Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

Department of Information Engineering

Master Thesis in ICT for Internet and Multimedia

Exploring the Emotional Spectrum. Investigating the Impact of Colour and Sounds on Human Emotions through Experimental Studies on Audiovisual Stimuli

Master Candidate

Martina Mazzucato

Student ID 2056095

Supervisor

Prof. Antonio Rodà

University of Padova

Academic Year: 2022/2023

Date of graduation: October 19th, 2023

To my grandparents

Abstract

In an era where effective communication is paramount, Sound Design is gaining every day more importance in the design of multimedia products, enabling them to engender profound emotional responses in their audiences, which are fundamental for its applications in marketing and multimedia, but also in more artistic fields like the world of cinema and video games industries.

Our objective is to delve deep into the world of sound, investigating the intricate dynamics of emotions it evokes when combined to other stimuli: colours, contents and camera motion techniques; basing on objective scientific observations. In particular, we want to verify which are the main types of stimuli that influence the human perception of emotions. To do so we built two questionnaires, in which participants were asked to watch some short videos containing the stimuli and to rate the emotions they felt. The first one focuses on the specific relationship between sounds, colours and evoked emotions; the second, instead, aims to recreate in a more complete way the set of stimuli that we can find in a cinematic scene: sounds, colours, contents, camera movements and their relationship with the emotions the scene evokes. We therefore collected and analyzed the data.
The results of the first experiment highlighted the almost complete dependence of perceived emotions on the sounds presented in the videos, while the influence of colours was, just in some cases, only reinforcing. The results of the second experiment demonstrated that the main influence on the evoked emotions when watching to a cinematic scene is given by the content of the scene itself, and a secondary influence is given by the presented auditory stimuli. Colours and Camera motion techniques resulted to have respectively a minimum influence and no influence on the overall emotional perception.

It is believed that the results of these experiments may contribute to a deeper understanding of the intricate interplay between stimuli in multimedia design, enabling professionals across various industries to harness the power of Sound Design more effectively, to elicit specific emotional responses from their audiences.

Sommario

In un'era in cui una comunicazione efficace è essenziale, il Sound Design sta acquisendo ogni giorno più importanza nella progettazione di prodotti multimediali, consentendo di generare profonde risposte emotive nel pubblico, attributo fondamentale per le sue applicazioni nel marketing e nell'industria, ma anche in campi più artistici, come nel mondo del cinema e dei videogiochi.

Il nostro obiettivo è quello di approfondire le intricate dinamiche delle emozioni che il suono evoca se combinato con altri stimoli: colori, contenuti e movimenti della macchina da presa; basandosi su osservazioni scientifiche oggettive. In particolare si vuole verificare quali siano le principali tipologie di stimoli che influenzano la percezione delle emozioni nell'uomo. Per fare ciò abbiamo creato due questionari, in cui ai partecipanti è stato chiesto di guardare alcuni brevi video contenenti gli stimoli e di valutare l'emozione provata dopo la visione. Il primo si concentra specificamente sulla relazione tra suoni, colori ed emozioni; il secondo, invece, ha lo scopo di ricreare in modo più completo l'insieme degli stimoli che è possibile trovare in una scena cinematografica: suoni, colori, contenuti, movimenti di macchina e il loro rapporto con le emozioni evocate dalla scena. Abbiamo quindi raccolto e analizzato i dati.
I risultati del primo esperimento hanno evidenziato una dipendenza quasi totale delle emozioni dai suoni presentati nei video; mentre l'influenza dei colori è, solo in alcuni casi, rafforzativa. I risultati del secondo esperimento hanno dimostrato che l'influenza principale sulle emozioni percepite dalla la visione di una scena cinematografica, è data dal contenuto della scena stessa. Un'influenza secondaria è, poi, data dagli stimoli uditivi. I movimenti della macchina e i colori, invece, risultano rispettivamente avere influenza minima e parziale e non avere alcuna influenza sulle emozioni percepite dall'audience.

Si ritiene che i risultati di questi esperimenti possano contribuire a una comprensione più profonda dell'intricata interazione tra gli stimoli nel contesto della progettazione multimediale, consentendo ai professionisti di vari settori di sfruttare la potenza del Sound Design in modo più efficace, per suscitare specifiche risposte emotive nel loro pubblico.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Theory of Emotions

Humans exhibit emotional responses to various sounds, encompassing everyday noises, sound effects, or musical compositions. These reactions are influenced by attributes such as pitch, tempo, intensity, frequency, regularity, and more. However, a central query arises when delving into this subject: what exactly constitutes an emotion?

According to the Cambridge Online Dictionary[1], an emotion is defined as *"a strong feeling such as love or anger, or strong feelings in general."* Conversely, the Collins Online Dictionary[2] characterizes an emotion as *"a feeling... which can be caused by the situation that you are in or the people you are with."*. While these definitions suffice for simple and casual discussions of emotions, they prove inadequate from a scientific standpoint.

Renowned American psychologist Robert Plutchik observed that, over the course of the 20th century, more than 90 different definitions of the term "emotion" were proposed by various researchers, psychologists, and theorists, highlighting the substantial lack of consensus on this intricate subject (Plutchik, Kellerman, 2013) [47].

[1] `https://dictionary.cambridge.org/`

[2] `https://www.collinsdictionary.com/`

Numerous researchers have endeavored to explore the realm of emotions, seeking to conceptualize a theory that garners scientific consensus. They have classified emotions as primary, basic, or along dimensions of valence (positivity or negativity) and arousal (excitement or relaxation). Yet, no definitive set of basic emotions has been universally agreed upon, as opinions diverge on the number, identification, and fundamental nature of these emotions.

This study find bases on the categorization of emotions as primary or basic, as elucidated by Charles Darwin's theories (1872)[12]. In particular, it examines the sets of emotions proposed by Robert Plutchik (1980)[46] and Paul Ekman (1992)[15], both of whom offered lists of fundamental emotions that remain influential in contemporary discourse. The theoretical foundations and findings of these researchers underpin the experimental research presented in this thesis.

In particular, Charles Darwin's seminal work, *"The Expression of the Emotions in Man and Animals"* (1872)[12], has played a pivotal role in shaping our understanding of emotional expressions. Darwin posited three principles to elucidate the origins and development of expressive actions and movements triggered by various stimuli. These principles are: the principle of serviceable habits, the principle of antithesis, and the principle of the direct action of the excited nervous system on the body.

The first principle underscores the influence of habit, explaining how certain movements become ingrained in our lives. These movements, initially performed consciously, gradually become habitual and can be triggered whenever corresponding desires or sensations arise. Additionally, actions are intertwined with other actions and mental states. Consequently, when a stimulus alters an individual's mental state, a sequence of movements ensues. Some of these actions may even occur involuntarily, resulting in expressive gestures. For example, when surprised, we instinctively raise our eyebrows to enhance our field of vision.

The second principle, the principle of antithesis, posits that certain movements are diametrically opposed to those generated by specific situations and purposes. Darwin illustrated this concept with the contrasting postures and movements of a dog and a cat preparing to attack or display affection. Darwin suggests that these antithetical actions were, at the beginning, performed voluntarily and, over time, they become habitual.

The third principle asserts that certain expressions serve as outlets for discharging excess nervous energy. For instance, laughter helps dissipate nervous energy stemming from tickling or humorous stimuli. Excitement in a sensitive nerve triggers a cascade of reactions involving the heart, brain, and the body's appearance and movements.

One effective means of discerning emotions is by categorizing them as either exciting (high in arousal) or depressing (low in arousal). When the body and mind are brimming with energy, they respond rapidly, as seen in states of happiness or anger. Conversely, in states of sadness, these responses slow down. These movements, stemming from the nervous system, are acquired through habit, as established by the first principle. Consequently, when faced with a previously encountered exciting or depressing stimulus, individuals unconsciously reproduce the associated reactions.

In Darwin's view, emotional expressions serve as manifestations of underlying emotional states aimed at enhancing survival. These expressions can be observed across various species, transcending age, race, and individual differences, further substantiating their innate and hereditary nature.

Although Charles Darwin's work significantly advanced our understanding of emotional expressions, he did not offer a precise definition of the term "emotion".

Robert Plutchik sought to clarify the concept of emotion and devised a method to measure it. Drawing inspiration from four intellectual traditions on emotions, including Darwin's evolutionary theory (1859) [40], James' psychophysiology (1884)[30], Cannon's neurology (1927)[10], and Freud's psychodynamics (1915) [19], Plutchik constructed a systematic structural model of emotions. This model outlined eight fundamental patterns of behaviour crucial to the process of evolution, grouped into four pairs of polar opposites: destruction and protection, incorporation and rejection, reproduction and deprivation, orientation and exploration. Each pattern of behaviour corresponds to a cluster of emotions essential for survival. Plutchik proposed that emotions represent body reactions to survival challenges, functioning as mechanisms to control specific triggering events.

While Plutchik posited the existence of eight primary emotions: fear, anger, joy, sadness, acceptance, disgust, anticipation, and surprise (well depicted in his Wheel of Emotions (1980)[63], reported in Fig. 1.2), other scholars proposed different numbers and definitions of basic or primary emotions, illustrating the **lack of a universally agreed-upon framework**.

Ekman, in 1992, identified six basic emotions [13]: anger, contempt, disgust, enjoyment, fear, sadness, and surprise. These emotions are considered basic because they are distinct from one another in terms of valence, arousal, reception, and response. They have evolved to address fundamental life tasks - such as fighting, escaping, and reproducing - and exhibit specific physiological patterns. Ekman's 1993 work emphasized the universality of facial expressions in conveying emotions[14], with distinct patterns of nervous system activity associated with specific emotions like anger, fear, disgust, and sadness. Examples are depicted in Fig. 1.1.

In summary, emotions are complex responses that encompass various interconnected elements, including stimulus processing, feelings, psychological changes, impulses to action, and goal-directed behaviour. They serve as reactions to significant situations in an individual's life, communicating information about intentions and facilitating survival. Nevertheless, emotions are rarely experienced in isolation, often coexisting and varying in intensity, similarity, and polarity. This complexity underscores the challenges inherent in studying and precisely defining emotions.

1.2 Colours and Emotions

The profound and intricate relationship between emotions and colours is a path into the crossroads of perception, psychology, and culture. Colours possess a remarkable ability to forge a direct connection with our innermost feeling, evoking a range of emotions that are as diverse as they are powerful. This fu-

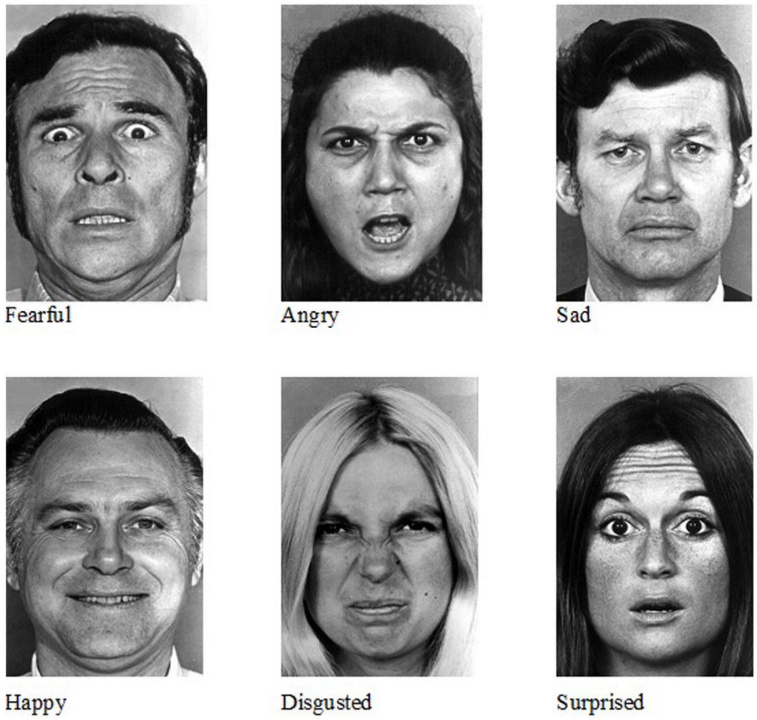


Figure 1.1: Ekman's facial expressions in conveying emotions (1993)[14], each one associated with distinct patterns of nervous system activity associated with specific emotions like anger, fear, disgust, and sadness.

sion of the visual and emotional domains has captivated human minds across time and space, prompting explorations into the phenomenon known as **colour psychology** (Elliot et al, 2016)[16].

From the serene calmness of a soothing blue to the passionate intensity of a vibrant red, each colour is endowed with its own unique ability to elicit specific emotional responses. Across cultures and epochs, colours have been imbued with symbolic meanings and emotional connotations (Mackenzie, 1922)[34], shaping perceptions and behaviours in ways that often escape conscious awareness.

Colour psychology delves into the mechanisms by which colours can evoke emotions and have an impact in human cognition. The exploration of colour symbolism and associations, as well as the examination of how individual differences and personal experiences shape colour preferences, adds depth to our understanding of the intricate interplay between the visual and emotional realms (Tham et al, 2020)[60].

Moreover, the relationship between emotions and colours is a vibrant tapestry woven into various aspects of human life, from art and design to marketing and therapy. Artists wield colours as a palette to convey and evoke emotions, while designers harness their power to create impactful visual experiences. Marketing experts strategically deploy colours to influence consumer perceptions and decisions, and psychologists explore how colour-based interventions can promote emotional well-being and healing (Falcinelli, 2017)[18].

In unraveling the connection between emotions and colours, we embark on a multidisciplinary voyage that traverses aesthetics, psychology, sociology, and beyond. This relationship serves as a testament to the intricate ways in which our senses, cognition, and emotions intertwine, ultimately shaping the rich tapestry of human experience, and continuing to fascinate, inspire, and challenge our understanding.

1.2.1 The Wheel of Emotions

The Wheel of Emotion, also known as Plutchik's Wheel of Emotions [45], is a model of human emotions spectrum developed by Robert Plutchik in 1980. This model is designed to help us understand and categorize a wide range of emotions based on their primary and secondary characteristics. Plutchik organizes these emotions into a circular diagram, which consists of eight primary emotions arranged in pairs of opposites, much like the colours on a colour wheel. These pairs of opposites represent contrasting emotional states.

The eight primary emotions in Plutchik's Wheel of Emotion are:

- Joy - Sadness
- Trust - Disgust
- Fear - Anger
- Surprise - Anticipation

Each of these primary emotions can combine with one another to create secondary emotions, resulting in a more nuanced understanding of human emotional experiences. For example, combining Joy and Trust might result in the secondary emotion of Love, while combining Fear and Surprise might lead to the secondary emotion of Alarm.

Plutchik's model also includes various degrees of intensity for each emotion, allowing for a wide spectrum of emotional experiences. The intensity can range from mild to extreme, which reflects the strength (or *arousal*) of a particular emotion.

It's important to note that Plutchik's Wheel of Emotions, in Fig. 1.2, is just one of many models designed to explain and categorize human emotions. While it provides a helpful framework for understanding emotions, it's essential to recognize that emotions are complex and multifaceted, and no single model can capture the full depth of human emotional experiences. Nonetheless, Plutchik's model remains a valuable tool for exploring and discussing the basic emotions and their interrelationships.

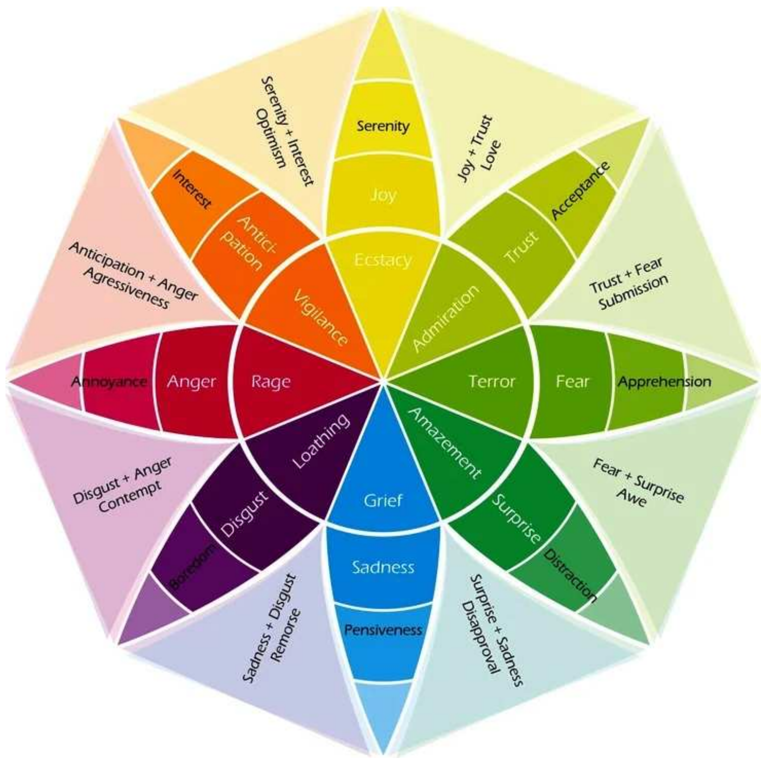


Figure 1.2: The Wheel of Emotions - Plutchik's three-dimensional circumplex model describing the relationships between colours, emotions and their relationship (Plutchik, 1980)[63].

1.3 Sounds and Emotions

The auditory stimuli can have a profound impact on our innermost feelings and states of being. Due to the natural way in which our ears are made and work, preventing sound perception is not possible. Unlike the eyes, which we can close to control visual input, our ears lack such mechanisms, forcing us to listen to sounds every day, in every moment. However, humans do possess the power to decide how to listen, as not all sounds receive the same level of attention.

Sound intention plays a significant role in understanding its meaning, particularly in emergency situations. While the source of a sound expresses intentions, the receiver also plays a role in sound perception.

The intricate interplay between emotions and sounds has given rise to a field of study known as **sound** or **auditory psychology**. The human auditory system can translate a wide range of sounds into nuanced emotional responses, from the melodic cadence of a joyful tune to the haunting echoes of a melancholic melody (Tan et al, 2017)[59]. Just as colours imbue visuals with meaning and resonance, sounds possess the ability to tap into human sentiment. The tones, rhythms, and timbres that make up music, speech, and ambient noise evoke a spectrum of feelings, often bypassing conscious thought and directly influencing emotional landscapes.
Cultural contexts intertwine with this relationship, shaping how sounds are interpreted and the emotions they evoke. Soundscapes of nature, urban environments, and personal surroundings contribute to forming emotional connections and memories, underscoring the fusion of auditory experiences and human emotions.

Researchers and artists seek to unravel the neurological, psychological, and cultural mechanisms underlying the relationship between emotions and sounds. This exploration spans psychology, neuroscience, musicology, and product design, aiming to comprehend how sounds influence mood, cognition, and behaviour. Technological advancements offer opportunities for therapeutic interventions, creative expression, and immersive experiences.

Sound perception's interplay with emotions is rooted in the brain's processing mechanisms. The subcortical pathway, or subcortical filter, categorizes and assigns meaning to sound. Emotional responses are mediated by the amygdala within the limbic system, which is highly sensitive to sound (McAdams & Bigand, 1993)[36].

In essence, the relationship between emotions and sounds is a multidimensional exploration that weaves together our auditory perceptions, emotional landscapes, and cognitive interpretations. As we navigate the intricate harmonies of this relationship, we uncover the profound influence that sounds exert on our emotions, beliefs, and ultimately, our humanity.

1.4 Colours in Cinema

Colour theory is a fundamental aspect of cinematic storytelling, encompassing a wide range of considerations, from hue to saturation and brightness (Coates, 2010)[11].
Filmmakers strategically employ colour to communicate emotions, enhance narrative depth, and create memorable visual experiences. **By carefully selecting colours for a film's palette, directors can evoke specific moods and themes** (Watkins, 2002)[68].
For instance, warm colours like reds and yellows often symbolize passion or danger, while cool blues and greens can evoke sadness or calmness. Moreover, colour can be a potent tool for symbolizing character development or transformation throughout a movie.
Directors frequently utilize contrasting colour schemes to emphasize pivotal moments and direct the viewer's attention to key elements within a scene.

Colours possess an extraordinary capacity to **wield emotional influence**. They have the innate power to evoke specific feelings and moods in the audience, becoming fundamental in the cinematic stage, working in tandem with the narrative to immerse viewers in the emotional landscape of the story.

Beyond eliciting emotions, colours can be harnessed to **symbolize characters and their journeys**. A character's clothing or surroundings may undergo subtle or dramatic colour changes to reflect their transformation or emotional state throughout the film. This symbolic use of colour adds depth to character development, allowing the audience to glean insights into the character's inner world and evolution.

Moreover, the careful coordination of colours contributes to the **visual cohesion** of a film. A harmonious colour scheme helps tie scenes and elements together, creating a more coherent and aesthetically pleasing narrative. This visual continuity enhances the viewer's engagement and aids in conveying the intended mood and atmosphere.

Colours also play a pivotal role in **setting the tone** of different scenes or the entire film. The chosen colour scheme can instantly establish the time period, genre, or atmosphere, providing vital contextual information to the audience. Whether it's the warm, earthy tones of a drama or the cold, metallic hues of a sci-fi thriller, colours serve as visual cues that guide viewers into the desired cinematic experience.

In addition, directors often strategically use colour to **highlight key elements within a scene or as a sign**. By drawing attention to specific objects or details through colour contrast, filmmakers guide the audience's focus and emphasize critical plot points or symbols. This subtle manipulation of colour directs the viewer's gaze and enhances their understanding of the narrative. A famous example of this technique is the red coated girl in *Schindler's List* (Spielberg, 1993)[51], depicted in Fig. 1.3.

As we already pointed out in the previous paragraph, colours also carry **cultural and symbolic significance**, which filmmakers can leverage to enrich their storytelling (Tham et al, 2020)[59]. Certain colours hold deep-rooted meanings in various cultures, and these associations can be harnessed to add layers of depth and nuance to the narrative.

Furthermore, colours are a subtle yet powerful tool for **visual storytelling**. They can convey information, themes, motifs, and even the dynamics of character relationships without the need for explicit dialogue. The choice of colours in a scene can communicate subtext, suggesting hidden emotions, conflicts, or narrative threads that enrich the viewing experience (Rothstein, 2020)[48].

Figure 1.3: Frame from the Steven Spielberg's movie *Schindler's List* (1993)[51]. It is the most famous and evident example of the technique of using colour contrast to highlight key elements within a scene, with the aim to guide the audiences focus and emphasize critical plot points or symbols.

Lastly, colours contribute significantly to the **overall aesthetic appeal** of a film. Beyond their narrative and emotional functions, a well-crafted colour palette enhances the visual beauty and memorability of a cinematic work (Bramesco, 2023)[7]. It leaves a lasting imprint on the viewer's mind, elevating the film from a mere visual experience to a work of art. Some good example of this are Wes Anderson's movies, i.e. *Grand Budapest Hotel* (Anderson, 2014)[25], in Fig 1.4, *The French Dispatch* (Anderson, 2021)[7], *The Wonderful Story of Henry Sugar* (Anderson, 2023)[61], in which the colour palette (together with the used camera motion techniques) is always similar, making his works immediately recognisable and therefore constituting a signature of the director on his artwork.

In essence, colour theory in cinema is a language of its own, enriching storytelling and leaving a lasting impression on the audience, not only as an embellishments but as an integral element of storytelling and artistic expression. It helps to set the tone for each scene and serves as a powerful tool for conveying emotions, defining character traits, and even indicating a character's transformation throughout the story. Different colour palettes, such as monochromatic, analogous, complementary, triadic, and tetradic, are chosen to align with the mood and narrative of the film, ensuring a balanced and meaningful use of colour.

Figure 1.4: Wes Anderson's signature colour palette in *Grand Budapest Hotel* (Anderson, 2014)[25], a master example of the aesthetic usage of colours in movies. Colours contribute significantly to the overall aesthetic appeal of a film, and a well-crafted colour palette enhances the visual beauty and memorability of a cinematic work (Bramesco, 2023)[7]. It leaves a lasting imprint on the viewers mind, elevating the film from a mere visual experience to a work of art.

1.5 Sounds in Cinema

The significance of sound in cinema is immeasurable. It serves as an indispensable component, comprising human voices, music, and sound effects, that profoundly impacts the viewer's emotional connection to a film. Sound in cinema is not merely an accompaniment but a storytelling partner, intricately linked to the visuals. Human voices, such as actor dialogue, must seamlessly synchronize with the film's action, conveying character motivations and plot developments. Music, often performed by a large orchestra, drives the narrative, eliciting emotional responses and aiding scene transitions. Even when the cinema was silent, there was a need to accompany the film with sounds and music that were produced live, in the theaters.

Sound effects add a layer of realism, breathing life into characters and scenes. Sound can be categorized as diegetic, occurring within the story's world, or non-diegetic, external to it, with both types adding vital context (Percheron & Butzel, 1980)[44].

Ultimately, sound in film is an essential element that complements the visuals, enhancing the overall cinematic experience (Watkins, 2002)[68].

First and foremost, sound possesses an extraordinary capacity to **evoke emotional responses** (Rothstein, 2021)[49]. Auditory cues, such as music, voice tone, and sound effects, have a direct and powerful impact on human emotions. Sound becomes a fundamental element in the storytelling process, helping to shape and enhance the emotional landscape of the narrative.

Moreover, the creation of **atmosphere and mood** in cinema is intrinsically linked to soundscapes (Rothstein, 2021)[49].
The choice of music, ambient sounds, and sound effects significantly influences how audiences perceive the setting and emotional tone of a scene or the entire film. For example, a serene forest scene filled with chirping birds and rustling leaves elicits a vastly different emotional response than an eerie soundscape featuring distant howls and creaking trees.

Sound also plays a pivotal role in forging a deeper **connection between the audience and the characters**. Dialogue, with its nuances of tone, pitch, and pauses between words, serves as a direct conduit for conveying characters' thoughts, emotions, and intentions (Rothstein, 2021)[49]. Additionally, background sounds and music work in tandem to underline the emotions a character is experiencing, reinforcing the audience's empathy and understanding of their inner world. In this way, sound acts as a bridge between the fictional plan of the film and the viewer's own emotional landscape, fostering a more profound and immersive connection.

Beyond surface-level emotions, sound can **convey subtext and symbolism** that may not be explicitly depicted visually (Rothstein, 2021)[49]. The choice of musical themes or motifs associated with characters or themes can suggest hidden emotions, past experiences, or underlying motivations. Famous examples of this characters-themes association are Angelo Badalamenti's characters' themes in the serial drama Twin Peaks (Lynch, 1990-2014)[64]; each theme is associated to a different character and plays whenever that character is the protagonist of a scene. These subtextual cues add layers of depth to the storytelling, inviting audiences to engage with the narrative on a more nuanced and interpretive level. This intricate use of sound encourages viewers to become active participants in unraveling the emotional complexities of the story.

Sound is also indispensable for creating a heightened **sense of realism and immersion** in cinematic worlds. Whether through the strategic use of silence to induce shock and suspense or the inclusion of subtle background sounds to make a scene feel more authentic, filmmakers rely on sound to transport audiences into the heart of the narrative. By manipulating sound, directors can control the emotional experience of the audience, amplifying the impact of pivotal moments and enhancing the overall cinematic journey.

In terms of **narrative pacing and flow**, sound is a critical tool. The tempo and intensity of music and sound effects can mirror the pace of the story, guiding audiences through moments of tension, action, and relaxation (Rothstein, 2021)[49]. Well-timed auditory cues contribute to the ebb and flow of emotions, aligning the viewer's emotional journey with that of the characters. This synchronization adds depth and resonance to the storytelling, making the emotional arcs of the characters more compelling and relatable.

Lastly, sound in cinema holds **cultural and psychological significance**. Different sounds, including specific musical scales, instruments, or tonalities, may evoke distinct emotions in various cultures. Moreover, sounds can trigger personal memories and associations unique to individual viewers, further intensifying their emotional engagement with the film. In this way, sound transcends linguistic and cultural barriers, offering a universal language of emotions that resonates with audiences on a profound level (Beck, 2010)[4].

In summary, the use of sound in cinema is a multi-dimensional art that intertwines with visuals to create a rich emotional tapestry. By leveraging auditory cues such as music, dialogue, and sound effects, filmmakers can evoke emotions, set tones, and craft immersive experiences that resonate deeply with audiences. The careful orchestration of soundscapes enhances the storytelling process, making cinema a holistic and impactful medium that appeals to both the intellect and the emotions of viewers.

1.6 Camera Movement in Cinema

The significance of camera movements in film production, encompassing their role in storytelling, mood creation, and character portrayal, is widely acknowledged. These movements, through their ability to convey realism and immerse the audience in a fictional world, stand as powerful cinematic tools, effectively bridging the gap between viewers and the mediated plan of fiction (Morgan, 2016) [39].

In the early 20th century, the first moviegoers marveled at "phantom rides," a term coined for the captivating effect produced by moving cameras that seemed to glide on train tracks and canals of Venice as if guided by an invisible force (Salt, 2009) [50].
In 1930, Slavko Vorkapich, a Serbian filmmaker who also found success in Hollywood, argued that humans, from newborns to adults, inherently appreciate and find pleasure in motion as conveyed through moving images (Vorkapich, 1930)[66]. He linked the perception of motion on screen directly to the sensory and motor experiences of the audience, suggesting that merely witnessing motion on screen can provoke similar reactions as active participation; in fact, functional neuroimaging studies by Jääskeläinen et al. (2020[28], 2021[27]) have shown that different viewers' brain activations can correlate when they watch the same films, indicating that cinema can transform individual perceptual experiences into a shared, extroverted vision. Moreover, 2012 studies by Gallese and Guerra suggests that viewers' minds simulate or mirror the events they observe on screen (Gallese & Guerra, 2012) [20].

Cognitive-perceptual theories applied to film studies often draw parallels also between the moving camera and the human eye (Bordwell, 1977 [6]; Sobchack, 1982 [55]; Barker, 2009 [3]; Guerra, 2015 [26]; Schonig, 2017 [52]). This involves considering that certain camera movement techniques, like Steadicam and handheld shots (taken into account in this thesis), mimic the dynamic movements of the human eye. Furthermore, camera movement on screen resembles the muscular movements of the human body, creating perceptual similarities between the viewer's experience and physical movement (Barker, 2009)[3].

Cinematic engagement, in essence, enables subjective movement through an objective world. This relationship between movement and space is so fundamental that, without it, cinema as we know it would not exist. In the words of Garrett Brown, the inventor of Steadicam, *"We are there"*, indicating that, much like the human eye, the moving camera explores the physical story space, moment by moment (Brown, 2003) [8] [3].

Furthermore, sound can trigger mental imagery and associations with visual entities. For instance, the Maluma/Takete experiment by Wolfgang Köhler in the '20s (Görne, 2019) [24] demonstrated that people tend to associate the word "Maluma" with round shapes and "Takete" with angular, edgy forms, illustrating how sound can evoke visual metaphors. This phenomenon, known as phonosymbolism, plays a role in crossmodal metaphors and can influence design choices, particularly in fields like marketing, where the sound of a brand name or product can impact consumer perception and decision-making.

1.7 Everyday Sounds

Our world is a rich tapestry of sounds, and our daily auditory experiences, commonly referred to as soundscapes, encompass a diverse range of elements, including music, speech, animal calls, and various environmental noises. These ambient sounds, often categorized as *everyday sounds*, can encompass actions as subtle as the rustling of pages in a book, the jingling of car keys, the resonance of a car door closing, or the rhythmic tap of fingers on a keyboard. It's worth noting that these everyday sounds are primarily generated by human activities (Susini, Hoiux & Misdariis, 2014) [58] , and this study primarily focuses on investigating these types of sounds.

Everyday sounds fall into the category of *impact sounds*, characterized by impulsive signals resulting from brief interactions between objects (Visell et al., 2009) [65]. When we hear a sound, our perception is shaped by its various perceptual and physical parameters, such as loudness, pitch, timbre, reverb, and frequency. Remarkably, humans possess the ability to discern the source or

[3] `https://www.garrettcam.com/the-moving-camera-part-1`

cause of a sound simply by listening, a skill rooted in evolutionary necessities such as danger awareness.

These sound characteristics convey a multitude of information, such as the source's spatial location, time, intent, emotional state, and interpersonal relationships. Regarding spatial information, louder sounds are generally perceived as originating closer to the listener, while reverb and echo provide insights into the size and acoustic properties of the space. Materials in the environment also influence these parameters, with different materials absorbing or transmitting sound waves differently.

In terms of time, the regularity or rhythm of a sound, like the ticking of seconds on a clock or the pace of music, helps us track time and sets the tempo for our perception of events. Additionally, specific combinations of physical and perceptual sound parameters can indicate intent. For instance, urgency is often conveyed through a fast rate, high pitch, and irregular harmonics (Susini, Hoiux and Misdariis, 2014) [58]. Consider, for instance, the urgency communicated by an alarm. Similarly, everyday actions like knocking on a door can reveal emotional states, with high-frequency, high-volume knocks possibly indicating fear and softer, less intense knocks suggesting sadness.

Voice plays a significant role in perceiving emotions, with variations in prosodic parameters like pitch, duration, loudness, and voice quality allowing us to gauge happiness, anger, sadness, or surprise. Pitch, in particular, carries substantial emotional connotations, with higher pitches often associated with happiness and lower pitches with anger or sadness (Latinus & Pascal, 2011)[32].

Knocking Sounds

These acoustic cues extend beyond emotions to provide information about physical attributes and interpersonal dynamics.

Knocking on a door, a common everyday sound, not only conveys information about the person knocking and the door's material but can also indicate the relationship between the knocker and the person inside the room. The number, duration, rate, and regularity of knocks can hint at familiarity or emotional states, facilitating non-verbal communication (Visell et al, 2009) [65].

In summary, the multifaceted nature of sound properties, characteristics, and information has far-reaching implications. These principles are essential not only in artistic and entertainment contexts, such as filmmaking and video game design but also in industrial product design. Sound can serve as a powerful tool for storytelling, as evidenced by the creation of audio films for the visually impaired, where sound effects, reverberation, and music convey intricate narratives and emotions without the need for visual cues. Thus, Sound Design is a crucial element in shaping our sensory experiences and perceptions in diverse aspects of our lives.

1.8 Related experiments

This paragraph delves into the context of previous experiments, which form the bedrock upon which our thesis stands. These experiments have paved the way for our own investigations and have significantly shaped the questions we seek to answer. In the chapter are therefore reported the key findings, methodologies, and insights from these seminal studies, underlining their significance in guiding our own research. Through a retrospective lens, we will not only acknowledge the intellectual debt we owe to these pioneering studies but also offer a valuable context for the work presented in this thesis.

1.8.1 Synthesising Knocking Sound Effects Using Conditional Wave GAN

The experiment was conducted by Adrián Barahona-Ríos and Sandra Pauletto in 2020 [1].
The aim was to synthesise new emotional knocking sounds that could potentially be used in the media industry, using GANs[23] and eGANs[38].
The training set was composed of 500 knocking sounds performed by the professional Foley Artist Ulf Olausson [2] [4], keeping in mind five basic emotions:

[4]https://doi.org/10.5281/zenodo.3668503

anger, fear, happiness, neutrality and sadness.
A listening test was therefore performed on the newly generated samples, to verify both if the generated sample would have been perceived as "fake" from the listeners and if the emotions the samples should have conveyed would have been correctly evoked in the listeners.
The results of the experiment was a proved association between knocking sounds and the five basing emotions, except for fear and anger that are frequently confused between each other.

This past study is important for our experiments because it provides us material to build the dataset (we used the same set of knocking sounds) and sets our assumptions that the knocking sounds correctly conveys the associated emotions.

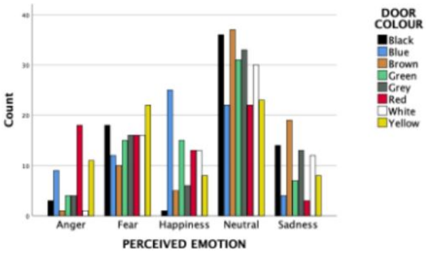
1.8.2 Perception of Emotions in Multimodal Stimuli: The Case of Knocking on a Door

The experiment was conducted by Alessandro Iop and Sandra Pauletto in 2021 [41].
Their aim was to investigate how the visual characteristics (colours, texture and materials) of a door, together with emotionally expressive knocking actions can affect the overall emotion evoked in the audience. This experiment find genesis from the fact that in media productions knocks represent a key moment for the storytelling, preluding to an important transition in the plot, that can be introduced in different ways depending on how the knock scene is performed.

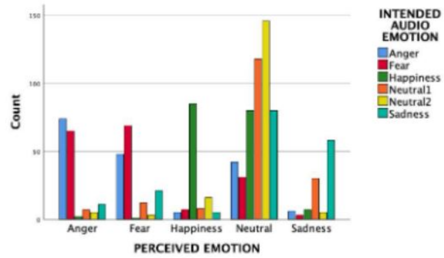
The results, reported in Fig. 1.5, show that the predominant influence on the perceived emotions is the one of the knocking sounds, i.e. the audio stimuli, while the influence of the visual characteristics were little and discardable.

1.8.3 Knocking on a yellow door: interactions among knocking sounds, colours, and emotions

The experiment was conducted by A. Turato, A. Rodà, S. Canazza, A. Chmiel, M. Murari, E. Schubert and J. D. Zhang in 2023[62].
The aim was to investigate the hypothesis that the association between knocking



(a) Door's Colour vs. Perceived Emotion



(b) Knocking Sound vs. Perceived Emotion

Figure 1.5: Pre-study of Iop and Pauletto's Experiment *Perception of Emotions in Multimodal Stimuli: The Case of Knocking on a Door*, highlighting how the the predominant influence on the perceived emotions is the one of the knocking sounds, i.e. the audio stimuli (they were all correctly and clearly associated to the intended emotion), while the influence of the visual characteristics, i.e. the colour of the door, were not clearly and uniquely associated with the intended emotion.

sounds and emotions change according to the colour of the door.

Participant were given a set of short clips of a hand knocking (with different emotional knocking sounds) on a door of various colours, and were asked to rate the emotion perceived after watching the videos. The collected data were then analyzed through a Three-Ways ANOVA.
Results showed no statistically significant effect of colours on the perception of emotions associated to the knocking sounds. Anyway, some interesting trends can be recognized, e.g., knocking sounds were perceived as happier if the hand knocks on a yellow door.

It's important for the present thesis because starting from its limits we built an experiment to deepen this same analysis, and highlighted a statistically significant dependence of the emotion perceived by the participants to the experiment on the knocking sound presented in the video clips, but not on the colours of the door.

Effect	DFn	DFd	F	Sig.	p<.05
sound	2.96	168.77	23.306	p < .001	*
colour	4.00	228.00	2.747	p = .029	*
emotion	3.00	171.00	7.969	p < .001	*
sound:colour	16.00	912.00	1.407	p = .13	
sound:emotion	6.19	352.84	121.845	p < .001	*
colour:emotion	8.94	509.47	1.733	p = .079	
sound:colour:emotion	21.62	1232.45	1.219	p = .222	

Figure 1.6: Results of Turato et al.'s experiment *Knocking on a yellow door: interactions among knocking sounds, colours, and emotions*[62]. The ANOVA analysis performed on the collected data, that will constitue part of the data taken into account in the Experiment #1 of the present thesis, highlighted a statistically significant dependence of the emotion perceived by the participants to the experiment on the knocking sound presented in the video clips, but not on the colours of the door.

1.8.4 An Embodiment of the Cinematographer: Emotional and Perceptual Responses to Different Camera Movement Techniques

The experiment was conducted by Burak Yilmaz M., Lotman E., Karjus A, and Tikka P. in 2023 [9].

The aim was to give an experimental approach to cinematography, enabling the systematic study of creative intuitions and audience responses in controlled settings, investigating the relationship between camera movement techniques and cognitive responses in audiences.
The results for this experiment were ambiguous: movement affects the sense of involvement but not necessarily the specific emotional response, as reported in Fig. 1.7.

The experiment is important for this thesis because it provided the clips used to build our dataset[5], provided us the idea to add two more cinematic factors (content and camera motion technique) to our Experiment #1 and sets our assumptions that the camera movement techniques influence in some way the perception of the scene.

[5]https://zenodo.org/record/7979713

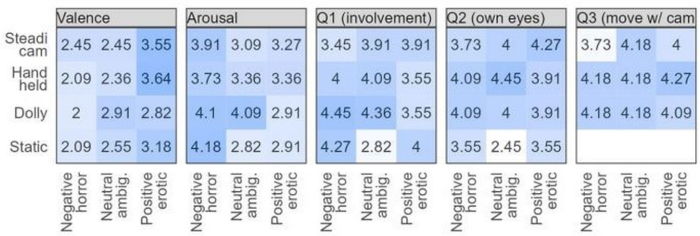


Figure 1.7: Results of Burak et al.'s experiment *Perception of Emotions in Multimodal Stimuli: The Case of Knocking on a Door*[9]. Mean rating values, ranging between 1 and 5, for combinations of camera movement and scene's content, averaged across participants. Values closer to 5 (darker background) correspond to a stronger response, e.g., the horror scene elicits stronger arousal than the erotic scene. In questions 1 and 2, moving cameras, in contrast to a static camera, lead to different immersion rating.

2
State of the Art

The research carried out for this thesis has been divided into two paths, intending to be a continuation of the thesis *"Emotions in everyday sounds. An experimental study on the influence of colour in emotional knocking sounds."*[63], by Asia Turato, presented the last academic year. The first has the objective of continuing the data collection of the experience on knocking sounds to verify whether a greater availability of data reports could lead to more significant results, while the second path has the objective of evolving the previous experiment to make up for the limits identified therein.

2.1 Relationships between Colours and Emotions

Colour psychology is a field that studies the impact of colours on human emotions and behavior. While specific emotional responses to colours can vary across cultures and individuals, certain trends have been identified.

Warm colours, such as red, orange, and yellow, tend to evoke feelings of warmth, vitality, and excitement. This phenomenon has even been explored in the research conducted by Satyendra Singh in his 2006 study, *"Impact of colour on Marketing"* [54], where he found that warm colours can nudge consumers toward impulse purchases, creating a sense of urgency.

On the flip side, colours like blue, green, and purple are often associated with serenity, tranquility, and relaxation. The connection between these colours and a sense of calmness was also noted in Robert Gerard's 1958 study, *"Effects of Four Psychological Primary colours on GSR, Heart Rate, and Skin Temperature"*[29].

However, it's essential to acknowledge that individual differences and cultural backgrounds play a significant role in shaping our colour-emotion associations. Andrew J. Elliot and his team emphasized this in their 2007 study, *"colour Psychology: Effects of Perceiving colour on Psychological Functioning in Humans*" [17]. This underscores that what one person finds soothing, another might find stimulating, based on personal preferences and cultural context.

Moreover, the context in which colours are used, such as in branding or design, can profoundly influence emotional reactions. Angela Wright's 1998 research on the colour Affects System[69] provides valuable insights into how colours can be categorized based on their emotional impact, shedding light on the intricate relationship between colour and our psychological responses.

2.2 Relationships between Sounds and Emotions

Psychoacoustics investigates how sound characteristics influence human emotional responses. Here are some significant concepts and findings.

One of the fundamental aspects is the sounds' pitch. High-pitched sounds often carry associations of excitement, happiness, and positive emotions, while their low-pitched counterparts can evoke feelings of sadness, seriousness, or even danger. Daniel Levitin's work in *"This Is Your Brain on Music"* (2006)[33] provides a comprehensive exploration of how variations in pitch influence our emotional engagement with music.

Other critical factors are tempo and rhythm. Faster tempos and rhythmic patterns tend to infuse music with a sense of energy and excitement, while slower tempos are more closely associated with feelings of calmness and relaxation. The research presented in *"Emotional Responses to Music: the Need to Consider Underlying Mechanisms"* by Juslin and Västfjäll (2008)[31] delves into the nuanced ways in which tempo contributes to emotional expression in music.

Timbre, the unique quality that distinguishes one sound from another, plays a significant role in our emotional perception of auditory stimuli. Peeters et al. explored this aspect in their 2011 study, *"The Timbre Toolbox: Extracting Audio Descriptors from Musical Signals"*[43], followed by Xin Wang et al.'s study *"A Cross-Cultural Analysis of the Influence of Timbre on Affect Perception in Western Classical Music and Chinese Music Traditions"* (2021)[67], shedding light on how timbre shapes listeners' emotional interpretations of music.

Cultural influences also come into play, with different cultures often exhibiting distinct emotional associations with particular sounds. Research such as *"Cultural Differeces in Auditory Ecology"* by Carlos R. Benítez-Barrera et al. (2023)[5] delves into the fascinating ways in which cultural backgrounds shape our emotional responses to auditory stimuli.

Furthermore, psychoacoustics intersects with cross-modal effects, where sensory modalities interact with one another. For instance, people tend to associate high-pitched sounds with light colours and low-pitched sounds with dark colours, forming intriguing cross-modal correspondences. Cesare Parise's study, *"Crossmodal Correspondences: Standing Issues and Experimental Guidelines"* (2016)[42], offers valuable insights into this area of perception, where sound and vision converge in our perception of the world.

2.2.1 Relationships between Camera Movements and Emotions

As previously pointed out, the relationship between camera movement techniques and emotions in cinema is a critical aspect of filmmaking that directors and cinematographers use to convey a wide range of feelings and immerse the audience in the story. The way a camera moves can greatly influence how viewers perceive characters, events, and the overall mood of a scene (Schonig, 2022)[53], (Gallese & Guerra, 2015)[21].

Let's explore this relationship (Gallese & Guerra, 2014)[22] in detail:

- **Static Shots:**
  A static shot, where the camera remains still, can create a sense of stability and calmness. It's often used for dialogues, introspective moments, or when the focus is entirely on the characters' expressions and dialogue. In emotional conversations or intimate moments, a static shot can allow

the audience to connect more deeply with the characters' emotions. It can also convey a sense of stagnation or contemplation when necessary.

- **Dolly/Tracking Shots:**
  Tracking shots involve moving the camera smoothly along a path, either following characters or objects. They can convey a sense of fluidity and engagement with the characters or the environment.
  As an example, in a romantic scene, a dolly shot following two characters walking together can emphasize their connection and intimacy. In suspenseful scenes, tracking shots can build tension by revealing information progressively or by creating a sense of pursuit.

- **Handheld Shots:**
  Handheld camera work adds a sense of realism, immediacy, and chaos. It can make the audience feel like they are part of the action and elicit emotions like anxiety, fear, or excitement.
  In action sequences, a shaky handheld camera can intensify the chaos and danger. In found footage or documentary-style films, it can create a sense of authenticity and vulnerability.

- **Steadicam Shots:**
  Steadicam shots are characterized by smooth, stable movements and are often used for long takes. They can convey a sense of grace, elegance, and continuity.
  A steadicam shot can, as an example, be used for a dance sequence to showcase the beauty of the character dynamic. In a continuous tracking shot, it can immerse the audience in a complex and emotionally charged scene, such as a heated argument or an emotional revelation.

- **Crane Shot:**
  Crane shots involve the camera being mounted on a crane or jib, allowing for sweeping movements from above or below. They can evoke a sense of grandeur, power, or on the other hand, vulnerability.
  A crane shot looking down on a character can convey their vulnerability or isolation. Conversely, a crane shot rising above a crowd can create a feeling of empowerment or hope.

- **Zooms and Push/Pull Focus:**
  Zooming in or out, along with changes in focus, can be used to highlight specific details, emotions, or perspectives within a scene.
  Zooming in on a character's face during a crucial moment can emphasize their emotional reaction. A pull focus from a background object to a character can shift the audience's attention and reveal something significant.

In summary, camera movement techniques are a powerful tool in filmmaking to manipulate the emotions of the audience. Directors and cinematographers choose specific camera movements and techniques to enhance the storytelling and evoke the desired emotional response, whether it's to create tension, intimacy, excitement, or a sense of calm. The art of cinematography lies in the ability to use these techniques effectively to convey the intended emotions and immerse the audience in the cinematic experience.

3

Experiment #1

This section elaborates an experiment conducted by Asia Turato in 2022, *"Emotions in everyday sounds. An experimental study on the influence of colour in emotional knocking sounds."*[63], with the collaboration of Prof. Emery Schubert and Diana Zhang from the University of New South Wales Sydney, in conjunction with Dr. Anthony Chmiel from Western Sydney University; serving as a foundational exploration that paves the way for the ongoing investigation within the present thesis. The central aim of this prior experiment was to examine how colours influence the perception of emotions conveyed through sound, with a specific focus on the context of knocking sounds.

The insights gleaned from this study carry significant relevance across a broad spectrum of communication domains where evoking emotions holds paramount importance. These domains range from cinematic productions and video games to music videos, video advertisements, personal and professional content shared on social media platforms, as well as immersive experiences in virtual reality.

Turato's experiment took place in June 2022, while our Experiment #1 took place from August to September 2023.

3.1 The Previous Experiment

The experiment was designed to investigate how specific colours, known to be associated with particular emotions, impact the emotional interpretation of knocking sounds, drawing upon the groundwork established by Pauletto and Iop (2021)[41], which validated the connection between knocking sounds and emotions within multimodal stimuli, and the research conducted by McDonald et al. (2022) [37], which introduced the variable of door colour to explore potential alterations in this emotional relationship. These prior studies underscored the dominance of auditory perception over visual perception in this context. While this investigation draws inspiration from the previously mentioned studies, it distinguishes itself by its unique colour-emotion associations and by its focus only on sounds and colours stimuli.

In summary, the shared elements between Turato's study and the aforementioned researches include the selection of specific emotions and the utilization of knocking sounds from the same dataset employed in Pauletto and Iop[41]. However, this study adopted a rating scale ranging from 0 to 10 and incorporated follow-up questions related to synesthesia, similar to the approach taken by McDonald et al.[37]. Notably, deviations arise in the use of video stimuli instead of still images as in Pauletto and Iop, as well as our utilization of distinct colour-emotion associations and the implementation of continuous sliders rather than integer scales. Consequently, the hypothesis posits the following colour-emotion associations:

- Yellow elicits Happiness
- Red elicits Anger
- Purple elicits Fear
- Grey elicits Sadness
- White elicits a Neutral response

3.1.1 Stimuli

For this study, we utilized audio stimuli from the dataset of emotional knocking actions compiled by Barahona-Ríos and Pauletto in 2020, which was also employed in the work of Pauletto and Iop in 2021[38]. These knocking actions were professionally performed by Ulf Olausson, an expert Foley artist[2].

The selection of this dataset was motivated by its established validation and the professional quality of the performances. As a result, the sound quality and overall performance were expected to be superior for analysis compared to knocking actions carried out by untrained volunteers.

We specifically chose five audio recordings, each representing a distinct emotion, along with one that was neutral, from the available options.

To create the visual stimuli, videos of a hand knocking on a white door have been recorded and the chosen knocking sounds have been synchronized to the videos.
These videos were captured using an external camera on an iPhone 8 and subsequently edited using HitFilm Express to incorporate the audio and trim them to a uniform length of 3 seconds. Further adjustments, including altering the colour of the door, were made using Adobe Premiere Pro.

The following colour-emotion associations were established (graphically reported in Table 3.1):

- Happiness was linked with yellow (hex code: #EDE939)
- Anger was associated with red (hex code: #D7414B)
- Fear was connected to purple (hex code: #A36AB8)
- Sadness was correlated with grey (hex code: #838686)
- Neutral emotions were represented by the natural colour of the door (average hex code: #E6DFDC)

Happiness	Sadness	Fear	Anger	Neutral
Yellow	Gray	Purple	Red	Neutral

Table 3.1: Table of colour-emotion associations for the Turato's Experiment (2022), *"Emotions in everyday sounds. An experimental study on the influence of colour in emotional knocking sounds."*[63]. This hypothesis on the colour-emotion association is shared with our Experiment #1.

In total, we produced twenty-five audiovisual materials, comprising five audio stimuli paired with five distinct colours.

3.1.2 Procedure

An online questionnaire was developed using the PsyToolkit online platform (Stoet, 2010, 2017)[56] [57] [1]. Participants were initially provided with an overview of the study's purpose and the privacy policy.

To allow participants to adjust the volume of their device's speakers or headphones, before commencing the experiment they were required to listen to two audio samples (different from the ones used in the survey to prevent any bias, but still taken from the Olausson's dataset[2]) - one with high intensity and the other with low intensity, and were asked to maintain the same volume level throughout the whole experiment.

Subsequently, they were acquainted with the survey's structure. The questionnaire consisted of twenty-five questions; each page featured a 3-seconds video clip followed by four sliders. After viewing the video, which they could replay as many times as they wished, participants were tasked with assessing the emotion conveyed in the video by adjusting the four sliders placed on a continuous scale spanning from 0 (indicating *complete disagreement*) to 10 (indicating *complete agreement*). Prior to each slider, participants encountered the following statements:

- The emotion expressed in the video is HAPPINESS.
- The emotion expressed in the video is SADNESS.
- The emotion expressed in the video is FEAR.
- The emotion expressed in the video is ANGER.

The sequence in which these stimuli were presented was randomized for each participant.

After completing the twenty-five questions, participants were asked to provide

[1] `https://www.psytoolkit.org/`

some personal information: their gender, age, and country of origin.
Additionally, information regarding their visual and auditory capabilities was collected. Given that the experiment involves the integration of two senses, participants were presented with a series of questions about synesthesia, adapted from McDonald et al. (2022)[37], and accompanied by some example to guarantee a better comprehension of the questions:

1. Do numbers or letters cause you to have a colour experience?
2. Do weekdays and months have specific colours?
3. Do you imagine or visualise weekdays, months and/or years as having a particular location in space around you?
4. Does hearing a sound make you perceive a colour?
5. Do certain words trigger a taste in your mouth?
6. Do you feel a sense of touch when you smell things?
7. Do you suspect that you experience an unusual blending that other people do not have (other than the one listed above)?

3.1.3 Gathering More Data

To gather more data for Turato's experiment (2022)[63], the online questionnaire from last year, created using the online platform PsyToolkit (Stoet, 2010, 2017)[56][57][2], was copied (to not lose 2022's data) and reopened.

Wanting to mash newly collected data with the last years ones, to gather more participants we decided to share the link to the survey in various social media platforms (Facebook, Reddit, Instagram), in groups specifically dedicated to surveys exchange. The number of participants still didn't increase enough, so we uploaded it to the dedicated website SurveyCircle (SurveyCircle, 2023)[35][3]. The website works by sharing a rank of the uploaded surveys, each one associated with an amount of points we could gather by answering it. The more the score associated with your account increases, the more your survey rises in the ranking and the more people would participate to it. As we could modify the

[2] https://www.psytoolkit.org/
[3] https://www.surveycircle.com/

questionnaire design, a code to redeem points in SurveyCircle was added at the end of the survey, in the thanks page.

3.1.4 Participants

One hundred and fifty-one participants completed the questionnaire. However, five of them have been removed from the final analysis because they reported having audiovisual disorders such as colour blindness, an essential factor for the success of the experiment. Moreover, 23 people declared to have not watched the videos in full.
In addiction, we applied a filter to exclude the participants who completed the survey in less then 6 minutes, calculated to be the minimum needed time to at least completely watch the whole set of videos plus a reasonable amount of time dedicated to inserting the answers.
Therefore, a total of 118 participants have correctly completed the experiment.

66 participants are females (55.3%), 51 males (43.9%) and 1 preferred not specify the gender (0.8%). A pie chart of the gender distribution is reported in Fig. 3.1.

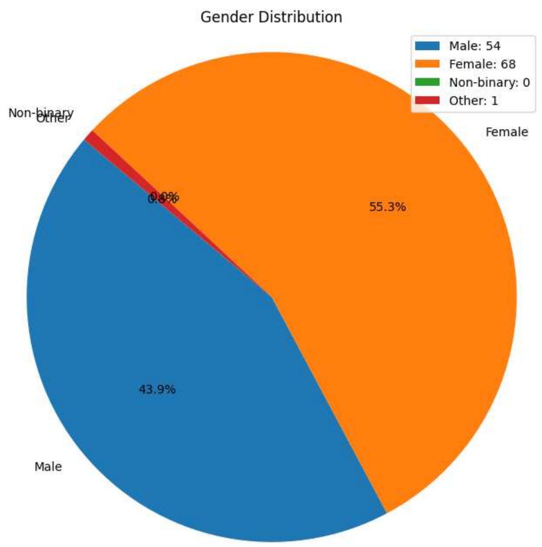


Figure 3.1: Pie chart of the gender distribution of the participants to Experiment #1 - 66 participants are females (55.3%), 51 males (43.9%) and 1 preferred not specify the gender (0.8%).

Concerning age, the overall range is 18-66, with 60 people in the 18-25 range

(51.3%), 26 in the 26-34 (22.2%), 7 in the 35-42 (6%), 17 in the 43-50 (14.5%), 2 in the 51-59 (1.7%) and 5 above 60 years old (4.3%). Two people wrote >45 and were therefore not taken into account in this specific section of the analysis.

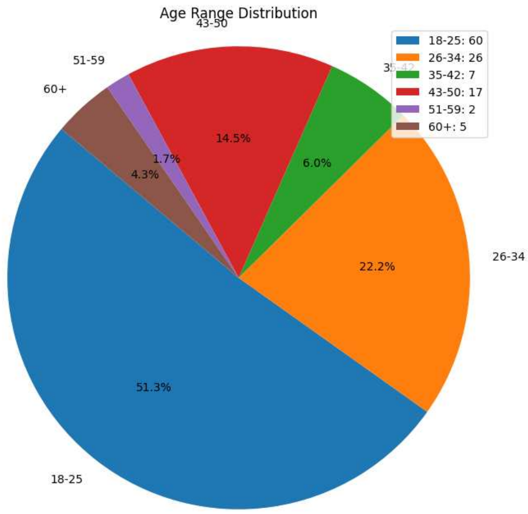


Figure 3.2: Pie chart of the age ranges of the participants to Experiment #1 - The overall age range is 18-66, with 60 people in the 18-25 range (51.3%), 26 in the 26-34 (22.2%), 7 in the 35-42 (6%), 17 in the 43-50 (14.5%), 2 in the 51-59 (1.7%) and 5 above 60 years old (4.3%). Two people wrote >45 and were therefore not taken into account in this specific section of analysis.

The majority of participants come from Italy (78) but we also had participants from the United States (10), the United Kingdom (5), Australia (2), Malaysia (2), Saudi Arabia (2), Brazil (2), Netherlands (1), Sweden (1), Kazakhstan (1), Hong Kong (1), Greece (1), India (1), South Africa (1), Japan (1), Canada (1), Germany (1), Jamaica (1), Philippines (1), Gabon (1), Spain (1), Poland (1), Romania (1). One participant wrote "Europe", so he won't be taken into account in the map of Fig. 3.3.

69 participants (58.47%) reported having normal sight, while 49 (41.53%) reported having corrected-to-normal sight, for instance with glasses or contact lenses. Furthermore, 100% of them reported to have normal hearing.

Concerning the device with which the participants watched the videos, 62 used their smartphones (52.54%), 2 their tablet/iPad (1.69%), 37 their laptop (31.36%)

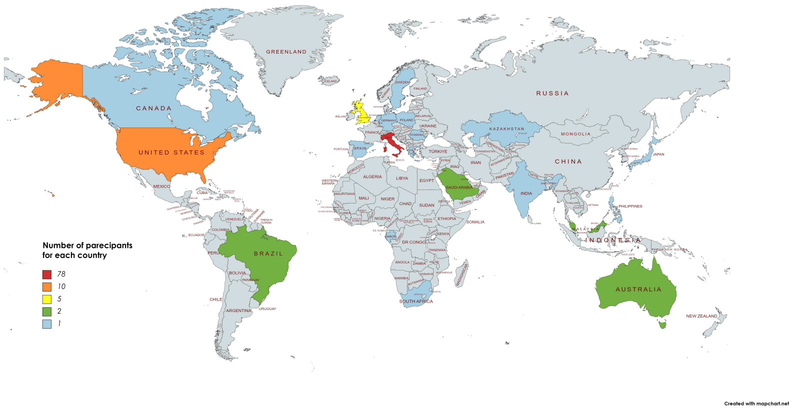


Figure 3.3: Map of the number of participants for each country in Experiment #1 - The most of them come from Italy (78), the US (10), the UK (5), but we had also 2 participants from the countries highlighted in green and 1 from the countries highlighted in light blue.

and 17 their computer monitor (14.41%).

To complete the whole experiment, participants spent an average of 15.4 minutes.

To conclude, since this experiment aims at investigating the effects of the union of two senses, we report a summary of the questions regarding synesthesia (McDonald et al., 2022)[37]. Each one of them was followed by a pertinent example to better explain the concept of synesthesia to participants who were not familiar with it.
Tables 3.2 from *a* to *g* report a summary of the answers at the seven questions, highlighting how females, in general, report being more sensitive than males to synesthesia.

3.1.5 Statistics

From the data we removed the entries that were not suitable for the analysis, i.e. the participants with sight or hearing disorders, the ones who didn't watch the videos presented in the survey in full, and the ones who spent less then 6 minutes or more then 100 minutes to complete the test.
We therefore computed data statistics that we will lately use to perform ANOVA: mean and standard deviation, reported in Tables 3.3 and 3.4.

(a) Intended emotion: HAPPINESS

Door Colour	Stat	Happiness	Sadness	Fear	Anger
White	Mean	7.57	0.84	0.47	0.82
	SD	2.98	1.84	1.03	2.14
Yellow	Mean	8.47	0.83	0.73	0.70
	SD	2.48	2.22	2.12	2.03
Gray	Mean	7.53	0.68	0.67	0.69
	SD	2.82	1.66	1.77	1.81
Red	Mean	8.10	0.78	0.70	1.12
	SD	2.29	1.63	1.63	2.24
Purple	Mean	7.22	1.12	0.32	0.74
	SD	2.88	2.17	0.95	1.80

(b) Intended emotion: SADNESS

Door Colour	Stat	Happiness	Sadness	Fear	Anger
White	Mean	0.69	6.15	2.73	0.47
	SD	1.64	3.65	3.51	1.24
Yellow	Mean	0.53	5.98	3.14	0.58
	SD	1.10	3.97	3.79	1.47
Gray	Mean	0.67	6.90	2.47	0.51
	SD	1.61	3.52	3.56	1.57
Red	Mean	1.07	6.18	3.17	0.84
	SD	2.14	3.97	3.81	1.87
Purple	Mean	0.43	6.62	2.33	0.64
	SD	1.05	3.96	3.38	1.66

Table 3.3: Tables *a* and *b* show the mean values and the standard deviations of the perception for each intended emotion. Yellow cells show the highest mean value, whereas green cells show our starting hypothesis. If there are no yellow cells, the highest mean value corresponds to the hypothesis.

(a) Intended emotion: FEAR

Door Colour	Stat	Happiness	Sadness	Fear	Anger
White	Mean	2.03	0.82	5.92	7.23
	SD	3.08	1.95	4.14	3.47
Yellow	Mean	1.98	0.67	5.96	6.10
	SD	2.98	0.54	4.23	3.74
Gray	Mean	1.35	0.72	5.54	7.20
	SD	2.27	1.63	4.09	3.09
Red	Mean	1.59	0.47	5.84	6.96
	SD	2.76	1.21	4.09	3.56
Purple	Mean	1.66	0.92	5.58	6.95
	SD	2.96	2.01	4.37	3.73

(b) Intended emotion: ANGER

Door Colour	Stat	Happiness	Sadness	Fear	Anger
White	Mean	3.56	0.88	1.44	3.89
	SD	4.44	1.95	2.65	4.41
Yellow	Mean	0.90	0.86	3.21	7.63
	SD	1.98	1.98	3.82	3.32
Gray	Mean	1.00	0.98	3.36	7.54
	SD	2.03	2.19	3.58	3.36
Red	Mean	0.77	1.02	3.07	7.93
	SD	1.65	2.19	3.40	3.09
Purple	Mean	1.27	1.09	2.18	7.98
	SD	2.59	2.25	3.12	2.81

Table 3.4: Tables *a* and *b* show the mean values and the standard deviations of the perception for each intended emotion. Yellow cells show the highest mean value, whereas green cells show our starting hypothesis. If there are no yellow cells, the highest mean value corresponds to the hypothesis.

By looking at the average values, it is possible to notice that both happiness and sadness confirm our initial hypothesis: **Happiness is mostly associated to the colour Yellow** with an average score of 8.47 (Table 3.3a), while **Sadness with the colour Gray** with an average of 6.90 (Table 3.3b).

On the other hand, it is significant to notice that, in general, **Fear has been widely perceived as Anger** (Table 3.4a) and has been more **associated with the**

colour White (7.23 average score).
In any case, the higher score for Fear is associated to the colour Yellow (5.96), instead that with the expected colour Purple (5.58).

Lastly, **Anger** has been in general **correctly detected** (Table 3.4b), but more **associated with the colour Purple** (7.98) instead of Red (7.93).

This result is consistent with that of Pauletto and Iop (2021)[41]: findings in their study underlined the absence of significant statistical difference between Fear and Anger, testifying the difficulty for the participants in correctly identifying these two emotions. However, in general was Fear to be mostly confused with Anger, rather than the vice versa, and Red and Purple are the two colours that reinforce the perception of these emotions.

Concerning Standard Deviation (σ, SD), its very high values highlight a great variability among the data, as we can clearly see also from the box plots of Fig. 3.4.

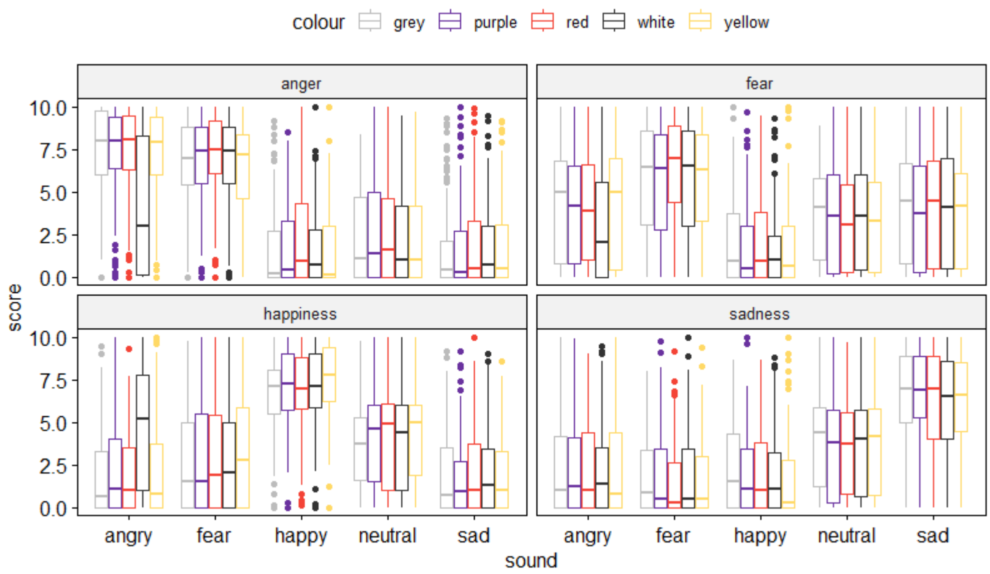


Figure 3.4: Boxplots of statistics of Experiment #1 - Happiness is mostly associated to the colour Yellow and Sadness with the colour Gray; Anger has been correctly detected but Fear tends to be mostly confused with Anger, and Red and Purple are the two colours that reinforce the perception of these emotions.

3.1.6 ANOVA

The ANOVA test is used to compare the variances between the means of different groups and to evaluate the effects on a variable of interest (in this case on the *score* associated with a specific emotion, which is a quantitative variable) of one or more factors (in in this case *sound* and *colour*, both qualitative variables). It is based on two hypothesis:

- *The Null Hypothesis* H_0, which predicts that the data of all groups in the population have the same mean, and that the differences observed in the sample between the group means are due only to chance.
  $H_0 : \mu_1 = \mu_2 = \mu_3 = ... = \mu$ i.e. $F = 1$
- *The Alternative Hypothesis* H_1, which predicts that there is a significant difference between the groups. In other words, that at least one group has an average value significantly different from the others.
  $H_1 : \mu_i$ are not all equals between each other (at least one of them is different from the others), i.e. $F > 1$

The objective of the ANOVA is therefore to decide whether or not it is better to reject the Null Hypothesis that the means of the groups in the population are all equal to each other, and it does so according to the following decision rule: if the calculated *F-value* is greater than a critical $F(F = 1)$, then the Null Hypothesis H_0 is rejected and hypothesis H_1 is accepted. In this case the test is significant, that is, it found a significant difference between the means in the groups, not due to chance.
Otherwise, do not reject H_0.

If the ANOVA is significant, then it is more likely that the differences observed between the means are a real difference between the groups, rather than just the result of their internal variability due to chance.

The value of F represents the ratio between the variance between the groups (called *between*) and the variance within the groups (called *within*) and, if it is greater than 1, it indicates the statistical significance of the ANOVA test, defined above.

This type of analysis can be used without problems when the data meets the following conditions:

- **Independence of the observations**: each subject should belong to only one group. There is no relationship between the observations in each group.
- **No significant outliers** in the data.
- **Normality**: the data should be approximately normally distributed.
- **Homogeneity of variances**: the variance of the outcome variable should be equal within the data.

In our case we will see that the homogeneity of the variances will not be satisfied, a condition that could lead to a higher probability of falsely rejecting the null hypothesis H_0.
To overcome the problem, we decided to proceed with the Three-Ways ANOVA analysis anyway, but with the foresight of also using Welch's ANOVA, a type of analysis (One-Way) that does not assume homoscedasticity between groups and is more robust when this assumption is violated, to verify in any case and in a more robust way the significance of the obtained results.

The significance of the interactions between the variables in the ANOVA is represented not only by F, but also by the *p-value*, which represents the probability that the variability in the means of the sample data is the result of pure chance (i.e. the probability of $F \leq 1$).
We can say that the lower **the *p-value*, the greater the statistical significance of the observed difference**, and therefore the lower the risk that the variability of the means is due to chance rather than an actual difference between the groups.
A *p-value* equal to or lower than $\alpha = 0.05$, where α represents the level of significance (that defines the set of values that belong to the region of rejection and acceptance of the Null Hypothesis), is generally considered statistically significant.

We will therefore verify with the Three-Way ANOVA the existence of significant interactions between the factors taken into consideration, and then verify whether Welch's ANOVA (that takes into account the non-verification of the hypothesis of homogeneity of variances), calculated on the factors that result dependent on each other, also provides us with a *p-value* lower than *0.05*, thus confirming or denying the significance of the obtained results.

3.1.7 The Analysis

A Three-Way ANOVA was conducted to determine the effects of Sound and Colour on the Emotions' scores given by the participants. To perform the analysis we utilized RStudio, an integrated development environment for R, a programming language for statistical computing and graphics.

Before performing the Three-Ways ANOVA, we had to check for its assumptions to be verified by our dataset.

Outliers

```
print((data_anova %>%
         group_by(sound, colour, emotion) %>%
         identify_outliers(score)), n=13300)
```

No extreme outliers, that could skew our computations, have been detected. Extreme outliers could be due to:

- data entry errors
- measurement errors
- unusual values

On the other hand multiple outliers have been detected, but we are going to keep them in our analysis because we believe they still might be meaningful for our scope.

After completing the analysis we also tried to remove outliers to check if the results (especially concerning ANOVA's significance) would change in a consistent way. The results didn't change much, so we definitively decided to keep them.

Normality

To verify normality we used the visual method of QQ Plots. The outcome is reported in Fig. 3.5.

```
#Check normality assumption by analyzing the model residuals

```

```
model  <- lm(score ~ sound*colour*emotion, data = data_anova)

# Create a QQ plot of residuals

ggqqplot(residuals(model))
```

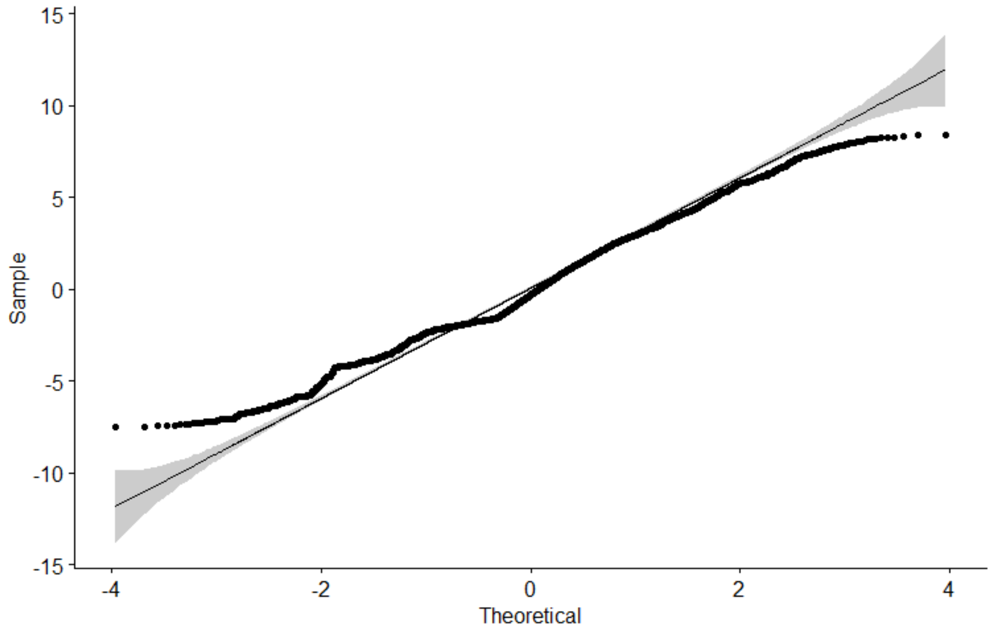


Figure 3.5: QQ Plot on the whole dataset, computed to verify the ANOVA's assumption of Normality on the collected data. We can see that, excluding the extremes, the points fall approximately along the reference line, so we can assume normality.

In the QQ plot, excluding the extremes, the points fall approximately along the reference line, so we can assume normality.

After normality on the whole dataset, we also checked the assumption for each combination of factor levels. The outcomes of the QQ plots has been reported in Fig. 3.6.

```
# QQ plot for each cell of design

ggqqplot(data_anova, "score", ggtheme = theme_bw()) +
  facet_grid(sound + colour ~ emotion, labeller = "label_both")
```

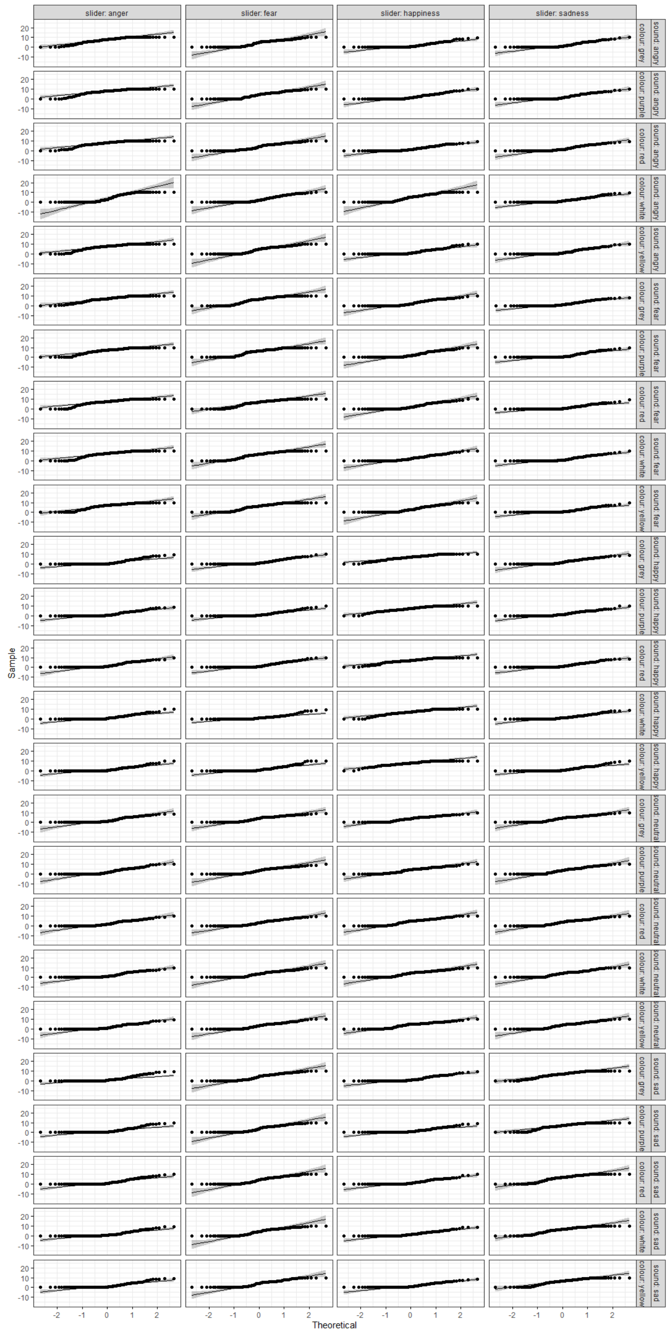


Figure 3.6: QQ Plots for each combination of factor levels - To verify assumption of normality, we check it also for each combination of factors level. The plots show that, excluding extremes, the points fall approximately along the reference line, so we can assume normality even for each factor levels' combinations.

Again, we can see from the plots that the points, excluding extremes, fall approximately in the region close to the reference line, we can therefore assume normality for each group of combinations.

Homogeneity of variances

Homogeneity of variances was investigated by Levene's test.

```
data_anova %>% levene_test(score ~ sound*colour*emotion)
```

The outcome of the test determined the non verification of the assumption ($p << 0.05$), meaning that the variance of outcome variables is not similar in each group. Consequently, there is grater probability of falsely rejecting the Null Hypothesis: *there is no difference among group means*.

Even if homogeneity was not verified, we decided to proceed with the Three-Ways ANOVA, but paying attention to run also Welch's ANOVA to the obtained results, that does not assume homoscedasticity between groups and is more robust when this assumption is violated, to check the significance.

Three-Ways ANOVA

```
res.3Waov <- data_anova %>% anova_test(score ~ sound*colour*emotion)
res.3Waov
```

The results of the Three-Ways ANOVA are reported in Table 3.5.

Effect	DFn	DFd	F	p	p<.05	ges
sound	4	13196	69.768	1.48e-58	*	0.021000
colour	4	13196	1.315	2.62e-01		0.000398
emotion	3	13196	40.390	5.67e-26	*	0.009000
sound:colour	16	13196	0.640	8.54e-01		0.000775
sound:emotion	12	13196	454.654	0.00e+00	*	0.293000
colour:emotion	12	13196	5.9151	2.25e-10	*	0.005000
sound:colour:emotion	48	13196	4.632	5.20e-24	*	0.017000

Table 3.5: Three-Way ANOVA test - The results highlight a statistically significant Three-Way interaction between Sound, Colour and Emotion (`sound:colour:emotion`), with $F(48, 13196) = 4.632, p = 5.20e - 24$.

From the table we can see that there was a statistically significant Three-Way interaction between Sound, Colour and Emotion, $F(48, 13196) = 4.632, p =$

$5.20e - 24$, but also significant Two-Ways interactions between both Sound-Emotion and Colour-Emotion.

As we can see from these results, ANOVA gives us a $p - factor$ way smaller than 0.05. This means that a statistically significant difference in score means was detected.

To go deeper into the interaction between the factors, we can decompose the significant Three-Ways interaction effect into simple Two-Way interactions at each level of the Emotion variable.
We decided to perform the Two-Ways interaction both between Sound and Emotion and between Colour and Emotion.

Data have been first grouped by Colour and the Two-Ways ANOVA was ran between Sound and Emotion.

```
# Group the data by colour and fit simple two-way interaction
model  <- lm(score ~ sound*colour*emotion, data = data_anova)
data_anova %>%
  group_by(colour) %>%
  anova_test(score ~ sound*emotion, error = model)
```

The result highlighted, as expected, that there was a highly statistically significant simple Two-Way interaction between Sound and Emotion independently from the presented door's Colour, as reported in Table 3.6. This result suggests that the effect that knocking sounds have on the perceived emotion is is primary and not much influenced by the colours of the door.

Colour	Effect	DFd	DFd	F	p	p<.05	ges
gray	sound:emotion	12	12989	119	1.94e-282	*	0.099
purple	sound:emotion	12	12989	125	4.64e-298	*	0.104
red	sound:emotion	12	12989	117	2.77e-277	*	0.097
white	sound:emotion	12	12989	96.5	5.07e-230	*	0.082
yellow	sound:emotion	12	12989	121	1.07e-287	*	0.101

Table 3.6: ANOVA Two-Ways test on Sound and Emotion - Here we reported only the interactions part of the whole table, that is the significant part for the current analysis. The results highlight a significant simple Two-Way interaction between sound and emotion (`sound:emotion`) for every presented colour; this means that the influence of Sound on the perceived emotion is primary and not much influenced by the colours of the door.

Even if we were sure about the significance of this result, we ran anyway the One-Way Welch's ANOVA on Sound, to inspect the direct effect of Sound on *score* and also to verify if following a procedure more robust to hypothesis relaxations, significance would still be confirmed.

```
# Group the data by emotion and colour, and fit Welch's anova, to see
    effect of only sound on score
sound.effect <- data_anova %>%
  group_by(emotion, colour) %>%
  welch_anova_test(score ~ sound)
sound.effect
```

Results are reported in Table 3.7.
The interactions inspected by the Welch's ANOVA results all highly significant, in fact we can see from the table that all the resulting $p - values$ are way smaller than the reference value 0.05.

colour	emotion	.y.	n	statistic	DFn	DFd	p	p<0.05
grey	anger	score	634	257.34	4	312.387	1.64e-97	*
purple	anger	score	633	316	4	313.007	1.56e-108	*
red	anger	score	643	238.66	4	318.942	1.54e-94	*
white	anger	score	644	134.65	4	316.746	4.76e-67	*
yellow	anger	score	649	198.5	4	320.985	1.86e-85	*
grey	fear	score	663	34.12	4	326.905	8.38e-24	*
purple	fear	score	660	51.52	4	323.180	1.55e-33	*
red	fear	score	665	34.28	4	329.202	6.32e-24	*
white	fear	score	657	71.9	4	315.363	3.15e-43	*
yellow	fear	score	659	51.56	4	322.979	1.49e-33	*
grey	happiness	score	656	126.97	4	324.568	4.13e-65	*
purple	happiness	score	656	153.27	4	323.865	2.1e-73	*
red	happiness	score	659	119.01	4	326.066	1.81e-62	*
white	happiness	score	657	116.15	4	324.193	2.44e-61	*
yellow	happiness	score	658	181.21	4	324.925	2.05e-81	*
grey	sadness	score	665	63.66	4	329.003	8.15e-40	*
purple	sadness	score	661	65.87	4	326.708	8.2e-41	*
red	sadness	score	661	59.79	4	325.523	8.07e-38	*
white	sadness	score	657	62.69	4	324.058	3.34e-39	*
yellow	sadness	score	656	72.5	4	322.595	9.24e-44	*

Table 3.7: Results of the Welch's ANOVA (One-Way) ran to inspect the effect of Sound on *score*, to verify if also following a procedure more robust to hypothesis relaxations significance would still be confirmed. We can see that the obtained $p-value$ is lower than 0.05 for every interaction, highlighting the significance of the dependence of the emotion's *score* on the emotional knocking sounds presented in the video clips.

The same analysis was ran grouping data by Sound and performing Two-Ways ANOVA between Colour and Emotion. The results are reported in Table 3.8.

```
# Group the data by sound and fit simple two-way interaction
model  <- lm(score ~ sound*colour*emotion, data = data_anova)
data_anova %>%
  group_by(sound) %>%
  anova_test(score ~ colour*emotion, error = model)
```

Sound	Effect	DFn	DFd	F	p	p<.05	ges
angry	colour	4	13196	2.52	3.9e-2	*	0.000763
angry	emotion	3	13196	358	9.25e-224	*	0.075
angry	colour:emotion	12	13196	20.8	4.22e-46	*	0.019
fear	colour	4	13196	0.26	9.03e-1		0.0000789
fear	emotion	3	13196	487	2.44e-300	*	0.1
fear	colour:emotion	12	13196	1.09	3.61e-1		0.000993
happy	colour	4	13196	0.628	6.42e-1		0.00019
happy	emotion	3	13196	580	0	*	0.117
happy	colour:emotion	12	1.34	40.390	1.88e-1		0.001
neutral	colour	4	13196	0.188	9.45e-1		0.000057
neutral	emotion	3	13196	45.8	2.05e-29	*	0.01
neutral	colour:emotion	12	13196	0.704	7.5e-1		0.000639
sad	colour	4	13196	0.279	8.92e-1		0.0000845
sad	emotion	3	13196	387	3.16e-241	*	0.081
sad	colour:emotion	12	13196	0.483	9.26e-1		0.000439

Table 3.8: ANOVA Two-Ways test on Colour and Emotion - The results highlight a significant simple two-way interaction between Colour and Emotion (`colour: emotion`) for the angry knocking sounds, $F(12, 13196) = 20.8, p = 4.22e - 46$, but not for the other sounds; in other words, the mean score in the Colour groups (Gray, Purple, Red, White, Yellow) was statistically significantly different only for videos that presented the Angry knocking sounds. This means that for the knocking sound "angry" the perception of the Emotion depends on the colour of the door in the clip. For the other knocking sounds, the influence of Colour was not significant for the participants' ratings.

The analysis underlined that there was a statistically significant simple two-way interaction between colour and emotion (`colour:emotion`) just for the Angry knocking sounds, $F(12, 13196) = 20.8, p = 4.22e - 46$, but not for the other sounds; in other words, the mean score in the Colour groups (Gray, Purple, Red, White, Yellow) was statistically significantly different only for videos that presented the Angry knocking sounds. This means that for the knocking sound "angry" the perception of the Emotion depends on the colour of the door in the clip. For the other knocking sounds, the influence of Colour was not significant for the participants' ratings.

To further verify the statistic significance of the found results we ran the One-

Way Welch's ANOVA on Colour for the Angry sound (as this was the only simple Two-Ways interaction that was statistically significant), to inspect the effect of Colour on *score* and also to verify whether when following a procedure more robust to hypothesis relaxations, significance would be confirmed.

```
# Group the data by emotion and sound, and fit Welch's ANOVA to see
   effect of only colour on score
welch_sound_score <- data_anova %>%
  group_by(sound, emotion) %>%
  welch_anova_test(score ~ colour)
welch_sound_score %>% filter(sound == "angry")
```

Results are reported in Table 3.9, and highlight that there were statistically significant simple simple main effect of Colour for videos with the "angry" knocking sound, but influenced only the score for the emotions Anger, Fear, and Happiness, while it made no difference for ratings of the slider "Sadness".
In this part of the analysis statistical significance was accepted at a Bonferroni-adjusted alpha level of $\alpha = 0.025$ (0.05 divided by 2, that is the number of simple simple main effects we are computing).

sound	emotion	.y.	n	statistic	DFn	DFd	p	p<0.025
angry	anger	score	665	18.8	4	329.	5.73e-14	*
angry	fear	score	665	3.22	4	330.	0.013	*
angry	happiness	score	665	17.5	4	329.	4.96e-13	*
angry	sadness	score	665	0.15	4	330.	0.962	

Table 3.9: Results of the Welch's ANOVA (One-Way) ran to inspect the effect of Colour on *score* and filtered with `sound == "angry"` (as this was the only simple Two-Way interaction that was statistically significant), to verify significance following a procedure more robust to hypothesis relaxations. The Colour factor influenced only the score for the emotions Anger, Fear, and Happiness (for which the obtained $p - value$ is lower than 0.025), while it made no difference for ratings of the slider "Sadness". NB: Statistical significance was accepted at a Bonferroni-adjusted alpha level of 0.025.

3.2 Results Summary

The experiment results suggest that the main influence on the association between the multimedia stimuli and the perceived emotion was given by the auditory stimuli presented in the clips (the knocking sound) more than by the visual stimuli (colours of the door).

We can see in Fig. 3.4, that represents the boxplots of the ratings given by participants, that stimuli intended to represent the emotions Happiness, Sadness and Anger have been correctly and uniquely classified. Fear, on the other hand, as we can see in Table 3.4a has been mostly wrongly detected as Anger and, as anger, enhanced when associated to White. In general, participants found difficult to discern between Fear and Anger.
From the same boxplot and from Tables 3.3 and 3.4, we can also spot some trends: the colour that enhanced the emotion Happiness has been Yellow, Gray for Sadness, and contrary to what was expected, Red for Fear (we associated it with Purple) and Purple for Anger (that we associated with Red). Has to be noticed also that, anyway, the responses contained a relatively high variance.

The effect of Colours on participants' ratings of the four basic emotions did not produce statistically significant results, but they resulted to be significant only to associate the Angry knocking sound to Anger, Fear or Happiness. Again the results contained very high variance in the responses.

In summary, the emotions intended by the knocking sounds have been mostly correctly detected by the audience, except for the fact that the "Fear" sound is many times confused with Anger, that in turn has been correctly detected when presented. Therefore, Sound appears to be the primary significant influence on the Emotions perceived by the participants.
The influence of the door colours is minimal, and, just in some cases, it only reinforces the emotion evoked by the knocking sounds.

3.3 Limits of the First Experiment

As one of the aims of this research is, with Experiment #2, to deepen the analysis we just performed, it is essential to acknowledge its inherent limitations.

Firstly, we could start by summarizing the limitations already suggested by Turato in 2022[63]:

- The study solely focuses on knocking sounds and their association with specific colours, excluding a comprehensive examination of the influence of other factors on emotion perception.
- The experiment's findings are based on a specific set of colour-emotion associations, which may not capture the full complexity of emotional responses to colours.
- The study employs a 0 to 10 rating scale and follow-up questions on synesthesia, potentially introducing subjectivity (not all the participants are equally open to perceive synesthesia) and response bias.
- The use of video stimuli, although a departure from previous studies, may not fully replicate real-world scenarios.
- The experiment was not conducted within a controlled laboratory setting, and the generalizability of the findings to real-world contexts warrants further investigation.

In addition to those limitations, we also identified that:

- The major influence of auditory stimuli over visual stimuli could be caused by the fact that while sound is dynamic, colours are static. Modifying the experiment introducing dynamicity to the visual stimuli has been a starting point for our second experiment.
- The questionnaire design was quite long and not very appealing, so that might have led the participants to close the survey before having completed it. This forced us to discard the data that where not associated to finished surveys.
- The usage of the platform PsyToolkit introduced some bugs that might have caused the increment of variability in the given responses. Sometimes the platform doesn't correctly lead the video stimuli, we therefore suspect that, when this bug occurred, some responses might have been given randomly, just to complete the survey. This is a major issue that might have skewed our analysis introducing so much difference in variances.

(a) Do numbers or letters cause you to have a colour experience?

Gender	Yes	No
Male	8	41
Female	13	47
Other	0	1

(b) Do weekdays and months have specific colours?

Gender	Yes	No
Male	10	39
Female	24	36
Other	0	1

(c) Do you imagine or visualise weekdays, months and/or years as having a particular location in space around you?

Gender	Yes	No
Male	8	41
Female	10	50
Other	0	1

(d) Does hearing a sound make you perceive a colour?

Gender	Yes	No
Male	7	42
Female	14	46
Other	0	1

(e) Do certain words trigger a taste in your mouth?

Gender	Yes	No
Male	9	40
Female	19	41
Other	0	1

(f) Do you feel a sense of touch when you smell things?

Gender	Yes	No
Male	11	38
Female	17	43
Other	0	1

(g) Do you suspect that you experience an unusual blending that other people do not have (other than the one listed above)?

Gender	Yes	No
Male	14	35
Female	22	38
Other	1	0

Table 3.2: These tables report a summary of the answers at the seven questions about synesthesia taken from McDonald et al. (2022)[37], for Experiment #1. They highlight how females, in general, report being more sensitive than males to synesthesia.

4

Experiment #2

4.1 Aim

The aim of the experiment was investigating the relationship between colours, sounds, camera movements, and their conveyed emotions when combined to form a cinematic scene, to gain a deeper understanding of how sensory elements in filmmaking work in concert to evoke specific emotional responses from the audience.
By systematically varying and controlling these variables, the experiment seeks to uncover patterns and correlations between contents, colour palettes, soundscapes, camera movements, and the emotional nuances they elicit.

This research not only contributes to the art and science of filmmaking but also aims to provide insights into the psychology of perception and emotional response.
Ultimately, the goal is to empower filmmakers with evidence-based tools to create more immersive and emotionally resonant cinematic experiences, where the synergy between visual, auditory, and kinetic elements can be harnessed to effectively convey the intended emotional states within a narrative, thereby enriching storytelling and audience engagement.

For the development of the questionnaire and other methodological matters, we collaborated with Prof. Emery Schubert and Diana Zhang of the University of New South Wales Sydney, and Dr. Anthony Chmiel of the Western Sydney

University.
The experiment was conducted from July 2023 to October 2023.

4.2 Stimuli

For this study, we chose four different visual stimuli from the dataset of camera movement scenes from the experiment of Yilmaz, Lotman , Karjus and Tikka (2023)[9]: the Pleasant and Unpleasant contents (representing a room with pleasant or unpleasant details and with a focus on the room's door), both shot in Static and Handheld techniques. The choice of discarding the neutral (ambiguous) content has been made to keep a reasonable number of stimuli, and the choice on the camera movements to include were made to take into consideration the most static and the most dynamic techniques. At each one of these four obtained videos, we added two different emotional knocking sounds, professionally performed by Ulf Olausson[2], a subset of the ones utilized by Barahona-Ríos and Pauletto in 2020[1], which was also employed in the work of Pauletto and Iop in 2021[41] and of Turato in 2022[63].
We therefore obtained eight videos, and for each of these videos we applied two colour filters to Red and Blue, to be added to the Neutrals (the original videos).

To prepare the dataset, we utilized the open source editor DaVinci Resolve. We trimmed the original video of the different scenes into the clips we wanted to use for the experiment, each one 26 seconds long to ensure consistency. We synchronized the knocking sounds at the end of the clips, when the door was framed, and we applied two different LUTs to edit the colours.

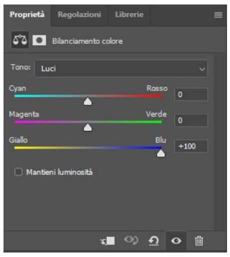
The LUTs were previously obtained with the program Adobe Photoshop 2024 and following the indications given by McDonald, Canazza, Chmiel, De Poli, Houbert, Murari, Rodà, Schubert, Zhang (2022)[37]: a colour balance adjustment layer was added to a chosen frame taken from the original clips by manipulating three properties: highlights, midtones, shadows. As the goal was to mimic a coloured spotlight, the levels were adjusted based on realism.

Here specific details on how we obtained the Red and the Blue LUTs.

4.2.1 Blue

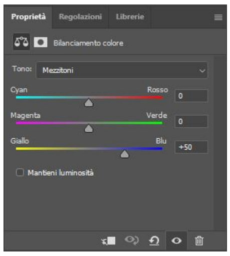
Highlights

Blue was adjusted to +100.
This was done because we are targeting the main source of light in the scene. Adjusting it to be fully blue lets us mimic a cinematographic blue-tinted singular light source.
Red and green were not adjusted, and were left at the default of 0.



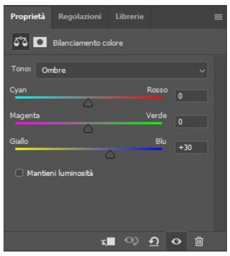
Midtones

Blue was adjusted to +50.
This level was chosen to simulate a natural light setting; we needed to account for areas that the light touches that would have a blue cast from a blue tinted spotlight, but would not be entirely saturated (i.e., at a level close to 100).
Red and green were not adjusted, and were left at the default of 0.



Shadows

Blue was to adjusted to +20.
This gave just a slight hint of blue to the shadows without appearing unnatural.
Red and green were not adjusted, and were left at the default of 0.



In addiction, we added a curves adjustment layer on top of the colour balance one, with the aim to add a little bit more of lighting as the clips were quite dark and the blue adjustment made it impossible to clearly see the scene content.

Both RGB's and Blue's highlights and shadows were made respectively brighter and darker of approximately the 5%.

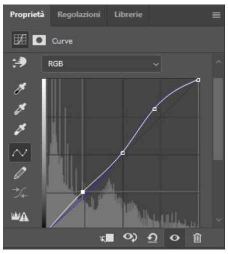


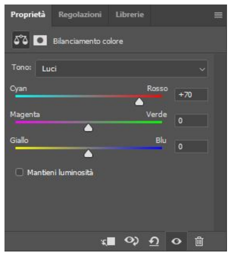
Figure 4.1: Curves adjustment layer added to the frame to obtain the Blue LUT.

We also visually increased a bit the overall lighting to increase visibility, but moving from the mentioned above guidelines.

4.2.2 Red

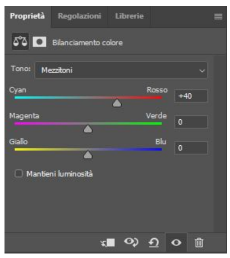
Highlights

Red was adjusted to +70.
This was done because we are targeting the main source of light in the scene. Adjusting it to be fully red lets us mimic a cinematographic red-tinted singular light source.
Blue and green were not adjusted, and were left at the default of 0.

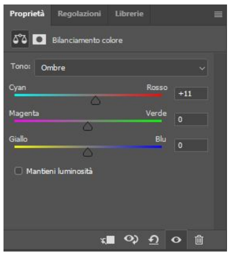


Midtones

Red was adjusted to +40.
This level was chosen to simulate a natural light setting; we needed to account for areas that the light touches that would have a red cast from a red tinted spotlight, but would not be entirely saturated.
Blue and green were not adjusted, and were left at the default of 0.



Shadows



Red was to adjusted to +11.
This gave just a slight hint of red to the shadows without appearing unnatural.
Blue and green were not adjusted, and were left at the default of 0.

In addiction, we added a curves adjustment layer on top of the colour balance one, with the aim to add a little bit more of lighting and made the clips look more natural.

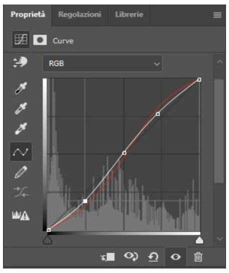


Figure 4.2: Curves adjustment layer added to the frame to obtain the Red LUT.

Both RGB's and Red's highlights and shadows were made respectively brighter and darker of approximately the 5% (it's an estimation, because the adjustment is made through a graphical regulator and not a numerical one).

The clips were already nice and clearly visible, so no overall lighting was added.

LUTs

From the rendered files in Adobe Photoshop 2024 we exported and saved the .CUBE LUTs (*"Look-Up Tables"*).
These were added to DaVinci Resolve's library and applied as an adjustment layer to the cut and trimmed clips.
The 24 so-obtained clips were then exported as low resolution files to better permit their loading during the online surveys, avoiding the lagging and bugs that were identified during the Experiment #1 (see Chapter 3).

All edits were made in a same session, on the same device and at the lighting conditions to ensure consistency.

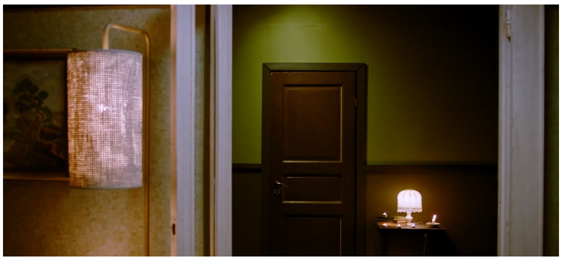


Figure 4.3: Reference shot taken from the stimuli clips, used to create the Blue and Red LUTs in Adobe Photoshop 2024 following the indications given by McDonald, Canazza, Chmiel, De Poli, Houbert, Murari, Rodà, Schubert, Zhang (2022)[37].

(a) Blue LUT (b) Red LUT

Figure 4.4: The files from which the LUTs were exported to be applied as adjustment layers to the clips, using DaVinci Resolve.

4.2.3 Procedure

An online questionnaire with the same structure of Experiment #1's one was developed using the PsyToolkit online platform (Stoet, 2010, 2017)[56][57][1]. Participants were initially provided with an overview of the study's purpose and the privacy policy.

To allow participants to adjust the volume of their device's speakers or headphones, before commencing the experiment participants were required to listen to two audio samples (the same ones of Experiment #1) - one with high intensity and the other with low intensity, and were asked to maintain the same volume level throughout the whole experiment.

Subsequently, they were acquainted with the survey's structure. The questionnaire consisted of eight questions; each page featured a 26-seconds video clip followed by four sliders. Participant were specifically asked to watch the

[1] `https://www.psytoolkit.org/`

clips in full. After watching the video, which they could replay as many times as they wished, participants were tasked with assessing the emotion conveyed in the video by adjusting four sliders placed on a continuous scale spanning from 0 (indicating *complete disagreement*) to 10 (indicating *complete agreement*). Prior to each slider, participants encountered the following statements:

- The emotion expressed in the video is HAPPINESS.
- The emotion expressed in the video is SADNESS.
- The emotion expressed in the video is FEAR.
- The emotion expressed in the video is ANGER.

After completing the eight questions, participants were asked to provide some personal information: their gender, age, and country of origin.
Additionally, information regarding their visual and auditory capabilities was collected. Given that the experiment involves the integration of two senses, participants were presented with a series of questions about synesthesia, adapted from McDonald et al. (2022)[37]:

1. Do numbers or letters cause you to have a colour experience?
2. Do weekdays and months have specific colours?
3. Do you imagine or visualise weekdays, months and/or years as having a particular location in space around you?
4. Does hearing a sound make you perceive a colour?
5. Do certain words trigger a taste in your mouth?
6. Do you feel a sense of touch when you smell things?
7. Do you suspect that you experience an unusual blending that other people do not have (other than the one listed above)?

To keep the questionnaire short enough (our target was to keep it under 10 minutes), we had to make the choice to not show the whole set of questions, that would have been 24 questions of about one minute each, more the final personal questions; we therefore chose to divide the 24 questions in 3 subsets of 8 questions each.

To choose which set would have been the shown one, at the beginning of each

survey a random number between 1 and 3 was generated and, depending on the number, the participants would have been redirected to the corresponding subset.
As we had to face some bugs of PsyToolkit that prevented us to use JavaScript, we couldn't be able to randomize the questions inside each subsets.
To increase randomization, we therefore decided to add 3 more subsets with the same compositions of the original three, but with the questions presented in the opposite order. The initial randomization was modified to be between 1 and 6.

To gather more participants we decided to share the link to the survey in various social media platforms (Facebook, Reddit, Instagram), in groups specifically dedicated to surveys exchange, and to the dedicated websites SurveyCircle (SurveyCircle, 2023)[35][2] and SurveySwap[3], after having added to the thanks page the *redeem codes* to permit to the participants to gather points on the two surveys exchange platforms through the participation in our research.

4.2.4 Participants

One hundred and four participants completed the questionnaire.
However, two of them have been removed from the final analysis because they reported having hearing disorders and one that reported to have sight disorders such as colour blindness, an essential factor for the success of the experiment. Moreover, 18 people did not watch all the videos in full.
In addiction, we applied a filter to exclude other five participants who completed the survey in less then 4 minutes, calculated to be the minimum needed time to at least completely watch the whole set of videos plus a reasonable amount of time dedicated to inserting the answers.
Therefore, a total of 78 participants have correctly completed the experiment.

As represented in the pie chart of Fig. 4.5, 51 participants are females (65.4%) and 27 are males (34.6%). None declared to be bisexual or preferred not to specify the gender.

[2] https://www.surveycircle.com/
[3] https://surveyswap.io/

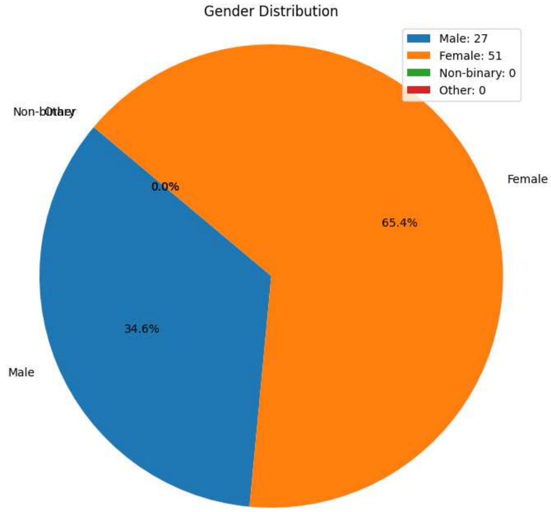


Figure 4.5: Pie Chart of the gender distribution of the participants to Experiment #2 - 51 participants are females (65.4%) and 27 are males (34.6%). None declared to be bisexual or preferred not to specify the gender.

Concerning age, the overall age range is 15-74, with 28 people in the 15-25 range (36.4%), 19 in the 26-34 (24.7%), 8 in the 35-42 (10.4%), 17 in the 43-50 (22.1%), 2 in the 51-59 (2.6%) and 3 above 60 years old (3.9%). One participant didn't write his age and were therefore not taken into account in this specific section of analysis represented by the pie chart in Fig. 4.6.

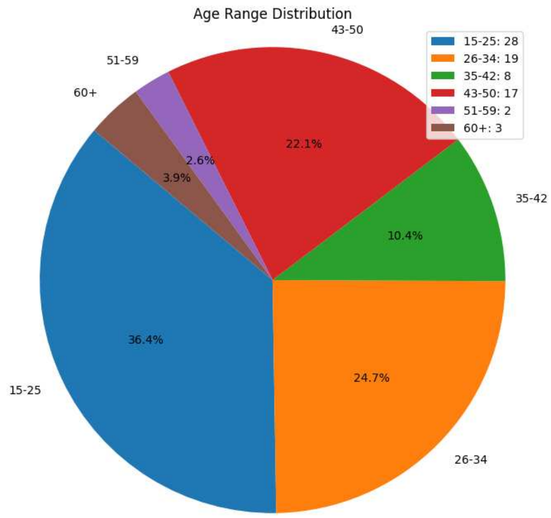


Figure 4.6: Pie chart of the age ranges of the participants to Experiment #2 - The overall age range is 15-74, with 28 people in the 15-25 range (36.4%), 19 in the 26-34 (24.7%), 8 in the 35-42 (10.4%), 17 in the 43-50 (22.1%), 2 in the 51-59 (2.6%) and 3 above 60 years old (3.9%).

The majority of participants come from Italy (36) but we also had participants from the United States (10), the United Kingdom (7), Canada (4), Australia (2), Germany (2), India (2), France (1), Saudi Arabia (1), Denmark (1), Portugal (1), Hungary (1), Ireland (1), Spain (1), Netherlands (1), Malaysia (1), Zambia (1), Colombia (1), China (1), Philippines (1), Vietnam (1). One participant preferred not to answer the question, so he won't be taken into account in the map of Fig. 4.7.

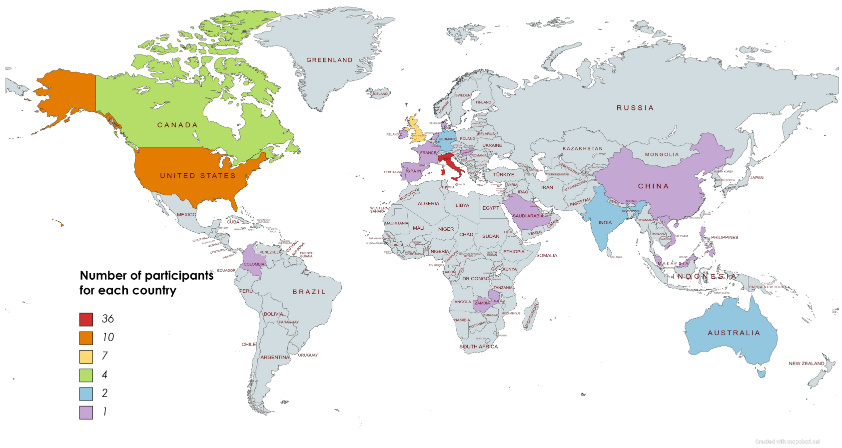


Figure 4.7: Map of the number of participants for each country of the participants to Experiment #2 - The most of them comes from Italy (36), the US (10), the UK (7), Canada (4) but we had also 2 participants from the countries highlighted in light blue and 1 from the countries highlighted in purple.

59 participants (75.64%) reported having normal sight, while 19 (24.36%) reported having corrected-to-normal sight, for instance with glasses or contact lenses. Furthermore, 77 participants (98.72%) reported having normal hearing, while 1 (1.28%) reported having corrected-to-normal hearing.

Concerning the device with which the participants watched the videos, 30 used their smartphones (38.46%), 1 their tablet/iPad (1.28%), 36 their laptop (46.15%) and 10 their computer monitor (12.82%). One participant wrote "Other".

To complete the whole experiment, participants spent an average of 11.2 minutes.

To conclude, since this experiment aims at investigating the effects of the union

of two senses, we report a summary of the questions regarding synesthesia (McDonald et al., 2022)[37]. Each one of them was followed by a pertinent example to better explain the concept of synesthesia to participants who were not familiar with it.
Tables 4.1 report a summary of the answers at the seven questions, highlighting how females, in general, report being more sensitive than males to synesthesia.

4.2.5 Statistics

From the data we removed the entries that were not suitable for the analysis, i.e. the participants with sight or hearing disorders, the ones who didn't watch the videos presented in the survey in full and the ones who spent less then 4 minutes or more then 100 minutes to complete the test.

Hypothesis Our hypothesis, classifying the intended emotions with reference to their levels of *valence* and *arousal*, and matching them with the video clips characteristics taken from Burak Yilmaz et al. (2023)[9], with reference to Fig. 1.7, were:

- HAPPINESS: ↑ valence, ↑ arousal ⇌ Pleasant Handheld Happy Neutral
- SADNESS: ↓ valence, ↓ arousal ⇌ Unpleasant Static Sad Blue
- FEAR: ↑ valence, → arousal ⇌ Pleasant Handheld Happy Neutral
- ANGER: ↓ valence, ↑ arousal ⇌ Unpleasant Handheld Happy Red

To match sounds we referred to the fact that the Happy knocking sound is higher in rhythm, leading us to associate it to higher arousal intended emotions, while Sad knocking sound, lower in rhythm, was associated to a lower arousal. Lastly, to match colour filters we hypothesized that high valence intended emotions might be enhanced by a more realistic look of the scene. On the other hand, low valence emotions might be enhanced by a more "unnatural" look of the scene. We associated the Blue filter with Sadness (being a colder colour we matched it with low arousal), and Red filter with Anger (hotter colour matched with higher arousal).

We therefore computed data statistics that we will lately use to perform ANOVA, reported in three different boxplots, in Fig. 4.8, 4.9, and 4.10 .

(a) Do numbers or letters cause you to have a colour experience?

Gender	Yes	No
Male	1	26
Female	11	39

(b) Do weekdays and months have specific colours?

Gender	Yes	No
Male	1	26
Female	16	34

(c) Do you imagine or visualise weekdays, months and/or years as having a particular location in space around you?

Gender	Yes	No
Male	1	26
Female	9	41

(d) Does hearing a sound make you perceive a colour?

Gender	Yes	No
Male	0	27
Female	9	41

(e) Do certain words trigger a taste in your mouth?

Gender	Yes	No
Male	0	27
Female	5	44

(f) Do you feel a sense of touch when you smell things?

Gender	Yes	No
Male	2	25
Female	8	41

(g) Do you suspect that you experience an unusual blending that other people do not have (other than the one listed above)?

Gender	Yes	No
Male	1	26
Female	12	37

Table 4.1: These tables report a summary of the answers at the seven questions about synesthesia taken from McDonald et al. (2022)[37], for Experiment #2. They highlight how females, in general, report being more sensitive than males to synesthesia.

From the boxplots in Fig. 4.8, that refers to the Content of the presented scenes (Pleasant or Unpleasant), we can see that the emotion Anger has been perceived mostly when the Upleasant content was presented, and when it was non-filtered with a LUT. We can notice, anyway, that also the Red-filtered Un-

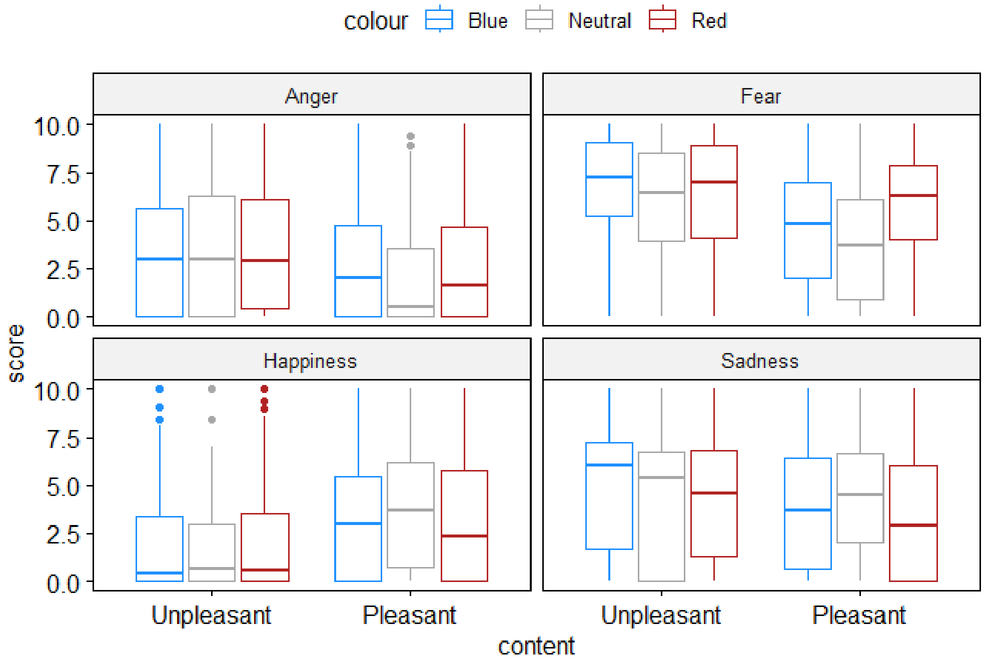


Figure 4.8: Boxplots of statistics of Experiment #2, with reference to the Content of the video clips - On the x-axis we find the factor content, on the y-axis the dependent variable *score*; the boxplots are classified per emotion and the boxes by colour.

pleasant content has received high ratings as Anger.
Fear has received higher ratings when the Unpleasant content was presented and filtered with the Blue LUT.
Happiness has been mostly evoked by Pleasant, Neutral content.
Lastly, Sadness has been detected mostly with the Unpleasant, Blue filtered content, but there is not that much difference with the Pleasant content.
The Content appears to have greatly influenced the perceived emotion.

The boxplots in Fig. 4.9 refers to the camera technique used to shoot the scene, that can be Static (less dynamicity, i.e. less arousal) or Handheld (highest in dynamicity, i.e. high in arousal). We can immediately notice that different Camera Movement techniques didn't influenced much the choice of the perceived emotion in the participants, as the ratings for every emotions are more or less equally distributed between both the presented techniques.
In any case, we can still notice some trends: Anger has received highest ratings with the Handheld Technique with the Red filter and Static with the Blue filter.
Fear has in general received higher ratings with the Handheld techniques, but

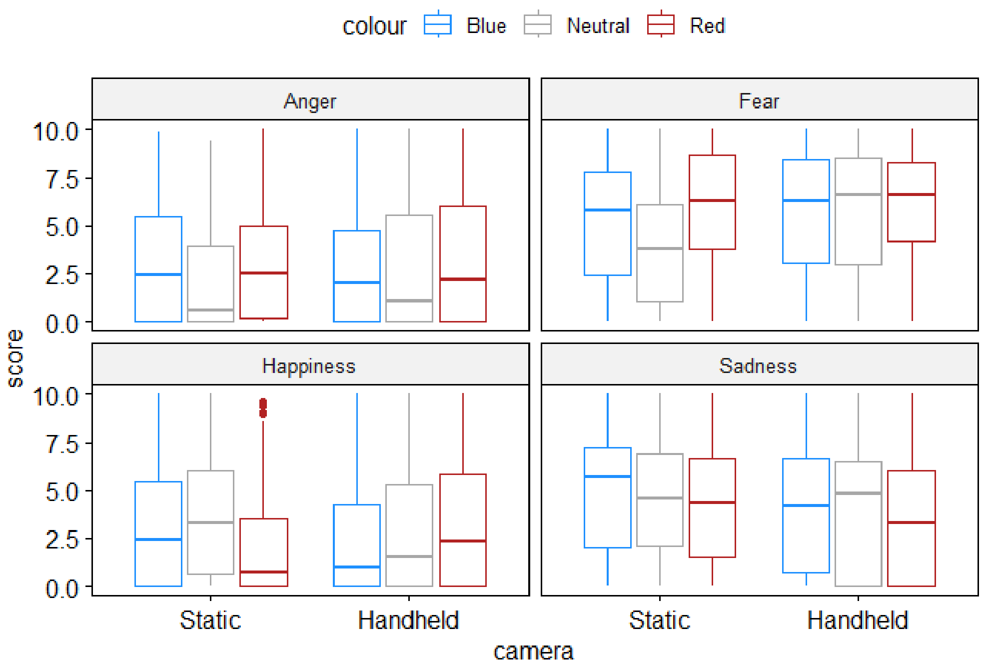


Figure 4.9: Boxplots of statistics of Experiment #2, with reference to the Camera movement techniques used in the video clips - On the x-axis we find the factor camera, on the y-axis the dependent variable *score*; the boxplots are classified per emotion and the boxes by colour.

the highest rating has been received with Static, Red filtered technique. Happiness has been mostly associated with the Static technique, and Sadness with Blue filtered, Static scenes.

The boxplots in Fig. 4.10, that refers to the emotional knocking sounds present in the clips (could be Happy or Sad), highlight that Sound, even if presented at the end of the clips, influenced quite a lot the perception of the conveyed emotion. Fear, that is also influenced by the Red and Blue filtered, i.e. the alterations from the most realistic scene, is mostly associated with the Sad knocking sound. Happiness has been clearly associated with the Happy knocking sound (which is higher in rhythm) and again with the Neutral colour, more close to what we are used to frequently see in real life. Sadness has been mostly associated, as expected, with the Sad knocking sound and mostly to the Blue filter. Anger, on the other hand, appears to not having been influenced much from the sound presence, even if the trend has been to associate it with

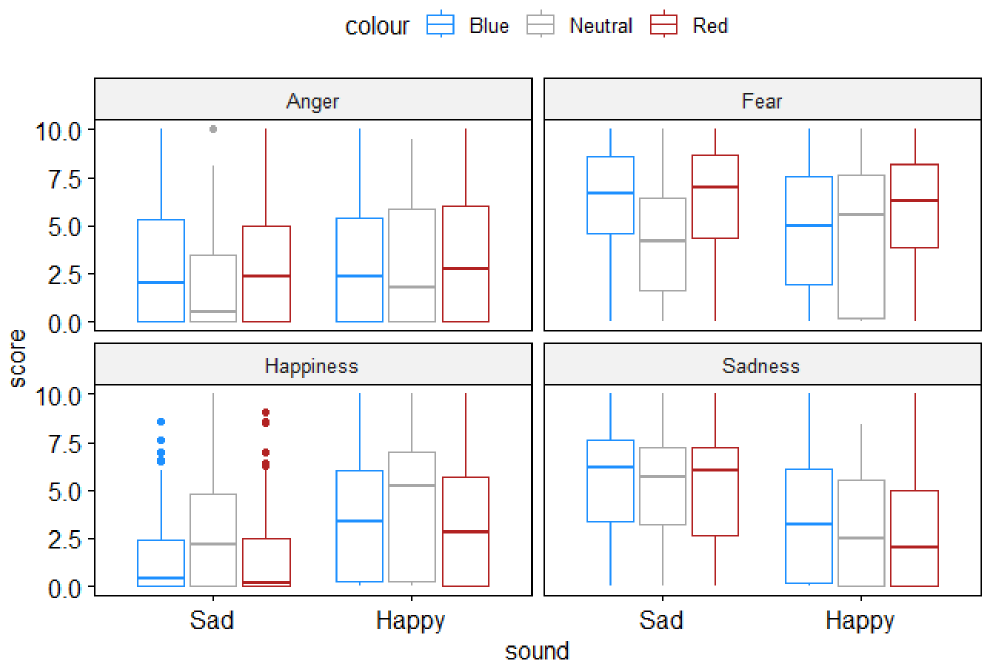


Figure 4.10: Boxplots of statistics of Experiment #2, with reference to the knocking Sound presented in the video clips.

happier knocking sounds.

From the observation of these boxplots we can also conclude that the coloured filter applied to the video clips does not have a great influence on the perceived emotion.
In the case of Sadness, however, the blue filter has a strengthening function and, in general, the presence of coloured filters (without significant differences between Red and Blue) amplifies the emotions of Anger and Fear.

To summarize, Content and Sound factors appear to be the two main influences on perceived emotions; Colour, also in light of the results of Experiment #1, is confirmed as a merely reinforcing influence, while Camera movements do not appear to have any significant influence on the perceived emotions.
These conclusions will be taken as hypothesis to help us interpret the following 5-Ways ANOVA analysis.

4.3 Analysis

Keeping in mind the hypothesis derived from the analysis of the statistics, we want to verify with a Five-Ways ANOVA the existence of significant interactions between the factors taken into consideration, and then verify whether Welch's ANOVA (that takes into account the non-verification of the hypothesis of homogeneity of variances), calculated on the factors that result dependent on each other, also provides us with a *p-value* lower than *0.05*, thus confirming or denying the significance of the obtained results.

4.3.1 Five-Ways ANOVA

A Five-Ways ANOVA was conducted to determine the effects of Content, Camera, Sound and Colour on the Emotions' scores given by the participants This was done using RStudio, an integrated development environment for R, a programming language for statistical computing and graphics.

Before performing ANOVA, we had to check for its assumptions to be verified in our dataset.

Outliers

```
print((data_anova %>%
         group_by(content, camera, sound, colour, emotion) %>%
         identify_outliers(score)), n=2840)
```

Four extreme outliers, that could skew our computations, have been detected and removed from the dataset.

```
# Function to filter out outliers
remove_outliers <- function(data_anova, outliers) {
  data_anova %>%
    filter(!(participant %in% outliers$participant & emotion %in%
   outliers$emotion))
}

# Remove outliers from the data frame
data_anova <- remove_outliers(data_anova, outliers)
```

We decided to maintain the identified multiple non-extreme outliers, believing on their meaningfulness for the scope of our analysis.

Normality

To verify normality we used the visual method of QQ Plots, reported in Fig. 4.11.

```
#Normality assumption by analyzing the model residuals

model  <- lm(score ~ content*camera*sound*colour*emotion, data = data
    _anova)
# Create a QQ plot of residuals
ggqqplot(residuals(model))
```

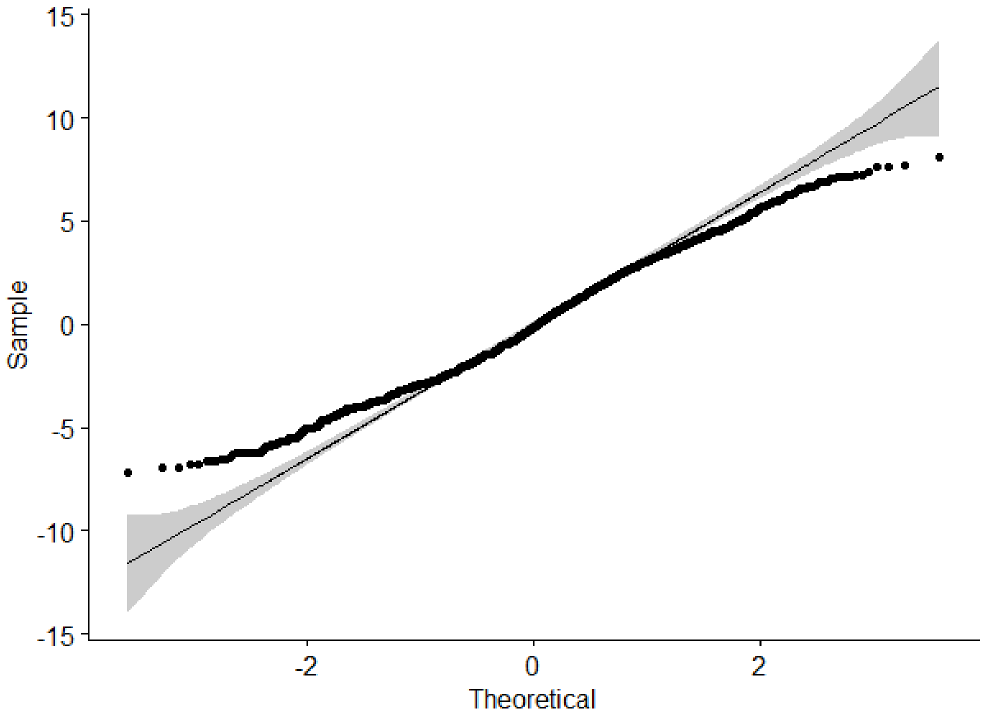


Figure 4.11: QQ Plot on the whole dataset of Experiment #2, computed to verify the ANOVA's assumption of normality on the collected data. We can see that, excluding the extremes, the points fall approximately along the reference line, so we can assume normality.

In the QQ plot 4.11, excluding the extremes, the points fall approximately along the reference line, so we can assume normality.

After normality on the whole dataset, we also checked the assumption for each combination of factor levels. The outcomes of the QQ plots has been reported in Fig. 4.12.

```
# QQ plot for each cell of design

ggqqplot(data_anova, "score", ggtheme = theme_bw()) +
  facet_grid(content + camera + sound + colour ~ emotion, labeller =
   "label_both")
```

Again, we can see from the plots that the points, excluding extremes, fall approximately in the region close to the reference line, we can therefore assume normality also for each group of combinations.

Homogneity of variances

Homogeneity of variances was investigated by Levene's test.

```
data_anova %>% levene_test(score ~ content*camera*sound*colour*
   emotion)
```

The outcome of the test is reported in Table 4.2, and it determined the non verification of the assumption ($p << 0.05$), meaning that the variance of outcome variables is not similar in each group. Consequently, there is grater probability of falsely rejecting the Null Hypothesis that *there is no difference among group means*.

df1	df2	statistic	p
87	2748	2.21	0.00000000159

Table 4.2: Results of the Levene's test on Experiment #2 dataset, ran to verify homogeneity of variances. The *p-value* of $1.59e-9$ states the non-verification of such hypothesis.

Even if homogeneity was not verified, we decided to proceed with the Five-Ways ANOVA, but paying attention to run also Welch's ANOVA to the obtained results, that does not assume homoscedasticity between groups and is more robust when this assumption is violated, to check the significance.

Five-Ways ANOVA

```
res.aov5w <- data_anova %>% anova_test(score ~ content*camera*sound*
    colour*emotion)
res.aov5w
```

The results of the Five-Ways ANOVA are reported in Table 4.3.

From the table we can see that there was a statistically significant Five-Ways interaction between Content, Camera, Sound, Colour and Emotion (`content:camera:sound:colour:emotion`), with $F(3, 2748) = 6.251, p = 3.12e-4$, this means that a statistically significant difference in score means was detected.
Furthermore we can notice than four and three-ways interactions result highly significant when the factors Content and Sound are taken into account (we don't give any interpretation to the significance of the factor Emotion because in our specific analysis it is not a real independent variable, but a factor useful to classify the given scores). The factor Camera result in significant interactions only when matched with Content, Sound or Emotion and, because of this, we can confirm our previous hypothesis that its influence on the perceived emotion is not statistically significant. This conclusion can help us to discard some post-hoc analysis.

As the influence of the used Camera technique and of the filter Colour result to be discardable, to deepen into the analysis we decided to run a Three-Way ANOVA to check the significance of the influence of Content and Sound on the Emotion's score. The results are reported in Table 4.4.

```
#3-WAYS ANOVA: compute ANOVA on content*sound*emotion interaction on
    score

model  <- lm(score ~ content*camera*sound*colour*emotion, data = data
    _anova)
print((data_anova %>%
         anova_test(score ~ content*sound*emotion, error = model)), n
    =7)
```

Effect	DFn	DFd	F	p	sig	ges
content	1	2748	18.224	2.03e-05	*	7.00e-03
sound	1	2748	0.163	6.86e-01		5.94e-05
slider	3	2748	138.669	1.16e-83	*	1.31e-01
content:sound	1	2748	0.250	6.17e-01		9.08e-05
content:slider	3	2748	37.517	9.33e-24	*	3.90e-02
sound:slider	3	2748	61.890	1.02e-38	*	6.30e-02
content:sound:slider	3	2748	4.194	6.00e-03	*	5.00e-03

Table 4.4: Three-Ways ANOVA test - The results highlight a statistically significant Three-Ways interaction between Content, Sound, and Emotion (`content:sound:emotion`), with $F(3, 2748) = 4.194, p = 0.006$.

To better understand the interaction between the factors, we can decompose the significant Three-Way interaction effect into simple Two-Ways interactions at each level of the Emotion variable.
We therefore decided to perform the Two-Ways ANOVA both on Content and Emotion and on Sound and Emotion.

Data have been first grouped by Sound and the Two-Ways ANOVA was ran between Content and Emotion.

```
#2-WAYS ANOVA: Group the data by content and fit simple two-ways interaction
model  <- lm(score ~ content*sound*emotion, data = data_anova)
data_anova %>%
  group_by(sound) %>%
  anova_test(score ~ content*emotion, error = model)
```

The result was, as expected, that there was a statistically significant simple Two-Ways interaction between Content and Emotion for both the sounds Happy and Sad, as reported in Table 4.5. This result suggests that the perceived emotion when watching to a video is primarily led by its content.

sound	Effect	DFn	DFd	F	p	p<.05	ges
Happy	content	1	2820	10.7	1e-3	*	0.004
Happy	emotion	3	2820	39.1	9.96e-25	*	0.04
Happy	content:emotion	3	2820	13.6	8.5e-9	*	0.014
Sad	content	1	2820	7.16	7e-3	*	0.003
Sad	emotion	3	2820	158	6.55e-95	*	0.144
Sad	content:emotion	3	2820	26.7	5.03e-17	*	0.028

Table 4.5: ANOVA Two-Ways test on Content and Emotion - The results highlight a significant simple two-ways interaction between content and emotion (`content :emotion`) for every group of sounds; this means that the mean score in the Content groups (Pleasant, Unpleasant) were statistically significantly different regardless the knocking sound presented in the videos. In other words, the principal influence on the perceived emotion is the one given by the content of the scene.

Even if we were sure about the significance of this result, we ran anyway the One-Way Welch's ANOVA on Content, to inspect its effect on *score* and also to verify if also following a procedure more robust to hypothesis relaxations, significance would be confirmed.

```
# Group the data by sound and emotion, and fit Welch's anova, to see
    effect of only content on score
content.effect <- data_anova %>%
  group_by(sound, emotion) %>%
  welch_anova_test(score ~ content)
content.effect
```

Results are reported in Table 4.6.

The interactions inspected by the Welch's ANOVA results all highly significant, in fact we can see from the table that all the resulting $p - values$ are way smaller than the reference value 0.05, except for the association between the Sad sound and the emotion Sadness (that might be led by the sound itself), highlighting the strong significance of the dependence of the emotion's *score* on the content of the scene presented in the video clips.

sound	emotion	.y.	n	statistic	DFn	DFd	p	p<0.05
Happy	Fear	score	350	9.09	1	301.459	0.003	*
Happy	Happiness	score	350	12.45	1	310.304	0.000481	*
Happy	Sadness	score	350	5.97	1	274.492	0.015	*
Sad	Anger	score	360	6.09	1	344.989	0.014	*
Sad	Fear	score	360	53.61	1	357.835	1.63e-12	*
Sad	Happiness	score	358	42.12	1	302.171	3.51e-10	*
Sad	Sadness	score	358	0.5	1	354.349	0.479	

Table 4.6: Results of the Welch's ANOVA (One-Way) ran to inspect the effect of Content on *score*, to verify if also following a procedure more robust to hypothesis relaxations significance would be confirmed. We can see that the obtained $p - value$ is lower than 0.05 for every interaction except for the association between the Sad sound and the emotion Sadness (that might be led by the sound itself), highlighting the strong significance of the dependence of the emotion's *score* on the content of the scene presented in the video clips.

The same analysis was ran grouping data by Content and performing Two-Ways ANOVA between Sound and Emotion. The results are reported in Table 4.7.

```
#2-WAYS ANOVA: Group the data by content and fit simple two-way
   interaction
model  <- lm(score ~ content*sound*emotion, data = data_anova)
data_anova %>%
  group_by(content) %>%
  anova_test(score ~ sound*emotion, error = model)
```

Content	Effect	DFn	DFd	F	p	p<.05	ges
Pleasant	sound	1	2820	0.392	5.31e-1		0.000139
Pleasant	emotion	3	2820	41.7	2.32e-26	*	0.042
Pleasant	sound:emotion	3	2820	34.0	1.47e-21	*	0.035
Unpleasant	sound	1	2820	0.01	9.19e-1		0.00000366
Unpleasant	emotion	3	2820	132	3.56e-80	*	0.123
Unpleasant	sound:emotion	3	2820	29.9	5.32e-19	*	0.031

Table 4.7: ANOVA Two-Ways test on Sound and Emotion - The results highlight a significant simple two-ways interaction between Sound and Emotion (`sound:emotion`) for both the groups of contents, meaning that Sound, along with Content, influenced in a statistically significant way the emotions' scores given by the participants.

The analysis underlined significant two-ways interaction between Sound and Emotion (`sound:emotion`) for both the groups of contents, meaning that Sound, along with Content, influenced in a statistically significant way the emotions' scores given by the participants.

To further verify the statistic significance of the found results we ran the One-Way Welch's ANOVA on Sound, to inspect its effect on the dependent variable *score* and also to verify if following a procedure more robust to hypothesis relaxations, significance would still be confirmed.

```
# Group the data by content and emotion, and fit Welch's anova, to
   see effect of only sound on score
sound.effect <- data_anova %>%
  group_by(content, emotion) %>%
  welch_anova_test(score ~ sound)
sound.effect
```

Results are reported in Table 4.8, and highlight that there were statistically significant simple simple main effect of Sound for all the videos that presented the Unpleasant content. For the Pleasant content's videos, the influence of Sound results to be significant only to discern between Happiness and Sadness.
In this part of the analysis statistical significance was accepted at a Bonferroni-adjusted alpha level of 0.025 (0.05 divided by 2, that is the number of simple simple main effects we are computing).

content	emotion	.y.	n	statistic	DFn	DFd	p	sig
Pleasant	Anger	score	390	3.84	1	387.803	0.051	
Pleasant	Fear	score	390	0.73	1	387.5587	0.392	
Pleasant	Happiness	score	390	22.04	1	387.999	3.71e-06	*
Pleasant	Sadness	score	390	79.98	1	368.475	1.86e-17	*
Unpleasant	Anger	score	320	12.33	1	281.030	0.000519	*
Unpleasant	Fear	score	320	6.9	1	286.172	0.009	*
Unpleasant	Happiness	score	318	38.85	1	210.078	2.47e-09	*
Unpleasant	Sadness	score	318	33.79	1	300.970	1.56e-08	*

Table 4.8: Results of the Welch's ANOVA (One-Way) ran to inspect the effect of Sound on *score*, to verify if also following a procedure more robust to hypothesis relaxations, significance would be confirmed. We can observe, looking at the $p-values$ that the influence of Sound has been always significant except for the association between the Pleasant content to the emotions Fear and Anger.

Lastly, we decided to perform a Welch's ANOVA test (taking into account the complete dataset) on Colour, to further check our conclusion of its non-significance on the scores given by the participants.

```
#WELCH'S ANOVA ON COLOUR
# Group the data by content, camera, sound and emotion, and fit Welch
    's anova, to see effect of only colour on score
print((colour.effect <- data_anova %>%
  group_by(content, camera, sound, slider) %>%
  welch_anova_test(score ~ colour)), n=32)
colour.effect
```

The results of the Welch's ANOVA test for Colour are reported in Table 4.9.

The results highlight that the mean score in the Colour groups (Red, Neutral, Blue) was statistically significantly different only for Pleasant Handheld Sad stimulus, when associated to the emotions Fear and Sadness, for Pleasant Static Happy stimulus, when associated to Happiness, for Unpleasant Handheld Happy stimulus, associated to Anger, and Unpleasant Handheld Sad stimulus when associated to Sadness, but not for for the other stimuli. Therefore we can confirm that, in general, a statistically significant influence of Colour is sporadic and not systematic.

4.4 Results Summary

The experiment results suggest that the main statistically significant influences on the association between the multimedia stimuli and the perceived emotion was given firstly by the Content of the scene presented in the clips (that could be Pleasant or Unpleasant), as we can deduce from the results of the Two-Ways ANOVA in Table 4.5 and of the Welch's ANOVA of Table 4.6; and secondly by the auditory stimuli presented in the clips (the knocking Sound), as we can deduce from the results of the Two-Ways ANOVA in Table 4.7 and of the Welch's ANOVA in Table 4.8.
Colour had low statistically significant influence on the perceived emotion (the *p-value* of the few influenced stimuli was always quite close to the significance level α) and just in some sporadic cases, so it cannot be considered a systematic and consolidated influence.
Furthermore, the Camera movement technique used to shot the videos seems to have absolutely no influence on the perceived emotion, as we deduced from the boxplots in Fig. 4.9 and from the Five-Ways ANOVA's results in Table 4.3.

Going back to our initial hypothesis of the associations between perceived emotions and video stimuli, referring the boxplots in Fig. from 4.8 to 4.10, we can deduce that:

- Happiness has been correctly mostly associated to Pleasant content, Happy Sound and Neutral colour, but wrongly associated to Static camera movement technique.
- Sadness has been mostly correctly associated to Unpleasant Static Sad Blue filtered videos, just as we had assumed.
- Fear was wrongly (with reference to our hypothesis) associated to Unpleasant content and Sad knocking sound(as common sense would also suggest), but correctly associated with Handheld technique. Furthermore, it appears to be enhanced by the presence of a colour filter, without a significant difference between the Red or the Blue one.
- Anger has been correctly associated to Unpleasant content and to Happy knocking sound, while camera movement techniques and colour filter don't seem to have influenced much the participants' ratings.

4.5 Limits of the Second Experiment

It is crucial to acknowledge limitations of the conducted experiment, both to delineate the boundaries within which our research findings can be interpreted, and to point out some starting point for further researches on this theme.

- The number of participant should be increased to check if a bigger amount of data would lead to different conclusions. Out target was to reach 90 participants. At the closing of the survey we obtained 104 completed surveys, but the initial data filtering led us to have to delete the data relating to 26 participants, leaving us with just 78 questionnaires suitable for the analysis.

- The study focus just the main elements that compose a cinematic scene, and each these elements are represented only by a few values they could assume, which may not capture the full complexity of the art of cinema.

- We think that the Pleasant video stimuli content was not different enough from the Unpleasant one, even though it still resulted to have highly influenced the participants' ratings. We would suggest producing stimuli with more diversified contents, so as to avoid the possibility of these being confused with each other.

- The usage of very similar video stimuli might have introduced biases in the participants' rating. Influenced from the previous videos, participants might have rated the following videos with some prejudices.

- The experiment's findings refers to a to a specific and narrow set of emotion, which may not capture the full complexity of emotional responses, we therefore suggest, in future analysis, to use as responses' structure something more similar, for example, to the Plutchik's Wheel of Emotions (Plutchick, 1980)[45] instead of multiple continuous sliders associated to different basic emotions.

- The study employs a 0 to 10 rating scale and follow-up questions on synesthesia, potentially introducing subjectivity (not all the participants are equally open to perceive synesthesia) and response bias.

- The usage of the sliders associated to the name of the emotions as responses' structure might be not very intuitive for the participant. We therefore suggest to use a more graphical structure for the responses, for example using facial expressions from Ekman's faces dataset of facial expressions (1970)[14] (even though the gender of the people presented in the pictures could influence the participants' rating), or again black and white emojis could be used (removing colour is fundamental so that they do not interfere with the experiment target visual stimuli).

- The survey is completely in English, but the participant are mostly from Italy and belongs to the most different age ranges. It could be helpful, to prevent misunderstandings due to the foreign language of the questionnaire, to produce an italian alternative version of the survey.

- The usage of such video stimuli, although complex, may not fully replicate real-world scenarios.

- The experiment was not conducted within a controlled laboratory setting, and the generalizability of the findings to real-world contexts warrants further investigation.

- The video stimuli were clips of 26 seconds each. We think that the length of the video might be too long to keep the attention of the participants until the very end, when the knocking sounds were presented. We therefore suggest that the factors important for our analysis, when the experiment is not conducted in a controlled environment, should be central in the development of the stimuli to be presented to the participants.

- The usage of the platform PsyToolkit(Stoet, 2010, 2017)[56][57][4] introduced some bugs that might have caused the increment of variability in the given responses. Sometimes the platform doesn't correctly lead the video stimuli, we therefore suspect that, when this bug occurred, some responses might have been given randomly, just to complete the survey. This is a major issue that might have skewed our analysis introducing so much difference in variances.

[4] https://www.psytoolkit.org/

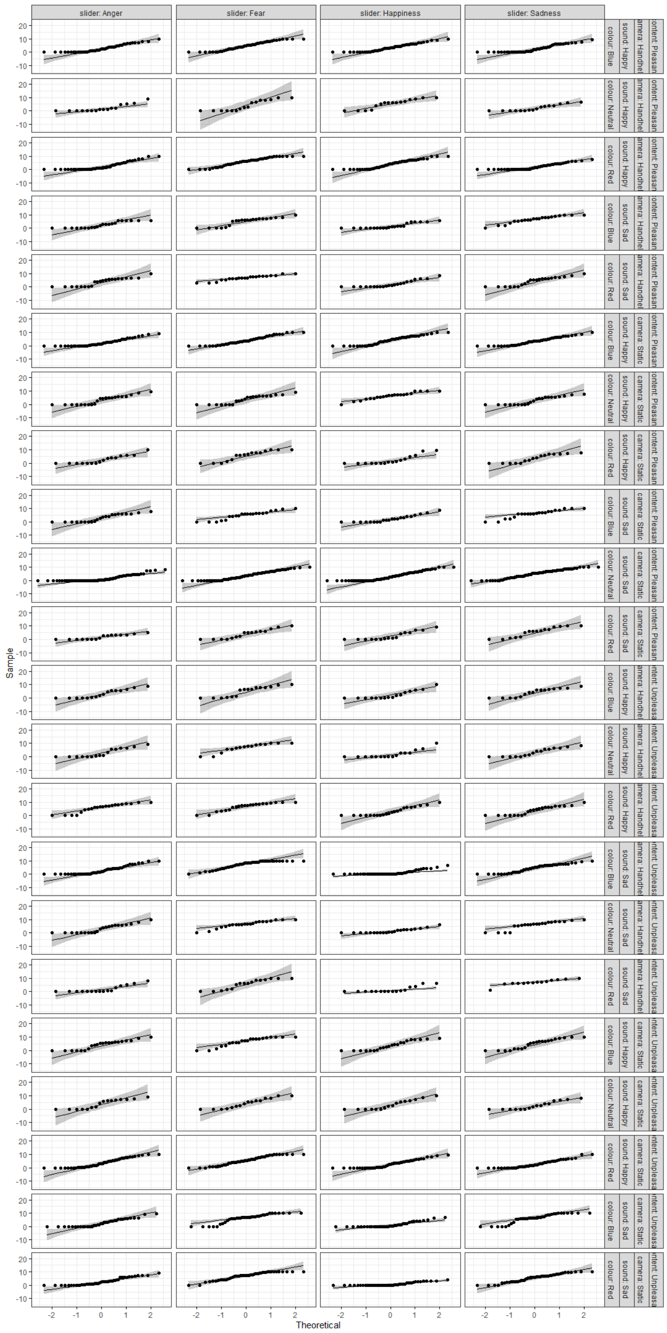


Figure 4.12: QQ Plot for each combination of factor levels of Experiment #2 - To further verify the assumption, we checked normality also for each combination of factors level. The plots show that, excluding extremes, the points fall approximately along the reference line, so we can assume normality even for each factor levels' combinations.

Effect	DFn	DFd	F	p	sig	ges
content	1	2748	9.04	0.003	*	0.003
camera	1	2748	0.087	0.768		3.18e-05
sound	1	2748	0.439	0.508		0.00016
colour	2	2748	0.372	0.69		0.00027
emotion	3	2748	138.669	1.16e-83	*	0.131
content:camera	1	2748	0.033	0.855		1.21e-05
content:sound	1	2748	6.749	0.009	*	0.002
camera:sound	1	2748	0.164	0.685		5.97e-05
content:colour	2	2748	2.798	0.061		0.002
camera:colour	2	2748	4.03	0.018	*	0.003
sound:colour	2	2748	0.803	0.448		0.000584
content:emotion	3	2748	16.163	2.05e-10	*	0.017
camera:emotion	3	2748	3.669	0.012	*	0.004
sound:emotion	3	2748	54.659	2.59e-34	*	0.056
colour:emotion	6	2748	3.942	0.000627	*	0.009
content:camera:sound	1	2748	0.353	0.553		0.000128
content:camera:colour	2	2748	1.58	0.206		0.001
content:sound:colour	1	2748	0.379	0.538		0.000138
camera:sound:colour	1	2748	1.359	0.244		0.000494
content:camera:emotion	3	2748	1.004	0.39		0.001
content:sound:emotion	3	2748	2.62	0.049	*	0.003
camera:sound:emotion	3	2748	0.362	0.781		0.000395
content:colour:emotion	6	2748	2.168	0.043	*	0.005
camera:colour:emotion	6	2748	1.575	0.15		0.003
sound:colour:emotion	6	2748	0.507	0.803		0.001
content:camera:sound:colour	1	2748	6.391	0.012	*	0.002
content:camera:sound:emotion	3	2748	1.587	0.19		0.002
content:camera:colour:emotion	6	2748	1.759	0.104		0.004
content:sound:colour:emotion	3	2748	3.924	0.008	*	0.004
camera:sound:colour:emotion	3	2748	0.375	0.771		0.000409
content:camera:sound:colour:emotion	3	2748	6.261	0.000312	*	0.007

Table 4.3: Five-Ways ANOVA test - The results highlight a statistically significant Five-Ways interaction between Content, Camera, Sound, Colour and Emotion (`content:camera:sound:colour:emotion`), with $F(3, 2748) = 6.251, p = 3.12e - 4$.

content	camera	sound	emotion	.y.	n	statistic	DFn	DFd	p	sig
Pleasant	Handheld	Happy	Anger	score	118	0.62	2	43.088	0.543	
Pleasant	Handheld	Happy	Fear	score	118	2.76	2	39.562	0.076	
Pleasant	Handheld	Happy	Happiness	score	118	2.16	2	40.359	0.129	
Pleasant	Handheld	Happy	Sadness	score	118	0.57	2	42.734	0.571	
Pleasant	Handheld	Sad	Anger	score	44	2.52	1	39.180	0.12	
Pleasant	Handheld	Sad	Fear	score	44	4.82	1	35.309	0.035	*
Pleasant	Handheld	Sad	Happiness	score	44	0.6	1	39.556	0.442	
Pleasant	Handheld	Sad	Sadness	score	44	6.81	1	41.006	0.013	*
Pleasant	Static	Happy	Anger	score	89	0.95	2	32.288	0.396	
Pleasant	Static	Happy	Fear	score	89	1.62	2	33.591	0.213	
Pleasant	Static	Happy	Happiness	score	89	8.83	2	35.573	0.000769	*
Pleasant	Static	Happy	Sadness	score	89	0.22	2	34.072	0.806	
Pleasant	Static	Sad	Anger	score	139	2.83	2	30.126	0.075	
Pleasant	Static	Sad	Fear	score	139	1.88	2	29.160	0.17	
Pleasant	Static	Sad	Happiness	score	139	0.9	2	30.999	0.416	
Pleasant	Static	Sad	Sadness	score	139	1.12	2	28.868	0.34	
Unpleasant	Handheld	Happy	Anger	score	54	4.21	2	32.356	0.024	*
Unpleasant	Handheld	Happy	Fear	score	54	1.26	2	30.989	0.299	
Unpleasant	Handheld	Happy	Happiness	score	54	0.54	2	32.559	0.587	
Unpleasant	Handheld	Happy	Sadness	score	54	0.29	2	32.416	0.749	
Unpleasant	Handheld	Sad	Anger	score	89	1.22	2	35.925	0.309	
Unpleasant	Handheld	Sad	Fear	score	89	1.41	2	33.282	0.259	
Unpleasant	Handheld	Sad	Happiness	score	88	0.45	2	32.848	0.639	
Unpleasant	Handheld	Sad	Sadness	score	87	6.83	2	35.374	0.003	*
Unpleasant	Static	Happy	Anger	score	89	0.2	2	35.069	0.823	
Unpleasant	Static	Happy	Fear	score	89	1.88	2	33.898	0.169	
Unpleasant	Static	Happy	Happiness	score	89	0.6	2	32.495	0.555	
Unpleasant	Static	Happy	Sadness	score	89	3.09	2	32.823	0.059	
Unpleasant	Static	Sad	Anger	score	88	0.001	1	75.459	0.975	
Unpleasant	Static	Sad	Fear	score	88	1.37	1	72.096	0.246	
Unpleasant	Static	Sad	Happiness	score	87	3.15	1	56.631	0.081	
Unpleasant	Static	Sad	Sadness	score	88	1.09	1	73.959	0.3	

Table 4.9: Results of the Welch's ANOVA (One-Way) ran to inspect the effect of Colour on *score*, to verify if following a procedure more robust to relaxations, our hypothesis of Colour as a generally non-significant factor, but only (just in some cases) reinforcing the emotion conveyed by other factors, would have been confirmed. By looking at the results we can confirm that, in general, a statistically significant influence of Colour is sporadic and not systematic.

5

Conclusions and Future Works

In the contemporary landscape, characterized by the paramount significance of effective communication, Sound Design has emerged as a critical element in the creation of multimedia products. Its ability to evoke profound emotional responses in audiences holds immense value for its applications in marketing and in multimedia production, as well as in artistic domains such as cinema and immersive video games.
Our journey into Sound Design allowed us to dismantle a cinematic scene into its main components, including scene content, camera movement techniques, sound landscape and colour filtering, and analyze how these variables interfere with each other to arouse specific emotions in audiences. We have endeavored to conduct an empirically rigorous investigation, grounded in objective scientific observations, to unveil the key factors influencing human emotional perception.

We divided our analysis in two different paths.
The first one had the aim to deepen other studies conducted by Turato et al. in 2022[63] and 2023[62]. The studies had the objective of investigating the influence of certain colours, proved to be associated with specific emotions, on emotionally performed knocking sounds. The hypothesis was that the association between knocking sounds and emotions would depend on the colour of the door on which the knocking sound is performed.
The second path had the aim of translating this same type of analysis onto a set of multiple factors, which together could come closer to composing a real cinematographic scene. We therefore added two further factors to the first anal-

ysis: the background content of the scene and the camera movement technique used. The hypothesis was that, classifying the factors that compose the stimuli depending on their levels of valence and arousal, and matching them with the emotion with correspondent valence and arousal levels, we could predict which emotion the participants would associate to a specific scene.

Both the researches have been carried out through the development of two different surveys, using the PsyToolkit online platform (Stoet, 2010, 2017)[56] [57] [1].

Experiment #1 The first survey was composed of 25 video stimuli of 3 seconds each, which consist of a video of a hand knocking on a door of different colours (white, yellow, red, purple, gray), synchronized with the audio of five different emotional knocking sounds taken from the same dataset used by Barahona-Ríos and Pauletto in 2020, which was also employed in the work of Pauletto and Iop in 2021[38]. These knocking actions were professionally performed by Ulf Olausson, an expert Foley artist[2].
The 25 videos were presented in the survey followed by four continuous sliders, each one associated to one of the four emotions: Happiness, Sadness, Fear and Anger. Participant were asked, after watching each video, to rate the emotion they perceived by moving the four sliders.
118 participants correctly completed this first survey. The data have been collected and elaborated. Firstly we ran a general analysis of the statistics (means and SDs) to obtain an overview of the given responses and to immediately highlight some trends. After this first step, we proceeded to perform a Three-Ways ANOVA test (and its post-hoc tests) on the results, to understand whether the scores given by the participant had been influenced by the emotional intentions of the knocking sounds presented in the videos and/or by the colours of the doors.

The experiment results suggest that the main influence on the association between the multimedia stimuli and the perceived emotion was given by the auditory stimuli presented in the clips (the knocking sounds) more than by the visual stimuli (colours of the door).

[1] `https://www.psytoolkit.org/`

We can see in Fig. 3.4, that represents the boxplots of the ratings given by participants, that stimuli intended to represent the emotions Happiness, Sadness and Anger have been correctly and uniquely classified. Fear, on the other hand, as we can see in Table 3.4a has been mostly wrongly detected as Anger. In general, participants found difficult to discern between Fear and Anger.
From the same boxplot and from Tables 3.3 and 3.4, we can also spot some trends: as we expected the colour that enhanced the emotion Happiness has been Yellow and Gray for Sadness, and contrary to what we expected, Red for Fear (we associated it with Purple) and Purple for Anger (that we associated with Red).
The effect of Colours on participants' ratings of the four basic emotions did not produce systematic statistically significant results.
Has to be noticed also that, in general, the responses contained a relatively high variance.

In summary, regarding Experiment #1, the emotions intended by the knocking sounds have been mostly correctly detected by the audience, except for the fact that the "fear" sound is many times confused with Anger, that in turn has been correctly detected when presented. Sound therefore appears to be the main influence on the emotions perceived by the participants.
The influence of the door colours is minimal, and, just in some cases, it only reinforces the emotion evoked by the knocking sounds.

Experiment #2 The second phase of our study comprised eight video stimuli chosen from a pool of 24 options, each lasting 26 seconds. The videos were created by picking four different visual stimuli from the dataset of camera movement scenes from Yilmaz, Lotman , Karjus and Tikka experiment (2023)[9]. The four stimuli encompassed both Pleasant and Unpleasant content, captured using both Static and Handheld techniques. At each one of these four obtained videos, we incorporated two different emotional knocking sounds, professionally performed by Ulf Olausson[2], a subset of the ones utilized by Barahona-Ríos and Pauletto in 2020[1], which was also employed in the work of Pauletto and Iop in 2021[41] and of Turato in 2022[63].
We therefore obtained eight videos, and for each of these videos we applied two

colour filters to Red and Blue, to be added to the Neutral videos (the original ones).
The 8 videos were presented in the survey followed by four continuous sliders, each one associated to one of the four emotions: Happiness, Sadness, Fear and Anger. Participant were asked, after watching each video its entirety, to rate the emotion they perceived by adjusting the four sliders.
78 participants correctly completed this first survey. We collected and analyzed the data.
Our analysis commenced with a general examination of the statistics, calculating means and standard deviations. We visualized the responses with informative boxplots to identify trends that would facilitate subsequent analysis. After this first step, we proceeded to perform a Five-Ways ANOVA test (and its post-hoc tests) on the results, to understand whether the scores given by the participant had been influenced by the intention of the scene content, the camera movement technique, the knocking sounds presented in the videos and/or by the colours of the doors.

The experiment results suggest that the main statistically significant influences on the association between the multimedia stimuli and the perceived emotion was given firstly by the Content of the scene presented in the clips (that could be Pleasant or Unpleasant), as we can deduce from the results of the Two-Ways ANOVA in Table 4.5 and of the Welch's ANOVA of Table 4.6; and secondly by the auditory stimuli presented in the clips (the knocking sound), as we can deduce from the results of the Two-Ways ANOVA in Table 4.7 and of the Welch's ANOVA in Table 4.8.
Colour had low statistically significant influence on the perceived emotion and just in some sporadic cases, so it cannot be considered a systematic and consolidated influence.
Furthermore, the Camera movement technique used to shot the videos appeared to have absolutely no substantial influence on the perceived emotion, as we deduced from the boxplots in Fig. 4.9 and from the Five-Ways ANOVA's results in Table 4.3.

Revisiting our initial hypothesis regarding the associations between perceived emotions and video stimuli, referring the boxplots in Fig. from 4.8 to 4.10, we observed the following:

- Happiness has been correctly mostly associated to Pleasant content, Happy Sound and Neutral colour, but wrongly associated to Static camera movement technique.
- Sadness has been mostly correctly associated to Unpleasant Static Sad Blue filtered videos, just as we had assumed.
- Fear was wrongly (with reference to our hypothesis) associated to Unpleasant content and Sad knocking sound(as common sense would also suggest), but correctly associated with Handheld technique. Furthermore, it appears to be enhanced by the presence of a colour filter, without a significant difference between the Red or the Blue one.
- Anger has been correctly associated to Unpleasant content and to Happy knocking sound, while camera movement techniques and colour filter don't seem to have influenced much the participants' ratings.

Limitations and Suggestions for Future Work It is imperative also to acknowledge the limitations that have emerged from our conducted experiment. These limitations serve a twofold purpose: first, they delineate the boundaries within which our research findings can be reasonably interpreted, and second, they provide a foundation for further investigations on this theme.

To begin with, we must address the issue of participant numbers. Although our initial target was to recruit 90 participants, we exceeded this number, receiving a total of 104 completed surveys. However, due to data filtering and quality control, we were forced to exclude data from 26 participants, leaving us with only 78 questionnaires suitable for our analysis. Increasing the number of participants could offer insights into whether a larger dataset would yield different conclusions.

Furthermore, our study primarily focuses on the fundamental elements that constitute a cinematic scene. Each of these elements is represented by only a limited range of values, which may not fully capture the intricacies and nuances of the art of cinema.

A significant consideration is the similarity between the Pleasant and Unpleasant video stimuli. While both had a notable impact on participants' ratings, it is possible that more diversified content would help avoid potential confusion between the two.

The introduction of biases in participant ratings due to the similarity of video stimuli is another concern. It is plausible that the influence of previous videos affected how participants rated subsequent ones, introducing a potential source of bias.

Our experiment's findings are confined to a specific and narrow set of emotions, potentially missing the complexity of emotional responses. Future analyses might benefit from utilizing a more comprehensive framework, such as Plutchik's Wheel of Emotions (Plutchick, 1980)[45], instead of multiple continuous sliders associated with basic emotions.
In general, the structure of responses using emotion names associated with sliders may not be intuitive for participants. Employing graphical representations, such as Ekman's facial expressions dataset (1970)[14] or black and white emojis, could be more user-friendly.

The use of a 0 to 10 rating scale and follow-up questions on synesthesia introduces subjectivity and response bias, as not all participants may be equally open to perceiving synesthesia.

Considering that participants primarily hail from Italy and represent diverse age groups, offering an Italian version of the survey may help mitigate language-related misunderstandings.

We must also acknowledge that the complexity of video stimuli, while valuable, may not fully replicate real-world scenarios.

It's important to note that our experiment was not conducted in a controlled laboratory setting, which raises questions about the generalizability of our findings to real-world contexts.

The length of our video stimuli, at 26 seconds each, may have posed a challenge in maintaining participants' attention until the presentation of the crucial knocking sounds. This issue underscores the need to consider the attention span of participants in future experimental designs, especially in less controlled settings.

Lastly, the use of the PsyToolkit online platform (Stoet, 2010, 2017)[56][57][2] introduced technical issues that may have increased response variability. These platform-related bugs, including video stimuli lagging and bugs, could have led to random responses in some instances, potentially introducing significant variances in our analysis. This is a notable concern that warrants further attention and resolution in future research endeavors.

An interesting evolution for this study could be to classify on the basis of *valence* and *arousal* the factors that compose actual clips of cinematic scenes taken from existing films and, on the basis of this classification, hypothesize their association with emotions.
These hypotheses could be investigated through a questionnaire similar to those developed for the present research, but I would suggest using a response's structure that can span the entire spectrum of emotions that are depicted in Plutchik's Wheel of Emotions (1980) [45], in order to allow participants to give only one vote for each stimulus, which may already contain the nuances of the emotions felt.
This type of analysis could certainly evaluate much more precisely the influences of the various factors that compose a real cinematic scene.

Contributions In an era where emotional engagement is central to successful communication, these experiments have generated insights that hold significant implications for professionals across diverse industries engaged in multimedia design, and can assist the creation of impactful and emotionally resonant multimedia products.
By unraveling the intricate interplay of stimuli, our research contributes to a deeper understanding of the art and science of Sound Design. This deeper understanding provides a valuable toolkit for these professionals, enabling them to harness the power of Sound more effectively in order to elicit specific emotional responses in their audiences.
As we conclude this study, it is clear that the field of Sound Design is poised

[2] https://www.psytoolkit.org/

to play an increasingly vital role in shaping the emotional landscapes of our multimedia creations.

Full of curiosity for the thriving future of Sound Design, we conclude this study, confident that Sound will continue to enrich the tapestry of human expression.

> The earth has music for those who listen.
>
> *William Shakespeare*

References

[1] Adrián Barahona-Ríos and Sandra Pauletto. "Synthesising Knocking Sound Effects Using Conditional Wave-GAN at Sound and Music Computing". In: Online, June 2020.

[2] Adrián Barahona-Ríos and Ulf Pauletto Sandra; Performed by Olausson. "Knocking Sound Effects With Emotional Intentions [Data set]". In: Zenodo, 2020. DOI: \url{https://doi.org/10.5281/zenodo.3668503}.

[3] Jennifer M. Barker. *The Tactile Eye: Touch and the Cinematic Experience*. University of California Press, 2009, p. 196.

[4] Jay Beck. *The Routledge Companion to Film History*. Routledge, 2010, pp. 78–90.

[5] Carlos R. Benítez-Barrera et al. "Cultural differences in auditory ecology". In: *JASA Express Letters* (2023). DOI: https://doi.org/10.1121/10.0020608.

[6] David Bordwell. "Camera movement and cinematic space". In: *Ciné-Tracts 1* (1977), pp. 19–25.

[7] Charles Bramesco. *Colours of Film: The Story of Cinema in 50 Palettes*. Frances Lincoln Publishers Ltd, 2023.

[8] Garrett Brown. "The Moving Camera, Part One". In: 2003. URL: https://api.semanticscholar.org/CorpusID:149375702.

[9] Mehmet Burak Yilmaz et al. "An embodiment of the cinematographer: emotional and perceptual responses to different camera movement techniques". In: *Frontiers in Neuroscience* 17 (2023). ISSN: 1662-453X. DOI: 10.3389/fnins.2023.1160843. URL: https://www.frontiersin.org/articles/10.3389/fnins.2023.1160843.

[10] William Cannon. “The James-Lange theory of emotions: A critical examination and an alternative theory”. In: *The American Journal of Psychology, 39* (1927), pp. 106–124.

[11] Paul Coates. *Cinema and Colour: The Saturated Image*. British Film Institute, 2010.

[12] Charles Darwin. *The expression of the emotions in man and animals*. John Murray, London, 1872.

[13] Paul Ekman. “Are there basic emotions?” In: *Psychological Review* 99(3) (1992), pp. 550–553.

[14] Paul Ekman. “Facial expression and emotion.” In: *American Psychologist* 48(4) (1993), pp. 384–392.

[15] Paul Ekman and Richard Davidson. *The nature of emotion: Fundamental questions*. Oxford University Press, 1994.

[16] Andrew J. Elliot, Mark D. Fairchild, and Anna Franklin. *Handbook of Color Psychology*. Cambridge University Press, Apr. 2016. DOI: https://doi.org/10.1017/CBO9781107337930.

[17] Andrew J. Elliot and Markus A. Maier. “Color psychology: effects of perceiving color on psychological functioning in humans”. In: *Annual Review of Psychology* (2014), pp. 95–120.

[18] Riccardo Falcinelli. *Cromorama, Come il colore ha cambiato il nostro sguardo*. Einaudi, 2017.

[19] Sigmund Freud. *The Unconscious*. Hogarth, London, 1915, 14: 159–204.

[20] Vittorio Gallese and Michele Guerra. “Embodying movies: embodied simulation and film studies”. In: *Embodying movies: embodied simulation and film studies*. 2012, pp. 183–201. URL: https://api.semanticscholar.org/CorpusID:12237211.

[21] Vittorio Gallese and Michele Guerra. *Lo Schermo Empatico. Cinema e neuroscienze*. Raffaello Cortina Editore, 2015, p. 315.

[22] Vittorio Gallese and Michele Guerra. *The Feeling of Motion : Camera Movements and Motor Cognition*. Mimesis Edizioni, 2014, pp. 103–112.

[23] Ian Goodfellow et al. "Generative Adversarial Nets". In: *Advances in Neural Information Processing Systems 27 (NIPS 2014)* 27 (2014). Ed. by Z. Ghahramani et al., pp. 2672–2680. URL: `https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf`.

[24] Thomas Görne. "The Emotional Impact of Sound: A Short Theory of Film Sound Design at klingt gut! 2017 International Symposium on Sound". In: Arts and Media Campus Hamburg, Germany, Jan. 2019. DOI: `10.29007/jk8h`.

[25] *Grand Budapest Hotel*. 2014.

[26] Michele Guerra. "Modes of Action at the Movies, or Re-Thinking Film Style From the Embodies Perspective". In: *Embodied cognition and cinema*. Ed. by Maarten Coëgnarts and Peter Kravanja. Leuven University Press, 2015.

[27] Iiro P. Jääskeläinen et al. "Movies and narratives as naturalistic stimuli in neuroimaging". In: *NeuroImage, Vol. 224* (2021).

[28] Iiro P. Jääskeläinen et al. "Neural Processing of Narratives: From Individual Processing to Viral Propagation, Vol. 14". In: *Frontiers in Human Neuroscience* (2020).

[29] Keith W. Jacobs and Frank E. Jr. Hustmyer. "Effects of four psychological primary colors on GSR, heart rate and respiration rate". In: *Perceptual and Motor Skills* 38 (1974), pp. 763–766. DOI: `https://psycnet.apa.org/doi/10.2466/pms.1974.38.3.763`.

[30] Williams James. *Mind, Vol.9, No. 34*. Oxford University Press, on behalf of the Mind Association, Apr., 1884, pp. 188–205.

[31] Patrik N. Juslin and Västfjäll Daniel. "Emotional responses to music: the need to consider underlying mechanisms". In: *Behavioral and Brain Sciences* (2008), pp. 575–621. DOI: `10.1017/S0140525X08005293`.

[32] Marianne Latinus and Belin Pascal. "Human voice perception." In: *Current Biology* 21(4) (2011), pp. 143–145.

[33] Daniel J. Levitin. *This Is Your Brain on Music: The Science of a Human Obsession*. New American Library, New York, 2006. ISBN: 9780452288522.

[34] Donald A. Mackenzie. "Colour Symbolism". In: *Folklore* 33 (1922), pp. 136–169. DOI: https://doi.org/10.1080/0015587X.1922.9719237.

[35] Germany Mannheim. "SurveyCircle". In: *Research website SurveyCircle. Published 2016.* 2023. URL: https://www.surveycircle.com.

[36] Stephen McAdams and Emmanuel Bigand. *Thinking in sound: The cognitive psychology of human audition*. Clarendon Press/Oxford University Press, 1993. DOI: https://doi.org/10.1093/acprof:oso/9780198522577.001.0001.

[37] J. McDonald et al. "Illuminating Music: Impact of Color Hue for Background Lighting on Emotional Arousal in Piano Performance Videos at International Symposium on Performance Science 2021". In: vol. 13. Mar. 2022. DOI: https://doi.org/10.3389/fpsyg.2022.828699.

[38] Mehdi Mirza and Simon Osindero. *Conditional Generative Adversarial Nets*. 2014. arXiv: 1411.1784 [cs.LG].

[39] Daniel Morgan. "Where are we?: camera movements and the problem of point of view". In: *New Review of Film and Television Studies, 14:2* (2016), pp. 222–248.

[40] The preservation of favoured races in the struggle for life. On the origin of species by means of natural selection or. *Darwin, Charles*. John Murray, London, 1859.

[41] Sandra Pauletto and Alessandro Iop. "Perception of Emotions in Multimodal Stimuli: the Case of Knocking on a Door at 18th Sound and Music Computing Conference, SMC". In: Online, July 2021.

[42] Cesare V. Pavise. "Crossmodal Correspondences: Standing Issues and Experimental Guidelines". In: *Multisensory Research* 29, 1-3 (2016), pp. 7–28. DOI: 10.1163/22134808-00002502.

[43] Geoffroy Peeters et al. "The Timbre Toolbox: Extracting audio descriptors from musical signals". In: *The Journal of the Acoustical Society of America* (Nov. 2011), pp. 2902–2916. DOI: 10.1121/1.3642604.

[44] Daniel Percheron and Marcia Butzel. "Sound in Cinema and its Relationship to Image and Diegesis". In: *Yale French Studies* 60 (1980), pp. 16–23. DOI: https://doi.org/10.2307/2930001.

[45] Robert Plutchick. “A psychoevolutionary theory of emotions”. In: *Science Information, 21(4-5)* (1982), pp. 529–553.

[46] Robert Plutchik. *Emotion: Theory, research, and experience: Vol. 1, Theories of emotion*. Academic Press, New York, 1980.

[47] Robert Plutchik and Henry Kellerman. *Theories of emotion. Vol. 1.* Academic Press, 2013.

[48] A. Rothstein. *Color Theory in Film: A Video Producers Guide*. https://www.ipr.edu/blogs/digital-video-and-media-production/color-theory-in-film-a-video-producers-guide/. Aug. 2020.

[49] A. Rothstein. *Why is Sound So Important in Film?* https://www.ipr.edu/blogs/film-and-sound-design-2/why-is-sound-so-important-in-film/. May 2021.

[50] Barry Salt. *Film style and technology: history and analysis*. Starword, London, 1992.

[51] *Schindler's List*. 1993.

[52] Jordan Schonig. “Seeing Aspects of the Moving Camera: On the Twofoldness of the Mobile Frame”. In: *Synoptique* (2017), pp. 57–78. URL: https://api.semanticscholar.org/CorpusID:149375702.

[53] Jordan Schonig. *The Shape of Motion: Cinema and the Aesthetics of Movement Cinema and the Aesthetics of Movement*. Oxford University Press, Mar. 2022. ISBN: 9780190093884. DOI: https://doi.org/10.1093/oso/9780190093884.001.0001. URL: https://doi.org/10.1093/oso/9780190093884.001.0001.

[54] Satyendra Singh. “Impact of color on marketing”. In: *Management Decision* 44, No. 6 (2006), pp. 783–789. DOI: https://doi.org/10.1108/00251740610673332.

[55] Vivian Sobchack. “Toward inhabited space: The semiotic structure of camera movement in the cinema”. In: *Semiotica, Vol. 41, no. 1-4* (1982), pp. 317–335.

[56] G. Stoet. “PsyToolkit - A software package for programming psychological experiments using Linux.” In: *Behavior Research Methods, 42(4)*. 2010, pp. 1096–1104.

[57] G. Stoet. "PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments." In: *Teaching of Psychology, 44(1)*. 2017, pp. 24–31.

[58] Patrick Susini, Olivier Houix, and Nicolas Misdariis. "Sound design: An applied, experimental framework to study the perception of everyday sounds". In: *The New Soundtrack* 4 (2014), pp. 103–121. DOI: 10.3366/sound.2014.0057.

[59] Siu-Lan Tan, Peter Pfordresher, and Rom Harré. *Psychology of music: From sound to significance*. Routledge, 2017.

[60] D. S. Y. Tham et al. "A systematic investigation of conceptual color associations". In: *Journal of Experimental Psychology: General* 149(7) (2020), pp. 1311–1332. DOI: https://psycnet.apa.org/doi/10.1037/xge0000703.

[61] *The French Dispatch of the Liberty, Kansas Evening Sun*. 2021.

[62] A. Turato et al. "Knocking on a yellow door: interactions among knocking sounds, colours, and emotions at Forum Acusticum 2023, 10th Convention of the European Acoustics Association". In: Torino, Italy, Sept. 2023.

[63] Asia Turato. *Emotions in everyday sounds. An experimental study on the influence of colour in emotional knocking sounds*. 2022.

[64] *Twin Peaks*. 1990-2014.

[65] Yon Visell et al. "Sound design and perception in walking interactions". In: *International Journal of Human-Computer Studies* 67, Iss. 11 (2009), pp. 947–954. DOI: https://doi.org/10.1016/j.ijhcs.2009.07.007.

[66] Slavko Vorkapich. "Cinematics: some principles underlying effective cinematography". In: *Hollywood Directors: 1914-1940* (1976).

[67] Xin Wang et al. "A Cross-Cultural Analysis of the Influence of Timbre on Affect Perception in Western Classical Music and Chinese Music Traditions". In: *Frontiers in Psychology* (2021). DOI: 10.3389/fpsyg.2021.732865.

[68] Liz Watkins. "Light, colour and sound in cinema." In: *Paragraph* 25(3) (2002), pp. 117–128.

[69] Angela B. Wright. "The Colour Affects System of Colour Psychology at AIC Quadrennial Congress". In: Sydney, Australia, Oct. 2009. DOI: 10.13140/2.1.4246.0489.

Acknowledgments

I would like to express my deepest gratitude to all the people who have accompanied me in these years. Your unwavering support and encouragement have been a beacon throughout my journey of academic and personal training.

First, I want to express my sincere appreciation to my thesis advisor, Professor Antonio Rodà. I am profoundly thankful for having introduced me to the melodic world of Sound Design, for his patience, and for guiding me during the last great adventure of my master's degree.

My heartfelt appreciation goes to my family. Mom and dad, your love, support, and sacrifices have been fundamental for my educational and personal pursuits. You have always been there for me, through the best times and through the most challenging times. I therefore thank you for your constant encouragement and for always pushing me to become the better version of myself.

To Ste', I want to express my deepest gratitude for always being there for me, from the beginning to the end of this journey, for believing in me, for your understanding, and for always being available and enthusiastic in helping me, no matter what. I thank you for making me feel happier, safer, stronger and more confident in myself. As I conclude this important chapter of my life, I can't wait to share with you all the successes, challenges, and the good and bad times that life has in store.

I am also grateful to Professor Badia for having failed me in the telecommunications exam during my Bachelor's Degree, pushing me to attend the summer courses in Bressanone where I was able to meet Noemi and Gianluca. I still didn't know that you two would become my greatest friends, sources of joy, laughter and comfort, and with whom I spent the most pleasant moments. May our friendship be a source of inspiration and support for each other as we embark on new paths in life.

I thank you, Ilaria, for sharing the most fun work evenings with me. Thank you for always being a great confidant and a great support, whenever I need it, but thank you also for all the laughs, experiences and funny stories we have exchanged over these years.

In parallel with my academic path, in the last years I have had the opportunity to follow an artistic path that took me to Verona, where I had the chance to meet two women who have been very important in my life in the last four years, Susanna and Roberta. Despite the vicissitudes we encountered on the road we traveled together, I am deeply grateful for having met you and being able to share with you passions, dreams, and many wonderful evenings.

To all the people I have met over the years, my journey would not have been the same without you.
One last profound thank you for being such a significant part of my life.

Martina