a story's climax and resolution (Bryant & Comisky, 1978). In a study examining perceptions of multimedia on a computer, immersion and entertainment were very highly intercorrelated (Karat et al., 2002). Additionally, users who reported higher mobile service quality had higher reported levels of playfulness, a measure synonymous with immersion (Tan & Chou, 2008). Immersion does not only affect enjoyment of the film, itself. In fact, it can induce a higher evaluation of main characters (Green & Brock, 2000) or advertisements viewed at the end of a story (Wang & Calder, 2006) or increase the desire to use or purchase a product (Jung, Perez-Mira, & Wiley-Patton, 2009).

## Affect

A number of stimuli can alter one's emotional state. This can take place through assessment of a personal experience or event that may induce a variety of emotions when judging an item separate from that which generated the affect. Schwarz and Clore's (1988) affect-as-information model and Forgas's (1995) affect infusion model indicated that users rely on affect when processing judgments heuristically but would not when evaluations are clear. Forgas further stated that heuristic processing utilized affect when the user had no previous experience or knowledge of the object being judged. Additionally, the user may be cognitively overloaded or under time pressure. For example, individuals whose mood was increased because their team just won a soccer match gave a better evaluation regarding satisfaction with their global well being, and a lower evaluation when their team tied (Schwarz, Strack, Kommer, & Wagner, 1987). This evaluation relied on the heuristic of affect as it was considered an abstract question that was hard to evaluate. Verifying mood as a heuristic, faster evaluations are made when the user relies on affect to judge an item. Pham, Cohen, Pracejus, and Hughes (2001) asked half of their participants to evaluate their feelings regarding a variety of pictures. The other half were asked to assess pictures without reliance on feelings. Evaluations regarding feelings took significantly less time than assessment evaluations. Yeung and Wyer (2004) showed participants pictures of shoes either before or after a positive or negative mood was induced. Mood affected evaluation when it was induced before being shown the products, such that individuals in a better mood liked the shoes more than those in a negative mood. There was no difference in evaluation when mood was induced after being shown the pictures, as mood was not attributed to the judgment of the shoes.

*Halo effect.*

One consequence that can occur as a result of heuristic processing is the halo effect. The halo effect indicates a tendency for a user to consider the item or person being rated as “all good”, resulting in a high correlation between responses. Heuristic processing may arise due to positive affect or lack of knowledge about the item or person being evaluated. For example, when appraising teachers’ behaviors on items that do not guarantee covariance such as giving lectures and friendliness, halo was greater for those in a positive mood compared to those in a negative mood (Sinclair, 1988). On the other hand, raters who lacked job knowledge for baseball players rated all items similarly, whereas those familiar with the sport and the players being rated were more accurate with their ratings (Kozlowski, Kirsch, & Chao, 1986).

*No Reliance on Affect*

In order to eliminate dependence on affect when making judgments, a direct access strategy can be utilized (Forgas, 1995). Direct access processing relies on direct retrieval of evaluation for a well-known target. For example, following a mood induction, managers familiar with providing feedback were equally polite when giving feedback whereas employees untrained in giving feedback were more polite in a positive mood and less polite in a negative mood (Forgas & Tehani, 2005). Therefore, a video technician who has existing schemas for video quality could make a quick, yet accurate, evaluation compared to a novice who would use a heuristic approach due to a lack of knowledge. Based on a proposed causal model (see Figure 1), the following questions were set forth:

1) Was there a positive correlation between affect and video quality ratings?

2) Did a halo effect occur when users in a good mood lacked knowledge about video quality?

3) Did knowledge about video quality assessment moderate the relationship between mood and video quality ratings?

4) Did media enjoyment mediate the relationship between affect and video quality ratings?

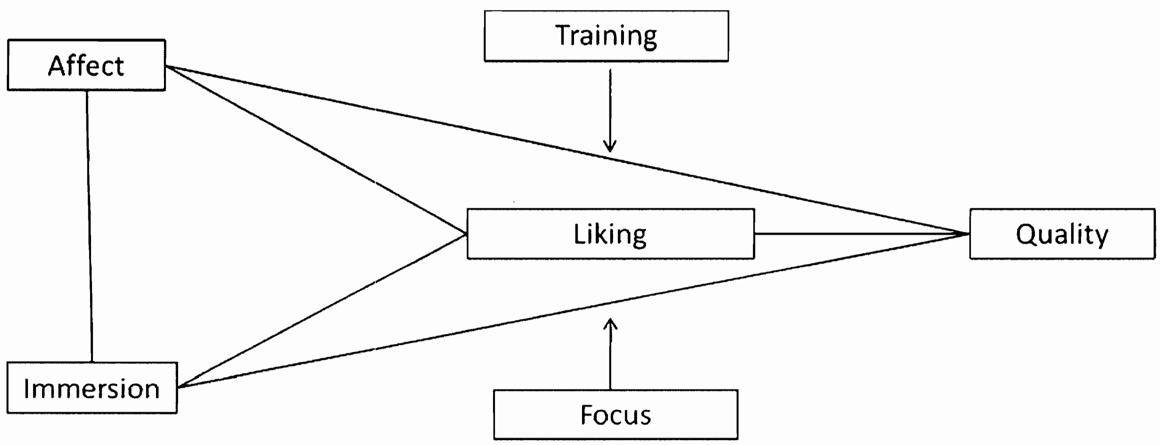



*Figure 1*. Model.

## Immersion

A flow state is a subjective experience derived from high levels of task involvement to the point where time and distractions are ignored (Csikszentmihalyi & Rathunde, 1992). However, flow requires high levels of skill and challenge, which are not necessarily present when watching television or a movie. In the absence of skill and challenge, the flow state, where a sense of time is distorted and concentration is maximized, can still be captured via synonymous concepts such as engagement, absorption, immersion, transportation, and presence (Green, Brock, & Kaufman, 2004;

Wiles & Danielova, 2009; Brown & Cairns, 2004; Csikszentmihalyi, 1990; Qin, Rau, & Salvendy, 2009). The term *immersion* will be utilized throughout the continuation of this thesis. Kortum and Sullivan (2010) found anecdotal evidence implicating immersion as a factor that may impact subjective video quality ratings when users observe film clips. It is indeed possible that greater levels of immersion in the film narrative may decrease awareness of other aspects of the film such as the video quality. Weber, Tamborini, Westcott-Baker, and Kantor (2009) described immersion as a state where all attention was focused on an activity and there was no remaining attention for another activity.

### *Attention and Immersion*

Attention is a key feature of immersion, as users must maintain attention on a medium in order to become immersed in that target. However, users cannot simultaneously attend to two tasks that interfere with each other. In the case of visual media, there are two pieces of information that are competing for attention. There is the content of the media, itself, and there is the quality of that media. An individual who wants to watch a movie for its content and nothing else would utilize selective attention, opting to focus on the content and ignoring all other information such as video quality. Broadbent (1958) would equate one's attention as a filter – only content would be filtered and therefore processed. Rock and Gutman (1981) found that individuals were unable to process unattended information. They had participants attend to one of two overlapping figures, one red and one green, and participants were informed that only one shape had to be rated on how pleasing it was. They were much more likely to recognize shapes that they increased attention to due to being told they would have to rate them. A more recent example demonstrating this involves users who listened to a short story while looking at advertisements (Shapiro & Krishnan, 2001). As users paid more attention to the short story, their recognition of advertisements viewed decreased. Additionally, users who talked on a telephone or generated new words verbally were slower and less accurate at tracking disks on a screen (Kunar, Carter, Cohen, & Horowitz, 2008). When a user tries to perform two tasks, performance on one or both tasks will be hindered, and performance particularly suffers when two tasks utilize the same modality (Pashler, 1994). Attending away from a film would decrease immersion in the film, as users are focusing their visual attention on another task and away from the film content. In a reading study, users asked to identify words that were hard to understand in a short story became less immersed in the story than users who were told to pay attention or act as if they were in the story (Green & Brock, 2000). Some of the users whose task involved word identification admitted that they failed this task when they started to focus on the story's plot, indicating an inability to focus on both the plot and word identification. Wang and Calder (2006) noted that an advertisement that appeared partway through a story in a magazine was considered more intrusive for individuals highly immersed, as it interrupted the story. Ad intrusiveness did not differ between immersed and non-immersed individuals when the ad was shown at the end, as it did not distract readers in the middle of a story.

In regards to this study, attention is operationally defined as focused attention, whereby attention is like a spotlight: Based on this spotlight theory (i.e, Posner, Snyder, & Davidson, 1980; LaBerge, 1983), the item or area within one's focus is attended to, and that outside of one's focus is not. Conceptualizing attention in this manner allows for cleaner understanding of the tasks utilized in this experiment. For this study, it was important to comprehend the extent to which participants could focus on one task (and in turn, ignore other tasks). Using a spotlight theory, one can say that participants focusing on video quality were directing their attention solely on quality, and those focusing on content where directing their attention solely on content. When asked to make judgments regarding quality, participants focusing on content would be highly inaccurate because this aspect was being ignored in order to focus on content; however, their recall of content from film clips would be very good because attention was directed on just content over the duration of the film clips. Were a divided attention theory utilized for this paper, it would indicate that participants who are focusing on quality could be simultaneously utilizing resources to focus on content. One could no longer state that a participant is focusing on content or quality; but would in turn have to look at the degree to which one task was being performed. This type of measurement was not implemented; however, a spotlight theory allowed for all-or-none measurement.

*Inattentional Blindness*

With a lapse of attention, details in a scene may go unnoticed. This concept of inattentional blindness occurs when users' attention is not equally dispersed over the visual field, but selectively focused on some information, thus dismissing other information. Increased attention toward the foreground may lead objects in the background to become overlooked. Attention paid to particular characters may make it possible for other characters or objects to go unnoticed or to change without detection (change blindness). Simons (1996) noted that changes to objects in a scene were not perceived during the dynamic event of a video. The author suggested that a participant does not form a fully accurate memory for a scene; therefore, changes may go undetected. Wiles and Danielova (2009) investigated the success of product placement in 31 popular movies from 2002. Returns from advertisement placement were decreased as audience immersion increased, indicating that less attention was paid to advertisements when users were focused on the narrative. Therefore, in order to increase attention to aspects of the film unrelated to the plot or narrative, users' attention must be intentionally redirected. For example, Levin and Simons (1997) instructed users to watch a film or to watch the film with the purpose of detecting changes. Half of the films contained gross changes, such as the main actor changing after a cut to a new scene. Users who were not instructed to detect changes had a 33% detection rate, whereas users attending to changes made less than one error each. Cater, Chalmers, and Ward (2003) were able to reduce the resolution of objects in the background of an animation without affecting users' ratings of overall quality of the animation. Because their attention was focused on the task of counting teapots, they were not aware of the decrease of graphical resolution in the background. Eye tracking supported the findings that users focused attention on the teapots and did not attend to the objects in the background where resolution was altered. Similarly, Sunstedt, Debattista, Longhurst, Chalmers, and Troscianko (2005) instructed users to count fire extinguishers in the foreground of a 17-second animation of an office building or view the animation without any given goal. Users who counted fire extinguishers failed to notice resolution degradation in the background. Interestingly, users who were viewing the animation without the goal of counting extinguishers also reported equal resolution between all animations, regardless of the true resolution displayed. Eye tracking revealed that the saliency of the fire extinguishers led users to focus on these objects although the task was not explicitly assigned. Therefore, users who are watching a movie may focus their attention on objects that a director wishes the public to focus on, to the point that a change in unattended objects or a change in resolution may go unnoticed.

*Affect and Immersion*

Affect may not only be necessary for media enjoyment, but it may also be considered a factor of presence, or immersion (Klimmt & Vorderer, 2003). Csikszentmihalyi (1997) believed that an assessment of happiness could appear after the flow state, as it would take away from flow to assess mood during a task. Following a box-clicking game, users in an experiment by Jennett et al. (2008) reported levels of affect and immersion. Positive affect and immersion were highly correlated with each other, where $\rho = 0.711$, $p < .001$. Becoming immersed in a movie can alter moods; for example, angry users became less angry the more they were immersed in what they were watching (Bryant & Zillmann, 1977). Referring again to the model in Figure 1 on page 12, the following questions were asked:

5) Was there a positive correlation between immersion and video quality ratings?

6) Did media enjoyment mediate the effect between immersion and video quality ratings?

7) Did focusing the user's attention moderate the relationship between immersion and video quality ratings?

## Experiment 1

information from these precursor affective responses. This information then influences overall product evaluation.

*Video Quality*

One area where assessment of quality is paramount is in the arena of video technology, as a high-quality visual display is of utmost importance to customer satisfaction. Video quality has been continually increasing to meet consumer demand. Televisions are increasing in resolution to the point where high-definition television (HDTV) is becoming the norm when purchasing a television. A Consumer Electronics Association (CEA) press release from 2007 indicated that 30% of homes in the United States contained a HDTV (Oxman & Pollack, 2007). The reason customers wanted these televisions was to increase movie and gaming experiences. Bracken (2005) supported this notion of increased media experience via immersion, where users become highly engrossed in what they are watching. Viewers watching a 35-minute film clip of natural scenery in HDTV reported higher levels of immersion than viewers watching the same clip in analog using the National Television System Committee (NTSC) standard. Additionally, higher image quality increased audience ratings of news credibility in addition to immersion (Bracken, 2006).

However, there is a high cost attached to sending a high-definition signal to consumers, and manufacturers can compress signals to the extent that perceived quality ratings are not reduced. Digital compression reduces the bit rate by removing redundancy in the video or audio signal. There are advantages to reducing bit rate, such as minimizing the amount of bandwidth utilized, maximizing space on a DVR, and

The goal of Experiment 1 was to select two types of film clips – those that elicit high levels of enjoyment and those that elicit low levels of enjoyment. By locating clips that induced predetermined likeability across all participants on average, the subsequent experiments carried out in the dissertation could have an independent variable of film clip enjoyment.

*Method*

*Participants.*

Twenty-six Rice University undergraduates with normal or corrected-to-normal vision and normal color vision were recruited to participate in Experiment 1. There were 24 females and 2 males with a mean age of 19.85 years (*SD* = 0.97). They participated for psychology experiment credit.

*Design.*

One within-subjects variable was used. This variable was the media enjoyment factor of the film clips. All participants viewed every film clip which varied in media enjoyment. The dependent variable was likeability. Participants rated the degree to which they liked a film clip on a 7-point Likert scale from 1 (not at all) to 7 (very much).

*Materials.*

Snellen eye chart. Participants read an eye chart from a distance of 20 feet, wearing corrective lenses if necessary, and required a passing vision score of 20/40.

Ishihara pseudoisochromatic plates. Participants viewed the Ishihara plates to test for normal color vision.

Film clips. Thirty-four 2-minute clips from motion pictures were tested in Experiment 1. The clips were selected by the experimenter to be either highly enjoyable or more neutral in content. The film clips were all released between 1994 and 2002. By selecting movies that were several years old, it was hoped that more of the participants (university undergraduates born around 1990) would not have seen the movies, however the movies were not so old that the film's quality would be poorer than current releases. The film clips were converted to 1100 kbps from original DVD quality (see Appendix A). While there are many ways to measure video quality, bit rate was chosen over others such as peak signal to noise ratio (PSNR; Huynh-Thu & Ghanbari, 2008) or video quality measure (VQM; Xiao, 2000), the choice to convert all clips to 1100 kbps was made based on Kortum and Sullivan (2010). While many bit rates were investigated, 1100 kbps demonstrated the cleanest monotonic relationship between liking of content and video quality ratings. 1100 kbps fell in the middle of the spectrum of bit rates investigated; therefore, a floor or ceiling effect was not expected due to video quality.

Film clips were placed into four separate Windows Media Player playlists in four random orders. An 8-second slide broke up each film clip that prompted the participant to pause the media player and ask the experimenter for the next questionnaire. Instruction slides were also inserted at the very beginning and end. The first slide had very basic instructions for the participant; the last slide instructed the participant to fill out the final questionnaire and that he/she would then be finished. These slides were created in Windows Movie Maker.

Experiment 1 film clip questionnaire. A form was created in which participants could determine how much they liked the film. In addition, the participant answered questions related to whether or not they had seen the film before in order to control for familiarity (see Appendix B).

End of experiment questionnaire. This form was used to collect demographic information and ask about the types of films the participant typically enjoyed (see Appendix C).

*Procedure.*

A playlist with 20 film clips was launched before a participant entered and paused so that the welcome screen was sitting on the monitor upon a participant's arrival. A participant entered the laboratory, signed the consent form, and then received experimental instructions. Before viewing any film clips, the participant was given a Snellen eye test and Ishihara test to determine if he/she had normal acuity and color vision. No participants failed either test. Following these screening tests, the participant sat in front of a computer monitor. The participant viewed all film clips sequentially. After viewing each clip, the participant filled out the Experiment 1 film clip questionnaire. When the rating sheets were completed, the experimenter started the next two-minute film clip, and this procedure continued until all twenty clips were presented. At the end of the experiment, the participant filled out the end of experiment questionnaire. He/she was then debriefed and dismissed. The experiment took one hour. The procedure was repeated for all twenty participants.

After eight participants were run, eight film clips were bimodal such that their distributions were not heavily liked or disliked. These film clips were excluded as they were not demonstrating a positive or negative rating skew towards enjoyment or the lack thereof, which the other 12 film clips were showing. They were exchanged for new film clips, and new playlists were created. After six more participants, six more film clips were replaced; therefore, 34 film clips were examined in total. Some clips were viewed by all 26 participants and some were viewed by as few as 12 by the end of the experiment.

*Results*

The 20 final film clips from the third iteration were analyzed. The film clips' ratings were averaged to arrive at a mean likeability score per clip. The six clips with the highest ($M = 5.21$ on a 7 point scale, $SD = 0.83$) and lowest ($M = 3.44$, $SD = 0.81$) means were placed into two groups, and a paired samples *t*-test revealed that the two groups did differ in likeability, where $t(25) = 7.99$, $p < .001$. Of these 12 film clips, very few had been seen before by any of the participants. Of the 312 film clips viewed in total (12 film clips viewed x 26 participants), only 16 (5%) were recognized, therefore it was not believed that any of these film clips was highly familiar across all individuals. The film clips with the highest ratings were selected as enjoyable clips to be utilized in the subsequent experiments. These clips were taken from: Analyze This, Cruel Intentions, Gattaca, Ocean's Eleven, The Impostors, and The Siege. The six clips given ratings demonstrating neutral enjoyment were also selected for use; these clips were taken from Fargo, Girlfight, Guarding Tess, Murder in the First, The Deep End, and The Hudsucker Proxy.

## Experiment 2

This experiment was designed to examine the model in Figure 1 (see page 12). Specifically, it explored: 1) the relationship between affect and video quality ratings 2) the relationship between immersion and video quality ratings 3) whether likeability mediates the relationship between affect and video quality ratings 4) whether likeability mediates the relationship between immersion and video quality ratings, and 5) if an increase in affect led to a halo effect in participants' ratings. Additionally, it examined if the influence of media enjoyment impacted participants' ratings on any of the variables in the model.

### *Method*

#### *Participants.*

Thirty Rice University undergraduates with normal or corrected-to-normal vision and normal color vision participated in the experiment. None of them participated in Experiment 1. There were nineteen females and 11 males, with a mean age of 18.93 years ($SD$ = 0.91). All participants received psychology experiment credit.

#### *Design.*

A repeated measures design was utilized. There was a 2-level within-subjects independent variable of enjoyment, as determined by Experiment 1. Half of the 12 film clips were enjoyable and half were neutral. There were four key dependent variables: quality (measured with the mean opinion score or MOS), affect, immersion, and liking. Six more dependent variables were measured as well in order to assess presence of a halo effect.

*Materials.*

Some of the materials (Snellen eye chart, Ishihara plates, and end of experiment questionnaire) were identical to those used in Experiment 1.

Affect grid. Participants rated their level of mood and arousal on the affect grid at baseline and following each film clip (see Appendix D). The affect grid, introduced by Russell, Weiss, and Mendelsohn (1989), is a one-item measure designed to quickly assess state levels of mood and arousal. It correlated strongly with other affect scales on both the affect and arousal dimensions. The grid has been used previously to determine participants' emotions when watching films (Sato, Noguchi, & Yoshikawa, 2007).

Immersion questionnaire. Participants rated their level of immersion using the 6-item engagement (mental immersion) scale of the Temple Presence Inventory (TPI; see Appendix E). Items are rated on a 1 to 7 Likert-type scale. This questionnaire has a Cronbach's alpha of 0.90 (Lombard, Ditton, & Weinstein, 2009).

Experiment 2 film clip questionnaire. Participants were asked questions regarding film quality (using a 9-point mean opinion score, or MOS). This scale uses 9 anchors: *excellent*, *very good*, *good*, *somewhat good*, *fair*, *somewhat poor*, *poor*, *very poor*, and *bad.* Participants were also asked questions related to media enjoyment, genre assessment, and film recognition (see Appendix F). Other questions such as acting quality and MPAA rating were included as filler questions to distract participants from the key questions being asked. They were also used to assess the presence of a halo effect. This questionnaire was given to participants to fill out after each film clip and integrated the affect grid and immersion questions.

Film clips. The twelve 2-minute clips located in the pilot study were utilized. Six of these clips induced high levels of enjoyment and six, low levels of enjoyment.

*Procedure.*

A 19" standard-size (4:3) LCD Dell monitor was connected to a desktop computer in a dim room (brightness ranging from 20.2 Lux – 111.3 Lux) for the experiment. Four randomized playlists were configured in Windows Media Player, and a playlist was selected for viewing before the participant entered the laboratory. Upon arrival, a participant was asked to sign a consent form. His/her vision was then tested using the Snellen eye chart and Ishihara plates; no participants failed either test. The participant then filled out a baseline measure of affect using the affect grid. Instructions on the experiment were then administered. The participant was informed that a number of film clips were to be viewed, and after each clip a questionnaire was to be completed.

Two sample film clips were shown in order to accustom participants with the procedure. These clips were neutral in enjoyment – they were selected from the results in Experiment 1. The clips were from The Jackal and What Dreams May Come. After viewing the sample clips, the participants could have any additional questions about the procedure clarified, and then the viewing of the twelve test clips commenced. They were viewed sequentially, with the participant filling out the Experiment 2 film clip questionnaire after each clip. At the end of all the test clips, participants filled out the end of experiment questionnaire. Participants were then debriefed and dismissed. The experiment took one hour.

*Results*

*Analysis.*

A repeated measures ANOVA would be suitable for the majority of the analyses. However, aggregation across film clips could reduce error variance and inflate $R^2$ (Walker & Catrambone, 1993), therefore benefits to multilevel modeling exist such that the change in measurement over time within participants could be detected. HLM 6.08 allows for this type of repeated measures analysis, as it gives each participant his/her own repeated measures design. However, HLM is limited in its graphing capabilities and is limited in the output provided and the customizations available compared to other programs such as PASW Statistics 18 that do not offer multilevel modeling. Therefore, analyses were first performed using a repeated measures ANOVA in PASW Statistics 18 and results were supported with multilevel modeling using HLM 6.08. It would be expected that hierarchical linear modeling could give a more accurate assessment of effects that may be inflated using an ANOVA due to aggregation. All time-variant variables (affect, quality, immersion, and liking) were utilized in the Level-1 model in HLM, while the time-invariant variable of experimental condition was used in the Level-2 model.

*Familiarity with film clips.*

It was necessary to determine if there was an effect of having seen a film clip before on the key dependent variables (liking, affect, immersion, and quality (MOS)). Of the 360 film clips seen (30 participants x 12 film clips), 65 (18.06%) had been seen before. An independent-variables *t*-test exploring all 360 rows of data investigated if the dependent variables were rated differently by participants if the film clips had or had not

In order to replicate the findings of Kortum and Sullivan (2004; 2010), a positive correlation was expected between ratings on liking and the mean opinion score (MOS) across all film clips. This was found to be the case, as $r(28) = .68, p < .001$ (see Figure 2).

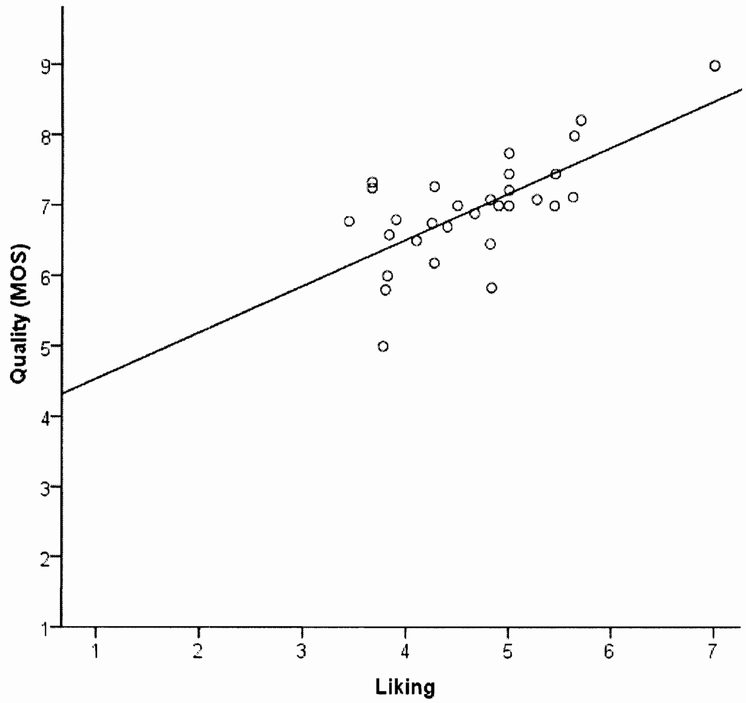



*Figure 2.* Correlation between liking and quality (MOS) in Experiment 2.

*Affect.*

The first relationship in the model predicted that participants in a good mood would give higher video quality ratings. The two dimensions of affect from the affect grid were combined into one score such that a high score of 10 would reveal positive mood and excitement and a low score of -10 would reveal negative mood and boredom. This final score was correlated with the MOS score. The correlation revealed that $r(28) =$

.60, $p < .001$, therefore the first portion of the model showing a direct relationship between affect and quality (MOS) was confirmed (see Figure 3).

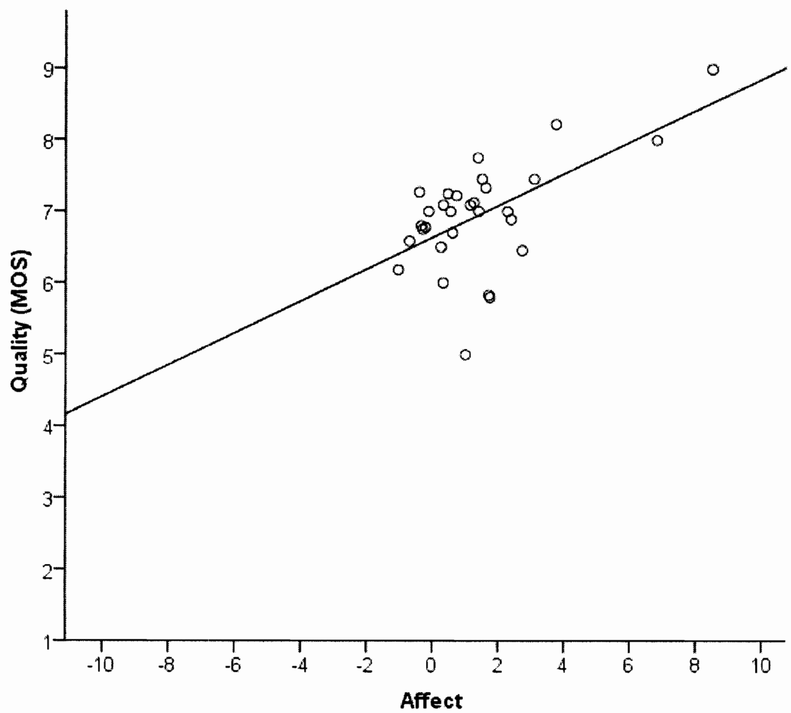



*Figure 3*. Correlation between affect and quality (MOS) in Experiment 2.

It was also examined if liking mediated the relationship between affect and quality ratings. In order to do so, liking was first regressed on affect (the predictor variable). Next, quality (the criterion variable) was regressed on affect, and finally, quality was regressed on both affect and liking. These steps were consistent with the three-step method outlined in Baron and Kenny (1986). The results depicted in Figure 4 indicated support for partial mediation. In the first stage (regressing liking on affect), a significant beta was found, where $\beta = .74, p < .001$. This was the case for the second stage as well where quality was regressed on affect ($\beta = .60, p < .001$). In the third stage, the beta weight for affect was no longer significant due to the presence of liking in the model (*β* = .20, *p* = .349). Following these regressions up with the Sobel test, the indirect effect of affect on quality (MOS) via liking was reliably greater than 0, $z = 2.37$, $p = .018$, indicating the presence of partial mediation for liking. Utilization of this meditational model implies causality via use of path analysis. The diagram in Figure 4 follows the assumption that the independent variable (affect) predicts the mediator (liking), and that both affect and liking predict the dependent variable (quality), as evaluated with the aforementioned regressions. This diagram is just a segment of that from the full model in Figure 1; if this model is incorrect, little meaning can be extracted from this study.

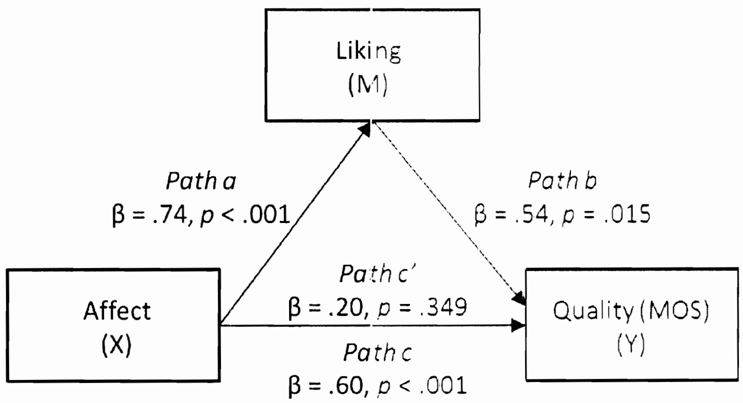



*Figure 4.* Testing liking as a mediator between affect and quality (MOS). *Note*. Path c = regression of criterion (Quality) on predictor (Affect). Path a = regression of mediator (Liking) on predictor. Path b = regression of criterion on mediator controlling for predictor. Path c' = regression of criterion on predictor controlling for mediator. Sobel test of mediation significance: $z = 2.37$, $p = .018$.

Because the first portion of the model was supported, a halo effect was sought such that high intercorrelations across all variables were expected for participants. This was performed using the 10 continuous variables from the Experiment 2 film clip questionnaire. Table 3 demonstrates the presence of positive correlations between all 10

**Table 3**: Intercorrelation matrix for Experiment 2

| | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. MOS | 6.95 | 0.77 | - | | | | | | | | |
| 2. Immersion | 4.52 | 0.77 | .73** | - | | | | | | | |
| 3. Liking | 4.66 | 0.80 | .68** | .84** | - | | | | | | |
| 4. Affect | 1.41 | 2.06 | .60** | .74** | .74** | - | | | | | |
| 5. Box Office Predictions[a] | 4.38 | 0.85 | .53* | .70** | .82** | .72** | - | | | | |
| 6. Resolution | 4.75 | 0.79 | .80** | .73** | .77** | .71** | .61** | - | | | |
| 7. Acting | 6.95 | 0.81 | .73** | .65** | .56** | .52** | .63** | .60** | - | | |
| 8. Lighting[a] | 5.07 | 0.66 | .60** | .56* | .69** | .59** | .73** | .62** | .60** | - | |
| 9. Curious | 4.80 | 0.79 | .72** | .80** | .80** | .56** | .61** | .64** | .53** | .53* | - |
| 10. Audio[a] | 7.04 | 0.79 | .75** | .78** | .74** | .59** | .44* | .67** | .77** | .55** | .69** |

items. Using 10 items from the film clip questionnaire, exploratory factor analysis was conducted using principal components. Promax roation was set at a Kappa of 4 due to the high correlations between items, and the factor analysis produced one factor, with 71.82% of the total variance explained (see Table 4). This indicated the presence of a halo effect, as all items were positively correlated and loaded onto the same factor. Participants did not distinguish items as separate factors, even though some items, such as immersion, predictions of how well the film did at the box office, and opinions of the cinematographer's choice for lighting had no reason to share a positive correlation.

**Table 4:** Exploratory factor analysis

| Item | Factor 1 |
|---|---|
| Liking | .933 |
| Immersion | .923 |
| Resolution | .872 |
| Video Quality (MOS) | .859 |
| Acting Quality | .848 |
| Curious What Happens Next | .830 |
| Audio Quality | .828 |
| Affect | .811 |
| Box Office Predictions | .798 |
| Cinematographer's Choice of Lighting | .757 |
| % of Variance | 71.82% |

*Immersion.*

The proposed model indicated that a positive correlation should be found between immersion and video quality ratings. The six items from the immersion questionnaire were highly reliable, ($\alpha = .961$). The correlation, $r(28) = .73, p < .001$, indicated increased attention was related to an increase in video quality ratings (see Figure 5).

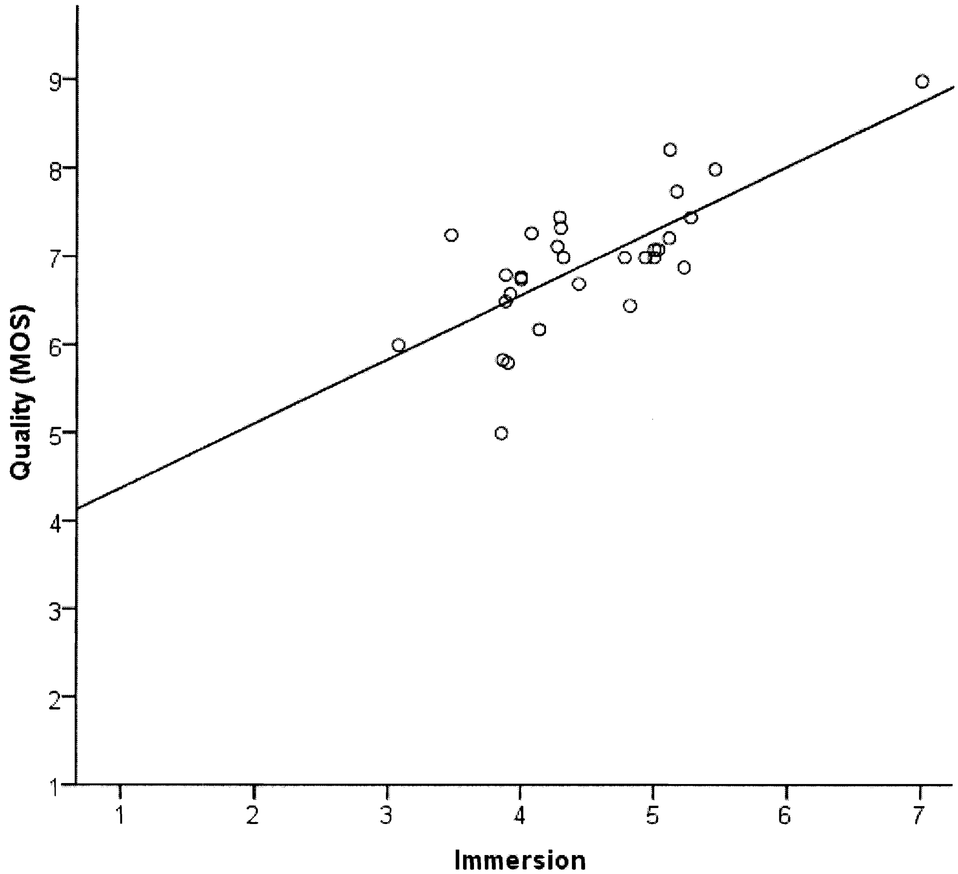



*Figure 5*. Correlation between immersion and quality (MOS) in Experiment 2.

Mediation was examined for this pathway as well – did liking mediate the relationship between immersion and quality ratings? Using the same three-step approach, liking was first regressed on immersion (β = .84, $p < .001$; see Figure 6). Next, quality was regressed on liking (β = .73, $p < .001$). Finally, quality was regressed on immersion and liking. The beta weight from immersion to quality only dropped a little, where β = .52, $p = .038$ when liking was controlled; but more importantly, the beta weight from liking to quality was not reliable when immersion was controlled, where β = .25, $p = .299$. Because of this, the Sobel test revealed that liking did not mediate the relationship between immersion and quality, $z = 1.05$, $p = .293$.

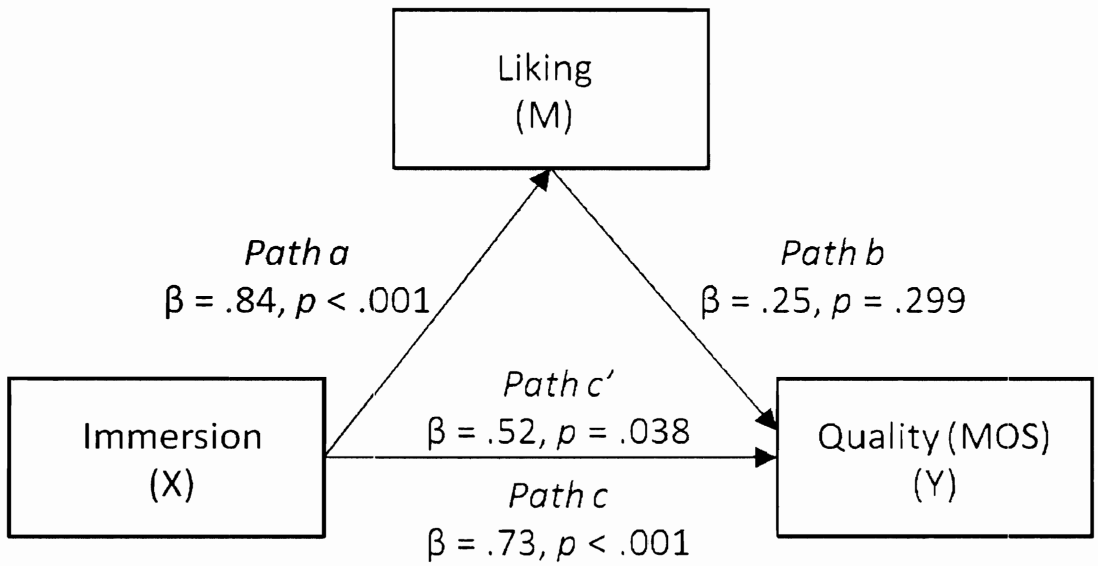



*Figure 6*. Testing liking as a mediator between immersion and quality (MOS). *Note*. Path c = regression of criterion (Quality) on predictor (Immersion). Path a = regression of mediator (Liking) on predictor. Path b = regression of criterion on mediator controlling for predictor. Path c' = regression of criterion on predictor controlling for mediator. Sobel test of mediation significance: $z = 1.05$, $p = .293$.

*Media Enjoyment*.

Figure 7 displays the liking scores for each film clip on a scale from 1 to 7. It was validated that the clips predetermined to elicit high or neutral enjoyment did so in this experiment. Similar to the pilot study, a paired-samples *t*-test was performed, which revealed that the six enjoyable clips ($M = 5.38$, $SD = 0.88$) were liked more than the six neutral clips ($M = 4.10$, $SD = 0.97$), where $t(29) = 7.78$, $p < .001$. The same analysis was performed to examine whether participants rated the enjoyable clips differently than the neutral clips on the other key dependent variables (affect, immersion, and quality (MOS)). Affect was rated from -10 to 10; immersion, from 1 to 7; and quality, from 1 to 9. For affect, participants were in a better mood when viewing the enjoyable clips ($M =$

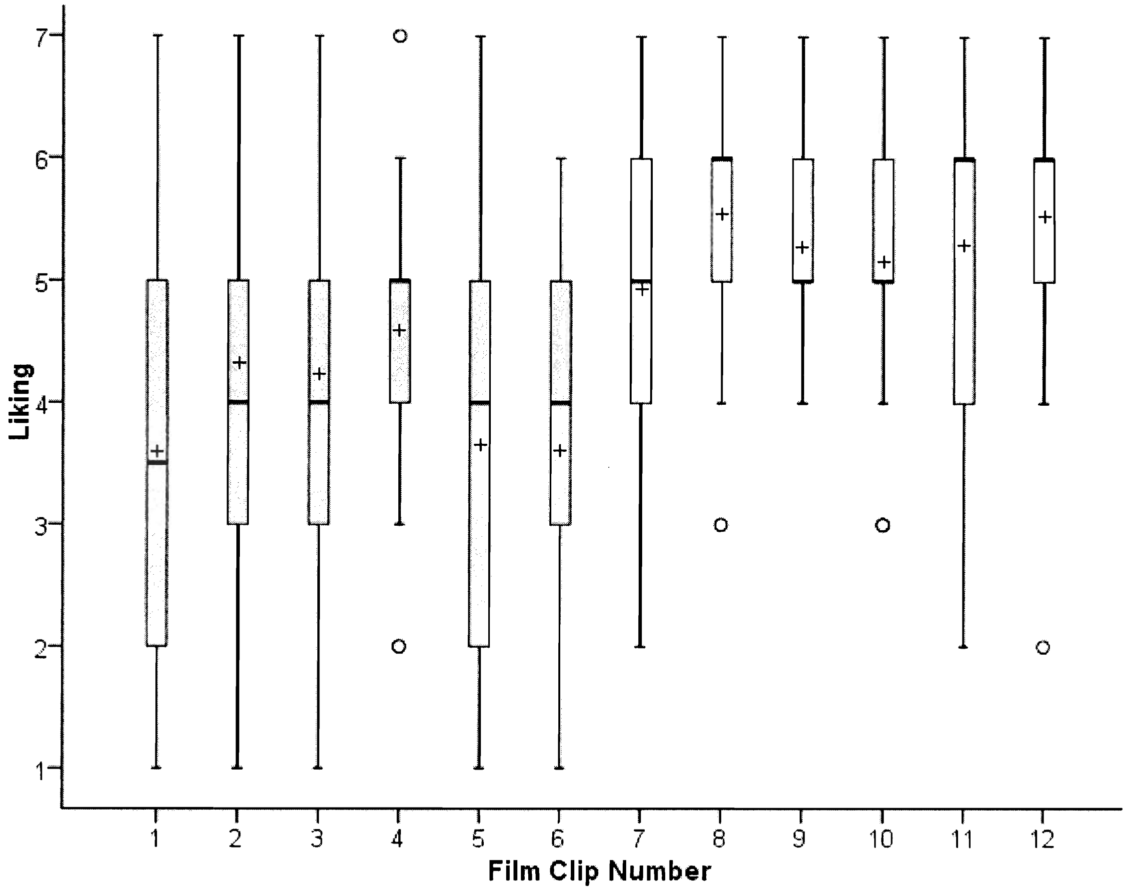



*Figure 7*. Rating of liking for each film clip in Experiment 2. *Note*. Film clips 1-6 were rated low in enjoyment in pilot study; clips 7-12 were rated high in enjoyment.

3.11, *SD* = 2.17) compared to the neutral clips (*M* = 0.13, *SD* = 2.42), $t(29) = 7.55$, $p < .001$. For immersion, participants were more immersed in content when viewing the enjoyable clips (*M* = 5.00, *SD* = 0.83) compared to the neutral clips (*M* = 4.12, *SD* = 0.95), $t(29) = 5.28$, $p < .001$. Finally, quality was rated higher for those viewing enjoyable clips (*M* = 7.28, *SD* = 0.82) than those viewing neutral clips (*M* = 6.69, *SD* = 0.87), where $t(29) = 4.70$, $p < .001$. Figures 8a – 8d show all four of these relationships between media enjoyment and the dependent variables.

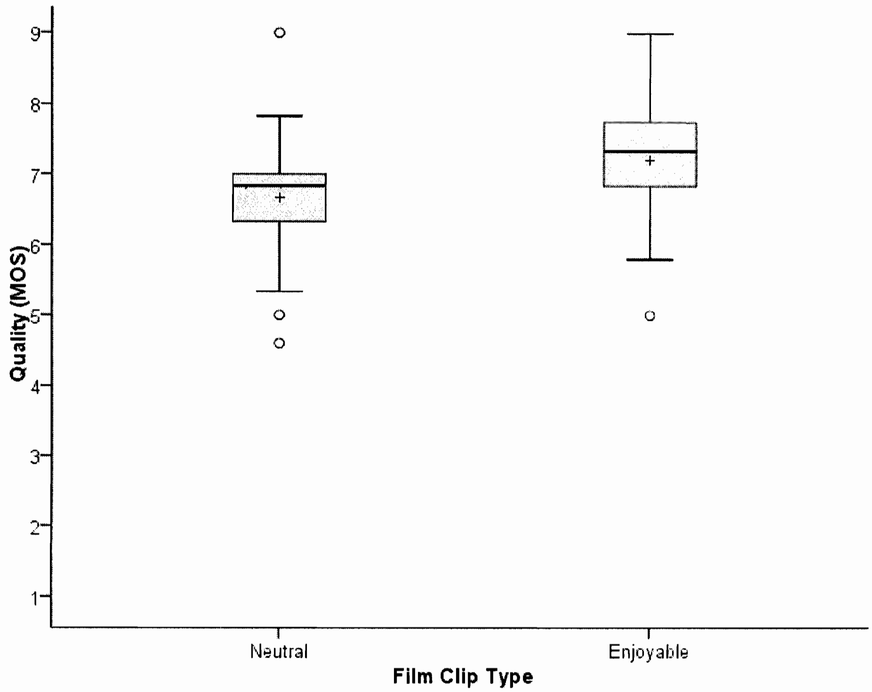



*Figure 8a*. Quality (MOS) rating for both film clip types (neutral and enjoyable).

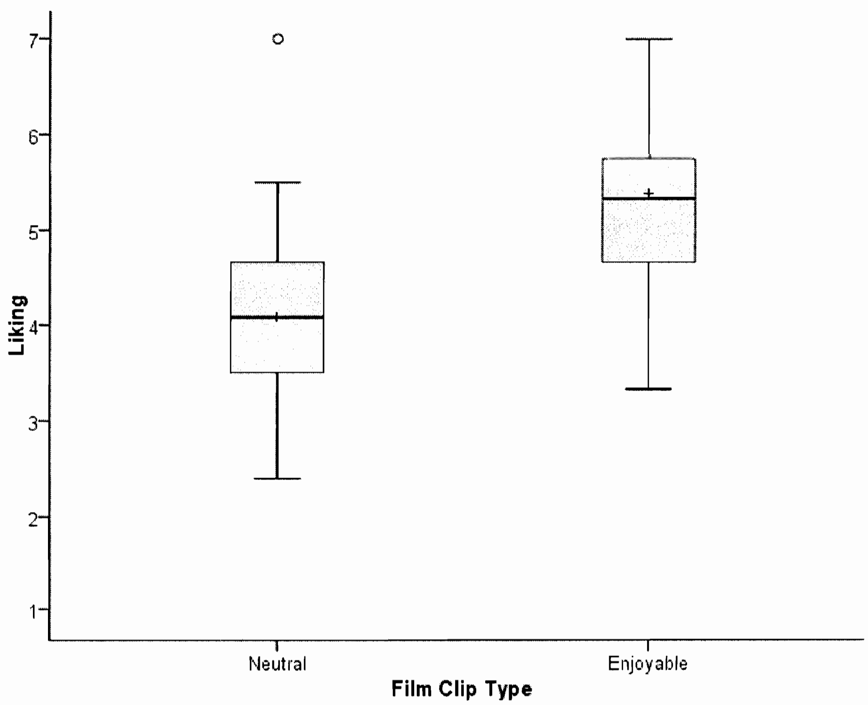



*Figure 8b*. Liking rating for both film clip types (neutral and enjoyable).

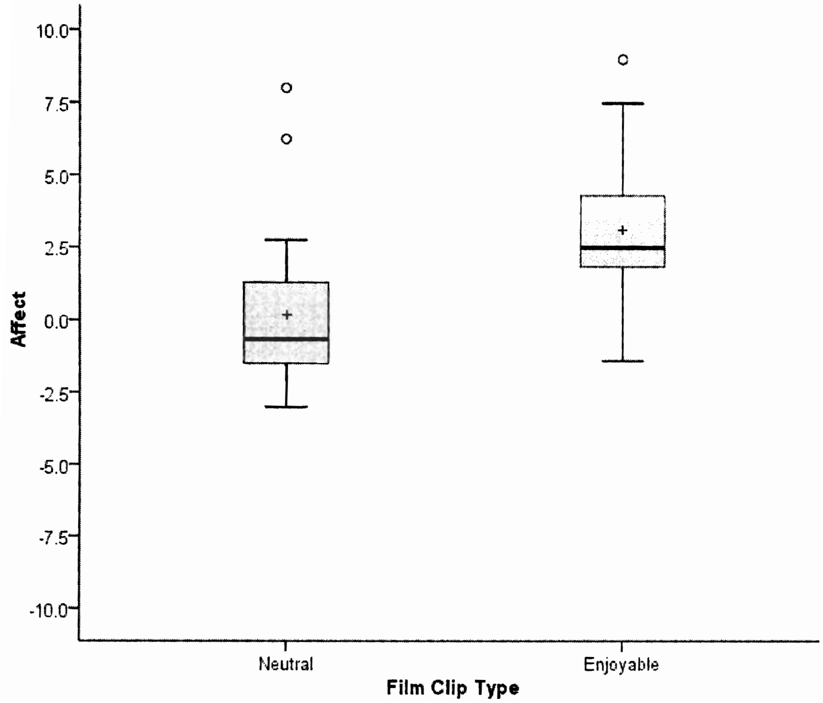



*Figure 8c*. Affect rating for both film clip types (neutral and enjoyable).

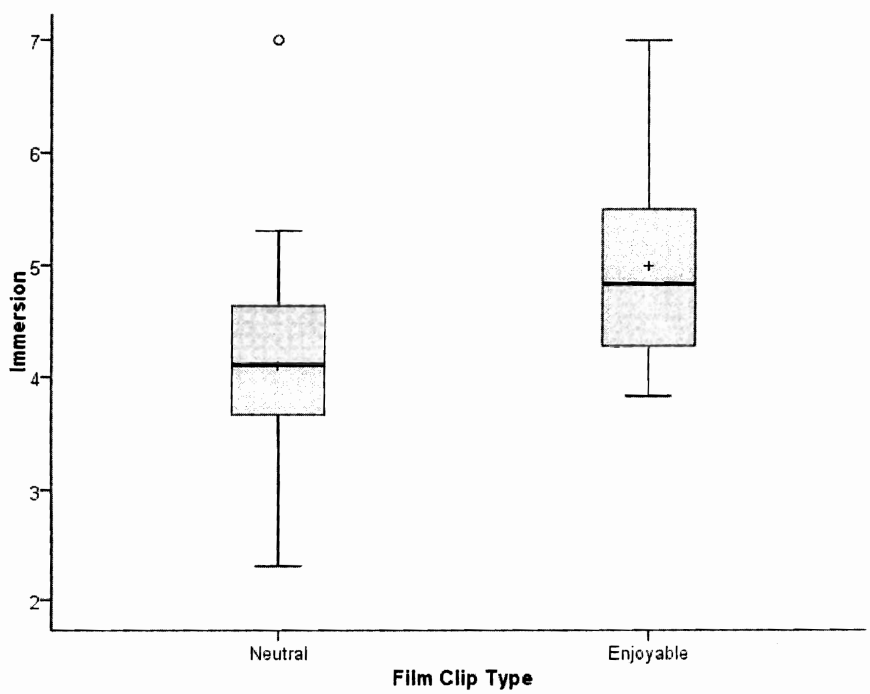



*Figure 8d*. Immersion rating for both film clip types (neutral and enjoyable).

*HLM.*

In order to determine if the full model from Figure 1 was supported using multilevel modeling, it was entered into HLM as the following: MOS = $\beta_{00} + \beta_{10}$*Liking + $\beta_{20}$*Immersion + $\beta_{30}$*Affect + $\beta_{40}$*Enjoyment + $\varepsilon$. The relationships from this model are listed in Table 5. It should be noted that in this full model, direct relationships such as those between affect and quality were not reliable when items such as liking were inserted into the model because of the mediating effect of liking on affect, whereas there was a direct effect of immersion on quality because there was no mediating effect of liking on immersion. Additionally, the effect of enjoyment on quality was not reliable due to the strong impact of liking on quality. These two constructs were very similar; however one measures the degree to which participants liked a film clip as an independent variable; and the other, as a dependent variable.

**Table 5:** Full model from HLM for Experiment 2

| Effect | Coefficient | SE | *t* | *df* |
|---|---|---|---|---|
| Intercept, $\beta_{00}$ | -0.01 | 0.08 | -0.16 | 29 |
| Liking, $\beta_{10}$ | 0.33 | 0.08 | 4.32** | 354 |
| Immersion, $\beta_{20}$ | 0.15 | 0.07 | 2.10* | 354 |
| Affect, $\beta_{30}$ | 0.00 | 0.07 | 0.01 | 354 |
| Enjoyment, $\beta_{40}$ | 0.04 | 0.05 | 0.96 | 354 |

*Note*. * $p < .05$; ** $p < .01$

*Discussion*

Experiment 2 was able to replicate and expand the work of Kortum and Sullivan (2010). First, it was able to determine that both affect and immersion were directly related to quality judgments. While liking partially mediated the relationship between affect and quality, it did not do so for the relationship between immersion and quality.

Therefore, the direct relationship of immersion on quality was not explained by liking. On the other hand, the direct relationship of affect on quality was better explained by liking – the effect of affect on quality was partially explained by participants' liking of a film clip. This makes sense because previous literature from the marketing and social psychology arenas have studied the powerful relationship between affect and liking, as noted in the introduction to this paper. It does appear that their relationship was quite strong and a causal issue arises of course – do participants use their mood to assess their liking or do they assess their liking which influences their mood state? This chicken-and-the-egg question was not answered here definitively, as no causal paths were explicitly drawn.

Second, while Kortum and Sullivan (2010) examined liking only as a dependent variable, Experiment 2 was also able to focus on liking as an independent variable through media enjoyment. Media enjoyment impacted participants' ratings on all variables in the model, such that ratings for quality, affect, immersion, and liking were reliably higher for enjoyable film clips than neutral clips. Because participants rated quality higher for enjoyment as an independent variable, this indicates that researchers can select film clips that elicit a particular level of enjoyment across a selected demographic for use in testing. In this experiment, a particular genre of film clip may have been considered enjoyable for undergraduates; however, different stimuli may be necessary for older or younger participants. Once a film clip has a predicted level of enjoyment for a particular demographic, one can predict a future participant's rating on quality based on the equation created for that film clip or enjoyment level. This implies that researchers can model quality scores in industry for a particular demographic upon determination of what is or is not enjoyable, and these researchers can determine an acceptable level of quality based on the range of enjoyment that a product will convey for its demographic.

## Experiment 3

In Experiment 3, participants were either trained or untrained on how to assess video quality. The affect infusion model discussed the utilization of training as one method to remove users' reliance on affect when making judgments, therefore it was hoped that there would no longer be a positive correlation between affect and quality (MOS) for trained participants, additionally removing a halo effect for these individuals. There were also three qualities of film clips utilized. It was believed that trained participants would be more accurate in rating the film clip quality compared to the untrained participants due to the increased subject knowledge that trained participants gain during the training protocol.

### *Method*

#### *Participants.*

Seventy-six Rice University undergraduates signed up to participate in the experiment. One participant who failed the Ishihara color test was excluded; therefore, 75 total undergraduates with normal or corrected-to-normal vision and normal color vision participated in the experiment. None of them participated in the prior experiments. There were 59 females and 16 males, with a mean age of 19.00 years (*SD* = 0.94); they participated for psychology experiment credit. Participants were divided across condition (training vs. no training) using random assignment, where 6 males and 31 females (37 individuals) were not trained, and 10 males and 28 females (38 individuals) were trained.

*Design.*

There were two within-subjects variables and one between-subjects variable – a 2-level variable measuring media enjoyment (neutral and enjoyable) and a 3-level variable measuring video quality (low, medium, and high quality). The between-subjects variable had 2 levels. Participants were either untrained or trained on video quality. There were four key dependent variables: quality (MOS), affect, immersion, and liking. Eight more dependent variables were measured as well in order to assess presence of a halo effect.

*Materials.*

Most materials used in Experiment 2 were used in Experiment 3. Only the differences are mentioned here.

Video Quality Assessment Overview. Half of the participants were trained using this overview of video quality terminology. It included descriptions of resolution, pixel drop-out, blockiness, motion blur, and mosquito noise, which were supported with demonstrations showing a low- and high-quality image of each artifact. The training materials are included in Appendix G.

Experiment 3 Film Clip Questionnaire. This questionnaire included the same 10 items as the Experiment 2 Film Clip Questionnaire, with the addition of 2 new questions related to specific image quality metrics (the sharpness of the image edges and the pixilation of the image). This was to determine if trained participants rated all video quality questions differently from untrained participants. The Experiment 3 film clip questionnaire is included in Appendix H.

Experiment 3 End of Experiment Questionnaire. The end of experiment questionnaire was altered to include additional questions to determine if trained participants' video quality knowledge differed from untrained participants. A 4-item multiple choice quiz was given at the end of the experiment to determine participants' knowledge of video quality. This quiz was utilized as the manipulation check. (see Appendix I).

Film Clip Quality. In Experiment 2, all 12 film clips were viewed at one bit rate (1100 kbps). In this experiment, the film clips were converted into three bit rates using DVDFab 8. A low (550 kbps), medium (1100 kbps), and high (6000 kbps) quality version of all 12 film clips was created, and the films were truncated into 2-minute clips as they were in Experiment 2 using Solveig Multimedia Video Splitter. Three playlists were created in Windows Media Player. Each playlist contained one quality version of a film clip, therefore across all three playlists, each quality version (low, medium, and high) was represented. Each of the 3 bit rates was represented 4 times on each playlist, as there were 12 film clips, and they were manipulated so that one quality type would not appear more than twice in a row.

*Procedure.*

The same setup utilized in Experiment 2 was used in Experiment 3. However, there were now two computer stations, allowing two participants to run simultaneously.

Before participants arrived, one of three playlists was loaded on each computer using random assignment. Participants were also randomly assigned to one of two conditions: video quality training or no training. For the training group, a brief overview of video quality terminology (resolution, pixel drop-out, blockiness, motion blur, and mosquito noise) was supported with demonstrations showing a low- and high-quality image of each artifact. For the no training group, participants were not given this overview.

One to two participants entered the laboratory at a time. They were first given a consent form to read and sign. Vision was tested using the Snellen eye chart and Ishihara plates. One individual failed the Ishihara plates–he was allowed to participate, however his data were discarded. Once both participants had arrived, broad-level instructions on the experiment were administered. The participants were informed that they would be watching a number of film clips, and after each clip they would complete a two-sided questionnaire. Depending on the condition (training or no training), participants were then trained on video quality; otherwise this step was skipped. Training involved the experimenter going over the slides demonstrating video quality artifacts. This step took an additional 5 minutes. All participants were given an overview of how to fill out the affect grid (collection of a baseline measure of affect was deemed unnecessary, as a baseline was collected during the sample film clip; however, the nontraditional format of the affect grid still warranted instructions).

One sample film clip (The Jackal) was shown in order to accustom participants with the procedure. Any remaining questions were answered after participants filled out the film clip questionnaire for this clip, and then participants viewed the twelve test clips sequentially, filling out the Experiment 3 film clip questionnaire after each clip. At the end of all test clips, participants filled out the video quality training assessment and end of experiment questionnaire. Participants were then debriefed and dismissed. The experiment took between 1 and 1.5 hours.

*Results*

The data were examined across all items to determine which clips had been seen before by participants, and 748 out of 900 clips viewed, or 83.1% of the clips, had not been seen before. An independent-samples *t*-test revealed that on all dependent variables (quality (MOS), affect, liking, and immersion), film clips that were seen before were given higher ratings than those that were not seen before (see Table 6). Therefore, 152 film clips were removed from the analyses due to their potential for inflation. No participants had seen all 12 film clips before. At most, a participant had seen 5 of the 12.

**Table 6:** Difference in ratings for film clips not seen before/seen before

| Variable | *M, SD, N*: Not Seen Before | *M, SD, N*: Seen Before | *t* | *df* |
|---|---|---|---|---|
| Quality (MOS) | 6.64, 1.42, 748 | 7.16, 1.07, 152 | 5.19* | 272.3 [a] |
| Immersion | 4.32, 1.24, 748 | 5.17, 1.00, 152 | 9.13* | 255.7[a] |
| Affect | 0.64, 3.11, 748 | 2.81, 2.59, 152 | 8.04* | 898 |
| Liking | 4.20, 1.50, 748 | 5.59, 1.23, 152 | 12.24* | 252.9 [a] |

*Note*. * = $p < .001$. [a] = unequal variance assumed, based on O'Brien's r (O'Brien, 1981).

*Manipulation check.*

Participants were either trained or untrained, and the end of experiment questionnaire included a 4-question multiple choice quiz to determine if trained participants performed better than untrained participants on video quality knowledge. An independent-samples *t*-test revealed that trained participants ($M = 3.63$, $SD = 0.71$) answered reliably more questions correctly than those who were not trained ($M = 1.24$, $SD = 0.95$), where $t(66.6) = 12.25$, $p < .001$, assuming unequal variance based on O'Brien's r, O'Brien, 1981. Therefore, the manipulation was successful.

*Affect.*

In order to determine if training moderated the relationship between affect and quality (MOS), the mean scores for affect and quality (MOS) were created in order to run a regression of quality on affect. There was a positive relationship between quality and affect when including in untrained participants ($\beta = .48$, $p = .003$), however there was no relationship when including trained participants ($\beta = -.03$, $p = .837$). The difference in slope is apparent in Figure 9. In testing for a difference between the two slopes using $t = \frac{b_1 - b_2}{s_{b1-b2}}$, $t = 3.35$ which exceeded the critical $t$ value of 2.02, therefore the slopes were considered different, and moderation was confirmed.

Because of the findings in the linear regression, it seemed that a halo effect may be present for those who were untrained yet absent for those trained. Table 7 demonstrates the presence of positive correlations between most variables (for 65/66, $p < .05$) for untrained participants, whereas Table 8 shows that there are not as many positive correlations when considering trained participants (for 37/66, $p < .05$). As there was no longer a relationship between affect and quality for trained participants, it could be said that liking would no longer mediate the relationship the relationship between affect and quality, as one requirement for mediation is the existence of a direct effect between the criterion and predictor. The correlation matrix also revealed that there was no longer a relationship between liking and quality for trained participants ($r = .16, p = .328$), although this relationship was still quite strong for untrained participants ($r = .73, p < .001$).

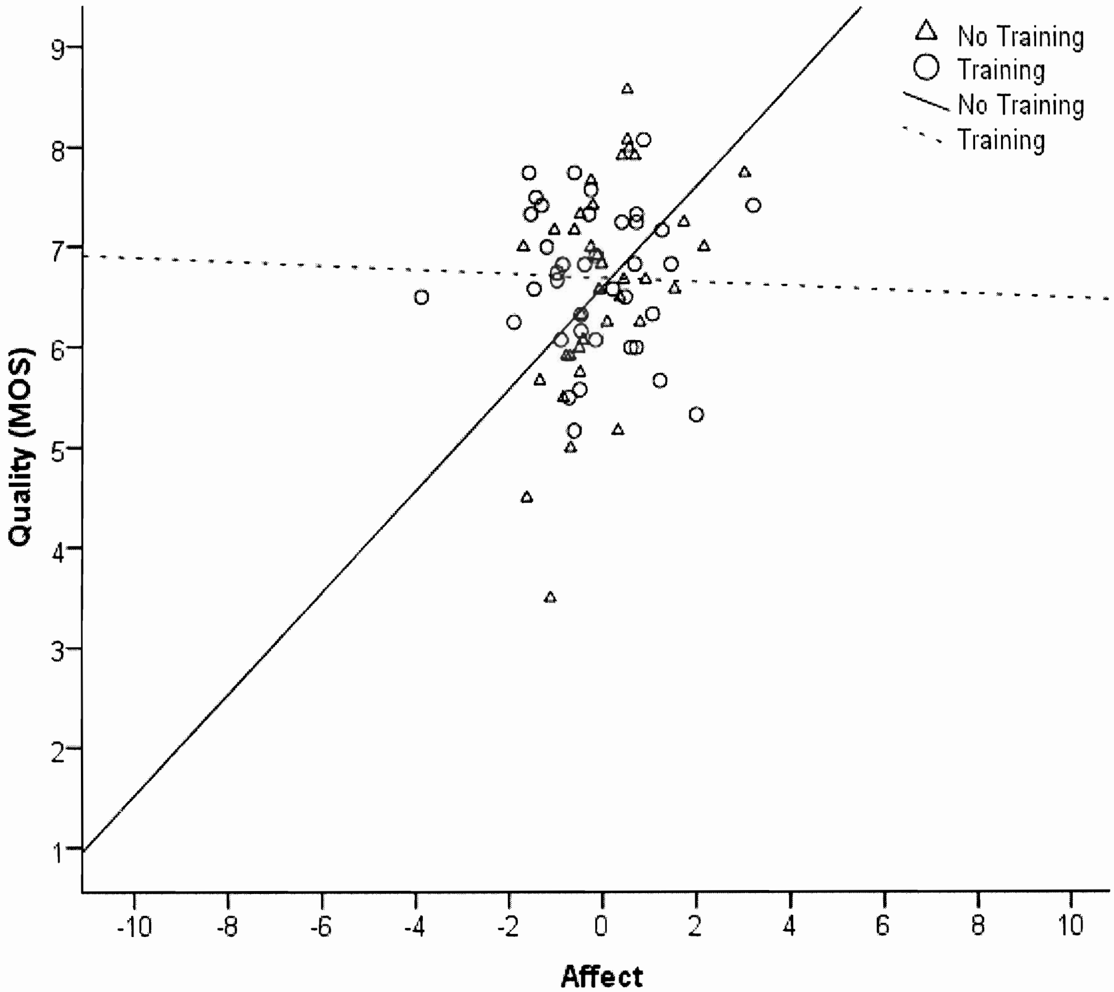



*Figure 9*. Interaction of training between affect and quality (MOS).

Using the 12 continuous variables from the film clip questionnaire, exploratory factor analysis was conducted using principal components. Promax roation was set at a Kappa of 4 due to the high correlations between all items, and the factor analysis

**Table 7:** Intercorrelation matrix for untrained participants in Experiment 3

| | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. MOS | 6.61 | 1.06 | - | | | | | | | | | | |
| 2. Immersion | 4.39 | 0.64 | .73** | - | | | | | | | | | |
| 3. Liking | 4.28 | 0.79 | .73** | .64** | - | | | | | | | | |
| 4. Affect | 0.91 | 1.07 | .48** | .34* | .51** | - | | | | | | | |
| 5. Box Office Predictions | 4.06 | 0.69 | .68** | .51** | .49** | .31* | - | | | | | | |
| 6. Curious | 4.71 | 0.86 | .68** | .74** | .71** | .25 | .58** | - | | | | | |
| 7. Acting | 6.72 | 0.91 | .85** | .70** | .60** | .36* | .78** | .70** | - | | | | |
| 8. Lighting | 4.85 | 0.88 | .73** | .69** | .56** | .43** | .57** | .64** | .67** | - | | | |
| 9. Audio | 6.86 | 1.03 | .92** | .66** | .61** | .41* | .58** | .60** | .83** | .66** | - | | |
| 10. Resolution | 4.42 | 0.93 | .92** | .66** | .60** | .42* | .58** | .58** | .76** | .70** | .87** | - | |
| 11. Edges | 4.55 | 0.91 | .90** | .56** | .68** | .45** | .54** | .53** | .71** | .66** | .84** | .92** | - |
| 12. Pixels | 4.52 | 0.96 | .89** | .66** | .51** | .35* | .54** | .57** | .70** | .71** | .85** | .95** | .86** |

*Note*. $N = 37$ for all items. * = $p < .05$; ** = $p < .01$.

**Table 8:** Intercorrelation matrix for trained participants in Experiment 3

| | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. MOS | 6.70 | 0.72 | - | | | | | | | | | | |
| 2. Immersion | 4.38 | 0.44 | .00 | - | | | | | | | | | |
| 3. Liking | 4.27 | 0.56 | .16 | .59** | - | | | | | | | | |
| 4. Affect | 0.71 | 1.36 | -.03 | .43** | .22 | - | | | | | | | |
| 5. Box Office Predictions | 4.02 | 0.53 | .22 | 59** | .47** | .36** | - | | | | | | |
| 6. Curious | 4.56 | 0.70 | .14 | .59** | .73** | .07 | .53** | - | | | | | |
| 7. Acting | 6.86 | 0.57 | .42** | .38* | .39* | .07 | .49** | .39* | - | | | | |
| 8. Lighting[a] | 4.83 | 0.67 | .53** | .46** | .42** | .05 | .47** | .42** | .61** | - | | | |
| 9. Audio | 7.03 | 0.79 | .75** | .11 | .26 | -.07 | .15 | .20 | .52** | .67** | - | | |
| 10. Resolution | 4.58 | 0.76 | .80** | .20 | .30 | -.14 | .44** | .35* | .29 | .60** | .57** | - | |
| 11. Edges | 4.64 | 0.64 | .70** | .20 | .22 | .07 | .36* | .20 | .22 | .42** | .41* | .75** | - |
| 12. Pixels | 4.69 | 0.64 | .72** | .10 | .05 | -.05 | .28 | .11 | .20 | .40* | .47** | .73** | .83** |

*Note*. $N = 38$ for all items. * = $p < .05$; ** = $p < .01$.

produced one factor with 68.53% of the total variance explained for untrained participants. For trained participants, the analysis loaded onto three factors, and 73.81% of the total variance was explained by them (see Table 9). This indicates the presence of a halo effect for the untrained participants, as all items were positively correlated and loaded onto the same factor. For the trained participants, there was no halo, as the correlations were not positive across all items. Additionally, the three factors loaded onto

**Table 9:** Exploratory factor analysis for Experiment 3

Untrained Participants

| Item: | Component 1 |
|---|---|
| Video Quality (MOS) | .970 |
| Resolution | .919 |
| Audio Quality | .908 |
| Edges | .886 |
| Pixels | .884 |
| Acting Quality | .882 |
| Cinematographer's Choice of Lighting | .816 |
| Immersion | .809 |
| Curious What Happens Next | .765 |
| Liking | .763 |
| Box Office Predictions | .722 |
| Affect | .507 |
| % of Variance | 68.56% |

Trained Participants

| Item: | Component 1 | Component 2 | Component 3 |
|---|---|---|---|
| Video Quality (MOS) | .907 | .433 | -.308 |
| Pixels | .896 | .239 | -.051 |
| Resolution | .895 | .517 | -.108 |
| Edges | .881 | .314 | .111 |
| Audio Quality | .689 | .594 | -.443 |
| Cinematographer's Choice of Lighting | .606 | .800 | -.094 |
| Acting Quality | .363 | .773 | -.093 |
| Curious What Happens Next | .190 | .767 | .323 |
| Liking | .177 | .762 | .360 |
| Immersion | .128 | .669 | .637 |
| Box Office Predictions | .364 | .644 | .553 |
| Affect | -.042 | .119 | .754 |
| % of Variance | 43.22% | 20.72% | 9.87% |

three distinct groups. The structure matrix in Table 9 shows that 5 items (quality (MOS), pixels, resolution, edges, and audio quality – Component 1) loaded onto one factor, which could be considered an overall audio and video quality. It makes sense that trained participants would rate these items in a distinct manner, and they were able to assess quality without the interference of affect, which loaded onto its own factor.

*Quality.*

In order to determine if training impacted the accuracy of quality ratings, a 3 (quality) x 2 (enjoyment) x 2 (condition) mixed design ANOVA was carried out, using MOS as a dependent variable. There was a reliable quality x condition interaction, $F(2, 146) = 4.04$, $p = .020$. There was also a main effect of quality, $F(2, 146) = 26.56$, $p < .001$, but no main effect of condition, $F(1, 73) = 0.19$, $p = .662$ (see Figure 10). A linear contrast on the interaction revealed that the linear effect of quality was different depending on the condition (trained versus untrained), $F(1, 73) = 8.65$, $p = .004$, passing the Scheffé criterion. Specifically, as can be seen in Figure 10, trained participants gave the low quality film clips a lower MOS score and the high quality film clips a higher MOS score than untrained participants. The linear contrast of quality was highly reliable, $F(1, 73) = 56.67$, $p < .001$, again passing the Scheffé criterion, indicating that high quality film clips were rated higher on the MOS variable than medium quality film clips, and medium quality film clips were rated higher than low quality film clips.

The enjoyment x condition interaction was not reliable, $F(1, 73) = 1.28$, $p = .261$, indicating that film clip enjoyment did not influence either trained or untrained participants differently when they rated MOS. There was a main effect of enjoyment,

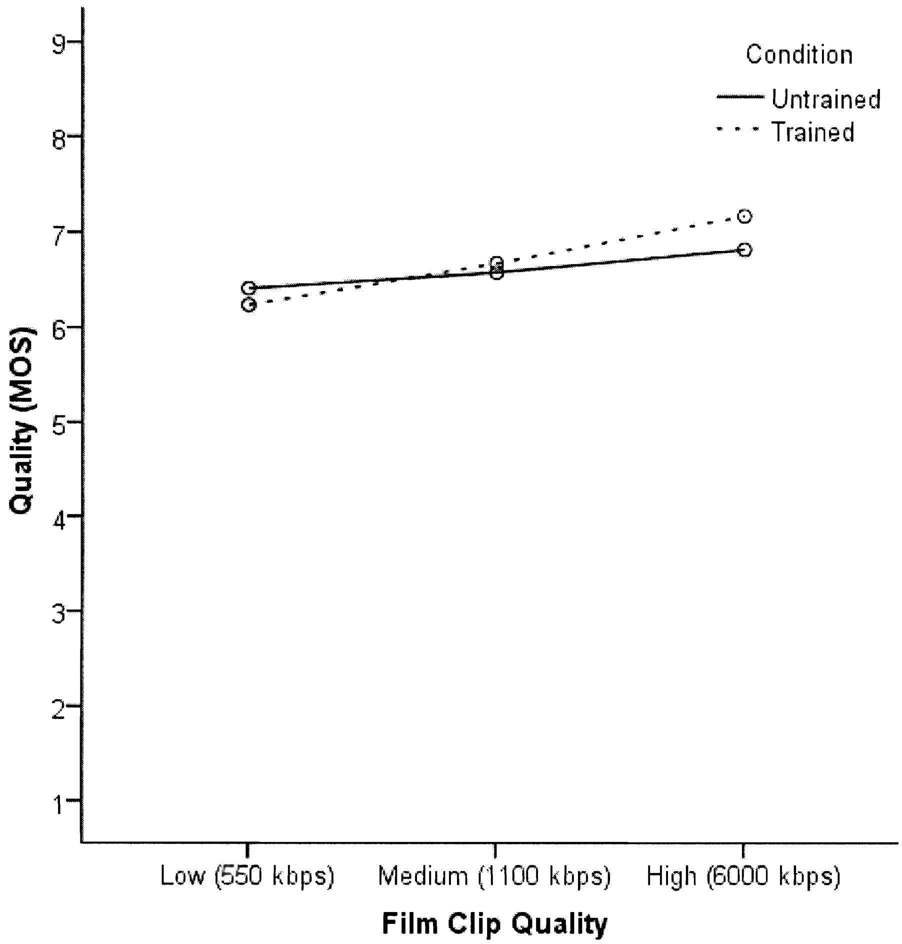



*Figure 10*. Quality x condition interaction

however, where enjoyable clips were rated higher on MOS than neutral film clips ($F(1, 73) = 22.36, p < .001$; see Figure 11). The quality x enjoyment interaction was reliable, where $F(2, 146) = 7.77, p = .001$ (see Figure 12). The linear interaction contrast indicated that the linear effect of quality differed depending on the level of enjoyment, $F(1, 73) = 11.88, p < .001$, passing the Scheffé criterion. From the figure, it can be seen that the slope for neutral film clips was relatively flat across film clip quality, whereas the slope was positive for enjoyable film clips across film clip quality, increasing as film clip quality increased. Finally, the 3-way interaction between condition, enjoyment, and quality was not reliable, $F(2, 146) = 1.48, p = .232$.

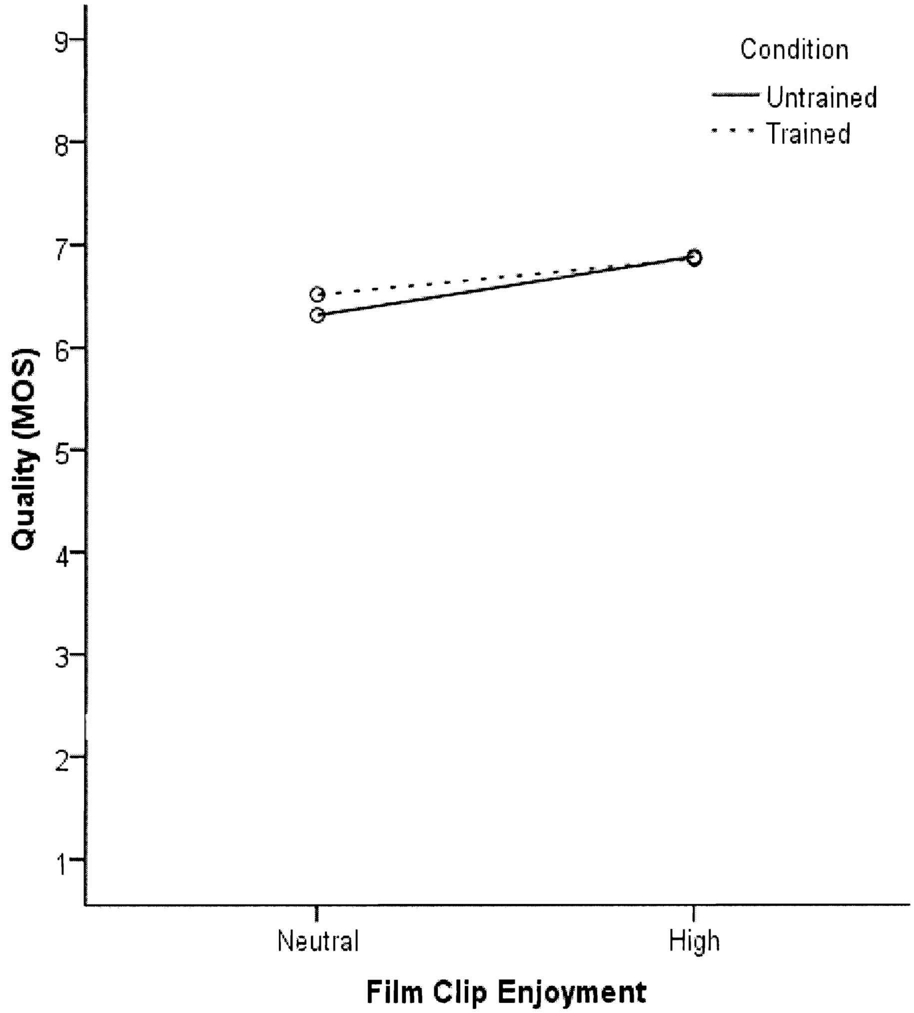



*Figure 11*. Main effect of enjoyment. *Note*. There was no interaction with condition.

In an attempt to generalize effects found from the 12 film clips utilized in the experiment to all film clips, both film clips and participants were held as random effects, and the main effects and interactions were then investigated for the fixed effects (condition, film clip quality, and film clip enjoyment) using JMP 9. Two main effects were still reliable, as they were above, when using the random effects model, however the interactions were no longer reliable (see Table 10).

**Table 10:** ANOVA for fixed effects when film clips and participants are random

| Source | *F* | *df* |
|---|---|---|
| Condition | 0.49 | 1, 59.31 |
| Film Clip Quality | 16.33** | 2, 512 |
| Film Clip Enjoyment | 4.93* | 1, 51.11 |
| Condition × Enjoyment | 0.00 | 1, 515.2 |
| Quality × Enoyment | 2.07 | 2, 521.7 |
| Condition × Quality | 2.29 | 2, 511.2 |
| Condition × Quality × Enjoyment | 2.73 | 2, 515.6 |

*Note*. * $p < .05$ ** $p < .01$

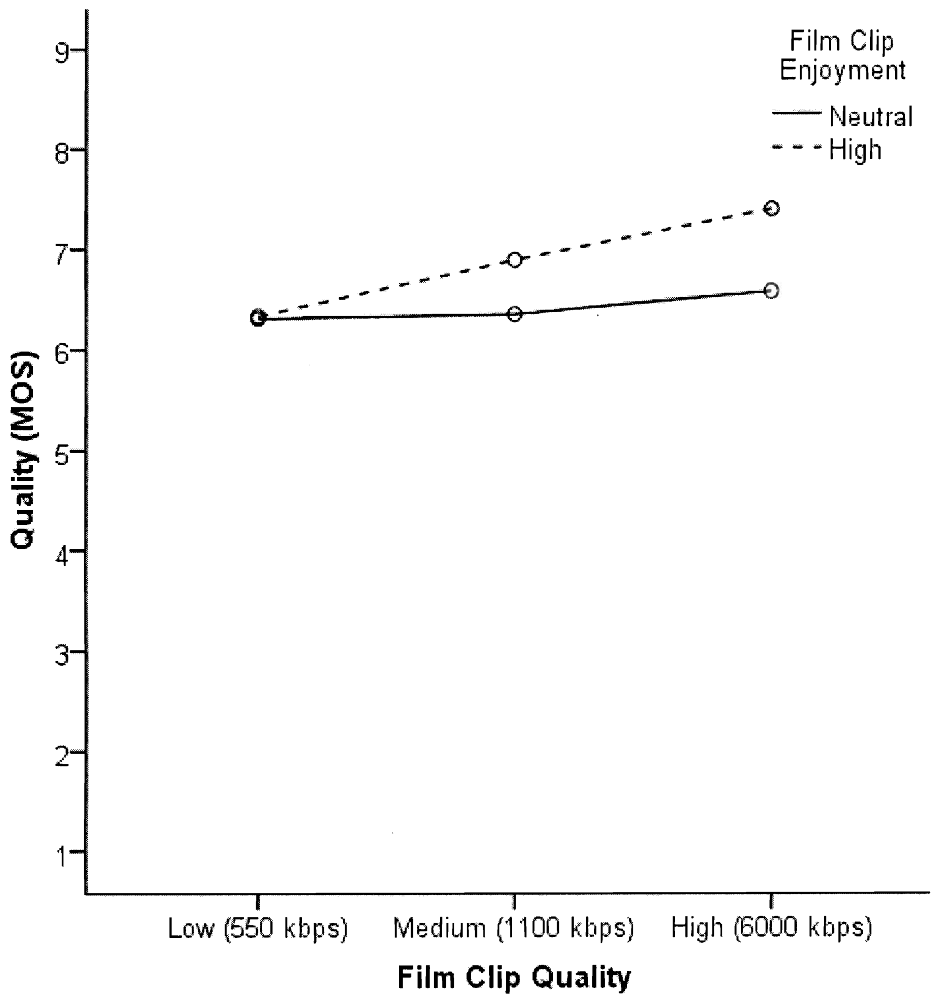



*Figure 12*. Quality x enjoyment interaction.

*HLM.*

In order to determine if the above effects tested in SPSS remained reliable when put into a multilevel model, HLM was utilized. In HLM, the z scores of all variables were entered in order to achieve the interactions. The model entered was as shown: MOS = $\beta_{00}$ + $\beta_{01}$*Condition + $\beta_{10}$*Film Quality + $\beta_{11}$*Condition*Film Quality + $\beta_{20}$*Affect + $\beta_{21}$*Condition*Affect + $\beta_{30}$*Enjoyment + $\beta_{31}$*Condition*Enjoyment + $\beta_{40}$*Film Quality*Enjoyment + $\beta_{41}$*Condition*Film Quality*Enjoyment + $\varepsilon$. All relationships in HLM agreed with those found in SPSS. The findings from HLM are listed in Table 11.

**Table 11:** Full model from HLM for Experiment 3

| Effect | Coefficient | SE | *t* | *df* |
|---|---|---|---|---|
| Intercept, $\beta_{00}$ | 0.02 | 0.07 | 0.23 | 73 |
| Condition, $\beta_{01}$ | 0.07 | 0.07 | 1.01 | 73 |
| Film Quality, $\beta_{10}$ | 0.16 | 0.03 | 6.37** | 886 |
| Film Quality ×Condition, $\beta_{11}$ | 0.06 | 0.03 | 2.37* | 886 |
| Affect, $\beta_{20}$ | 0.10 | 0.03 | 3.16** | 886 |
| Affect ×Condition, $\beta_{21}$ | 0.07 | 0.03 | 2.21* | 886 |
| Enjoyment, $\beta_{30}$ | 0.15 | 0.03 | 5.30** | 886 |
| Enjoyment ×Condition, $\beta_{31}$ | -0.05 | 0.03 | -1.75 | 886 |
| Film Quality × Enjoyment, $\beta_{40}$ | 0.06 | 0.03 | 1.98* | 886 |
| Film Quality × Enjoyment × Condition, $\beta_{41}$ | 0.03 | 0.03 | 1.00 | 886 |

*Note*. * $p < .05$; ** $p < .01$

*Discussion*

As trained participants did not rely on affect or liking when making quality judgments, it appears that one way to eliminate biases is to train users or select users who are knowledgeable on video quality. The training that the participants in Experiment 3 underwent was very brief and did not even show examples from video clips. However it was clearly effective enough to remove reliance on affect, and subsequently, liking. Further, participants were more accurate when rating quality via the MOS score as shown by the quality x condition interaction which indicated that trained participants were able to rate the different film clip qualities more distinctly compared to untrained participants. While untrained participants' ratings were clustered together across all film clip qualities, the trained participants' ratings fanned out such that there was a distinct linear separation from low to high film clip qualities. It is important to note that quality ratings were not impacted by training when film clips and participants were random; therefore, it is possible that the above findings may not generalize beyond those film clips and participants utilized in this experiment.

It is possible that it was not the training that assisted the participants who were in the experimental group; but instead, the fact that they may have taken on a goal of video quality assessment. The affect infusion model (Forgas, 1995) states that individuals will no longer rely on affect when making judgments either because they have gained content knowledge (i.e., have been trained or already possess knowledge in what is being evaluated) or because they are motivated to perform well (i.e., they are pursuing a goal). Therefore, even if training was not necessarily the driving force behind the success of affect extinction, the model would still succeed if participants were pursuing a goal of assessing video quality.

In regards to entertainment, the enjoyable clips were rated higher in quality (MOS) than the neutral clips, and training did not impact participants' ability to rate clips equally in quality regardless of their level of enjoyment. However, the quality x enjoyment interaction was less intuitive. It appeared that participants were able to rate film clips more accurately for the entertaining clips than for the neutral clips, as there was a positive slope from low (550 kbps) to high (6000 kbps) film clips for the enjoyable clips, yet a flat slope from low to high film clips for neutral clips. Therefore the neutral clips were all given the same quality rating, whereas enjoyable clips' ratings varied based on their quality. One would assume that neutral clips would be rated more accurately as liking would play less of a role; however, from this experiment, it would indicate that researchers should utilize enjoyable film clips and train participants to receive accurate feedback on video quality. It is possible that users pay less attention to neutral film clips and are therefore less accurate in rating them than enjoyable clips – if this is the case, it will surface in Experiment 4, where attention will be examined. It should again be pointed out that the quality x enjoyment interaction was not reliable when film clips and participants were random. Experiment 4 will also be able to support or deny the ability to generalize the above findings beyond the film clips used in this study.

## Experiment 4

In Experiment 4, participants were given different sets of directions to impact their focus of attention. As participants rated quality higher depending on their level of content immersion in Experiment 2, it was believed that reducing content immersion by focusing their attention directly on quality would help to extinguish the positive relationship. Inattentional blindness research gave reason to believe that participants focusing on the target of quality would be able to accurately respond to the film clip's quality, whereas those focusing elsewhere would not notice quality decrements. It was also believed that participants focusing on quality would be more accurate in rating the three levels of quality, as their attention will be directed away from content and onto quality.

*Method*

*Participants.*

Sixty-three Rice University undergraduates with normal or corrected-to-normal vision and normal color vision were recruited to participate in the experiment. None of them participated in previous experiments. One participant fell asleep partway through the experiment; therefore, her data were excluded as she was not attending to the film clip stimuli. Considering the 62 remaining participants, there were 19 males and 44 females with a mean age of 19.15 years ($SD = 1.02$). Participants were equally divided among the two experimental conditions (31 in each group – focus on content or focus on quality).

*Design.*

There were two within-subjects variables and one between-subjects variables. The within-subjects variables were identical to those in Experiment 3 – a 2-level variable measuring media enjoyment and a 3-level variable measuring video quality. The between-subjects variable had 2 levels. Participants were instructed to focus on either video content or video quality. Dependent variables measured on the film clip questionnaire were quality (MOS), immersion, affect, and liking. For this experiment, quality (MOS) was investigated as the dependent variable, whereas immersion was treated as an independent variable.

*Materials.*

All materials used in Experiment 2 were used in Experiment 4. The Experiment 4 film clip questionnaire was modified to include the manipulation check for direction of attention. This involved five custom questions per film clip pertaining to the content of each film clip, such as "What color jacket was the driver wearing?" and "What shape were the man's glasses?" In addition, there were five specific questions related to the quality of each film clip. These were inserted to motivate participants who were directed to focus on quality to do so. The Experiment 4 film clip questionnaire is included in Appendix J – as the manipulation check varied for each of the 12 film clips, the one included here is for the film clip from The Siege. The Experiment 4 end of experiment questionnaire is included in Appendix K.

*Procedure.*

The same setup utilized in Experiment 3 was used in Experiment 4. The 62 participants were equally divided across the two conditions. Participants were randomly assigned to focus on content or quality.

One or two participants entered the laboratory at a time and signed a consent form. Vision was then tested using the Snellen eye chart and Ishihara plates; no participant failed either test. Instructions on the experiment were then administered. The participant was informed that he/she would be watching a number of film clips, and after each clip he/she would complete a series of questionnaires. The participant then received additional instructions on where to focus attention–on video quality or on film content. He/she was told that there would be questions specifically directed towards the assigned task, and the first film clip could be treated as a practice run to familiarize oneself with the task and the types of questions being asked.

The sample film clip was shown in order to accustom participants with the procedure. Any remaining questions were answered, and then a participant viewed the twelve test clips sequentially, filling out the items from the Experiment 4 film clip questionnaire after each clip. The clips were presented in one of three random orders. At the end of the test clips, participants filled out the end of experiment questionnaire. They were then be debriefed and dismissed. The experiment took between 1 and 1.5 hours.

*Results*

The data were examined across all items to determine which clips had been seen before by participants, and 626 out of 744 clips viewed, or 84.1% of the clips, had not been seen before. An independent-samples *t*-test revealed that on all dependent variables (quality (MOS), affect, liking, and immersion), film clips that were seen before were given higher ratings than those that were not seen before (see Table 12). Therefore, 118 film clips were removed from the analyses due to their potential for inflation. No participants had seen all 12 film clips before. At most, a participant had seen 4 of the 12.

**Table 12:** Difference in ratings for film clips not seen before/seen before

| Variable | *M, SD, N*: Not Seen Before | *M, SD, N*: Seen Before | *t* | *df* |
|---|---|---|---|---|
| Quality (MOS) | 6.44, 1.36, 626 | 6.83, 1.34, 118 | 2.85* | 742 |
| Immersion | 4.22, 1.28, 626 | 4.92, 0.94, 118 | 7.04** | 208.3[a] |
| Affect | 0.45, 3.10, 623 | 2.25, 2.94, 118 | 5.85** | 739 |
| Liking | 4.10, 1.56, 626 | 5.31, 1.26, 118 | 9.14** | 190.4[a] |

*Note*. * = $p < .01$; ** = $p < .001$. [a] = unequal variance assumed (O'Brien's r, 1981).

*Manipulation check.*

The manipulation affected participants' attention: they were told to focus on video quality or content. Those who focused on quality should not have been able to divide attention between video quality and content, therefore a difference in correct responses between groups was not expected to surface for the content questions asked. There were 5 questions asked per film clip, and the number of questions was averaged across all 12 clips to arrive at a mean score per participant. An independent-samples *t*-test revealed that those focusing on content ($M = 2.97$, $SD = 0.52$) answered reliably more questions correctly than those focusing on quality ($M = 1.19$, $SD = 0.80$), where $t(51.5) = 10.45$, $p < .001$, unequal variance assumed (O'Brien, 1981). Thus, the manipulation was successful.

*Immersion.*

For the 6 immersion items, there was a Cronbach's alpha of $\alpha = .912$, indicating high reliability between items. In order to determine if focus of attention moderated the relationship between immersion and quality (MOS), the mean scores for immersion and quality (MOS) were created in order to run a regression of quality on immersion. There was a positive relationship between quality and immersion when examining participants who focused on content ($\beta = .49$, $p = .006$), and quality ($\beta = .42$, $p = .020$), and testing against a critical value of $t = 2.004$, the *t*-value was -0.33 using $t = \frac{b_1 - b_2}{s_{b1-b2}}$. Therefore, there was no difference between slopes, indicating no moderation (see Figure 13).

*Quality.*

A 3 (quality) x 2 (enjoyment) x 2 (condition) mixed design ANOVA was carried out, using MOS as a dependent variable in order to determine the effect of condition on quality. There was an interaction between quality and condition, $F(2, 120) = 6.67$, $p = .002$, however, no main effect of condition, $F(1, 60) = 1.55$, $p = .218$. (see Figure 14). A linear interaction contrast demonstrated that the slopes for quality differed depending on

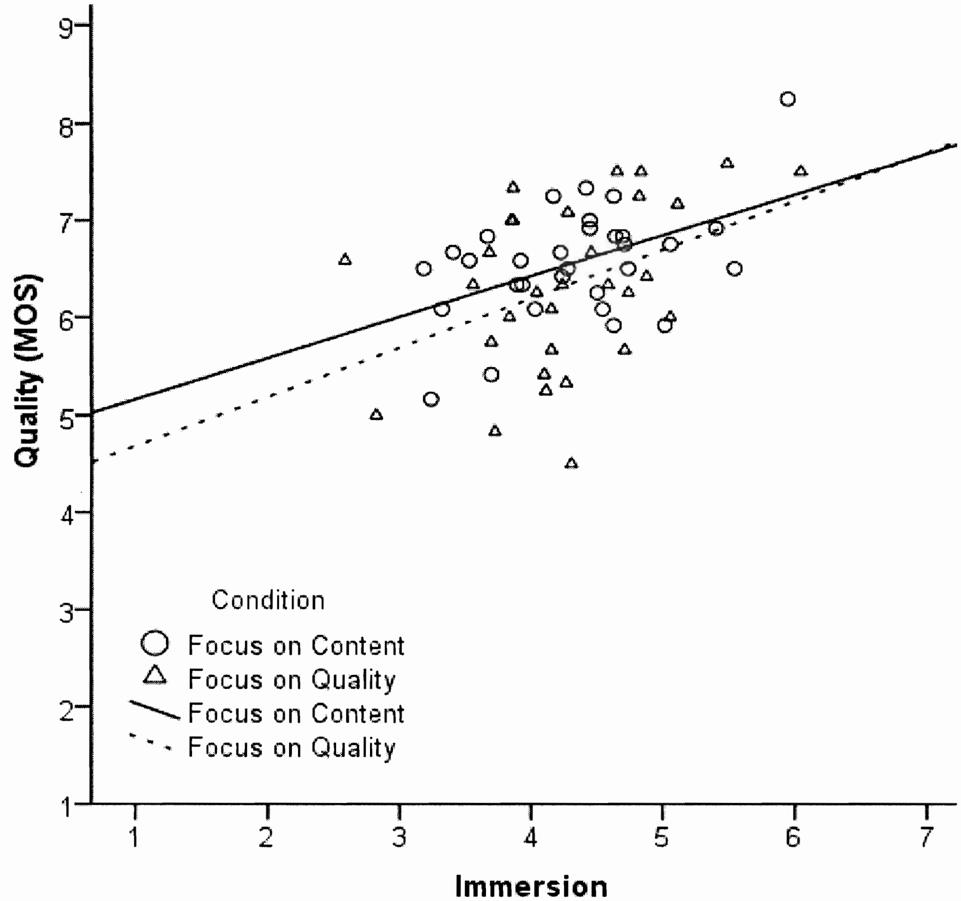



*Figure 13*. Lack of interaction for direction of focus between immersion and quality (MOS).

condition, $F(1, 60) = 10.19$, $p = .002$, and from the figure it can be seen that the slope is rather flat for participants who focused on content, indicating that they were less likely to discriminate between film quality types. However, the slope for those focusing on quality is more positive, indicating that they were more likely to give a higher MOS rating to a film clip with a higher quality. A custom contrast (-2 1 1 2 -1 -1) examined whether the low film quality pair differed from the medium and high film quality pairs, and this was the case, $F(1, 60) = 12.77$, $p = .001$, passing the Scheffé criterion. There was no interaction between enjoyment and condition, $F(1, 60) = 0.03$, $p = .870$. Similar with

Experiment 3 were an interaction between quality and enjoyment, $F(2, 120) = 6.47, p = .002$, and main effects for both quality, $F(2, 120) = 19.08, p < .001$, and enjoyment $F(1, 60) = 12.99, p = .001$ (see Figure 15). This figure was similar to that in Experiment 3, and a custom interaction contrast (-2 1 1 2 -1 -1) investigated whether the low film quality pair differed from the other two film quality pairs (medium and high) depending on film quality. This interaction contrast was reliable, $F(1, 60) = 13.55, p < .001$, passing the Scheffé criterion. From the figure, it can be seen that film clips at low quality (550 kbps) were not rated differently, regardless of their enjoyment. This may imply that there was a lower boundary at which video quality impacts participant ratings and satisfaction.

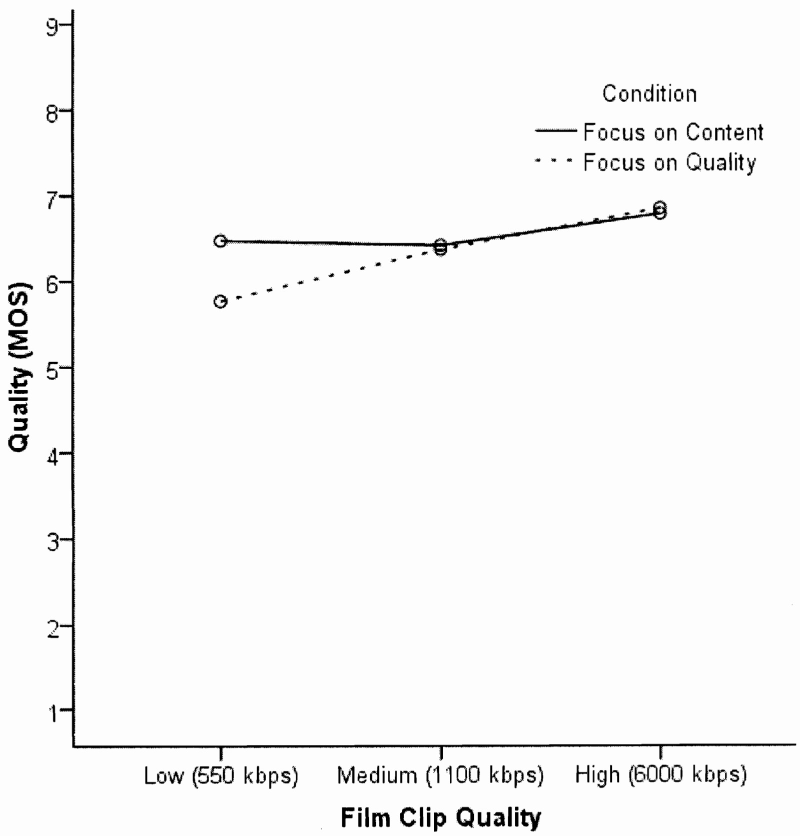



*Figure 14*. Quality x condition interaction.

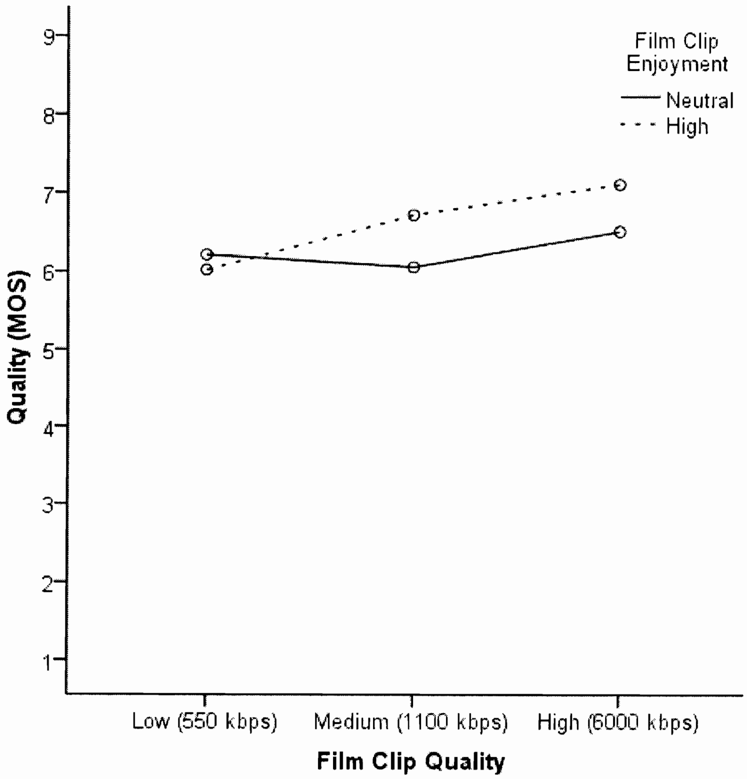



*Figure 15*. Quality x enjoyment interaction.

However, for both medium (1100 kbps) and high quality (6000 kbps) film clips, enjoyment drove the MOS score, such that the film clips were rated higher when the film clip was enjoyed. A three way interaction (enjoyment x quality x condition) was not reliable, $F(2, 120) = 0.36, p = .701$.

In an attempt to generalize effects found from the 12 film clips utilized in the experiment to all film clips, both film clips and participants were held as random effects, and the main effects and interactions were then investigated for the fixed effects (condition, film clip quality, and film clip enjoyment). Here, the same effects that were reliable above were reliable with random effects, indicating the generalizability of the findings beyond those 12 film clips viewed just in this experiment (see Table 13).

**Table 13:** ANOVA for fixed effects when film clips and participants are random

| Source | *F* | *df* |
|---|---|---|
| Condition | 1.13 | 1, 60.85 |
| Film Clip Quality | 18.12** | 2, 549.6 |
| Film Clip Enjoyment | 6.07* | 1, 10.52 |
| Condition × Enjoyment | 0.05 | 1, 549.8 |
| Quality × Enoyment | 6.60** | 2, 549.8 |
| Condition × Quality | 5.25** | 2, 551 |
| Condition × Quality × Enjoyment | 0.37 | 2, 551.1 |

*Note*. * $p < .05$ ** $p < .01$

*HLM.*

In order to determine if the above effects tested in SPSS remained significant when the error variance was increased using multilevel modeling, the full model was run in HLM. In order to achieve interactions, the z scores of all variables were entered. The model entered was as shown: MOS = $\beta_{00}$ + $\beta_{01}$*Condition + $\beta_{10}$*Film Quality + $\beta_{11}$*Condition*Film Quality + $\beta_{20}$*Immersion + $\beta_{21}$*Condition* Immersion + $\beta_{30}$*Enjoyment + $\beta_{31}$*Condition*Enjoyment + $\beta_{40}$*Film Quality*Enjoyment + $\beta_{41}$*Condition*Film Quality*Enjoyment + ε. All relationships in HLM agreed with those found in SPSS. The findings from HLM are listed in Table 14.

**Table 14:** Full model from HLM for Experiment 4

| Effect | Coefficient | SE | *t* | *df* |
|---|---|---|---|---|
| Intercept, $\beta_{00}$ | -0.00 | 0.06 | -0.03 | 60 |
| Condition, $\beta_{01}$ | -0.07 | 0.06 | -1.23 | 60 |
| Film Quality, $\beta_{10}$ | 0.20 | 0.03 | 6.67** | 734 |
| Film Quality ×Condition, $\beta_{11}$ | 0.12 | 0.03 | 3.99** | 734 |
| Immersion, $\beta_{20}$ | 0.28 | 0.04 | 7.70** | 734 |
| Immersion ×Condition, $\beta_{21}$ | -0.02 | 0.04 | -0.64 | 734 |
| Enjoyment, $\beta_{30}$ | 0.07 | 0.03 | 2.21* | 734 |
| Enjoyment ×Condition, $\beta_{31}$ | 0.01 | 0.03 | 0.30 | 734 |
| Film Quality × Enjoyment, $\beta_{40}$ | 0.11 | 0.03 | 3.55** | 734 |
| Film Quality × Enjoyment × Condition, $\beta_{41}$ | 0.02 | 0.03 | 0.60 | 734 |

*Note*. * $p < .05$; ** $p < .01$

*Discussion*

For all main effects and interactions found in this experiment, the ability to generalize beyond those film clips utilized in the study was supported. This indicates that any film clips can be selected for use, and the same findings in regards to immersion, enjoyment, and quality ratings will surface.

It was not found to be the case that focus of attention moderated the relationship between immersion and quality (MOS). However, participants focusing on quality were able to discern between the three quality levels of the film clips, rating the low-quality clips lower than the medium, and the medium, lower than the high, as demonstrated with a linear contrast. It was apparent in Figure 15 (see page 62) that the high quality film clips were rated higher than the low quality clips, therefore the large jump in bit rate for the high quality film clip (6000 kbps versus 1100 or 550) was perceptible.

Participants in both conditions reported high levels of immersion – it would have been expected that those focusing on quality would have experienced less content immersion. However it should be noted that the power of the media was quite strong in that participants in the quality condition noted the challenge to stay focused on the task. Many explicitly stated at the end of the experiment that exciting or engaging film clips made it hard to focus on quality, as this type of clip distracted them from the task. Therefore, it was surprising that there was no condition x enjoyment interaction, as it would seem that participants would have been able to focus on their tasks better for neutral clips and worse for enjoyable clips. It is possible that all clips were considered too entertaining for the purposes of this experiment, as all MOS scores rated quite high on average (although the full 1 – 9 range was utilized). It is also possible that 2 minutes was considered too long per film clip. An assessment of film quality was potentially made in a short period of time, at which point a participant could switch his/her attention to the content. Another study could replicate the experiment with shorter film clips to determine if attention span can be kept on task. An eye tracker can also determine if a user's attention is wandering after a certain amount of time, or whether participants did focus on quality and therefore away from content when they were instructed to do so. While the manipulation check demonstrated that participants focusing on quality were less able to accurately recall the content of the film clips compared to participants focusing on content, it does not indicate whether the two groups were immersed in different areas of interest of the film clips (i.e. the center versus the periphery).

The quality x enjoyment interaction was similar to that in Experiment 3, although in this experiment it was able to generalize beyond those film clips utilized in this study. There was a positive slope for enjoyable clips, however a flatter slope for neutral clips. It appears that participants were more accurate in rating quality, rating clips converted to 550 kbps lower on the MOS scale than those converted to 6000 kbps, whereas there was no difference for the rating of neutral clips. This would again indicate a benefit in selecting enjoyable clips and asking participants to focus on the quality of these clips in order to attain accurate quality ratings.

## General Discussion

### *Experiment 1*

In Experiment 1, film clips were divided based on their level of enjoyment. Enjoyable clips were liked more than neutral film clips. The main effect of enjoyment consistently found across all subsequent experiments demonstrated the power of film clip enjoyment on user ratings. As enjoyment increased, so did ratings of video quality.

*Experiment 2*

In Experiment 2, it was determined that both affect and immersion directly impacted video quality ratings, as they were both positively correlated with video quality (MOS). While liking did not partially mediate the relationship between immersion and video quality (MOS), it did for the affect-video quality (MOS) relationship. Therefore, video quality ratings were explained by participants' level of immersion as well as how much they liked the film clip. Because of the partial mediation, looking at why an individual likes a film clip is a better proxy than looking at their mood state. However, because of the lack of a true causal relationship, it possible that affect would mediate the relationship between liking and video quality, as the variables were so tied. The literature review indicated how individuals rely on their mood state when assessing their liking or satisfaction. Therefore, as liking only partially mediated the relationship, there may be benefits to understanding a user's affect, level of immersion, and liking in order to assess a more accurate subjective rating of video quality.

*Experiment 3*

In Experiment 3, video quality training increased participants' accuracy when rating video quality, as there was a positive slope from the lowest (550 kbps) to highest (6000 kbps) film clip type. Trained participants were better able to discriminate between different video qualities, whereas untrained participants were less able to do so. For the trained participants, liking no longer mediated the relationship between affect and video quality (MOS). A factor analysis also demonstrated that trained participants were able to relate variables to each other that should correlate (i.e., all variables assessing video quality fell onto one factor), whereas untrained participants did not discriminate when rating variables. This indicated that untrained participants succumbed to a halo effect, whereby they used heuristic processing because they lacked a mental model for the assessment of video quality. Their mood state guided their rating for all other variables such that all variables were highly intercorrelated. However trained participants had a mental model for video quality and were able to rely on that in lieu of a heuristic such as affect when rating items such as lighting, acting, and resolution. For these individuals, there was no presence of a halo effect. Training was able to remove the direct relationship between affect and video quality (MOS), as training partially moderated this relationship. These findings were in line with Forgas's affect infusion model (Forgas, 1995) – individuals who do not possess content knowledge of, in this case, video quality will rely on a heuristic and allow affect to color their judgments, whereas those with content knowledge will not let affect bias their judgments. Therefore, if a user researcher would like to increase the accuracy of subjective video quality assessment during product testing, one way to remove the bias of affect is to train users on video quality or seek out experts in lieu of novices. The affect infusion model is quite general: Regardless of the product being tested, training of the user on that product will remove biases due to affect.

The training in this experiment was quite minimal, however it was successful in increasing participants' mental models of video quality. Additionally, it is likely that one could rely on experts, who would have a much more complex mental model of video quality and whose accuracy may be even more demonstrative than that seen by trained undergraduates in Experiment 3. In user testing, there are scenarios where experts are already utilized in lieu of novices because their subject matter knowledge is desired. However, when experts are selected for user testing, their task becomes one of detection, where they simply find what is wrong with the system or product, working from the top down. This differs from the task of a novice, who attempts to notice anything that is disliked or challenging from the bottom up. If the majority of the end users do possess some expertise in regards to the product being tested, then it does make sense to test users who are knowledgeable. However, if the end users tend to be novices or if the product is completely novel, it begs the question of whether training users or selecting experts prior to testing is a best practice. Indeed, it is the case with many products, such as consumer electronics, that the average user will possess little to no technical knowledge about the product. While the feedback in regards to quality assessment may be more accurate when it comes from a trained or expert user, this feedback may skew that which would be received from the average consumer. If the majority of users would give high subjective ratings to the video quality of products that would otherwise be highly criticized by experts, this is important input to receive, as these lay consumers will be the ones ultimately purchasing the products. Therefore, it most certainly depends on the target market as well as the aspect(s) being assessed when making the decision of which users to select for product testing.

How should Experiment 3 have approached selecting experts prior to the task without revealing the purpose of the experiment? Participants were asked in the end of experiment questionnaire about their HD knowledge and what cables they utilized for connecting the video on their televisions. Could this information have been utilized to differentiate experts from novices? For example in Experiment 2, only 9 of 30 participants believed they were knowledgeable when it came to HD technology, and 9 of 30 acknowledged using HDMI or component cables, indicating that the rest were not receiving an HD quality image due to lower quality cables. There was very little overlap between those who believed they were knowledgeable and used better cables; therefore, knowledge was not supported with actual usage. This may help to explain why simply asking users about their habits or beliefs is not a perfect proxy in determining who is an expert, even moderately so, in assessing video quality. Even for those using the HDMI or component cables, they still relied heavily on liking and affect when assessing quality in the experiment. It may take a more in-depth conversation with a user to reveal content knowledge such as understanding of bit rate differences or video quality artifacts in order to locate competent users. While this method would suffice, it would not be advisable if the research being conducted required the user to be blind to the purpose of the experiment.

*Experiment 4*

In Experiment 4, focusing on quality increased participants' accuracy across the three levels of video quality. On the other hand, participants focusing on content did not discriminate well between the three levels of video quality, as there was not a positive slope from the lowest (550 kbps) to highest (6000 kbps) quality film clip. However, focus of attention did not moderate the direct relationship between content immersion and video quality – participants focusing on quality still rated immersion and video quality

(MOS) items similarly. This could have been due to the film clips, themselves. Many individuals explicitly stated that the film clips were enjoyable and that the film clip content did make it hard to focus on quality and stay on the task assigned to them. Additionally, each film clip lasted 2 minutes, and one can assess a film clip's quality in much less than 2 minutes. It is possible that a participant was able to assess the film clip quality and then opted to switch attention to the content, which would increase the participant's content immersion. In order to more accurately assess moderation here, it would be necessary to use eye tracking to determine if participants are focusing away from the center of the screen where the majority of the action is occurring. The length of the film clips should also be reduced, as it is possible to become immersed in less time for those focusing on content, while those focusing on quality would be less likely to become distracted from their own task. In regards to user testing, there are benefits to informing a user to focus on quality. As was demonstrated in Experiment 4, there would be more accurate ratings in video quality by telling the user to focus on video quality, as the user's attention would be devoted to the task of quality detection.

In this experiment, an early selection approach was assumed, indicating that the non-attended stimuli were not processed early on in perceptual processing (i.e. Broadbent, 1958; Treisman, 1969). Individuals focusing on content were performing well on the follow-up questions assessing performance on attention to content. Their performance on questions related to quality suffered either because they were unable or unwilling to focus on quality, selectively focusing on content. However, if a late selection approach was utilized (i.e. Deutsch & Deutsch, 1963; Norman, 1968), this would indicate that individuals had processed both content and quality in parallel, however they had chosen to dismiss the quality information before it entered their short term memory. This latter approach seems unlikely, as participants were asked if the non-assigned task (i.e. focus on quality) distracted them from their assigned task (i.e. content). Those focusing on content did not say that the film clip quality distracted them; therefore, it seems that they did not notice the quality to even answer the question. Along these lines, those focusing on quality did say that film clip content did, indeed, distract them from focusing on the task of evaluating quality. When asked why, many participants said that the entertaining clips, particularly, made it hard to keep attention fixed solely on quality.

Based on this, it seems that these participants were in more of a focused attention scenario as opposed to a selective attention scenario, as the nature of the content, due to the film clip enjoyment, was distracting users' attention away from quality assessment. If this was the case, this may imply two things. For one, it is possible that the interaction between condition (focus of attention) and film clip quality would have increased were the task guaranteed to have been a selection attention task. Those focusing on quality would have been even more accurate without the distractions of content. Second, there was potentially no moderation between immersion and quality based on focus of attention because those focusing on quality admitted that the film clip content was distracting. Over the course of the two-minute evaluation of a film clip, participants were able to focus on quality (demonstrated by their enhanced evaluations of film clip quality), however the distracting nature of the film clips raised their immersion levels as well. By performing this task with only neutral clips that would not divert the participants' attention away from quality assessment, it is possible that moderation would have been successful, and the task could have returned to a selective attention task. Shortening the task from two minutes to only a few seconds would be another solution to investigate if diverting attention to film clip quality could have moderated the relationship between immersion and video quality ratings. In this scenario, distraction from the content would have substantially decreased performance on quality evaluation, as there would have been much less time to view the film clip for visual artifacts.

While a divided attention scenario is also plausible, it is less likely due to the complexity of the task. No participant appeared to perform perfectly across both the content and quality questions that followed each film clip, and many focused their efforts on answering only the questions that applied to their experimental condition, ultimately skipping the other questions altogether, particularly as time progressed.

*Video Quality Discrimination*

For untrained participants and those who focused on content, ratings were more accurate for enjoyable clips compared to neutral clips, as seen with the quality x enjoyment interactions in both Experiments 3 and 4. This could be explained by the amount of movement in enjoyable clips compared to neutral clips. For example, many of the enjoyable clips were actions or comedies, where movement levels were high; whereas the neutral clips tended to be dramas, where there was considerably less movement. Video artifacts are much more noticeable when movement is higher (Corriveau, 2006), therefore accuracy would be increased with enjoyable clips simply because the blockiness, edges, et cetera, would have been more prominent in the lower quality video clips and less so in the higher quality clips. Therefore, it may be more crucial to find a clip with high levels of movement in lieu of high levels of enjoyment in order to achieve accuracy in ratings. However, there are some artifacts such as gradient blockiness that are particularly noticeable in video with low movement; therefore, low-motion clips may be useful tools for artifact detection as well.

In both Experiments 3 and 4, there was an inflation in quality ratings that was particularly noticeable for the lowest film clip quality (550 kbps) for participants who were neither trained nor focusing on quality. Essentially, these participants found this film clip quality to be quite satisfactory, rating it similarly to their rating of medium-quality film clips (1100 kbps). As the average viewer of media would likely fall into both groups, being both untrained and being likely to focus on the content of the medium being watched, it can be presumed that media viewed with a low bit rate will be perceived as higher in quality than it actually is. Industry professionals could interpret this as a way to save in bandwidth, providing lower quality media to an audience that will not notice. This is particularly beneficial when the consumer's hardware cannot accommodate a high bit rate. For example, Netflix selects a bit rate to provide to the consumer when he/she is streaming media based on his/her configuration. However, individuals are demanding higher quality technology, despite research backing up their lack of ability to discriminate between high and low quality video. Offering the choice to the consumer allows those who prefer HD and who can handle its size to select it, however those who do not know the difference or who may have a slower connection may select a lower quality video. For example, YouTube offers a variety of quality selections, from the default (360 pixels) to several high-definition selections (the maximum being 1080p).

*Limitations and Future Directions*

As the study was only able to capture subjective data, there would be benefits to adding in objective measurements via an eye tracker. While it is complex to utilize an eye tracker for video, an area of interest could capture whether participants are focusing on the content (directed to look in the center of the screen) or on the quality (directed to look at the periphery of the screen). There would also be benefits to drawing definite causal paths – the model in Figure 1 should be interpreted with caution. As many of the variables were so highly correlated it was hard to determine what variable preceded the other. Specifically, it was indicated that the causal direction went from affect to liking; however, the case could easily be made that liking predicts affect. Further, while the path analyses from mediation indicated that liking, affect, and immersion predicted quality ratings, it is possible to find a reverse causal effect, where quality would predict liking, affect, or immersion. While these reverse effects are often meaningless in research, they are plausible in this study. For example, a participant who views a clip that has high video quality may increase their mood, immersion, and/or enjoyment for the clip due to the quality, or reduce mood, immersion, and/or enjoyment when faced with a clip that has very low quality. The model also assumes that there are no other variables impacting quality that are correlated with either affect, immersion, or liking. Because it was hard to tease apart the variables along the spectrum of time, better work in ecological momentary assessment in order to receive feedback in regards to affect, liking, immersion, and quality as they are interpreted by the user would improve confidence in the model and its causal pathways. Further, Experiment 1 implied that very few film clips had been seen before by the undergraduate students who participated, as the films were not recent, however the subsequent experiments did reveal that some films (such as Oceans 11 and

Gattaca) were popular. There were benefits to removing data where a participant had seen the film clip before, however future research could put more effort into finding more obscure films. It may be beneficial to control for film clip familiarity when testing video quality in industry as well.

The ability to discern easily if affect or other ratings are independent is indeed muddled when groups are utilized in lieu of individual ratings. Therefore it should be noted that participants were run individually and not in groups or pairs on purpose, even when there were two computer stations in a room at a time. Many precautions were taken to make sure that individuals were unable to influence each other, such as using headphones for listening to the film clips which had the added bonus of blocking out ambient noise in the room. The two computer stations were also positioned so that the two participants could not see each others' screens. This was deemed important to help eliminate any contagion effect that could have occurred were individuals able to observe each other. This would indicate that one participant would mimic and then take on the affective state of the other participant (Barsade, 2002). Affecting a participant's mood ancillary to that induced from the film clip could act as a third variable that was otherwise unaccounted for. If the two participants were working as a group, mood would again factor in as a strong influence. Within a group, the leader's mood is particularly contagious (Sy, Cote, & Saavedra, 2005). Beyond mood, one's performance would also be influenced by the intelligence or personality of the other individual or by other social influence. Therefore, it could be expected that an expert's opinions would influence a novice's ratings. For example, Day, Arthur Jr., Bell, Edwards, Bennett Jr., et al. (2005) found dyadic team performance to increase on a complex computer game as team intelligence increased. Additionally, simply observing someone else who is training is beneficial in boosting your own performance (Shebilske, Jordan, Goettl, & Day, 1999). It can be concluded that being able to communicate with or watch an expert who is assessing video quality would be quite helpful in advancing one's own video quality accuracy, however it would not accurately capture one's own performance.

As this study focused on video quality, it ought to generalize to other products such as mobile devices, vehicles, or clothing. While individuals tend to use a heuristic such as brand name or price when assessing the quality of a product, training users on how to judge the quality of a particular product would increase their accuracy. For example, users could override their reliance on affect and more accurately judge vehicles or clothing by receiving training that would help them learn what to look for under the hood of a car or what type of stitching is more durable for a hemline. Additionally, having users focus on quality would help them to notice any flaws that may exist in a product, as they may otherwise be distracted by other items that do not impact quality such as color or body style, succumbing to inattentional blindness. By undergoing user testing and arriving at an accurate assessment of subjective quality before releasing a product to the market, one will help to eliminate the chance that a product is perceived differently by the population than it was during testing. Additionally, there is a high manufacturing cost associated with the production of all products, and detecting the point at which quality plateaus will enable a company to satisfy a consumer with a product that is perceived as high in quality while saving on the bottom line with a product that is not too costly to produce.

*Conclusions*

This study was able to reestablish the finding by Kortum and Sullivan (2010), as a positive correlation existed between liking and subjective quality ratings. In addition, both affect and immersion were positively correlated with subjective video quality ratings. Liking partially mediated the relationship between affect and video quality, however it did not do so for the relationship between immersion and video quality. Because of the increase in affect, participants rated all items highly, as a reflection of their good mood. This demonstration of a halo effect was eradicated when participants were trained on the assessment of video quality, as they were able to bypass reliance on their affective state and could instead more accurately rate video quality. Therefore, video quality training moderated the relationship between affect and video quality ratings. Participants who focused on video quality were still highly immersed in the content of the film clips, therefore focus of attention did not moderate the relationship between immersion and video quality ratings. However, both training and focus of attention did increase participants' accuracy when rating video quality. Therefore, both affect and immersion influence video quality ratings and there are benefits to training participants on quality and/or directing participants on the task of focusing on quality in order to increase their accuracy on video quality. Additionally, training has the added benefit of removing participants' reliance on their affective state during product assessment.

References

Anderson, D. R., & Kirkorian, H. L. (2006). Attention and television. In J. Bryant and P. Vorderer (Eds.), *Psychology of Entertainment* (pp. 35–54). Mahwah, NJ: LEA Press.

Baron, R. M., & Kenny, D. A., (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47,* 644–675.

Beatty, J. (1988). Pupillometric signs of selective attention in man. In G. C. Galbraith, M. L., Kietzman, and E. Donchin (Eds.), *Neurophysiology and Psychophysiology: Experimental and Clinical Aplications* (pp. 138–143). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Publishers.

Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal, 4,* 115–150.

Bracken, C. C. (2005). Presence and image quality: The case of high-definition television. *Media Psychology, 7,* 191–205.

Bracken, C. C. (2006). Perceived source credibility of local television news: The impact of television form and presence. *Journal of Broadcasting and Electronic Media, 50,* 723–741.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45,* 602–607.

Brasel, S. A., & Gips, J. (2008). Breaking through fast-forwarding: Brand information and visual attention. *Journal of Marketing, 72,* 31–48.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. *CHI 2004,* 1297–1300.

Bryant, J., & Comisky, P. W. (1978). The effect of positioning a message within differentially cognitively involving portions of a television segment on recall of the message. *Human Communication Research, 5,* 63–75.

Bryant, J., & Zillmann, D. (1977). The mediating effect of the intervention potential of communications on displaced aggressiveness and retaliatory behavior. In B. D. Ruben (Ed.), *Communication Yearbook 1* (pp. 291–306). New Brunswick, NJ: ICA-Transaction Press.

Cater, K., Chalmers, A., & Ward, G. (2003). Detail to attention: Exploiting visual tasks for selective rendering. *Eurographics Symposium on Rendering 2003, 14,* 270–281.

Compeau, L. D., Grewal, D., & Monroe, K. B. (1998). Role of prior affect and sensory cues on consumers' affective and cognitive responses and overall perceptions of quality. *Journal of Business Research, 42,* 295–308.

Corriveau, P. (2006). Video quality testing. In H. R. Wu and K. R. Rao (Eds.), Digital Video Image Quality and Perceptual Coding (pp. 125–153) Boca Raton, FL: CRC Press.

Cronin Jr., J. J., & Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *The Journal of Marketing, 56,* 55–68.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

Csikszentmihalyi, M. (1997). *Finding Flow: The Psychology of Engagement with Everyday Life.* New York, NY: Basic Books.

Csikszentmihalyi, M., & Rathunde, K. (1992). The measurement of flow in everyday life: Toward a theory of emergent motivation. In R. Dienstbier & R. M. Ryan (Eds.), *Nebraska Symposium on Motivation* (pp. 57–97). Lincoln, NE: University of Nebraska Press.

David, P., Horton, B., & German, T. (2008). Dynamics of entertainment and affect in a Super Bowl audience: A multilevel approach. *Communication Research, 35,* 398–420.

Day, E. A., Arthur Jr., W., Bell, S. T., Edwards, B. D., Bennett, W., Jr., Mendoza, J. L., & Tubré, T. C. (2005). Ability-based pairing strategies in the team-based training of a complex skill: Does the intelligence of your training partner matter? *Intelligence, 33,* 39–65.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Eliashberg, J., & Sawhney, M. S. (1994). Modeling goes to Hollywood: Predicting individual differences in movie enjoyment. *Management Science, 40,* 1151–1173.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117,* 39–66.

Forgas, J. P., & Moylan, S. (1987). After the movies: Transient mood and social judgments. *Personality and Social Psychology Bulletin, 13,* 467–477.

Forgas, J. P., & Tehani, G. (2005). Affective influences on language use: Mood effects on performance feedback by experts and novices. *Journal of Language and Social Psychology, 24,* 269–284.

Giese, J. L., & Cote, J. A. (2002). Defining consumer satisfaction. *Academy of Marketing Science Review, 2000*, 1-24.

Gouaux, C. (1971). Induced affective states and interpersonal attraction. *Journal of Personality and Social Psychology, 20,* 37–43.

Green, M. C. & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79,* 701–721.

Ha, J., & Jang, S. (2010). Effects of service quality and food quality: The moderating role of atmospherics in an ethnic restaurant segment. *International Journal of Hospitality Mangement, 29,* 520–529.

Hands, D. S. (2004). A basic multimedia quality model. *IEEE Transactions on Multimedia, 6,* 806–816.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science, 132,* 349–350.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143,* 1190–1192.

Hupe, J-M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision, 9,* 1–19.

Huynh-Thu, Q., & Ghanbari, M. (2008). Scope of validity of PSNR in image/video quality assessment. *Electronics Letters, 44,* 800–801.

Hwang, Y., & Kim, D. J. (2007). Customer self-service systems: The effects of perceived web quality with service contents on enjoyment, anxiety, and e-trust. *Decision Support Systems, 43,* 746–760.

International Committee for Information Technology Standards. ISO 9241-11. (1998). Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11. Guidance on usability.

Ivory, J. D., & Kalyanaraman, S. (2007). The effects of technological advancement and violent content in video games on players' feelings of presence, involvement, physiological arousal, and aggression. *Journal of Communication, 57,* 532–555.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*, 641–661.

Jumisko-Pyykko, S., & Hakkinen, J. (2005). Evaluation of subjective video quality of mobile devices. *MM'05,* 535–538.

Jung, Y., Perez-Mira, B., & Wiley-Patton, S. (2009). Consumer adoption of mobile TV: Examining psychological flow and media content. *Computers in Human Behavior, 25,* 123–129.

Karat, C-M., Karat, J., Vergo, J., Pinhanez, C., Riecken, D., & Cofino, T. (2002). That's entertainment! Designing streaming, multimedia web experiences. *International Journal of Human-Computer Interaction, 14,* 369–384.

Kim, J. & Rubin, A. M. (1997). The variable influence of audience activity on media effects. *Communication Research, 24,* 107–135.

Klimmt, C., & Vorderer, P. (2003). Media psychology "is not yet there": Introducing theories on media entertainment to the presence debate. *Presence, 12,* 346–359.

Kortum, P., & Sullivan, M. (2004). Content *is* king: The effect of content on the perception of video quality. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting,* 1910–1914.

Kortum, P., & Sullivan, M. (2010). The effect of content desirability on subjective video quality ratings. *Human Factors, 52,* 105–118.

Kozlowski, S. W. J., Kirsch, M. P., & Chao, G. T. (1986). Job knowledge, ratee familiarity, conceptual similarity, and halo error: An exploration. *Journal of Applied Psychology, 71,* 45–49.

Krugman, H. E. (1965). A comparison of physical and verbal responses to television commercials. *Public Opinion Quarterly, 29,* 323–325.

Kunar, M. A., Carter, R., Cohen, M., & Horowitz, T. S. (2008). Telephone conversation impairs sustained visual attention via a central bottleneck. *Psychonomic Bulletin & Review, 15,* 1135–1140.

LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 371–379.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review, 4,* 501–506.

Lin, P.-C., & Chien, L.-W. (2010). The effects of gender differences on operational performance and satisfaction with car navigation systems. *International Journal of Human-Computer Studies, 68,* 777–787.

Lombard, M., Ditton, T. B., & Weinstein, L. (2009). *Measuring (tele)presence: The Temple Presence Inventory.* Presented at the Twelfth International Workshop on Presence, Los Angeles, California, USA.

Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space, and Environmental Medicine, 78,* B165–B175.

Monroe, K. B., & Krishnan, R. (1985). The effect of price on subjective product evaluations. In J. Jacoby & J. C. Olson (Eds.), *Perceived Quality: How Consumers View Stores and Merchandise*. (pp. 209–232). Lexington, MA: Lexington Books.

Morgan, L. A. (1985). The importance of quality. In J. Jacoby and J. C. Olson (Eds.), *Perceived Quality: How Consumers View Stores and Merchandise.* (pp. 61–64). Lexington, MA: Lexington Books.

Nabi, R. L., & Krcmar, M. (2004). Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory, 14,* 288–310.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review, 75,* 522–536.

O'Brien, R. G. (1981). A simple test for variance effects in experimental designs. *Psychological Bulletin, 89,* 570–574.

Oliver, M. B., & Bartsch, A. (2010). Appreciation as audience response: Exploring entertainment gratifications. *Human Communication Research, 36,* 53–81.

Oxenfeldt, A. R. (1950). Consumer knowledge: Its measurement and extent. *The Review of Economics and Statistics, 32,* 300–314.

Oxman, J., & Pollock, M. (2007). 30 percent of U.S. households own an HDTV, CEA research finds. Retrieved from http://www.ce.org/Press/CurrentNews/press_release_detail.asp?id=11309 on January 14, 2010.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies, 59,* 185–198.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244.

Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research, 28,* 167–188.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160–174.

Qin, H., Rau, P-L. P., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human-Computer Interaction, 25,* 107–133.

Pascoe, G. C. (1984). Patient satisfaction in primary health care: A literature review and analysis. *Evaluation and Program Planning, 6,* 185–210.

Riley, M. J., & Richardson, I. E. G. (1997). *Digital Video Communications.* Norwood, MA: Artech House.

Rimmell, A. N., Mansfield, N. J., Hands, D. (2008). The influence of content, task and sensory interaction on multimedia quality perception. *Ergonomics, 51,* 85–97.

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Human Perception and Performance, 7,* 275–285.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57,* 493–502.

Sato, W., Noguchi, M., & Yoshikawa, S. (2007). Emotion elicitation effect of films in a Japanese sample. *Social Behavior and Personality, 35,* 863–874.

Schubert, E. (2010). Affective, evaluative, and collative responses to hated and loved music. *Psychology of Aesthetics, Creativity, and the Arts, 4,* 36–46.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? The informative function of affective state. In K. Fiedler and J. Forgas (Eds.), *Affect, Cognition, and Social Behavior* (pp. 44–62). Lewiston, NY: C. J. Hogrefe.

Schwarz, N., Strack, F., Kommer, D., & Wagner, D. (1987). Soccer, rooms, and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific domains. *European Journal of Social Psychology, 17,* 69–79.

Shapiro, S., & Krishnan, H. S. (2001). Memory-based measures for assessing advertising effects: A comparison of explicit and implicit memory effects. *Journal of Advertising, 30,* 1–13.

Shebilske, W. L., Jordan, J. A., Goettl, B. P., & Day, E. A. (1999). Cognitive and Social Influences in Training Teams for Complex Skills. *Journal of Experimental Psychology: Applied, 5,* 227–249.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science, 7,* 301–305.

Sinclair, R. C. (1988). Mood, categorization breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations. *Organizational Behavior and Human Decision Processes, 42,* 22–46.

Steenkamp, J.-B. E. M. (1990). Conceptual model of the quality perception process. *Journal of Business Research, 21,* 309–333.

Sunstedt, V., Debattista, K., Longhurst, P., Chalmers, A., & Troscianko. (2005). Visual attention for efficient high-fidelity graphics. *Proceedings of the 21st Spring Conference on Computer Graphics, 21*, 162–168.

Sy, T., Cote, S., & Saavedra, R. (2005). The contagious leader: Impact of the leader's mood on the mood of group members , group affective tone, and group processes. *Journal of Applied Psychology, 90,* 295–305.

Tan, F. B., & Chou, J. P. C. (2008). The relationship between mobile service quality, perceived technology compatibility, and users' perceived playfulness in the context of mobile information and entertainment services. *International Journal of Human-Computer Interaction, 24,* 649–671.

Wang, Y. J., & Minor, M. S. (2008). Validity, reliability, and applicability of psychophysiological techniques in marketing research. *Psychology & Marketing, 25,* 197–232.

Weber, R., Tamborini, R., Westcott-Baker, A., & Kantor, B. (2009). Theorizing flow and media enjoyment as cognitive synchronization of Attentional and reward networks. *Communication Theory, 19,* 397–422.

Weir, C. S., Anderson, J. N., & Jack, M. A. (2006). On the role of metaphor and language in design of third party payments in eBanking: Usability and quality. *International Journal of Human-Computer Studies, 64,* 770–784.

Wessman, A. E. (1979). Moods: Their personal dynamics and significance. In C. E. Izard (Ed.), *Emotions in Personality and Psychopathology* (pp. 73–102). New York, NY: Plenum Press.

Westbrook, R. A., & Oliver, R. L. (1991). The dimensionality of consumption emotion patterns and consumer satisfaction. *The Journal of Consumer Research, 18,* 84–91.

Wiles, M. A., & Danielova, A. (2009). The worth of product placement in successful films: An event study analysis. *Journal of Marketing, 73,* 44–63.

Wissmath, B., Weibel, D., & Groner, R. (2009). Dubbing or subtitling? Effects on spatial presence, transportation, flow, and enjoyment. *Journal of Media Psychology, 21,* 114–125.

Xiao, F. (2000). DCT-based video quality evaluation. *Final Project for EE392J,* 1–10.

Yeung, C. W. M., & Wyer Jr., R. S. (2004). Affect, appraisal, and consumer judgment. *Journal of Consumer Research, 31,* 412–424.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.

Zhang, T., Rau, P.-L. P., & Salvendy, G. (2010). Exploring critical usability factors for handsets. *Behavior & Information Technology, 29,* 45–55.

Zillmann, D. (1988). Mood management: Using entertainment to full advantage. L. Donohew, H. E. Sypher, and E. T. Higgins (Eds.), *Communication, Social Cognition, and Affect* (pp. 147–171). Hillsdale, NJ: LEA Publishers.

Zillmann, D. (2003). Theory of affective dynamics: Emotions and moods. In J. Bryant, D. Roskos-Ewoldsen & J. Cantor (Eds.), *Communication and Emotion: Essays in Honor of Dolf Zillmann* (pp. 533–567). Mahwah, NJ: Erlbaum.

Zillmann, D., & Bryant, J. (1994). Entertainment as media effect. In J. Bryant and D. Zillmann (Eds.), *Media Effects: Advances in Theory and Research* (pp. 437–461). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

# Appendix A

## Film clips utilized in Experiment 1

| Movie Title | Release Year | Chapter Used for Film Clip |
|---|---|---|
| 28 Days Later | 2002 | 2 |
| Ace Ventura: Pet Detective | 1994 | 14 |
| Ace Ventura: Pet Detective | 1994 | 22 |
| Analyze This* | 1999 | 17 |
| Boondock Saints, The | 1999 | 9 |
| Cruel Intentions* | 1999 | 2-3 |
| Deep End, The* | 2001 | 23 |
| Fargo* | 1996 | 10 |
| Fifth Element, The | 1997 | 19 |
| Finding Forrester | 2000 | 4 |
| French Kiss | 1995 | 14 |
| Gattaca* | 1997 | 20 |
| Go | 1999 | 16 |
| Girlfight* | 2000 | 27 |
| Guarding Tess* | 1994 | 1 |
| Guarding Tess | 1994 | 18 |
| Hudsucker Proxy, The* | 1994 | 31 |
| Impostors, The* | 1998 | 8 |
| Impostors, The | 1998 | 11 |
| Jackal, The | 1997 | 19 |
| Keeping the Faith | 2000 | 26 |
| Kissing Jessica Stein | 2001 | 21 |
| Lost Highway | 1997 | 5 |
| Lost Highway | 1997 | 10 |
| Mars Attacks! | 1996 | 6 |
| Mercury Rising | 1998 | 14 |
| Moulin Rouge! | 2001 | 2 |
| Murder in the First* | 1995 | 9 |
| Murder in the First | 1995 | 22 |
| Ocean's Eleven* | 2001 | 23 |
| Road to Perdition | 2002 | 5 |
| Siege, The* | 1998 | 4 |
| Stargate | 1994 | 7 |
| What Dreams May Come | 1998 | 16 |

*Film clips selected for use in Experiments 2-4.

# Appendix B

## Experiment 1 Film Clip Questionnaire

Film Clip # _______

1) Have you seen this film before?

Yes No Not Sure

2) If yes, what is the name of this film?

______________________________________________

3) How much did you like the film clip?

(not at all) (very much)

1 2 3 4 5 6 7

# Appendix C

## Experiment 1 & 2 End of Experiment Questionnaire

1) What is your gender? (circle one)

Male Female

2) What is your age? ______

3) What types of movies do you typically enjoy watching? (circle all that apply)

Action Comedy Drama Horror Science Fiction War Western

4) What types of movies do you typically dislike watching? (circle all that apply)

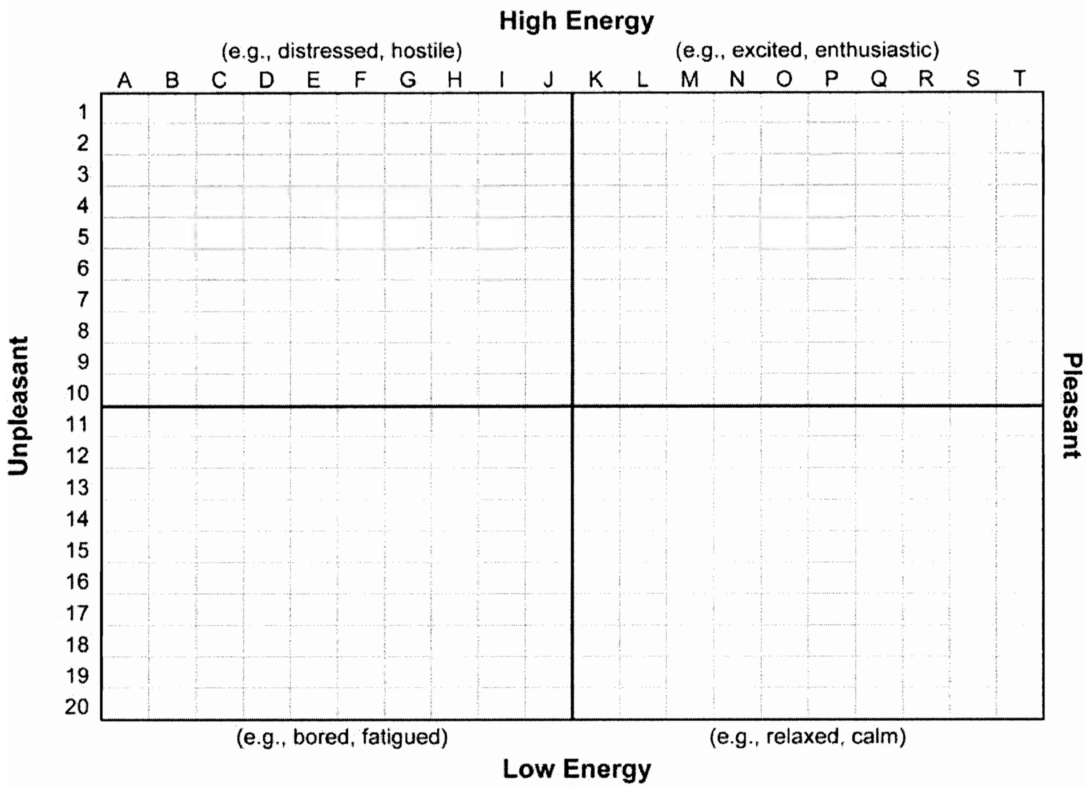
Action Comedy Drama Horror Science Fiction War Western

5) Is your cable service high-definition (1080i or 1080p)?

Yes No I'm not sure

# Appendix D

## Affect Grid



# Appendix E

## Immersion Scale

| | |
|---|---|
| 1) To what extent did you feel mentally immersed in the experience? | Not at all Very much<br>1 2 3 4 5 6 7 |
| 2) How involving was the experience? | Not at all Very much<br>1 2 3 4 5 6 7 |
| 3) How completely were your senses engaged? | Not at all Very much<br>1 2 3 4 5 6 7 |
| 4) To what extent did you experience a sensation of reality? | Not at all Very much<br>1 2 3 4 5 6 7 |
| 5) How relaxing or exciting was the experience? | Very relaxing Very exciting<br>1 2 3 4 5 6 7 |
| 6) How engaging was the story? | Not at all Very much<br>1 2 3 4 5 6 7 |

# Appendix F

## Experiment 2 Film Clip Questionnaire

1) Have you seen this film before? □ Yes □ No □ Not Sure

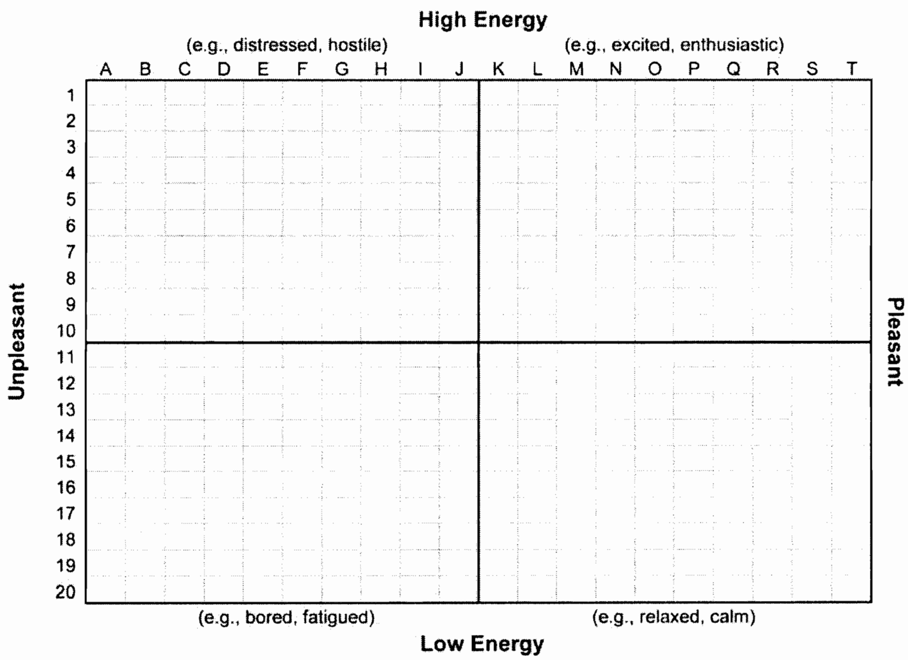
1a) If yes, what is the name of this film (if you remember)?

______________________________________

2) What genre best categorizes the film clip? □ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction □ Suspense □ War □ Western □ Other

______________

3) Please rate how you are feeling RIGHT NOW in the grid below.




4) How would you rate the overall video quality of the picture for this film clip? □Excellent □Very Good □Good □Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad

5) How much did you like the film clip? Please circle a number from 1 to 7 below.
(Not at all) 1 2 3 4 5 6 7 (Very much)

6) What Motion Picture Association of America (MPAA) rating would you give to this film clip? □ G □ PG □ PG-13 □ R □ I am not familiar with this rating system

| |
|---|
| 7) How well do you think this film did at the box office during its opening week?<br>(Worse than predicted) 1 2 3 4 5 6 7 (Better than predicted) |
| 8) To what extent did you feel mentally immersed in the film clip?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 9) How involving was the film clip?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 10) How completely were your senses engaged?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 11) To what extent did you experience a sensation of reality?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 12) How relaxing or exciting was the film clip?<br>(Very relaxing) 1 2 3 4 5 6 7 (Very exciting) |
| 13) How engaging was the story?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 14) How was the resolution in the film clip?<br>(Very blurry) 1 2 3 4 5 6 7 (Very sharp/clear) |
| 15) How would you rate the overall quality of the acting? □Excellent □Very Good □Good<br>□Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad |
| 16) What did you think of the cinematographer's choice for lighting for this film clip?<br>(Not good) 1 2 3 4 5 6 7 (Excellent) |
| 17) How curious are you as to what happens next in the film clip?<br>(Not curious) 1 2 3 4 5 6 7 (Very curious) |
| 18) How would you rate the overall audio quality of the film clip? □Excellent □Very Good<br>□Good □Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor<br>□Bad |
| 19) Have you seen any of the actors from the film clip in other movies/TV shows?<br>□ Yes □ No □ Not Sure |

# Appendix G

## Video Quality Assessment Overview

An Overview to Video Quality Assessment

### Introduction

- When you watch digital media on the TV or computer, you may occasionally notice degradation or artifacts when the media is not very high quality.
- This overview will teach you what these artifacts are so that you can identify them in the film clips you will soon watch.

## Resolution

- Brightness and contrast will make a film clip with a higher bit rate look sharper and therefore less blurry.

Low Quality High Quality

## Pixel Drop-Out

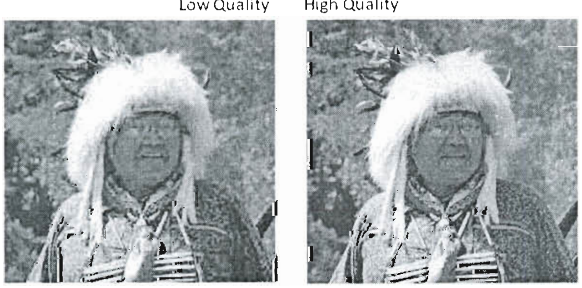
- A pixel may be missing from a low-quality image
- Can appear black, white, or any color


Before Pixel Drop-out After Pixel Drop-out

## Blockiness

- Lower quality clips may appear less smooth and blocks can appear, such as around the man's face

Low Quality High Quality



## Motion Blur

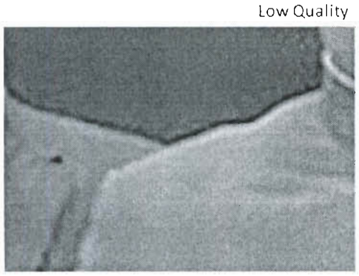
- When an object is moving, it looks blurrier if the quality is worse

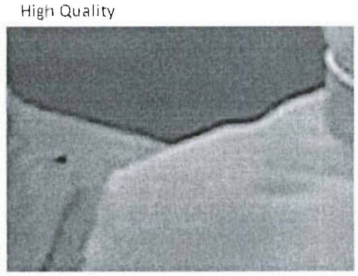


## Mosquito Noise

- Edges may appear blurry or speckled as image quality gets worse, and this is more noticeable as the object moves.

Low Quality High Quality





## Summary

- Resolution
  - Brightness & contrast are worse with low quality
- Pixel Drop-Out
  - A missing pixel affects the image
- Blockiness
  - Quality not as smooth because of blocks
- Motion Blur
  - Moving images are blurry
- Mosquito Noise
  - Edges of objects look blurry, especially when moving

# Appendix H

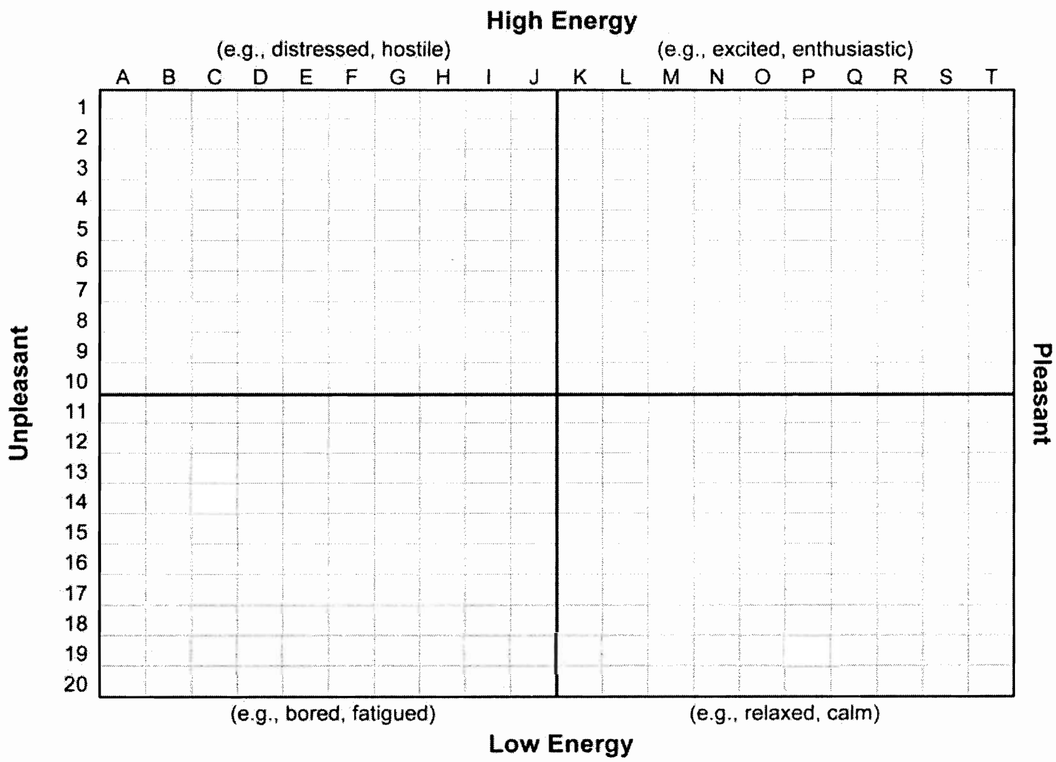
## Experiment 3 Film Clip Questionnaire

1) Have you seen this film before? □ Yes □ No □ Not Sure

1a) If yes, what is the name of this film (if you remember)? ______________________________

2) What genre best categorizes the film clip? □ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction □ Suspense □ War □ Western □ Other ____________

3) Please rate how you are feeling RIGHT NOW in the grid below.




4) How much did you like the film clip? Please circle a number from 1 to 7 below.
(Not at all) 1 2 3 4 5 6 7 (Very much)

5) How would you rate the overall video quality of the picture for this film clip? □Excellent □Very Good □Good □Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad

6) How would you rate the overall audio quality of the film clip? □Excellent □Very Good □Good □Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad

| |
|---|
| 7) How curious are you as to what happens next in the film clip?<br>(Not curious) 1 2 3 4 5 6 7 (Very curious) |
| 8) How were the edges in the film clip<br>(Very blurry) 1 2 3 4 5 6 7 (Very sharp) |
| 9) To what extent did you feel mentally immersed in the film clip?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 10) How involving was the film clip?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 11) How completely were your senses engaged?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 12) To what extent did you experience a sensation of reality?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 13) How relaxing or exciting was the film clip?<br>(Very relaxing) 1 2 3 4 5 6 7 (Very exciting) |
| 14) How engaging was the story?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 15) How well do you think this film did at the box office during its opening week?<br>(Worse than predicted) 1 2 3 4 5 6 7 (Better than predicted) |
| 16) How pixelated was the film clip?<br>(Very blocky) 1 2 3 4 5 6 7 (Very smooth) |
| 17) How would you rate the overall quality of the acting? □Excellent □Very Good □Good<br>□Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad |
| 18) How was the resolution in the film clip?<br>(Very blurry) 1 2 3 4 5 6 7 (Very sharp) |
| 19) What did you think of the cinematographer's choice for lighting for this film clip?<br>(Not good) 1 2 3 4 5 6 7 (Excellent) |
| 20) Have you seen any of the actors from the film clip in other movies/TV shows?<br>□ Yes □ No □ Not Sure |

# Appendix I

## Experiment 3 End of Experiment Questionnaire

1) What is your gender?

□ Male □ Female

2) What is your age? ______

3) What types of movies do you typically enjoy watching? (mark all that apply)

□ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction □ Suspense □ War □ Western □ Other ______________________

4) What types of movies do you typically dislike watching? (mark all that apply)

□ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction □ Suspense □ War □ Western □ Other ______________________

**Questions 5-8 test your knowledge of video and audio quality. Please circle the answer that you believe fits the question best.**

5) "Mosquito Noise" occurs when:

a) The video clip looks very grainy

b) There is a humming over the auditory channel

c) The edges of objects are fuzzy

d) The speakers make a crackling noise

6) If a spot on the film clip appeared to be missing or differed in color from the surrounding video, this would be an example of:

a) Pixel Drop-Out

b) Hue Desaturation

c) Video Lossiness

d) Damaged Video File

7) Motion blur occurs when:

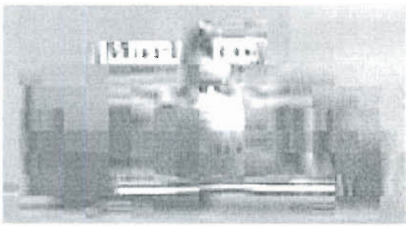
a) The videographer moves his/her camera across an object too quickly

b) A computer's video card cannot handle the fast-paced motion in a video

c) A low quality camera cannot capture a moving object without it appearing blurry

d) Objects look blurrier if a video is low quality

8) The image below has what video artifact:



a) Pixelization

b) Motion Smear

c) Blockiness

d) Video Lossiness

9) Do you prefer to view media in high-definition?

(Not at all) 1 2 3 4 5 6 7 (Very much)

10) If you own a computer, does it have high-definition (HD) video capabilities? It would have an HDMI, DVI, or mini DVI port if it had an HD video card.

□ Yes □ No □ Not Sure □ I do not own a computer

11) If you own a TV, is it HD-ready?

□ Yes □ No □ Not Sure □ I do not own a television

12) If you have cable service, do you subscribe to the high-definition channels?

□ Yes □ No □ Not Sure □ I do not have cable service

13) How do you connect the video from your cable box to your television?

□ A/V Cable (yellow) □ Component Cable (red, green, and blue) □ S-Video Cable □ HDMI Cable □ I do not know □ N/A □ Other ____________________

14) How knowledgeable do you think you are when it comes to HD technology?

(Not at all) 1 2 3 4 5 6 7 (Extremely)

15) Do you think video quality is an important factor for the gaming/media experience?

(Not at all) 1 2 3 4 5 6 7 (Very much)

# Appendix J

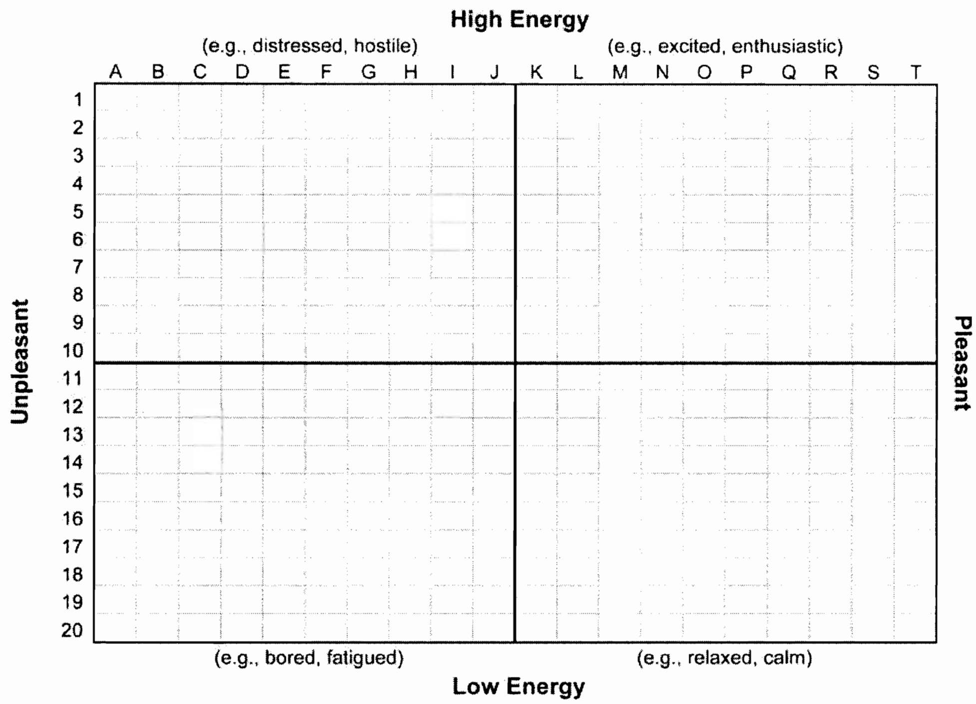
## Experiment 4 Film Clip Questionnaire

1) Have you seen this film before? □ Yes □ No □ Not Sure

1a) If yes, what is the name of this film (if you remember)? ______________________________

2) What genre best categorizes the film clip? (select ONE) □ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction □ Suspense □ War □ Western □ Other ____________

3) Please rate how you are feeling RIGHT NOW in the grid below.




4) How much did you like the film clip? Please circle a number from 1 to 7 below.
(Not at all) 1 2 3 4 5 6 7 (Very much)

5) How would you rate the overall quality of the acting for this film clip? □Excellent □Very Good □Good □Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad

6) How would you rate the overall video quality for this film clip? □Excellent □Very Good □Good □Somewhat Good □Fair □Somewhat Poor □Poor □Very Poor □Bad

7) To what extent did you feel mentally immersed in the film clip?
(Not at all) 1 2 3 4 5 6 7 (Very much)

8) How involving was the film clip?
(Not at all) 1 2 3 4 5 6 7 (Very much)

9) How completely were your senses engaged?

| |
|---|
| (Not at all) 1 2 3 4 5 6 7 (Very much) |
| 10) To what extent did you experience a sensation of reality?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| 11) How relaxing or exciting was the film clip?<br>(Very relaxing) 1 2 3 4 5 6 7 (Very exciting) |
| 12) How engaging was the story?<br>(Not at all) 1 2 3 4 5 6 7 (Very much) |
| **The following questions pertain to QUALITY. If you do not know an answer, you may leave the question blank.** |
| 13) How was the resolution of the words on the building signs that the suspect walks past?<br>□Very Blurry □Somewhat Blurry □Fair □Somewhat Sharp □Very Sharp □Did Not Notice |
| 14) How pixelated were the clothes on the clotheslines that the suspect runs underneath?<br>□Very Blocky □Somewhat Blocky □Fair □Somewhat Smooth □Very Smooth □Did Not Notice |
| 15) How pixelated was the subway/train that the suspect ran underneath?<br>□Very Blocky □Somewhat Blocky □Fair □Somewhat Smooth □Very Smooth □Did Not Notice |
| 16) How clear were the edges of the gate that the suspect jumped over?<br>□Very Blurry □Somewhat Blurry □Fair □Somewhat Clear □Very Clear □Did Not Notice |
| 17) How was the resolution of the bricks on the wall in the alleyway?<br>□Very Blurry □Somewhat Blurry □Fair □Somewhat Sharp □Very Sharp □Did Not Notice |
| **The following questions pertain to CONTENT. Please answer the questions in the remaining space in the box. If you do not know an answer, you may leave the question blank** |
| 18) What object(s) was/were the passenger holding in the car? |
| 19) What color jacket was the driver wearing? |
| 20) What kind of car did the suspect run into? |
| 21) What color shoes was the suspect wearing? |
| 22) What color bag was the suspect holding? |

# Appendix K

## Experiment 4 End of Experiment Questionnaire

1) What is your gender?

□ Male □ Female

2) What is your age? ______

3) What types of movies do you typically enjoy watching? (mark all that apply)

□ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction
□ Suspense □ War □ Western □ Other ____________________

4) What types of movies do you typically dislike watching? (mark all that apply)

□ Action □ Comedy □ Drama □ Horror □ Romance □ Science Fiction
□ Suspense □ War □ Western □ Other ____________________

5) Do you prefer to view media in high-definition?

(Not at all) 1 2 3 4 5 6 7 (Very much)

6) If you own a computer, does it have high-definition (HD) video capabilities? It would have an HDMI, DVI, or mini DVI port if it had an HD video card.

□ Yes □ No □ Not Sure □ I do not own a computer

7) If you own a TV, is it HD-ready?

□ Yes □ No □ Not Sure □ I do not own a television

8) If you have cable service, do you subscribe to the high-definition channels?

□ Yes □ No □ Not Sure □ I do not have cable service

9) How do you connect the video from your cable box to your television?

□ A/V Cable (yellow) □ Component Cable (red, green, and blue) □ S-Video Cable □ HDMI Cable □ I do not know □ N/A □ Other ______________________

10) Do you have other components connected to your TV/entertainment center (i.e. DVD player, gaming systems...)

□ Yes □ No □ Not Sure □ Not Applicable

12) How knowledgeable do you think you are when it comes to HD technology?

(Not at all) 1 2 3 4 5 6 7 (Extremely)

13) Do you think video quality is an important factor for the gaming/media experience?

(Not at all) 1 2 3 4 5 6 7 (Very much)

14) How easy was it to remember the goal you were given at the beginning (focus on quality or content)

(Very hard) 1 2 3 4 5 6 7 (Very easy)

15) Did the presence of questions on content/quality (whichever you were not told to focus on) lead you to focus on content/quality more than you otherwise would have (i.e. maybe you didn't want to leave them blank)?

(Not at all) 1 2 3 4 5 6 7 (Very much)

**If you were told to focus on quality, please answer questions 16 & 16a. Otherwise, you may skip them:**

16) Did the content of the film clips affect your ability to focus on video quality?

□ Yes □ No □ It depended on the film clip

16a) If you answered "It depended on the film clip", what factors made it hard/easy to focus on quality?

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________