Universiteit Utrecht

Utrecht University
18-08-2021

The influence of screen size and personality on the emotional response to film

A.F. Blok (5701880)
Supervisor: E. van den Broek
Daily supervisor: L. Arts
Second reader: A. Feelders

Master thesis, Utrecht University
Msc Human Computer Interaction (HCI)
Department of Information and Computing Sciences
Graduate School of Natural Sciences (GSNS)
Phase 1: 15 ECT, Phase 2: 30 ECT
INFOMHCI2

Abstract

Emotions guide us through our lives and are influenced by stimuli around us. How media stimuli, such as films, are presented on a device may influence the emotional response. Stimuli can be presented on various screen sizes, where on a larger screen, objects can appear closer and have a greater emotional response compared to smaller screens. An emotional response differs per the individual due to each individual's personality that plays an important role in the emotional response. This Thesis investigates the relationship between screen size, personality, and emotional response towards film stimuli. A survey was conducted with 72 participants in natural settings that examined the emotional response towards film clips presented on three different screen size conditions, TV, laptop, and mobile. The emotional response was measured in the two emotion dimensions of valence and arousal. The results suggest that the film clips were effective in eliciting the intended emotional states and the temporal spacing between the film clips was shown to be effective in decreasing the emotional states with 20 seconds of black screen. The significant results confirmed that the experiment's design was effective in eliciting emotional states and decreased the carryover effect between those emotional states. In addition, the results suggest that different screen size categories and personality traits, extraversion and neuroticism, have no effect on the intensity of the emotional response to film clips. The non-significant result was hypothesized to be caused by the screen resulting field of view, which was estimated to be roughly similar across the screen size conditions. The hypothesis arose that field of view might be a more determining factor than screen size for the emotional response to emotional stimuli concerning stimuli presented through a media form.

Contents

1 Introduction

These days it is not unusual to own numerous devices that give multiple options to view media content on, such as film. Each device has its own characteristic, one of which is mobility. With increasing mobility the screen size is reduced and restrict the space on which content is displayed on. Devices that had small screens grew in size to accommodate more functionalities on the screen. Where decades ago TV was the sole choice of watching media content, laptop and mobiles are now popular alternatives [49] and increasingly more films are streamed on mobile phones and laptops [81]. The interaction with these devices happens daily and can influence emotional responses through the design of devices and the way media stimuli is portrayed on a device. The media form on which stimuli is presented can be something that usually goes unnoticed with daily interactions, and choices to use certain devices are often made out of practical considerations. Such as using a mobile when moving around. Although, certain characteristics in the media forms can have the potential to influence us while we interact with those devices. The interaction may be adjusted to fit the objective of specific media by being aware of the emotional responses certain media can elicit.

It is critical to understand how content is delivered when emotion is a major notion. When content is presented on different screen sizes it may be perceived differently, which may impact the viewer's emotional response. Media with emotion as central concept, such as film, where emotions are essential to provide entertainment [101] or where emotion is needed as an activator for behavior, may be have a powerful motivator tool, to use in media, to persuade or teach. Emotion has the ability to chance the content of cognitive memory by cueing memories and thoughts that are associated with the motivational effect. Because emotion impacts perception and cognition, each emotion effects behavior and actions [48]. Especially arousal has an effect on memory performance, where highly arousing content is better remembered than low arousing content [15], and high arousal leads to increase in involvement [80]. The persuasion level of messages can be altered through specific emotions by putting overtones of certain emotions in a message [23]. Therefore, factors in media form that might impact the emotional response are important be considered for their function to be able to activate behaviour or influence the perception and cognition. Viewing media content on different screen sizes may impact on the emotional response, and consequently influence memory, involvement, behavior and persuasion.

Each individual has their unique personality, which may cause differences in emotional states that arise from a virtual experience. Emotion and personality affect human interactions [63], just as human computer interaction can affect emotion [62]. There is currently little data on how screen size and personality influence the emotional response. Either personality is left outside the scope or the experiment is conducted in a lab setting. While interaction with media content on devices are daily occurrences in our lives this has been left unexplored. This thesis aims to evaluate the effect on how screen size and personality influence the emotional response while watching film clips by answering the following research question: "Do personality and screen sizes have an effect on the intensity of the emotional state elicited by film stimuli?".

To provide some insight in the current state of research, the next section discusses the key works in different emotion theories, the influence of personality on emotions, and certain aspects of screen size that play a role in the emotional response. Based on the literature study a study experiment is designed to be conducted in natural settings, to evaluate whether screen size and personality have an influence on emotion intensity. A description of the experiment is presented in the methodology section. The results that were gathered are presented in the results section and interpreted in the discussion section. The discussion section also entails the limitations and opportunities of the current study and provides possibilities for further research.

2 Related work

2.1 Emotion

At the present moment, there is not an universally recognized theory that explains the functions of emotion, therefore the concepts of emotion also vary. A typical description is that emotion is a response to a situational event. That event is analysed by the individual for its meaning, hence not every event receives the same response since interpretation can differ per individual. Therefore, emotion is not solely the response to the event but also depends on the relationship an individual has with the situational event [36]. A common description of emotion is that it consists of the components subjective feeling, psychological response, and a set of action tendencies. These components cause humans to respond to an event that is biologically and/or individually significant [78, 74, 76]. The subjective feeling is how an individual perceives his current consciousness state, the psychological response is the body's automatic reaction to a stimulus and the action tendencies are the urge to execute expressive or instrumental behavior that is characterized by different emotions [36]. Thereby, differentiate the subjective feeling and psychological response of just feeling pleasant or unpleasant. Besides subjective feeling, psychological response, and action tendencies, Scherer [90] recognized two additional components of emotion, a cognitive component, such as appraisal and a motor expression component, like face, voice, gestures. There are a total of five components that have been established to be uniform for each individual. Two emotion theories, the categorical emotion theory and the dimensional emotion theory, that represent emotion in a theoretical framework are discussed.

2.1.1 Categorical emotion theory

Categorical theories regard each emotion as a discrete classification. The emotions are classified as independent entities and are typically distinguished by a smaller set of basic emotions created from a larger pool of complex emotions. A well known basic emotion set is designed by Ekman [28]. He defined six basic emotions: happiness, surprise, fear, sadness, anger and disgust, with specific facial compositions. Characteristics of basic emotions are inheritable, cross-cultural, and expressed by specific physiological reaction patterns and facial composition [46, 47, 29]. Hence, the basic emotions are hardwired in the biology of humans and as a result everyone possess these emotions that have distinct biobehavioral signatures. Whereas, complex emotions are learned, socially and culturally shaped, evolutionary new, and typically expressed by combinations of basic emotions along their characterized response physiological pattern. Complex emotions have the tendency to be influenced by language usage and emerge later in development [78, 12].

Emotions are described and expressed with words that are familiar and used in their everyday language. Language has a powerful role in emotion perception since it gives conceptual knowledge that influences the emotion perception [9]. The subjective experience is represented with a semantic term, that creates easy distinguishable categories for the emotion to be classified in [90]. The emotion term that is used for people to describe their subjective feeling is emotion recognition, the awareness of recognizing the emotional state that they are in [55]. Emotion recognition presumes that emotions are discrete events that fall in certain categories that are labeled with semantic terms [7] for recognition in other people but also in ourselves.

The basic emotion theory is prominent in psychological emotion theories but also criticised. In natural settings the basic emotion theory, with the associated signature facial composition, does not always hold up. A meta-analysis showed that the specific facial compositions that characterize a basic emotion in natural settings was only encountered in a minority of cases, with the exception of amusement [33]. Another critic of basic emotion categories is that researchers have a varying number of discrete qualifications of emotion, ranging from 6 to 22 emotions [17]. Further, emotion terms can differ across cultures and the same emotion terms can have a different value across cultures [93].

2.1.2 Dimensional theory

In contrast to categorical theories, dimensional theories consider each emotion as a point or zone within a complex space [86]. The emotions are measured along a continuous defined axis. Rather than emotion being classified in categories they are organized in a circular arrangement represented by two axis that represent two dimensions pleasantness - unpleasantness (valence), and arousal - relaxation (arousal), see Figure 1. Each emotion has these two qualities of valence and arousal [7], and these dimensions are the building blocks of the dimensional theory and of the core affect [87]. The dimensional model allows to measure between distinct emotional states and between the level of that given state. Russell [86] mapped 28 emotions terms as points into the VA space. The boundaries between the categories of emotional states are fuzzy because the emotional states have elements that can overlap with one or more categories. The fuzzy boundaries are specific to particular groups of categories. While each emotional state category is associated with a specific affective dimension rating, these ratings vary smoothly along the axis associated with each emotion category and do not shift abruptly [22].

Valence represents a degree of emotion's pleasantness ranging from unpleasantness to pleasantness. Arousal represents a degree of intensity or activation, ranging from deactivation to activation [78]. In addition to valence and arousal, the dimensional model has an extension to a third dimension: dominance that reflects the level of control of the emotional state, from submissive to dominant [88]. The dominance dimension makes it possible to distinguish from dominant and submissive, for instance angry from anxious,

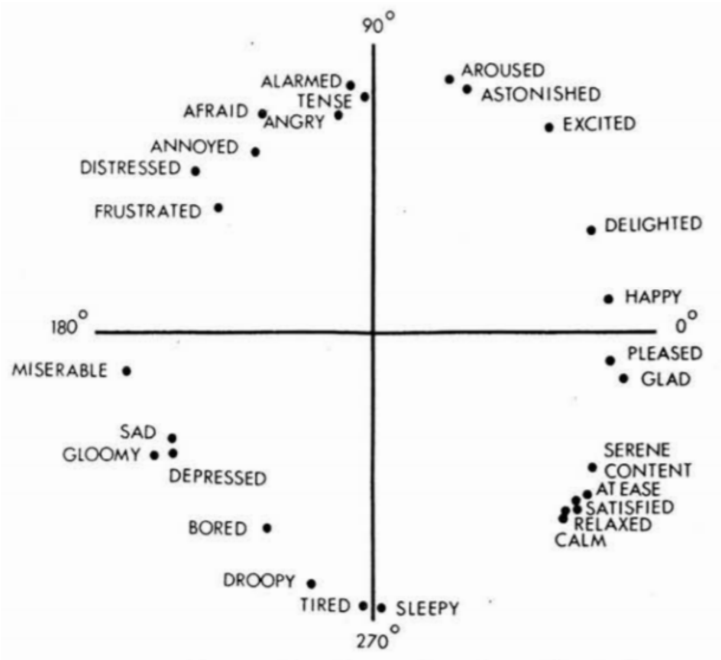


Figure 1: The circumplex model of [86], with the discrete emotion mapped into the two dimensional Valence-Arousal model, with on the x-axis valence and on the y-axis arousal.

alert from surprise, and relaxed from protected. The three dimensions can be viewed as crucial for an adequate description of emotion, and it is even argued that three dimensions are insufficient to account for all the emotion components [35]. The valence and arousal dimensions are not claimed to be sufficient to differentiate equally between all emotions. Nonetheless, dominance is difficult to measure and is often omitted, indicating that valence and arousal are sufficient to represent certain perceived emotion. The two dimensions can capture global and basic elements of emotions [40, 72, 14, 27, 108].

2.1.3 Combination in theories

Longstanding controversy in the field of emotion research is about whether emotions are better conceptualized in terms of discrete categories, or in continuous dimensions. The dimensional view proposes that a basic emotion such as fear emerges from a combination of negative valence and high arousal. While the discrete theory classify emotion in discrete semantic terms. The latter appears to be easier due to the simplicity that humans recognize emotional states and describe it in discrete terms. However, the dimensional model is considered to provide a more comprehensive description of emotional state since the variance of emotion can be explained on a continuous scale, therefore the dimensional model is more flexible than discrete categories. The flexibility is useful when a wide range of emotional states occurs. The dimensional can describe emotion with arousal and valence, which can be used to recognize dynamics and allow personal differences. As a result, dimensional models are better in handling variations in emotional states over time and also deals with non-discrete emotion [91, 20]. Although the continuous dimensional space may embody the dynamics better, categorical labels are better in capturing subjective experience because it provides a simplistic label, such as joy [108].

Despite the exclusive use of either one, the categorical and dimensional definitions of emotions are related [44]. The theoretical frameworks are rather complementary than contradictory [48]. The dimensional model is integrated with the discrete perspective [87] in order to adequately explain how fear, jealously, anger and shame differ and how observers may distinguish between them. Contrary to the discrete theory, where categories of emotions are separate from each other, in the dimensional theory these categories of emotions are organised in the dimensional space in a cohesive fashion were the categories have fuzzy boundaries and may overlap [22, 76].

2.2 Emotion assessment

Emotional states or responses can be assessed in three different methods, through overt behavioral acts, physiological reactivity, and affective report[16]. Overt behavioral acts includes facial and and body gestures, computer vision has been applied to recognize human facial expression to assess which emotional state is experienced. This is criticised because the emotion assessment originates from the assumption that facial expressions are unique per emotion category, as described in the discrete theory by Ekman [29]. However, emotional states tend to share common facial expressions, e.g. surprise and sadness can be difficult to differentiate when recognizing through facial expression [11]. Furthermore, in a review [11] of state-of-the-art face expression recognition it is also stated that there is different accuracy in the recognition of various emotional states. Another critic is that people are capable to manipulate their face expression to not let it represent their true emotional state, e.g. smiling while being uncomfortable. As described earlier, it has been proven that stereotypical facial expression only appear a small part of the time in the real world [33].

Emotional states can be assessed through physiological reactivity, which is the response of the body to a certain emotion or stimuli that can be classified as a certain emotional state. There are several physiological signals that can be measured to classify an emotional state, such as electromyograph, skin temperature, electroencephalography, and electrodermal activity or electrocardiogram [98]. Advantages of this method are that the signals from the physiological reaction are not manipulable, has to possibility to do continuous emotion assessment, and does not need manual annotation to identify which emotional state is present as a result does not interrupt the subject's attention to the presented stimuli. These methods are usually applied in the lab where there is a minimum of noise and the gathered data in return is more stable. They do not apply for studies conducted outside the lab where a lot of noise may be present and possibilities for using sensors are limited. Again, for this method the critic is the same as with facial expression recognition: various emotional states have different recognition accuracy rates. In addition to physiological signals, speech also contains reactivity to emotion. Emotion recognition through speech signals is also a prominent emotion recognition method [30]. This method is prone to the same critics as physiological signals, with the exception that no sensors are required, only a proper microphone and recorder. This method also requires participants to speak out loud which might be unnatural in their own home during or after viewing stimuli.

These above two methods are beneficial when applied in the right settings; locations outside the lab are less suited. Given that the required technology is not available in natural settings, and various privacy concerns may occur during the recording and collection of data. As opposed to previous methods, affective report can be done in discrete terms or dimensional scales and therefore not limited by technology requirements. In daily life, words that represent discrete emotional states are used to self-rapport emotional states to one and other. Disadvantages of semantic terms that represent emotional states are that the recognition of the true emotion might be reduced [9], the same term can be used to communicate different emotional states [8], and a term can entail and carry different value per individual [7, 32]. Whereas dimensional terms does not have the boundaries of categorical classification, and can surpass cultural, language and individual differences

To assess emotional states in a multiple dimensional space a measurement tool called Self-assessment Manikin (SAM) was designed. SAM is an imaginary-based measurement tool for the self-report of emotion in response to a stimuli [16], constructed by [51] initially as computer program. It was designed as an alternative for verbal self-reports, i.e. the discrete emotional terms, that can be ambiguous. SAM contains valence and arousal dimensions that the manikins represent. For the pleasure dimension the figure ranges from a unhappy frowning figure to a happy smiling figure and for the arousal dimension the figure ranges from a relaxed sleepy figure to an excited, wide-eyed figure. SAM scales allow quick assessment, which will wear out participants less than verbal self-report, as is done with speech emotion recognition [73].

Important to consider is the moment of self-assessment, which can be during or after viewing stimuli, or done continuously. Annotation during the stimulus can distract from the stimuli and influence the emotional response to the stimulus. Annotation directly after the stimuli the emotional response may be faded or disappeared. Thus, the resulting self-assessment may not be a true reflection of the emotional response during the stimuli [85], as it can aggregate the emotion felt during the whole stimuli into on singular measurement [53]. Continuous self-report can be applied to capture the change of emotion during the whole stimuli. However, this constant self report can give a cognitive constrain which can influence the emotional state when done actively. Self assessment directly after stimuli, is a common practice in emotional state assessment in response to stimuli and easy to apply to studies with natural settings.

2.3 Personality

Referring back to [36], that stated the emotional response is a reaction to a situational event and the individual's relation to that event. an individual has his own personality with certain personality traits that are correlated with emotional states [66]. Different personality traits have different activation thresholds for emotional states [67]. Each emotion also has an activation threshold. In addition, each individual has their own distinctive threshold for each emotion. These thresholds activation characteristics result in individual difference in the experience of emotional states [48].

Personality is described in dimensions, where scholars argue how many dimensions are necessary to describe emotions. The two main building blocks are extraversion and neuroticism. Extraverts receive more pleasure out of social a interactions, more sensitive towards positive effects, and have lower arousal levels during resting states when compared to introverts, which is the other end of extravert [64, 100]. Neurotics are more sensitive towards negative affects, and more likely to perceive situations as negative or stressful [41]. The other end of the neuroticism dimension is stablility. It is a widely held view that individuals with neuroticism are more likely to experience negative affect, whereas those with extraversion are more likely to experience positive affect [54, 21, 69]. In terms of discrete emotion terms, fear and sadness are related to neuroticism, and happy, joyful, excited, lively and energetic are strongly related to extraversion and agreeableness [109]. That individuals with extraversion are sensitive towards positive affect does not indicate a diminished sensitivity towards negative affect, as neuroticism does not indicate a diminished sensitivity towards positive affect.

The personality traits are influenced by a combination of genetic and environmental factors, with extraversion influenced relatively more by family environment than neuroticism [102]. As a result, personality is unique for the individual, making personality a crucial aspect to examine. Two common personality models are the Big-Five factor model [24, 70], that besides the two building blocks of extraversion and neuroticism, included three more dimensions: agreeableness (compassionate-dispassionate), conscientious (dutiful or easy-going), and Intellect (curious-cautious). The other model is the Eysenck Personality Questionnaire revised short scale (EPQ-RSS), Eysenck argued that three dimensions are sufficient, and only included one more dimension psychotism (impulsive-controlled) [31]. Compared to the Big-Five model, EPQ measures four dimensions instead of five with both measuring two of the same dimensions, neuroticism and extraversion, the advantage of EPQ is that this survey is shorter while measuring extraversion and neuroticism similarly [113]. The EPQ is an applicable and valid personality measure across diverse cultures in different countries [10, 13] and it accounts for the personality traits that cause most variance in emotion.

Personality can cause a bias towards positive or negative affect. Therefore emotional response can not be viewed as a secluded variable, without the relation with personality. Personality is taken into consideration in order to account for interpersonal variations on the emotional response. Personality traits are a key topic within the field of emotion, but lack a systematic understanding how different personality can have an effect in the emotional response to various representations of media content.

2.4 Screen size

The representation of media content is displayed daily through devices that have various screen sizes. Mobile phones, laptops, and TVs can be respectively categorized in small, medium and large screens. These devices are accessible, and provide sufficient options of viewing media content on a variety of screen sizes.

2.4.1 Screen size influence on presence

Media can create an experience of being 'there' in the media, a feeling of being present in the environment that is created through the media. This perception creates a sense of presence [38]. Presence has a relation with the emotional response that the viewer has to media content. A stronger presence can elicit a stronger emotional response, vice versa, a stronger emotional response can elicit a stronger sensation of presence. Two media characteristics that can influence the presence are media form and media content. Media form are the properties of the display medium, such as device or screen size, and media content represents the objects, actors and events that are displayed in the media stimuli [6]. The emotional response to media content depends to an extent on the media form it is presented on [45]. Media form can be manipulated to increase presence and screen size is one of the determinant media form variables that can influence presence [61]. The presence can be manipulated through device design. However, the presence is perceived by the viewer and is influenced by their own personality and functions of the mind.

Meaning that presence is an user experience and a product of the mind, not a direct product of the technology [84].

Presence is a sensation that occurs within a virtual experience [56]. Presence is a multidimensional construct that is conceptualized through social richness, realism, transportation, immersion, social actor within medium, and medium as social actor [59]. Transportation and immersion are two important presence concepts for media viewing, particularly for film stimuli. Transportation is the illusion of the viewer being transported in the fictional world, as if viewers are witnesses of events in that world and being 'there' in the media's environment. Immersion can be either perceptual or physiological immersion. Perceptual immersion refers to the number of senses that receive input from stimuli as well as the amount of input that is blocked from the physical environment. With physiological immersion, viewers feel involved, absorbed, engaged or engrossed in the media.

Strengthening the presence could be accomplished by a larger screen size to produce a larger visual presentation for greater visual immersion. A larger screen size could increase arousal as the objects on the screen appear to be closer to the viewer [80, 106]. In [106] it was demonstrated that with higher immersion the emotion intensified, although the comparison was made between 3D viewing and CAVE conditions. These are environments where more senses receive input than through 2D viewing that everyday devices provide. Still a sense of presence can be achieved even when the majority majority of the sensory is not engaged, due to imagination that can simulate the remaining sensory and create a convincing sense of presence [56].

A previous study has researched the difference between 56-inch, 13-inch and 2-inch picture heights that fitted onto the screen. Findings were that, the largest screen size had the most arousal response compared to the medium and small screen size for all types of content, especially exciting content [80]. These findings are unlike [60] which compared 46-inch to 12-inch screens it was reported that no greater enjoyment was experienced when watching the larger screen, although they did report a more intense response to the images on the large screen. On a larger screen movement is experienced as faster, the sense of the physical movement is greater and the viewing experience is more exciting resulting in a higher arousal [61]. Larger screen sizes would create in greater immersion and smaller screen sizes would make immersion more difficult [83]. If not decrease immersion, [81] showed that larger screens do not increase immersion but reduce immersion with comparing 4.5 inch to 13-inch and 30-inch

Since the conduction of these studies technology has progressed and the standard screen sizes of devices today differ, e.g. 2-inch screens are not to be found anymore on mobile phones. Also, the used stimul were pictures, videos of objects that did not leave the center of the screen [80], or broadcast stimuli [60], or parts of a movie where the film clips were not validated to a certain emotion but rather a general genre [81]. Considering that these studies were conducted in a lab setting with a fixed viewing distance rather than in the real world where the participant can hold the device in any comfortable position.

2.4.2 Field of View

Field of view (FOV) is defined as the visual area humans see while our eyes remain fixed. The nasal field is 60 degree, with a binocular field of 120 degree and a temporal field of 105 degree that together result in a visual field of 210 degree in a horizontal plane, see Figure 2. Additionally, the vertical view is a vertical plane of 135 degrees [3, 2]. Field of view can be defined as the whole visual field, or the functional field of view as the visual field were information can be gather with one eye fixation [4, 92].

FOV can be adjusted by screen size, viewing distance, and angle view that together influence how an object is perceived [25]. With a larger screen size the content is easier visible in the outer corners of the vision. Smaller screen sizes on mobile phones or other handheld devices are easier to move closer to the face, therefore achieve the same FOV as a larger screen by adjusting the viewing distance. Presence increases with a higher field of view until 140 degrees [58] and decrease with a restricting field of view [2]. The viewing distance that determines the FOV appears to change with the screen size. As demonstrated in Lund's work where images from 11 to 123 inches were shown, and a viewing distance to image height declined from 7.4 to 3.1 [65]. A larger screen is placed farther away and lower, while keeping the eye sight near the display top [96]. Viewers also use the viewing distance to increase sharpness by sitting close to the stimuli, using the viewing distance as a zoom in function [65]. All the studies presented so far had the viewing distance fluctuate based on the screen size.

On mobile phones, to be precise an iphone 5 with 4 inch screen, when looking at a picture viewing distance while sitting was found to range between 13.3 to 32.9 cm with a mean of 20.3 cm. Compared to a laying down position where the distance was reduced to 9.9 to 21.3 cm with a mean of 16.4 cm [111]. For monitors of 19 inch the preferred distance while sitting at desk was 68.0 cm, [96], in alliance with another

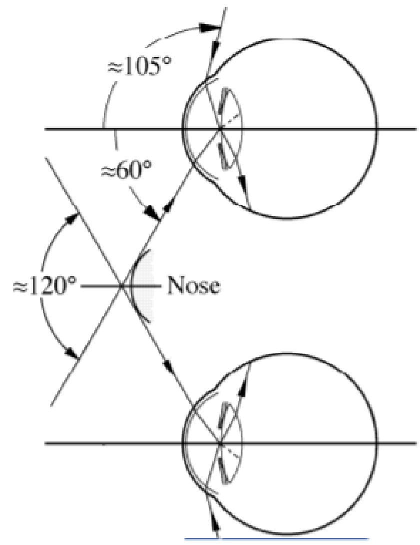


Figure 2: Horizontal field of view for humans, adopted from [3].

study that found an average viewing distance of 68 cm [19]. With the assumption that with normal or correct to normal vision, when sitting an average of 20.3 cm for the mobile phone viewing distance and an average of 68 cm for the laptop viewing distance are the established values used to calculate the FOV. Both horizontal and vertical FOV with a fixed head position, for both eyes, can be obtained through equation 1, adapted from [5, 104]:

$$FOV = 2\arctan\left(\frac{d}{2l}\right), \tag{1}$$

where: FOV = field of view in degree, d = screen width in cm for FOV horizontal or screen length in cm for FOV vertical , and l = viewing distance in cm.

2.4.3 Film as emotion stimuli

In the quantitative research several kinds of stimuli are used to elicit emotion on screen, most common are games, pictures and film clips. The three devices taken into considering to compare screen size can all be used to play games on, share pictures and watch films in the real world. For this Thesis stimuli are required to be visually consistent on each screen size and common in daily life to simulate real-life experiences. However, games are not visually consistent on different screen size because games are interactive and require an input interface to play the game. Therefore, the design of the game interface is fitted to the screen size and results in different interfaces between different devices. Film and pictures are both common to elicit emotions, films offer a dynamic audiovisual stimulus while pictures are static and visual. A common picture stimuli dataset is the IAPS [52], however validated to be effective the emotion elicited through static stimuli tends to be lower than elicited through film. Films provide dynamic stimuli which is better to capture attention, can elicit a higher intensity emotion [50].

Films are frequently being watched for entertainment, as they provide a method of escaping from daily life through a continuous led imagination that can evoke an emotional response [101, 107]. Films can make viewers cry as a response to sadness or laugh as a response to amusement, regardless of film being a virtual or fictitious representation of reality [85]. Due to the response that films evoke, film clips serve as common stimuli to elicit emotion because they are widely available, generally brief, instinctively powerful and can be standardized among participants [57, 85, 89]. An entire film typically includes a mixture of emotional states that can be cut up into short clips that are more homogeneous, and are therefore more effective to elicit a certain emotional state [85]. Emotion recognition research makes great use of film as emotion elicit stimuli to provoke a physiological response. The film clips are proven to be effective as emotion elicit stimuli for a wide variety of population [39, 42]. Multiple researchers developed databases of film clips to elicit certain discrete emotions [39, 89], since not every film elicits what is intended to elicit such databases are a necessity.

The film clips in the database are validated as a specific discrete emotion. However, common in films is the sad-film enjoyment paradox [75]. Horror and drama films are watched for entertainment while the these film elicit a negative valence emotion, they also elicit a positive valence emotion due to entertainment or being moved by the film [43]. The positive emotion does not necessarily replace the negative emotion [1]. Emotions are not mutually exclusive which can result in feeling mixed emotions where emotion of opposite valence co-occur, e.g. bittersweet situations or guilty pleasures [53]. Besides the paradox, due to the narrative of film it is also capable to elicit cognitively complex emotional states, such as nostalgia [85], that can also result in a mixed emotional state.

A drawback of the film stimuli is that individual differences in prior view of the film can modify the reactivity to the film clip [57]. Although, film clips can be standardized in presentation there is not a standardization for the content of the film clips since every film uses a different camera angle (i.e. framing, viewpoint, angle of the shot, movement), sound, number of people and colour scheme or light levels [57, 18]. There is also a risk of habituation, fatigue and desensitisation when film clips are shown for a longer period of time. Furthermore, consecutive shown film clips a carryover effect of emotion might be present. Which means that two consequence film clips can influence the emotional response to the second film clip, whether it causes a more intense emotional response or decrease the emotional response. To prevent this a temporal spacing should be used to minimize the carryover of the emotion of one video to another video [85]. Moreover, the physical setting where film clip is viewed can have different characteristics such as, room lighting, temperature, and space. These are factors likely to differentiate in the real world.

Previous research has studies the influence of screen size on emotional response or immersion. However, it failed to consider personality or was conducted with screen sizes that do not often appear in technology today. Therefore, it is not yet determined if screen size and personality traits influence the intensity of the emotional response. At the same time, it is crucial to establish whether screen size is an important factor in media form to take into consideration when presenting media stimuli. Because establishing factors in media form that influence the emotional response to media content can provide tailored designs to increase emotional response to enhance the goal of the media content. Whether it is to entertain or persuade, or even educate.

3 Methodology

A quantitative research study was conducted in natural settings with a survey that contained a personality questionnaire and film stimuli. This section elaborates on the materials that were used, the experiment design, and the procedure.

3.1 Participants

A total of 72 participants participated in the between-subject design experiment; condition TV had 24 participants (50% male, 50% female), condition laptop had 24 participants (37.5% male, 62,5% female), and condition mobile had 24 participants (62,5% male, 37.5% female). The participants were between 19-59 years ($mean = 26.01$), 87,5% of the participants had a Dutch nationality, and the other participants had nationalities of Colombian, Italian, and Polish. One participant had a dual nationality. Personality traits were assessed, which resulted in 22 extraverts and 11 neurotics in condition TV, 16 extraverts and 16 neurotics in laptop, and 9 extraverts and 16 neurotics in mobile. Since the personality traits are not mutually exclusive, participants can have both traits. All communication was in English. Participants were recruited online through an anonymous survey link to participate in the condition Mobile and Laptop and approached on location condition TV. No compensations were handed out for participation.

3.2 Materials

3.2.1 Hardware

The screen sizes were divided into three conditions to obtain a larger sample size in natural settings because screen sizes vary greatly across the same type of devices and simultaneously take into account three standard devices used in daily life. Therefore, the three conditions are TV screen, laptop screen, and Mobile phone screen. Correspondingly the screen sizes range from large, medium, and small. For the TV condition, the same TV screen was used for each participant in that condition. In the laptop and mobile conditions, the participants used their own devices.

Figure 3: The TV Condition setting, the participants set besides the Laptop in front of the TV where the survey is displayed.

TV: The TV is a Sony KD-55XF9005 with a 55-inch screen with 4K resolution connected with two Dali Ikon 5 speakers. The TV screen was connected via an HDMI cable to an Apple Macbook Pro with a 16-inch screen and usb connected mouse. The keyboard of the laptop and external mouse was used as input devices. The screen of the Apple Macbook was dimmed to black to prevent any distraction. The participants sat on a couch at a 235 cm from the TV, Figure 3. The laptop was placed on the participant's left side, and the mouse was connected with a USB cable to the laptop. The participants were free to place the mouse where they felt most comfortable, most common was on their right side. However, this setup would be tedious for participants to do at home. Therefore, the researcher did the set up in a home setting to reach more participants and ensure the participants were able to fill in the survey correctly with the corresponding input devices.

Laptop: The participant used their own Laptop. A laptop was required to have a diagonal screen sizeranging from 11 to 19 inch. The participants had a diagonal screen size ranging from 13 to 17.3 inch (mean = 14.4 inch). The input device is a mouse, this could be on a touchpad or an external mouse.

Mobile: The participants used their own Mobile. A mobile phone was required to have a diagonal screen size ranging 4 to 7 inches. The participants had a diagonal screen size ranging from 4 to 6.7 inch screens (mean = 6.01 inch). The input is done through a touchscreen with their finger.

FOV: FOV was calculated with the average values of viewing distance based on previous research, mentioned in section 2.4.2, for mobile and laptop. For TV, the viewing distance was measured to be 235 cm. The viewing distance (l) is constant across various screen sizes (d) in the same condition. In Figure 4 FOV horizontal and FOV vertical are presented per condition, where the maximum degree on the y-axis represents the binocular vision degree. The FOV horizontal represents the width of a screen size and the FOV vertical represents the length of a screen size. For the condition TV, the mean FOV vertical is 16.58 degrees, respectively laptop has a mean of 15.02 degrees, and mobile has a mean of 14.22 degrees. For the FOV horizontal, TV has a mean of 29.05 degrees, laptop of 26.38 degrees, and mobile of 24.99 degrees. Thus, all screen size conditions take up roughly one-eighth of the FOV horizontal (binocular vision, without the temporal field) and one-fifth of the FOV vertical. Both FOV horizontal and FOV vertical represent the overall screen size FOV and they both are roughly similar across conditions.

3.2.2 Software

The surveys were designed in Qualtrics [79], a software service provided by the Utrecht University. The distribution was with an anonymous link to let participants conduct the experiment at home. In total, three surveys were created for the three conditions TV, laptop, and mobile. Each survey was identical except when the device condition was mentioned.

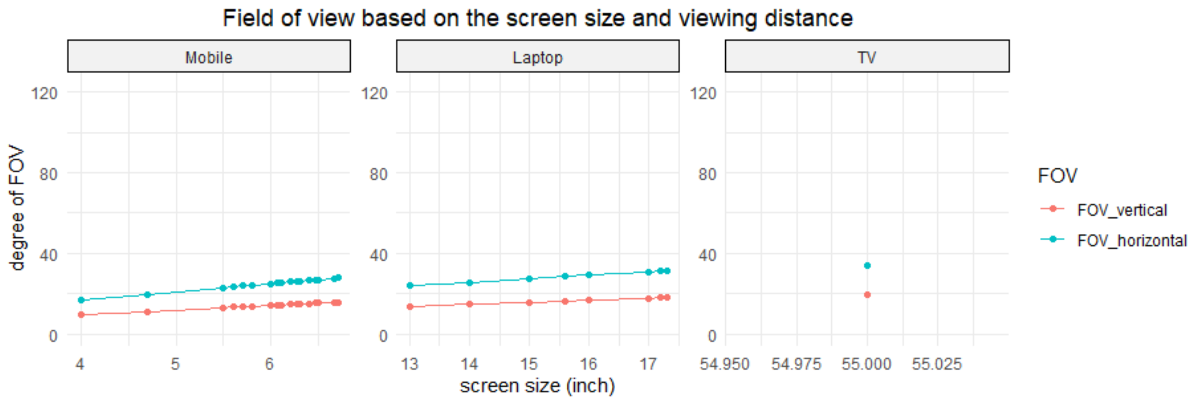


Figure 4: The degree of field of view horizontal and vertical per screen size condition. The y-axis is until 120 degree which is the degree of the binocular view. For TV there is one point because there was one screen size of 55 inch.

3.2.3 Personality Questionnaire

The short-scale Eysenck Personality Questionnaire-Revised (EPQ-R) [31] was implemented to determine the personality traits of the participants. EPQ-R included 48 questions that could be answered with yes or no, and assessed extraversion/introversion and neuroticism/stability. The EPQ-R evaluates personality traits on a dimensional scale, where someone could be somewhat or highly neurotic. For each personality trait, a maximum of twelve points could be measured. A higher score indicates a higher extreme in the personality trait. For this analysis, an artificial dichotomy was created with a median split on a score of 6. If a score was higher than 6 the participant was assigned as extravert and neurotic. Meaning that the somewhat neurotic were grouped with the highly neurotic, likewise for extraverts.

3.2.4 Film Stimuli

Eight emotion elicitation videos were extracted from YouTube to elicit four different emotional responses, amusement, relaxation, fear, and sadness, for each emotion two videos, see Table 1. Each emotion corresponds with a quadrant of the valence-arousal model to include each combination of high/low valence and high/low arousal. The emotion videos contain film scenes that were validated to elicit the intended emotion [85, 39, 95, 89]. Except video Spring was not previously validated but was a substitute for an expired video link of a relaxed video in the dataset from [95]. The replacement contained similar scenes as the expired video. The film clip of the Hangover missed the first 10 seconds of the clip that was validated in previous research [95] that video due to the low quality is replaced with another clip with higher quality that did not contain the first 10 seconds. These 10 seconds included a shot of the four men toasting with their shots on a rooftop. The videos were embedded in the survey from YouTube to prevent any copyright violation, the language in the videos was English, and the videos have a duration between 90-273 seconds. All the videos have a quality of 720 pixels or higher to prevent pixelated content.

3.3 Design

The study had a mixed subject design, each participant participated in one of the three conditions but viewed all the emotion videos. Each condition had a separate survey. The survey was required to be accessible on three different devices, therefore a browser-based application was used to design the survey to be opened on multiple browsers on various devices. The browser-based survey provided participants with easier access to the experiment in their home environment. This study design was adopted to allow a deeper insight into the real world dynamics and gain new insight into how screen size and personality influence the intensity of the emotional responses. The survey can be divided into two parts. The first part was the EPQ-R [31]. The second part used SAM [16] to annotate emotion that participants felt during the emotion video by putting the manikin above the Likert scale that ranged from 0 to 10. To provide a clear understanding of the emotional states that were portrayed by the manikins, a textual explanation was given before viewing the videos. The explanation stated that the valence scale, in Figure 5, the left manikin portrays the end of a scale that is an extremely unhappy feeling. If the reaction to the film scene was unhappy, annoyed, or unsatisfied, an answer could be placed on the scale where this manikin

Table 1: The eight films applied in the survey.

Film	**Emotion**	**Valence-Arousal**	**Film length (min:sec)**	**Quality**
Harry	Amusement	high-high	2:52	1080p
Hangover I	Amusement	high-high	2:38	720p
Bambi	Sadness	low-low	2:20	720p
Philadelphia	Sadness	low-low	4:33	1080p
Copycat	Fear	low-high	3:18	720p
Shining	Fear	low-high	2:34	1080p
Beach	Relaxing	high-low	2:55	1080p
Spring	Relaxing	high-low	1:38	1080p

Row Item: Harry = When Harry met Sally: Discussion of orgasm in restaurant, (from When Harry Met Sally, Castle Rock Entertainment, 1989, accessible through [97] used by [85, 39]); Hangover = Hangover I: waking up after a night of drinking to find a tiger in the bathroom, (from The Hangover I, Warner Bros. Pictures, 2009, accessible through [26], used by [95]); Bambi = Bambi: The mother of Bambi got shot and dies, (from Bambi, Walt Disney Animation Studios, 1942, accessible through [71] used by [39]; Philadelphia = Philadelphia: A man described the pain and passion felt by the opera singer while listening to the opera song on a stereo, to another man, (from Philadelphia, TriStar Pictures, 1993 [77] used by [89]); Copycat = Copycat: one of the characters gets murdered in the toilet, (from Copycat, Regency Enterprises, 1995 [105] used by [89]); Shining = The shining: Character played by Jack Nicholson chases his wife with an axe, (from The Shining, Warner Bros., 1980 [68] used by [89]). Beach = Relaxing music on beach: a still shot of a swing on the beach with relaxing music on the background (accessible through [99] used by [95]. Spring = Spring: the return of life: man voice explains the return of spring while mountains and the sun shining through trees are shown (accessible through [103]).

was placed. The right manikin on the other end of the scale portrays an extremely happy feeling. If the reaction to the film scene was happy, satisfied, pleased, or relaxed, an answer could be placed on the scale where this manikin was placed. Concerning the arousal scale, in Figure 6, the left manikin portrays the end of the scale that was extremely unaroused. If the reaction to the video was relaxed, calm, dull, or sleepy, an answer could be placed on the scale where this manikin was placed. The right manikin portraits the other end of the scale, namely an extremely aroused feeling. If the reaction to the film scene was intense, excited, stimulated, or wide awake, an answer could be placed on the scale where this manikin was placed. The participant was asked after watching a video to "*please rate your feeling according to the manikins*" to self assess their emotional response to film clips on valence and arousal scales, that ranged from 0 to 10. The participants also reported if they looked away during the video.

Temporal spacing between the videos was included with an embedded YouTube video that played 20 seconds of black screen. Before the first emotion video and after every following emotion video, a black screen was shown for 20 seconds, Figure 7. Instruction was given at each video to manually click to play the video, click to fullscreen, and after the video automatically stopped playing to click to exit the fullscreen. To ensure that participants looked at the black screen for 20 seconds the next page button appeared after 20 seconds spent on that page. The SAM annotation, emotion video, and black screen were displayed in a specific order visualized in Figure 7. The SAM annotation was required before viewing any emotion video and after the last emotion video had been viewed to include a baseline measurement. After the first SAM annotation, a black screen for 20 seconds was shown then the first emotion video was displayed. Consecutively after each emotion video, SAM annotation was required, and a black screen was shown. This was repeated until the last video, and black screen had been viewed, then the last SAM annotation was filled in. Context or familiarity can influence the emotional response, hence the participants were asked if they had seen the movie before.

Each emotional state is elicited through two videos because they are validated to elicit the same emotional state. The two videos are seen as repeated measures of the same emotional state, despite containing different content. The emotion videos are divided into two blocks that contain each four emotion videos, one of each emotional state. The first block contained: Harry, Bambi, Copycat, and Beach, and the second block contained: Hangover, Philadelphia, Shining, and Spring. Within these blocks, the order of the videos was randomized, resulting in 24 permutations within each block. Each condition contained 24 participants, as a result, each variation order of film clips could have been completed once. However, Qualtrics chooses randomly from the pool of variations. Consequently, the chosen variation, when completed by the participant is put back into the pool of variations. There was no simple solution to customize the randomization into a pseudo-randomization. As a result, of the 24 variations, in condition

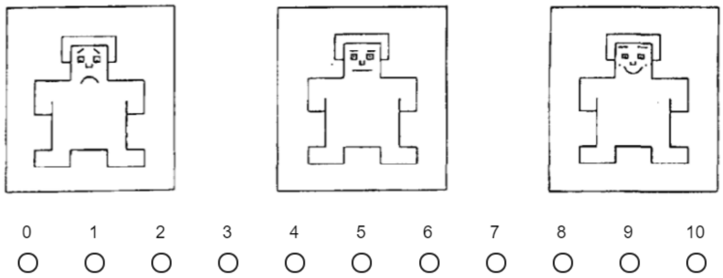


Figure 5: The SAM valence scale at 0 the manikin display an extremely unhappy feeling and at 10 the manikin displays a extremely happy feeling. SAM scale adopted from [16].

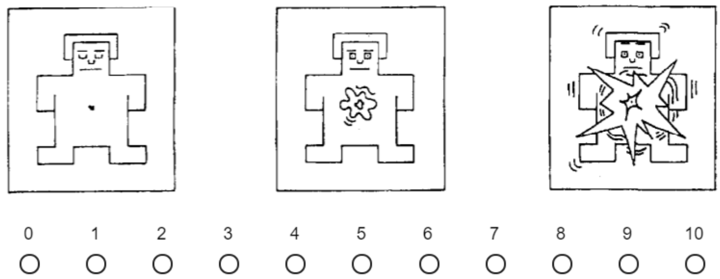


Figure 6: The SAM arousal scale at 0 it displays an extremely unaroused feeling and at 10 the manikin displays a extremely aroused feeling. SAM scale adopted from [16]

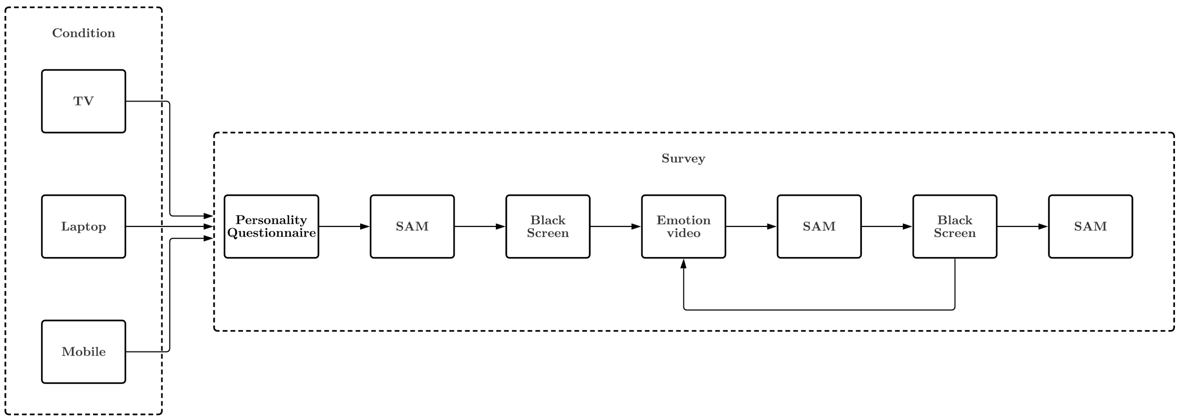


Figure 7: The experimental Design that used in each condition. This visualisation of the design demonstrates the order in which the personality questionnaire and the SAM annotation, 20 seconds of black screen and the emotion video appear in the experiment. It starts with the S-EPQ questionnaire. Then the first SAM annotation is required before watching any video or black screen. The black screen that follows is to help clear the mind of the participant before watching the first emotion video. After watching the emotion video SAM annotation is required, followed by a black screen for 20 seconds to help clear the mind again. Then the next emotion video is shown. This process repeats itself until the eight and last emotion video has been watch. When the black screen after the last video has been watched, the last SAM annotation is required.

TV 9 variations were not done, likewise for laptop 9 variations were not done, and in mobile 7 were not done. In the condition TV and Laptop 2 variations were executed 3 times, all the other included variations were executed 1 or 2 times.

Another note is that both blocks randomize separately, and the variation of film clips in the first block does not influence the variation in the second block. Therefore, the last emotion video in the first block could belong to the same emotional category as the first in the second block. The probability that this would occur was 6.25% per participant. For this to not occur across 24 trials is 22.25%. In condition TV and laptop, however, this occurred 10 times, and in mobile, it occurred only once.

3.4 Procedure

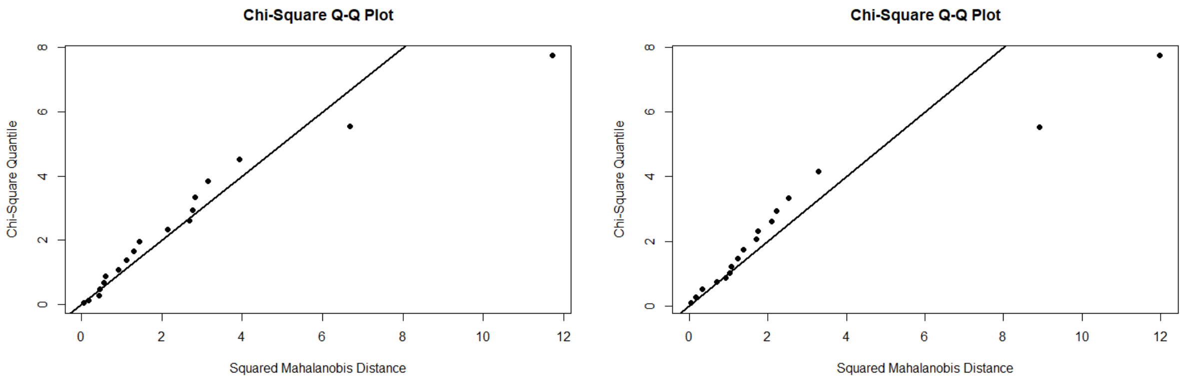
Before starting the survey, the participants were asked to seek and sit in a quiet surrounding where sounds could be played aloud. Also, they were asked to turn off notifications to prevent any distractions. The message sent out to recruit participants declared twice the device on which the experiment should be conducted. This message was not sent for condition TV as those participants were approached on location. When participants had the survey link on the appropriate device, they could start the experiment by clicking on the link. The survey started with an explanation about the experiment's layout, the experiment's aim, and requested consent for the usage of the provided data, given that it was used for this research purposes only and presented anonymously. After the survey is completed the participant is thanked for their time and participation.

A different approach was taken in the procedure of condition TV, then in the procedure of laptop and mobile. In the TV procedure, after being approached and consented to participate in the experiment, participants were asked to take place on the left side of the couch to position themselves in front of the TV screen. They asked to make themselves comfortable and at home. A short verbal explanation of the experiment's process was given, identical to the description written on the first page of the survey. The verbal explanation sought to make the participants feel more at home. If they have questions about the hardware or software, they could ask, but not during a video, only after or before. The researcher sat in the same room outside the participant's peripheral vision to supervise the sessions in case of technical problems. The researcher could not see the TV screen and did not respond to out loud thoughts but did intervene when the participants asked for help. This procedure differentiates from the laptop and mobile condition due to the additional verbal explanation and the possibility of asking questions about hardware or software. Since the laptop and mobile conditions were conducted in the participant's environment, hardware and software questions were thought to be unlikely to occur when participants conducted the experiment on their own device.

4 Results

4.1 Normality

Mardia's coefficient was calculated as an indication of multivariate nonnormality. Mardia coefficient showed that the assumption of multivariate normality was violated. The distributions were visually examined to concluded that violations were caused to outliers. To correct for the violation, outliers were detected and removed by using Mahalanobis Distance Test with a chi-square at $p = 0.001$ to assess for multivariate outliers. The Mahalanobis distance identified showed three multivariate outliers, and these were removed from the dataset. Two of theses outliers came from the same participants, therefore when performing the MANOVA, the participants with missing values excluded, 70 participants are included in the analysis instead of 72. Mardia's coefficient was calculated again and showed that almost all of the distribution was multivariate normal after the outlier removal, except for two distributions. Since two out of the 24 showed a violation of the normality assumption, the data was not further transformed. A square root transformation was performed to correct for the nonnormal distributions, which improved the normality of the distribution for laptop-amusement and mobile-fear. However, other normal distributions were transformed to nonnormal distributions. To compromise, the better option of the two was taken, which is leaving the data untransformed. This allowed fewer distribution groups to be nonnormal and allows more accessible interpretation of the results. The multivariate distribution is visualized in a chi-square QQ plot in Figure 8. It illustrates that the multivariate distribution was caused by two multivariate values that are outliers present in the distribution even after the outlier removal, but since they were not detected using the Mahalanobis distance they are kept in the data.



(a) fear-laptop (b) amusement-mobile

Figure 8: Multivariate distribution displayed in chi-square QQ plot (a) is untransformed data of group fear-laptop, (b) is untransformed data of group amusement-mobile.

4.2 Temporal spacing

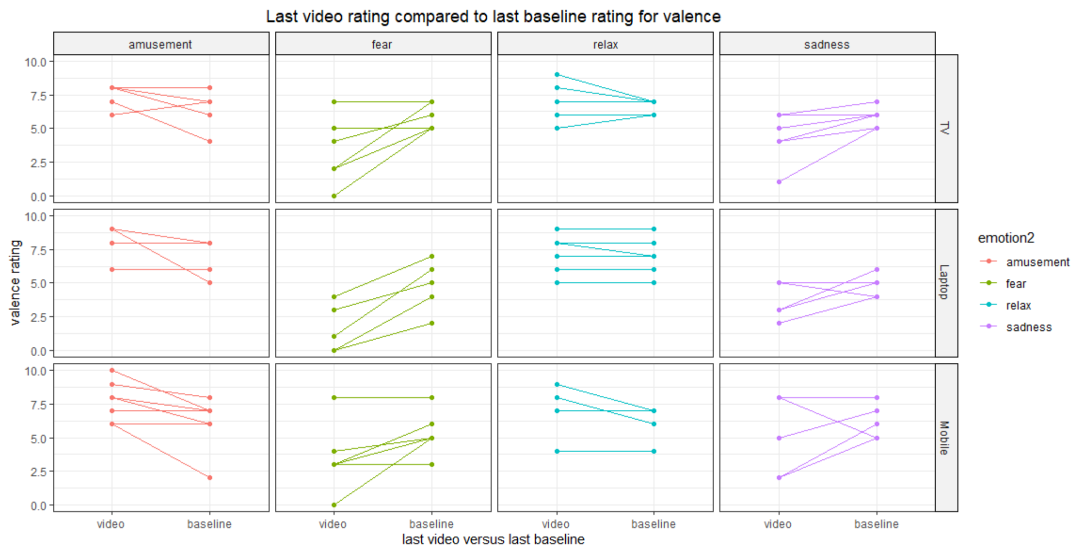
To assess if the carryover effect was present, the effectiveness of the temporal spacing was analyzed. A significant difference in valence and arousal rating between the last video viewed and the last baseline measurement can indicate if the temporal spacing was effective. Homogeneity was met (Levene's $p > 0.05$). A repeated measures MANOVA examined valence and arousal as dependent variables predicting within subject factor time of two levels (rating after last video and rating after 20 seconds of black screen) and between subject factor emotion (sadness, fear, amusement, relax). The analysis was conducted on all the emotion data points of the last video jointly because data points were limited for each emotional state per condition due to the randomization of video orders. As a result, each emotion category had a different sample size. A significant multivariate main effect was found for emotion ($F(6, 134) = 12.38, p < 0.001, \eta_p^2 = 0.357$) and a significant multivariate effect was found for time ($F(2, 67) = 3.38, p = 0.040, \eta_p^2 = 0.092$). Also, a significant multivariate interaction between emotion and time was found ($F(6, 134) = 13.68, p < 0.001, \eta_p^2 = 0.380$). Univariate ANOVAs was used to examine individual dependent variable contributions to the interaction. Valence showed a significant difference in the emotion and time interaction ($F(3, 68) = 19.29, p < 0.001, \eta_p^2 = 0.460$), as does arousal ($F(3, 68) = 12.08, p < 0.001, \eta_p^2 = 0.348$). Table 2 presents an overview of the last video and last baseline means, standard deviation, and sample size.

Post-hoc dependent t-test with a Bonferroni correction was performed to examine if the valence rating decreased for amusement and relax, and increased for fear and sadness after the 20 seconds black screen. For amusement ($p = 0.004, d = 0.91$), fear ($p < 0.001, d = 1.19$), and sadness ($p < 0.001, d = 0.80$) the valence rating of before the 20 seconds were found to be significantly different than the mean after 20 seconds. For relax the difference of valence rating before and after 20 seconds were found to be nonsignificant. Likewise, for arousal, post-hoc dependent t-test with a Bonferroni correction was performed to examine if the arousal rating decreased for amusement and fear, and increased for relax and sadness after 20 seconds black screen. For fear ($p < 0.001, d = 1.53$) and relax ($p = 0.013, d = 0.73$) the arousal rating of before temporal spacing were found to be significantly different than after temporal spacing. For amusement and sadness the difference of arousal rating between before and after temporal spacing was found to be nonsignificant. Table 2 presents an overview of before temporal spacing (i.e, the last video), and after temporal spacing (i.e. the last baseline) means, standard deviation, and sample size.

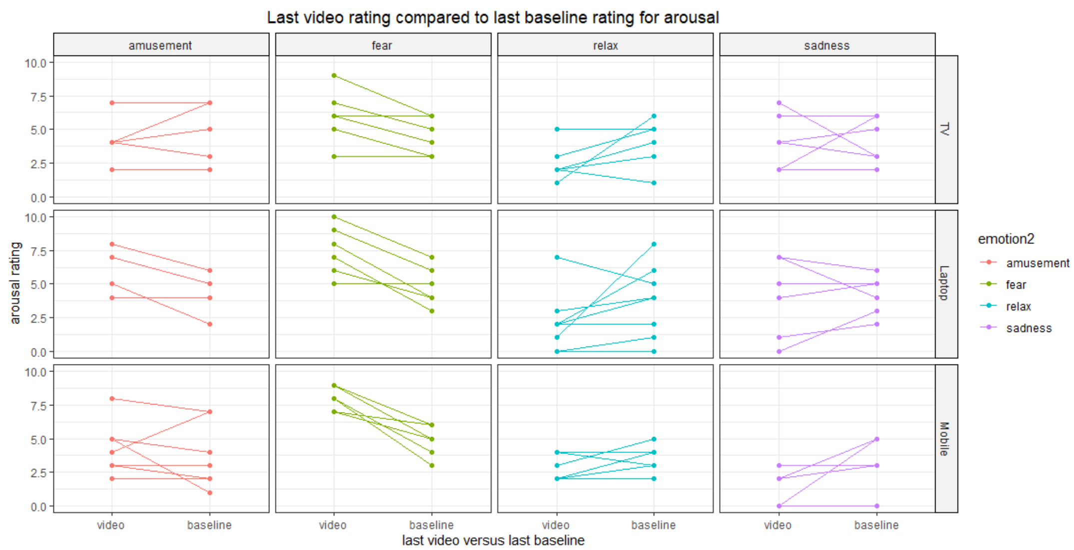
The statistical analysis does not distinguish between the conditions due to a lack of data. To demonstrate these effects per condition, the difference in rating of valence and arousal between the last video and the last baseline per condition and emotion category was visualized per screen size condition in Figure 9. In Figure 9a it is visible that the significant effects of valence rating for amusement, seems to be present in all conditions, however stronger for mobile and half of the ratings of laptop and mobile did not stagnate. For fear, the effect of the temporal spacing seems to be present in all conditions, especially in the laptop condition. For sadness, the effect of temporal spacing seems to be present in the condition TV, for mobile and laptop the temporal spacing seems to cause a higher valence rating however not for all participants. In Figure 9b the significant effects temporal spacing in arousal rating seems to present in all

Table 2: Mean and standard deviation of the valence and arousal rating per emotion category before and after temporal spacing. Before temporal spacing is the measurement of the last video that is viewed by the participant and after temporal spacing is the last baseline measurement.

		Before temporal spacing		**After temporal spacing**		
		Mean	*SD*	*Mean*	*SD*	*Sample size*
Sadness	Valence	4.35	2.03	5.59	1.94	17
	Arousal	3.29	2.47	3.88	2.26	17
Fear	Valence	3.11	2.42	5.42	1.34	19
	Arousal	7.11	1.76	4.80	1.74	19
Amusement	Valence	7.69	1.25	6.38	1.26	16
	Arousal	4.69	2.14	4.19	1.70	16
Relax	Valence	6.95	1.40	6.60	1.30	20
	Arousal	2.45	1.64	3.75	2.54	20

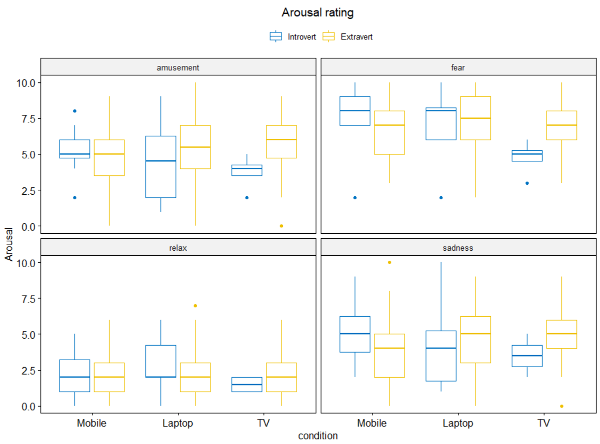


(a) Valence rating of the last video and the last baseline rating

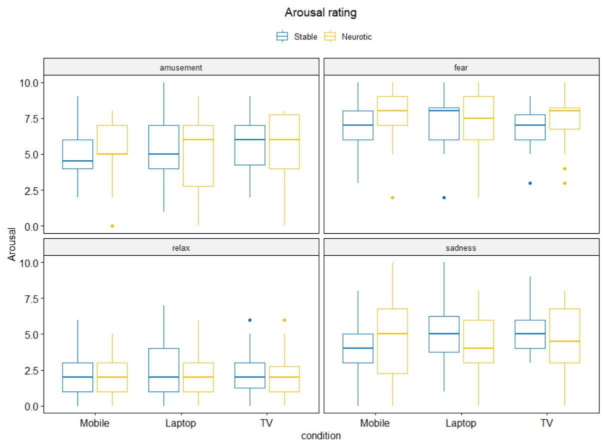


(b) Arousal rating of the last video and the last baseline rating

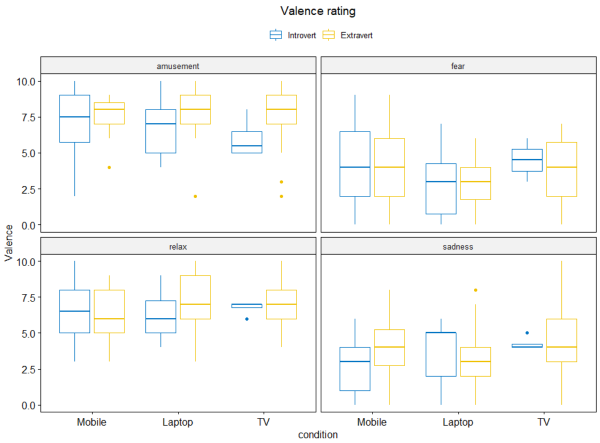
Figure 9: Last video and last baseline rating of valence (a), and arousal (b). The last video rating was filled in directly after the last video was viewed, then 20 seconds of black screen was viewed, afterwards the last baseline measurement was filled in. The difference between these two points display the change in rating after viewing 20 seconds of black screen.



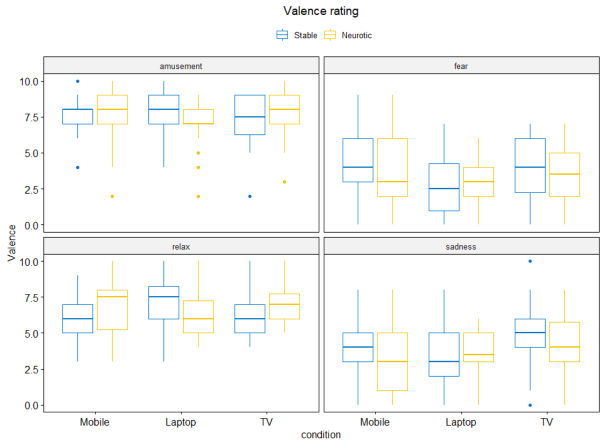
(a)



(b)



(c)



(d)

Figure 10: Boxplot of self-assessment ratings for valence per emotion, comparing screen size of condition TV (left), Laptop (middle), and Mobile (left). Figure (a) display arousal rating of extraversion (blue) versus introversion (red), (b) display arousal rating of stable versus (blue) neuroticism (red), (c) display valence rating of extraversion (blue) versus introversion (red) and (d) display valence rating of stable versus (blue) neuroticism (red).

conditions for fear and for relax it seems to be present in condition laptop and slightly in condition TV.

4.3 Emotion, screen size and personality

Three outliers were removed resulted in 70 participants that were retained for this analysis. TV and Mobile resulted in a remaining sample size of 23, and laptop of 24. In Figure 10, the distribution per emotional state, condition and personality trait is displayed. Homogeneity was met (Levene's $p > 0.05$). A repeated measures MANOVA was used to examine valence and arousal as dependent variables, predicting within subject factor emotion of four levels (sadness, fear, amusement, relax) and within subject factor video of two levels due to each emotional state elicited twice with two different film clips, and between subject screen size condition (TV, Laptop, Mobile), and personality trait, extraversion (versus introversion), and neuroticism (versus stability). Figure 10 depict the scores per conditions of valence and arousal between the personality traits. Significant multivariate main effect was found for emotion ($F(6, 53) = 27.777, p < 0.001, \eta_p^2 = 0.759$), but not for video, condition, extraversion or neuroticism. All interactions with emotions were found to be not significant.

Further, univariate ANOVAs were used to examine individual dependent variable contributions to main effects. Mauchly's test indicated that the assumption of sphericity has been violated for emotion, specifically for valence ($p = .013$) but not for arousal ($p = .076$), therefore degree of freedom was correct using Huyn-Feldt estimates of sphericity for valence ($\varepsilon = 0.86$). Valence rating showed to a significant difference in emotion ($F(3, 174) = 46.631, p < 0.001, \eta_p^2 = 0.446$), where post-hoc dependent t-test

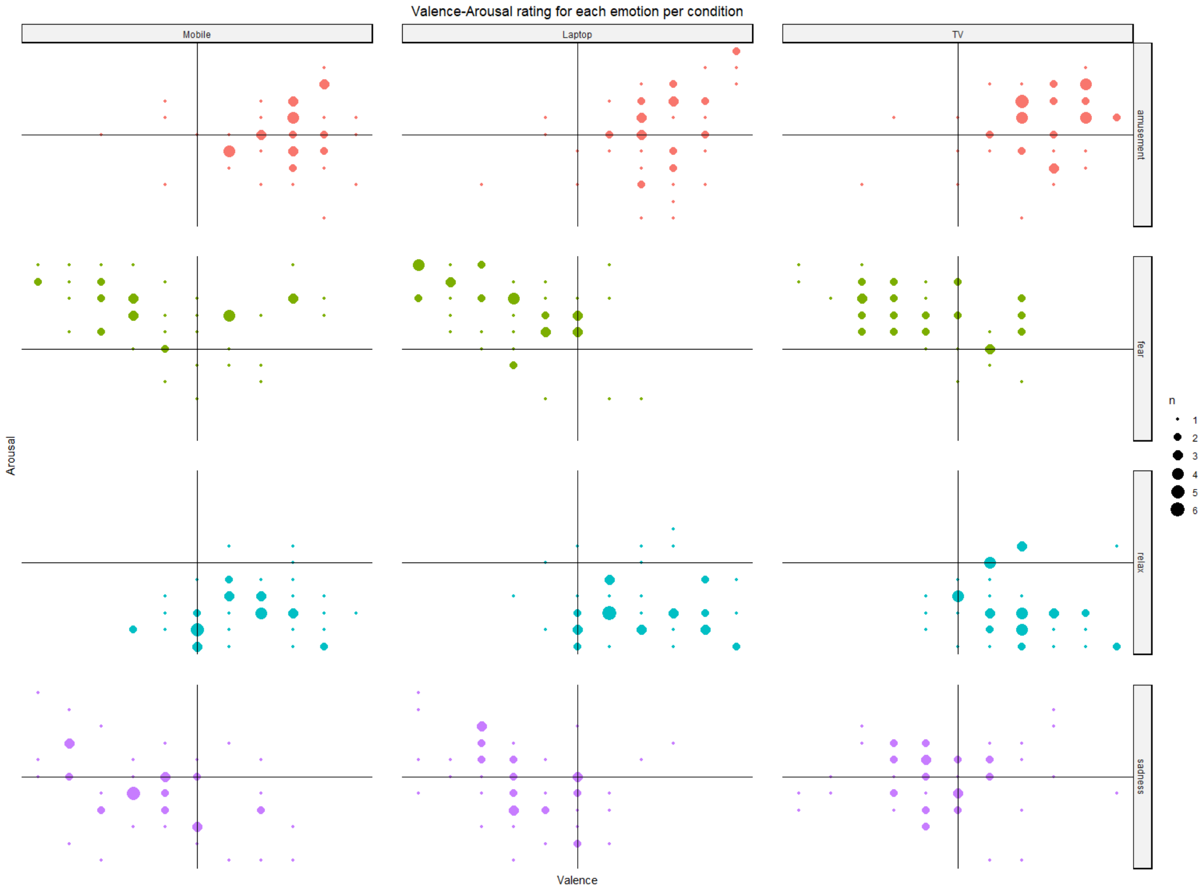


Figure 11: The rating of valence, on the x axis, and arousal, on the y axis, plotted per condition and emotion category. Each point represents a participant's rating of the video, the bigger the point the more participants rated the video the same.

with a Bonferroni correction was performed to examine which different emotional states were rated significantly different. For valence ratings, amusement was found to be significantly higher average than relax ($p < 0.001, d = 0.48$), sadness ($p < 0.001, d = 1.95$) and fear ($p < 0.001, d = 1.99$). Relax was found to have a significantly higher average than fear ($p < 0.001, d = 1.56$) and sadness ($p < 0.001, d = 1.51$). Lastly, sadness was not found to be significantly different than fear. The arousal rating showed a significant difference in emotion ($F(3, 174) = 63.200, p < 0.001, \eta_p^2 = 0.521$). Where again post-hoc dependent t-test with a Bonferroni correction was performed. Fear was found to have a significantly higher average than amusement ($p < 0.001, d = 0.92$), sadness ($p < 0.001, d = 1.23$) and relax ($p < 0.001, d = 2.69$). Amusement was found to be significantly have a higher average than sadness ($p = 0.010, d = 0.30$) and relax ($p < 0.001, d = 1.56$). Lastly, sadness was found to be significantly have a higher average than relax ($p < 0.001, d = 1.20$).

The self assessment of the emotional response was plotted in Figure 11 with valence on the x-axis and arousal on the y-axis. Visible here is the emotional states are self-assessed in the quartiles that the circumplex model described they would be, although also visible is that the self-assessment is not restricted to their described quartile. With the exception of relax, the self-assessment seems to spread out over two quartiles, and in case of sadness three quartiles. Note, in this figure it is also visible that sadness and fear videos evoked both a positive and a negative valence state.

5 Discussion

An experiment was conducted in natural settings where 72 participants viewed film clips in three different screen size categories, TV, laptop, and mobile. After each film clip, their emotional response during the viewing was self-assessed on a SAM scale. Two repeated-measures MANOVAs were conducted, the first for assessing if the temporal spacing was effective and the second to assess if screen size, extraversion, and neuroticism influenced the emotional response to film stimuli. The multivariate normality distribution was violated for amusement-laptop and fear-mobile, other groups were multivariate normal distributed. As a result an outlier was removed, and due to the normal distribution groups being in the majority, the data was not further transformed.

5.1 Interpretations

Different interpretations are made for emotion, personality, and screen size. The temporal spacing significantly reduced various emotional responses to the film stimuli. Indicated that 20 seconds of black screen effectively decreased the emotional state induced by the film clips. Post-hoc tests showed that positive valence/arousal decreased and negative valence/arousal increased after 20 seconds. The temporal spacing did not influence valence when a relaxed emotional state was induced, and arousal was not influenced by the temporal spacing with emotional states of amusement and sadness. Both valence and arousal determine what emotional state was felt and at which intensity. When an emotional state was further from neutral, the emotional state has a stronger intensity, and when an emotional state changes direction, it transfers into a different emotional state. Valence and arousal together form polar coordinates in the circumplex system that represents a point in the circumplex space that indicates an emotional state. A change in either of the dimensions would be a change in emotion. Therefore, if the temporal spacing did not influence one of the emotional dimensions, the other dimension could change emotional response. Therefore, the results showed a successful decrease of the various emotional response after 20 seconds of temporal spacing when all conditions are considered together. However, when visualising this effect per condition the effect of temporal spacing for each emotional state did not clearly appear for each condition. The influence of temporal space was visible for emotional state fear, in all conditions for both emotion dimension and for sadness it was visible in all condition for valence.

The results showed a significant influence of emotional state on valence and arousal. Both valence and arousal, shown to have a good predictive power and was able to explain 44% and 49% of all variance for the valence and arousal dimensions of VA model. Illustrate that both arousal and valence have equal strength in the account of the proportion of variance and have contributed equally to the significant result. This illustrates that emotion was equally explained by valence and arousal. The various film clips induced valence and arousal self-assessment that differed significantly per emotional state. Where amusement and relax evoked a higher valence response than fear and sadness, and fear and amusement evoked a higher arousal response than sadness and relax. The self assessment of the emotional states were in the quartile that was coherent with the circumplex model [86], and film clips elicited the emotional states validated by previous research [85, 39, 95, 89]. This indicates that emotional self-assessment through the SAM scale is an effective method to report emotion. In accordance with prior studies that stated singular measurements as the SAM scale was a sufficiently reliable measurement of emotion. [73, 43].

The assessment of valence for fear and sadness was somewhat surprising because in the circumplex model, the fear was placed in the quartile of high arousal and low valence. While, at the same time, visible in Figure 11, also place in the circumplex of high valence. It was indicating that the self-assessment of fear was spread out over low and high valence. This could indicate that the sad/horror film enjoyment paradox was experienced while watching film clips that are sad or scary. The dominant emotional response could be pleasant due to the film stimuli evoking emotional response by not living through an experience itself but being guided through an experience with imagination by watching a scene with emotional content. Different authors argue to why there is enjoyment in negative affect films [37, 75, 43], hypothesize that it was not the experience of negative emotional states that results in a positive emotional state, but rather the experience of negative emotional states are viewed as gratifying or as suspense. There was not a guarantee that negative affect films are viewed as entertaining, and when viewed as entertaining, there was not guaranteed to be predominant of fear or sadness, as previously stated by [1].

The interpretations of personality are in contrast to earlier findings that stated that extraversion were more likely to experience positive emotions and neuroticism were more likely to experience negative emotions [54, 21, 69]. No evidence was found towards personality traits, extraversion and neuroticism, have a sensitivity towards negative or positive emotions. An explanation for this might be that personality trait was measured in a point system, where a higher number of points represent a higher degree in a certain personality trait. The dimensional scale was divided into two groups, and the degree of the trait was disregarded to perform a categorical comparison. It cannot be determined whether there was a difference in emotional response between those who are moderately extravert and those who are highly extravert. Solely the possession of the trait was used as a variable.

This study has been unable to demonstrate that screen size influence valence and arousal. Additionally, extraversion and neuroticism were not of influence on the non-significant relation between screensize and the two emotion dimension. These study findings contradict previous research that states that screen size does influence valence and arousal. While previous research has focused on screen size in lab settings where the viewing distance is controlled, these results are not replicated in natural settings, where the interaction with the device was nonrestricted. These results may be explained by the estimated viewing distance at which the devices are held from the face, due to the absence of restrictions for viewing distance. Therefore the FOV was estimated with an average of the viewing [111, 96, 19], which gave on indication that the FOV could be roughly equal between the screen size conditions. The estimation resulted in degrees of horizontal and vertical FOV that are in close range of each other. Thus, despite the different screen sizes, the degree of FOV could be similar. FOV could be a more determining factor in media form presentation than screen size was.

5.2 Implications

With all the conditions combined, there was a significant effect of temporal spacing on emotional states. The temporal spacing between studies seems to differentiate. Where [39] also applied temporal spacing of 20 seconds and [94] applied a temporal spacing of 2 minutes. Based on the study results, it can be implicated that 20 seconds are effective in diminishing emotional states but it does not indicate elimination of the carryover effect of emotion. There was however, no consensus on the length of one phenomenon ending and moving on to the next phenomenon. Theses findings should be interpreted with caution since the effectiveness of temporal spacing was not visually present in all conditions when presented in Figure 11. This effect of temporal spacing can influence the emotional states, but also fatigue and boredom can influence emotional states. The duration of an experiment can be shortened by applying short-term temporal spacing, such as 20 seconds, which can may prevent fatigue or boredom. However, for experiments where the experiment's duration can be more flexible, longer temporal spacing can be applied to ensure a greater removal of the carryover effect. But 20 seconds with black screen has been shown to be effective to reduce emotional states which means a balance between burden and boredom can be found through temporal spacing by decreasing the experiments duration without having to cutback on the number of stimuli.

Emotion can be powerful tool to motivate, persuade, and guide people, therefore necessary to research which factors of media form can elicit a more intense emotional response. The results implied that screen size did not have an influence on the emotional response to film stimuli, in contrast to previous research [80, 60, 82]. Due to the freedom the participants had, they could move their device to their preferred viewing distance and change the FOV to their liking. With the exception of TV, which was wall-mounted. Where in previous research the viewing distance was kept constant, the screen size showed an influence on the immersion or emotion. One particular research [104], calculated the FOV for the different screen sizes on constant viewing distance, and it appeared that FOV has an influence on synergy. Indicating that the FOV for screens can have an influence on the interaction that humans have with devices. Naturally, small screen sizes can easily have smaller FOV than laptop or TV, but in natural settings the restriction on viewing distance disappears, allowing different screen size to have same FOV. Indicating that the FOV of screens can contribute to the emotional response. The results of this study implied that the effect may be attributed to FOV instead of screen size. A question that arose out of the results was whether FOV could be a factor in the intensity of the emotional response. Which led to the following questions that need to be asked regarding everyday 2D screen devices; if in the real world viewing distance is adjusted as a response to screen size to obtain a certain FOV; if the viewing distance adjustments that are made result in equivalent FOV between individuals, and thereby can minimize differences between media from presentation, and could minimize differences between emotional responses across individuals; and whether FOV has an influence on the intensity of emotional responses.

5.3 Limitations

5.3.1 Experiment design

There are several limitations in the experimental design and the software used to conduct the experiment. Due to the Qualtrics randomization tool, the video orders resulted in uncontrolled randomization and resulted in variation sequences that were of awkward nature. This is a result of Qualtrics functionalities. Despite, Qualtrics being a popular and widely used tool, mainly in social sciences, it unfortunately misses a basic function for pseudo-randomization. Alternatively this can only be achieved by expressing every variation of variable orders in a separate survey. Qualtrics offers a function that counts how many times a survey has been completed, for which a control function can be used to make the count equal among surveys. However, a large number of possible variations for multiple conditions would result in a greater number of unique surveys needed, which was time-consuming and raises the possibility of errors and inconsistencies in the surveys. Another issue was that when the obtained data was downloaded, the order of the variable variations resulting from the randomization were not automatically included in the dataset. The option to include the variations in the dataset was not directly visible and was tucked away under two clicks. Important, while the order in how variables are presented to the participants was a crucial factor in the interpretation of results. This creates a chance of misinterpreting the data due to possibly missing knowledge about the data structure. Because of to this function, the interpretation of the results must be taken with caution since the order of the videos can have an influence on each other, and the order effects are not counterbalanced since not every variation happened once.

Furthermore, the experimental had a mixed design conducted with 72 participants, each between-subject group contained 24 participants. This reduced to 69 participants after the removal of outliers due to the violation of multivariate normality, which let to an even smaller sample size per condition. Using distributions that do not conform to the assumption of a MANOVA can result in an increase of Type I errors [34].

Experiments conducted outside the lab take advantage of observing real world relations. However, but the real world brings factors in the equation that cannot be controlled, such as temperature, lighting, screen quality, and background noise. These are examples of variables that have known potential to influence an emotional response [57, 18]. However, these factors will always have a presence in the real world, so there is need to examine if a relation between variables is still present in an environment where these factors and noise could be present. To conduct the experiment in natural settings it allows to assess an emotional response representative of how it might likely happen in the real world. However, experiments not conducted in the lab also limit the possibilities to observe participants and ensure that the experiment is carried out correctly.

5.3.2 Emotion

When the self-assessment is done is a longstanding controversy for the most optimal moment and least disruptive timing to perform self-assessment. Therefore, a single momentary measurement of valence and arousal after the film clip can result in a loss of knowledge of the emotional trajectory during the clip. Although continuous measurement techniques are in existence [94, 112], these techniques are present on the screen or need external devices that can be disruptive in the measurement and therefore question the ecological validation. On the other hand, self-assessment of emotion after the video does not disrupt or distract away from the stimuli and the single measurement of SAM scale has been proven to be effective [73, 43].

The arousal and valence self-assessment was designed to record a single emotional state. There was no method to report mixed emotions. This allowed to capture the emotional state that was most prominent, or in case of mixed emotions, the emotional state with the upper hand. The drawback of determining the predominant emotional state is that it remains unclear if mixed emotions were experienced. While mixed feelings are a well-known phenomenon in films, i.e., the sad film paradox and nostalgia [37, 75, 43, 85]. This restricts the interpretation of the results since achieving a pure emotional state can be a difficult notion to achieve with stimuli as complex as film, and emotion itself being a complex concept.

Although the results indicate that the film elicits the intended emotional states, there are several limitations to how the film clips are embedded in the survey. First of all, the title of the YouTube video was visible, which gave extra context to the viewer while watching the film clips. Additionally, at the end of each YouTube video, suggested videos for further watching are provided. This occurs at the end of the video and between the self-assessment. It is possible that these pop-ups can distract away from the emotional state. Although these are major limitations, they are outweighed by the fact that YouTube videos can be distributed freely and provide easy access to film stimuli for conducting an experiment in the real world.

Other uncontrollable but essential factors important to mention are that every device can have different brightness and volume settings. Brightness and color are factors that reduce the standardization across the film clips [57, 18]. Moreover, for sound, the participants were requested to play sound out loud to reduce any auditory immersion caused by headphones however, the volume or quality of the sound is uncontrolled. Audio is a sense that can increase the perceptual immersion by absorbing an additional sensor beside the visual sensor [59, 110]. Since film is an audiovisual stimuli, immersion increase caused by audio is unpreventable. Volume can be adjusted during the experiment, and the volume level can differ per individual but in a real world situation behavior aspects are more challenging to control than design static aspects. Therefore, this factor is left out of the analysis since it was the aim of this thesis to measure interaction in the real world as it is likely to happen, including the volume control.

5.3.3 Personality

The personality questionnaire did not have limitations since it was established to be effective in measuring personality traits across multiple nationalities and relatively short. However, an artificial dichotomy was created that divided the personality traits into two groups instead of on a dimensional scale to be applied for the RM MANOVA. This is a loss of information about the degree of the personality trait and if that affects the emotional response. This data analysis is imitated by the general view applied to personality traits.

5.3.4 Screen size

Another note that is important to bear in mind is screen sizes are divided into the categories of big, medium, and small, based on popular devices used in daily life. The actual screen size is not taken into account. Due to the lack of various screen sizes in TV.

The viewing distance and the resulting FOV could be a more vital factor in the emotional response to film stimuli. However, the calculation where the hypothesis originates from is limited to the average of viewing distance that was presented in previous research [111, 96, 19]. While viewing distances are estimated to be dynamic over time, phones tend to move closer to the face the longer the phone is held. Also, the viewing distance can be used as zoom-in function when images are not clear. Likewise the equation used was based on a fixed head position [104]. Further, the viewing angle can influence the perception of the object [25]. Additionally, personal factors influence the viewing distance, thereby the FOV, such as age or vision that influence the functional field of view [4, 92], or the body's position that can influence the viewing distance [111]. These are factors not taken into consideration with the estimation of viewing distance. Therefore a basal estimation assessed the viewing distance across all participants, and screen sizes were based on an average. Since the experiment was conducted in natural settings technology options were limited to measure very precise FOV variables, such as viewing distance continuously. The estimation does give an indication towards the importance of FOV and can give direction for future research.

5.4 Recommendations/future work

Further research should be undertaken to investigate the influence of FOV and personality traits on emotion. The hypothesized FOV that was estimated to be roughly equal among different conditions, future research is needed to evaluate how the viewing distance originates in natural settings. FOV is dependent on the viewing distance that needs to be assessed if the distance is dynamically or static and whether this depends on the emotional state or is influenced by the emotional state. The various aspects that determine the actual viewing distance should be taken into consideration to assess how FOV comes about in the real world. The possibility of using viewing distance as a zoom-in function, which is easier for mobile and laptops than for TV, can relate to FOV and emotion intensity.

Additionally, to develop a complete picture of the intensity of emotional response elicited through stimuli presented through media forms, sound can be taken into consideration to get a full view of the perceptual immersion. Since, this Thesis focus was on the visual aspect in the virtual experience and not on the auditory experience.

6 Conclusion

This Thesis aimed to assess if personality traits, extraversion and neuroticism, and screen size had an influence on the intensity of the emotional response to film stimuli. Based on quantitative analysis that compared the intensity of emotional responses of amusement, relax, sadness and fear, to different screen size conditions. None of the emotional states had a different response in intensity when viewing film stimuli on various screen sizes. Both personality traits, extraversion and neuroticism, appeared to have a nonsignificant influence on positive and negative emotional states. On these findings, it can be concluded that screen size and personality do not have a relation with the intensity of emotional responses in natural settings. It was, however, assessed that 20 seconds of temporal spacing did reduce the emotional state that was elicited by the film stimuli. Although, the significant effect became less visible per condition when the valence and arousal assessment was visualized in a graph. The question arose that in natural settings, different screen sizes might result in different viewing distances. Therefore it was hypothesized that FOV could be a more crucial factor than screen size in the intensity of emotional response. To better understand the implications of these results, future studies should address the possible relation FOV can have with the intensity of emotional responses to media presented on everyday media forms.

7 Acknowledgements

Throughout this Thesis I received a great deal of assistance. I would like to thank my supervisor, Egon van den Broek, for his guidance and advise through each stage of the process, my daily supervisor Lukas Arts for his constructive suggestions and assistance during this process and my second supervisor A. Feelders for his contribution and time. I would also like to thank my friend who allowed me to use his hardware to conduct this experiment and all who participated in the experiment.

References

[1] E. B. Andrade and J. B. Cohen. On the consumption of negative feelings. *Journal of Consumer Research*, 34(3):283–300, 2007.

[2] K. W. Arthur. *Effects of field of view on performance with head-mounted displays.* The University of North Carolina at Chapel Hill, 2000.

[3] D. A. Atchison, G. Smith, and G. Smith. *Optics of the human eye*, volume 2. Butterworth-Heinemann Oxford, 2000.

[4] K. K. Ball, B. L. Beard, D. L. Roenker, R. L. Miller, and D. S. Griggs. Age and visual search: Expanding the useful field of view. *JOSA A*, 5(12):2210–2219, 1988.

[5] J. Banbury. Wide field of view head-up displays. *Displays*, 4(2):89–96, 1983.

[6] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, and B. Rey. Immersion and emotion: their impact on the sense of presence. *Cyberpsychology & behavior*, 7(6):734–741, 2004.

[7] L. F. Barrett. Discrete emotions or dimensions? the role of valence focus and arousal focus. *Cognition & Emotion*, 12(4):579–599, 1998.

[8] L. F. Barrett. Feelings or words? understanding the content in self-report ratings of experienced emotion. *Journal of personality and social psychology*, 87(2):266, 2004.

[9] L. F. Barrett, K. A. Lindquist, and M. Gendron. Language as context for the perception of emotion. *Trends in cognitive sciences*, 11(8):327–332, 2007.

[10] P. T. Barrett, K. V. Petrides, S. B. Eysenck, and H. J. Eysenck. The eysenck personality questionnaire: An examination of the factorial similarity of p, e, n, and l across 34 countries. *Personality and individual differences*, 25(5):805–819, 1998.

[11] V. Bettadapura. Face expression recognition and analysis: the state of the art. *ArXiv preprint arXiv:1203.6722*, 2012.

[12] S. L. Bosacki and C. Moore. Preschoolers' understanding of simple and complex emotions: Links with gender and language. *Sex roles*, 50(9-10):659–675, 2004.

[13] S. Bowden, D. Saklofske, F. Van de Vijver, N. Sudarshan, and S. Eysenck. Cross-cultural measurement invariance of the eysenck personality questionnaire across 33 countries. *Personality and Individual Differences*, 103:53–60, 2016.

[14] M. M. Bradley, M. Codispoti, B. N. Cuthbert, and P. J. Lang. Emotion and motivation i: defensive and appetitive reactions in picture processing. *Emotion*, 1(3):276–298, 2001.

[15] M. M. Bradley, M. K. Greenwald, M. C. Petry, and P. J. Lang. Remembering pictures: pleasure and arousal in memory. *Journal of experimental psychology: Learning, Memory, and Cognition*, 18(2):379–390, 1992.

[16] M. M. Bradley and P. J. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[17] S. Buechel and U. Hahn. Emotion analysis as a regression problem—dimensional models and their implications on emotion representation and metrical evaluation. In *Proceedings of the Twenty-second European Conference on Artificial Intelligence*, pages 1114–1122, 2016.

[18] S. Carvalho, J. Leite, S. Galdo-Álvarez, and O. F. Gonçalves. The emotional movie database (emdb): A self-report and psychophysiological study. *Applied psychophysiology and biofeedback*, 37(4):279–294, 2012.

[19] N. Charness, K. Dijkstra, T. Jastrzembski, S. Weaver, and M. Champion. Monitor viewing distance for younger and older workers. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 1614–1617. SAGE Publications Sage CA: Los Angeles, CA, 2008.

[20] T. Cochrane. Eight dimensions for the emotions. *Social Science Information*, 48(3):379–420, 2009.

[21] P. T. Costa and R. R. McCrae. Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. *Journal of personality and social psychology*, 38(4):668, 1980.

[22] A. S. Cowen and D. Keltner. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, 114(38):E7900–E7909, 2017.

[23] D. DeSteno, R. E. Petty, D. D. Rucker, D. T. Wegener, and J. Braverman. Discrete emotions and persuasion: the role of emotion-induced expectancies. *Journal of personality and social psychology*, 86(1):43–56, 2004.

[24] J. M. Digman. Higher-order factors of the big five. *Journal of personality and social psychology*, 73(6):1246–1256, 1997.

[25] A. Dillon, C. McKnight, and J. Richardson. Reading from paper versus reading from screen. *The computer journal*, 31(5):457–464, 1988.

[26] Eagle Burger. The hangover: Wake up scene (movie clip), 2018. Retrieved 07-07-2021, `https://www.youtube.com/watch?app=desktop&v=Zdlq4VQDnwI&ab_channel=EagleBurger`.

[27] T. Eerola and J. K. Vuoskoski. A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39(1):18–49, 2011.

[28] P. Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[29] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124–129, 1971.

[30] M. El Ayadi, M. S. Kamel, and F. Karray. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern recognition*, 44(3):572–587, 2011.

[31] S. B. Eysenck, H. J. Eysenck, and P. Barrett. A revised version of the psychoticism scale. *Personality and individual differences*, 6(1):21–29, 1985.

[32] L. A. Feldman. Valence focus and arousal focus: individual differences in the structure of affective experience. *Journal of personality and social psychology*, 69(1):153–166, 1995.

[33] J.-M. Fernández-Dols and C. Crivelli. Emotion and expression: Naturalistic studies. *Emotion Review*, 5(1):24–29, 2013.

[34] H. Finch. Comparison of the performance of nonparametric and parametric manova test statistics when assumptions are violated. *Methodology*, 1(1):27–38, 2005.

[35] J. R. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellsworth. The world of emotions is not two-dimensional. *Psychological science*, 18(12):1050–1057, 2007.

[36] N. H. Frijda. The laws of emotion. *American psychologist*, 43(5):349–358, 1988.

[37] B. Gaut. The paradox of horror. *Arguing about art: Contemporary philosophical debates*, pages 317–329, 1993.

[38] M. E. Grabe, M. Lombard, R. D. Reich, C. C. Bracken, and T. B. Ditton. The role of screen size in viewer experiences of media content. *Visual Communication Quarterly*, 6(2):4–9, 1999.

[39] J. J. Gross and R. W. Levenson. Emotion elicitation using films. *Cognition & emotion*, 9(1):87–108, 1995.

[40] H. Gunes and M. Pantic. Automatic, dimensional and continuous emotion recognition. *International Journal of Synthetic Emotions (IJSE)*, 1(1):68–99, 2010.

[41] K. C. Gunthert, L. H. Cohen, and S. Armeli. The role of neuroticism in daily stress and coping. *Journal of personality and social psychology*, 77(5):1087–1100, 1999.

[42] D. Hagemann, E. Naumann, S. Maier, G. Becker, A. Lürken, and D. Bartussek. The assessment of affective reactivity using films: Validity, reliability and sex differences. *Personality and Individual differences*, 26(4):627–639, 1999.

[43] J. Hanich, V. Wagner, M. Shah, T. Jacobsen, and W. Menninghaus. Why we like to watch sad films. the pleasure of being moved in aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts*, 8(2):120–129, 2014.

[44] E. Harmon-Jones, C. Harmon-Jones, and E. Summerell. On the importance of both dimensional and discrete models of emotion. *Behavioral sciences*, 7(4):66–82, 2017.

[45] W. IJsselsteijn. Presence in the past: what can we learn from media history? In *Being there: Concepts, effects and measurements of user presence in synthetic environments*, pages 17–40. Ios Press, 2003.

[46] C. E. Izard. Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, 99(3):561–565, 1992.

[47] C. E. Izard. Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115(2):288–299, 1994.

[48] C. E. Izard, D. Z. Libero, P. Putnam, and O. M. Haynes. Stability of emotion experiences and their relations to traits of personality. *Journal of personality and social psychology*, 64(5):847–860, 1993.

[49] H. Jenkins. *Convergence culture*. new york university press, 2006.

[50] J. M. Kory and S. K. D'Mello. Affect elicitation for affective. *The Oxford handbook of affective computing*, pages 371–383, 2015.

[51] P. Lang. Behavioral treatment and bio-behavioral assessment: Computer applications. *Technology in mental health care delivery systems*, pages 119–137, 1980.

[52] P. J. Lang, M. M. Bradley, B. N. Cuthbert, et al. International affective picture system (iaps): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 1(39-58):3–8, 1997.

[53] J. T. Larsen and A. P. McGraw. The case for mixed emotions. *Social and Personality Psychology Compass*, 8(6):263–274, 2014.

[54] R. J. Larsen and T. Ketelaar. Extraversion, neuroticism and susceptibility to positive and negative mood induction procedures. *Personality and individual Differences*, 10(12):1221–1228, 1989.

[55] J. E. LeDoux and R. Brown. A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences*, 114(10):E2016–E2025, 2017.

[56] C. M. Lee, S. S. Narayanan, and R. Pieraccini. Classifying emotions in human-machine spoken dialogs. In *Proceedings. IEEE International Conference on Multimedia and Expo*, volume 1, pages 737–740. IEEE, 2002.

[57] H. C. Lench, S. A. Flores, and S. W. Bench. Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychological bulletin*, 137(5):834–855, 2011.

[58] J.-W. Lin, H. B.-L. Duh, D. E. Parker, H. Abi-Rached, and T. A. Furness. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings ieee virtual reality 2002*, pages 164–171. IEEE, 2002.

[59] M. Lombard and T. Ditton. At the heart of it all: The concept of presence. *Journal of computer-mediated communication*, 3(2):JCMC321, 1997.

[60] M. Lombard, T. B. Ditton, M. E. Grabe, and R. D. Reich. The role of screen size in viewer responses to television fare. *Communication reports*, 10(1):95–106, 1997.

[61] M. Lombard, R. D. Reich, M. E. Grabe, C. C. Bracken, and T. B. Ditton. Presence and television. the role of screen size. *Human communication research*, 26(1):75–98, 2000.

[62] I. Lopatovska and I. Arapakis. Theories, methods and current research on emotions in library and information science, information retrieval and human–computer interaction. *Information Processing & Management*, 47(4):575–592, 2011.

[63] P. N. Lopes, P. Salovey, S. Côté, M. Beers, and R. E. Petty. Emotion regulation abilities and the quality of social interaction. *Emotion*, 5(1):113–118, 2005.

[64] R. E. Lucas, E. Diener, A. Grob, E. M. Suh, and L. Shao. Cross-cultural evidence for the fundamental features of extraversion. *Journal of personality and social psychology*, 79(3):452–468, 2000.

[65] A. M. Lund. The influence of video image size and resolution on viewing-distance preferences. *SMPTE journal*, 102(5):406–415, 1993.

[66] C. Z. Malatesta. The role of emotions in development and organization of personality. *Socioemotional development*, 36:1–56, 1990.

[67] K. A. Matthews and J. Angulo. Measurement of the type a behavior pattern in children: Assessment of children's competitiveness, impatience-anger, and aggression. *Child Development*, pages 466–475, 1980.

[68] Matthia Ellinger. The shining - best scene (hd), 2012. Retrieved 07-07-2021, `https://www.youtube.com/watch?v=19gfhnvtH4c`.

[69] K. Matzler, R. Faullant, B. Renzl, and V. Leiter. The relationship between personality traits (extraversion and neuroticism), emotions and customer self-satisfaction. *Innovative Marketing*, 1(2):32–39, 2005.

[70] D. P. McAdams. The five-factor model in personality: A critical appraisal. *Journal of personality*, 60(2):329–361, 1992.

[71] Measahipster. Bambi mom dies :(, 2016. Retrieved 07-07-2021, `https://www.youtube.com/watch?v=7sCqMEkgLIw&ab_channel=Measahipster`.

[72] J. A. Mikels, B. L. Fredrickson, G. R. Larkin, C. M. Lindberg, S. J. Maglio, and P. A. Reuter-Lorenz. Emotional category data on images from the international affective picture system. *Behavior research methods*, 37(4):626–630, 2005.

[73] J. D. Morris. Observations: Sam: the self-assessment manikin; an efficient cross-cultural measurement of emotional response. *Journal of advertising research*, 35(6):63–68, 1995.

[74] K. Mulligan and K. R. Scherer. Toward a working definition of emotion. *Emotion Review*, 4(4):345–357, 2012.

[75] M. B. Oliver. Exploring the paradox of the enjoyment of sad films. *Human Communication Research*, 19(3):315–342, 1993.

[76] R. W. Picard. *Affective computing*. MIT press, 2000.

[77] PlayNowPlayL8tr. Philadelphia - opera scene(la mamma morta), 2016. Retrieved 07-07-2021, `https://www.youtube.com/watch?app=desktop&v=DwRHwKZSu-w&ab_channel=fightoplanktonfightoplankton&pli=1`.

[78] D. Purves, R. Cabeza, S. A. Huettel, K. S. LaBar, M. L. Platt, M. G. Woldorff, and E. M. Brannon. *Cognitive neuroscience*. Sunderland: Sinauer Associates, Inc, 2008.

[79] Qualtrics. Qualtrics xm - experience management software, 2005. Accessed on 07.2021, `http://www.qualtrics.com/`.

[80] B. Reeves, A. Lang, E. Y. Kim, and D. Tatar. The effects of screen size and message content on attention and arousal. *Media psychology*, 1(1):49–67, 1999.

[81] J. M. Rigby, D. P. Brumby, A. L. Cox, and S. J. Gould. Watching movies on netflix: investigating the effect of screen size on viewer immersion. In *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services adjunct*, pages 714–721, 2016.

[82] J. M. Rigby, D. P. Brumby, S. J. Gould, and A. L. Cox. " i can watch what i want" a diary study of on-demand and cross-device viewing. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, pages 69–80, 2018.

[83] J. M. Rigby, D. P. Brumby, S. J. Gould, and A. L. Cox. Development of a questionnaire to measure immersion in video media: The film ieq. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, pages 35–46, 2019.

[84] G. Riva, F. Davide, and W. A. IJsselsteijn. Being there: The experience of presence in mediated environments. *Being there: Concepts, effects and measurement of user presence in synthetic environments*, 5, 2003.

[85] J. Rottenberg and D. Rebecca. Ray, and james j. gross. *Emotion elicitation using films.," in James A. Coan and John JB Allen, eds., Handbook of emotion elicitation and assessment. Series in affective science., Oxford University Press*, 2007.

[86] J. A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178, 1980.

[87] J. A. Russell. Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145–172, 2003.

[88] J. A. Russell and A. Mehrabian. Evidence for a three-factor theory of emotions. *Journal of research in Personality*, 11(3):273–294, 1977.

[89] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot. Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and emotion*, 24(7):1153–1172, 2010.

[90] K. R. Scherer. What are emotions? and how can they be measured? *Social science information*, 44(4):695–729, 2005.

[91] B. Schuller. Recognizing affect from linguistic information in 3d continuous space. *IEEE Transactions on Affective computing*, 2(4):192–205, 2011.

[92] B. Sekuler, P. J. Bennett, and A. Mortimer Mamelak. Effects of aging on the useful field of view. *Experimental aging research*, 26(2):103–120, 2000.

[93] G. R. Semin, C. A. Görts, S. Nandram, and A. Semin-Goossens. Cultural perspectives on the linguistic representation of emotion and emotion events. *Cognition & Emotion*, 16(1):11–28, 2002.

[94] K. Sharma, C. Castellini, F. Stulp, and E. L. Van den Broek. Continuous, real-time emotion annotation: A novel joystick-based analysis framework. *IEEE Transactions on Affective Computing*, 11(1):78–84, 2017.

[95] K. Sharma, C. Castellini, E. L. van den Broek, A. Albu-Schaeffer, and F. Schwenker. A dataset of continuous affect annotations and physiological signals for emotion analysis. *Scientific data*, 6(1):1–13, 2019.

[96] G. Shin and S. Hegde. User-preferred position of computer displays: effects of display size. *Human Factors*, 52(5):574–585, 2010.

[97] Shout! Factory. When harry met sally (1989) - clip: I'll have what she's having (hd), 2019. Retrieved 07-07-2021, https://www.youtube.com/watch?v=6pQgbEEFPq0&ab_channel=Shout%21Factory.

[98] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang. A review of emotion recognition using physiological signals. *Sensors*, 18(7):2074–2115, 2018.

[99] Spiritual Healing. Relaxing music therapy - relaxing piano music - nature scenes, 2011. Retrieved 07-07-2021, https://www.youtube.com/watch?v=tnHlB8aQp3Y&ab_channel=SpiritualHealing.

[100] G. Stemmler and J. Wacker. Personality, emotion, and individual differences in physiological responses. *Biological psychology*, 84(3):541–551, 2010.

[101] E. S.-H. Tan. Entertainment is emotion: The functional architecture of the entertainment experience. *Media psychology*, 11(1):28–51, 2008.

[102] A. Tellegen, D. T. Lykken, T. J. Bouchard, K. J. Wilcox, N. L. Segal, and S. Rich. Personality similarity in twins reared apart and together. *Journal of personality and social psychology*, 54(6):1031–1039, 1988.

[103] The secrets of Nature. Bambi mom dies :(, 2019. Retrieved 07-07-2021, https://www.youtube.com/watch?app=desktop&v=xm47coWDtb0&ab_channel=TheSecretsofNatureTheSecretsofNatureGeverifieerd&pli=1.

[104] F. van der Sluis, E. L. van den Broek, A. van Drunen, and J. Beerends. Enhancing the quality of service of mobile video technology by increasing multimodal synergy. *Behaviour & Information Technology*, 37(9):874–883, 2018.

[105] Veronique Laurent. Attack scene - "copycat" - sigourney weaver, 2017. Retrieved 07-07-2021, https://www.youtube.com/watch?v=uYNo3sGkYGs&ab_channel=VeroniqueLaurent.

[106] V. T. Visch, E. S. Tan, and D. Molenaar. The emotional and cognitive effect of immersion in film viewing. *Cognition and Emotion*, 24(8):1439–1445, 2010.

[107] P. Vorderer. It's all entertainment—sure. but what exactly is entertainment? communication research, media psychology, and the explanation of entertainment experiences. *Poetics*, 29(4-5):247–261, 2001.

[108] S. Wang and Q. Ji. Video affective content analysis: a survey of state-of-the-art methods. *IEEE Transactions on Affective Computing*, 6(4):410–430, 2015.

[109] D. Watson and L. A. Clark. On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of personality*, 60(2):441–476, 1992.

[110] N. Wilken and J.-L. Kruger. Putting the audience in the picture: Mise-en-shot and psychological immersion in audio described film. *Across Languages and Cultures*, 17(2):251–270, 2016.

[111] M. Yoshimura, M. Kitazawa, Y. Maeda, M. Mimura, K. Tsubota, and T. Kishimoto. Smartphone viewing distance and sleep: an experimental study utilizing motion capture technology. *Nature and science of sleep*, 9:59–65, 2017.

[112] T. Zhang, A. El Ali, C. Wang, A. Hanjalic, and P. Cesar. Rcea: Real-time, continuous emotion annotation for collecting precise mobile video ground truth labels. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2020.

[113] M. Zuckerman, D. M. Kuhlman, J. Joireman, P. Teta, and M. Kraft. A comparison of three structural models for personality: the big three, the big five, and the alternative five. *Journal of personality and social psychology*, 65(4):757–768, 1993.