This is a repository copy of *Understanding Immersion in the Context of Films with Spatial Audio*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/190351/

Version: Accepted Version

Proceedings Paper:

Understanding Immersion in the Context of Films with Spatial Audio

Joseph Williams*
School of Physics, Engineering and Technology
University of York
York, UK
jw1858@york.ac.uk

Sven Shepstone
Bang & Olufsen a/s
Struer, Denmark
ssh@bang-olufsen.dk

Damian Murphy
School of Physics, Engineering and Technology
University of York
York, UK
damian.murphy@york.ac.uk

August 15, 2022

ABSTRACT

As spatial audio technologies become more prevalent, more individuals are being exposed to film content via a traditional visual monitor, delivered in a way which makes them feel as though the film's soundtrack is emanating from all around them. In describing such experiences, the term immersion has become a buzzword both in academia and elsewhere. Through a review of relevant literature and qualitative analysis of descriptions of film immersion from a survey, an understanding of this ill-defined term is developed and the influence of spatial audio on film immersion is discussed. This work leads to a series of suggestions for future studies in this domain, notably highlighting the need for the assessment of the perception of *auditory immersion* and support for the use of physiological methods for assessing a subject's *absorption* in the film world.

Keywords Conceptual Model, Auditory Immersion, Absorption

1 Introduction

Immersion is a hard to navigate concept, with researchers within different domains, such as virtual reality and gaming research, putting forward many different interpretations for the same term. Within these interpretations, a common taxonomy exists in which the concept is considered some combination of a property of a system, the perception of being spatially surrounded by digitally mediated content (*spatial immersion*), and/or a deep sense of, cognitive and emotional, connection and involvement (*absorption*). Furthermore, the terms immersion and immersive have become popular buzzwords in marketing, notably in the context of immersive technologies such as virtual reality and spatial audio systems, adding further ambiguity to the meaning.

---

*Corresponding Author. Preprint version. Original version to be found on: http://www.aes.org/e-lib/browse.cfm?elib=21878.

In recent years purveyors of non-participatory audiovisual media have started to implement immersive audio technologies such as Dolby Atmos[1]: in cinemas through multichannel systems, in living rooms through soundbars, and elsewhere through head-tracked binaural renderings solutions such as those provided by Apple [2]. This motivates research exploring how these technologies influence the film viewing experience, apply this concept of immersion. However, before this task is approached it is important to answer a few prerequisite questions. Firstly, what existing research has been done in this area and how does this influence what research should be done in the future? and secondly, how may immersion be defined and conceptualized unambiguously in this context - where spatial audio is paired with a traditional visual display? To answer these questions, this paper presents a review of relevant literature and a survey study in which the understanding of the term immersion in the context of film is assessed for 36 participants through qualitative content analysis. Following this, a discussion is presented leading to a set of suggestions for how immersion may be interpreted in ongoing research.

1.1 Defining Immersion

Immersion has been used as a term for describing a figurative sense of involvement, or absorption of, some task, experience, condition or interest for centuries, and since the 1960s, for conceptualising a pedagogy for learning foreign languages [3, 4]. In the 1990s the term also became associated with new immersive technologies, and subjective responses to experiences mediated by these technologies [5, 6, 7]. Analysis of the terms use over time even suggests that such technologies may have caused the reintroduction of this now common term immersive into circulation in the English language [8, 9, 10].

Beyond the understanding, that immersion relates to a metaphorical concept, and that immersion can be considered a property of a medium, and/or a subjective experience in response to specific characteristics of medium or media, little consensus lies in this domain [5, 6, 7]. In a literal sense, immersion is defined as the plunging of something into water, or some other liquid [4]. In Hamlet on The Holodeck, published in 1997, this metaphorically informed interpretation of immersion is presented in the following extract:

> "Immersion is a metaphorical term derived from the physical experience of being submerged in water. We seek the same feeling from a psychologically immersive experience that we do from a plunge in the ocean or swimming pool: the sensation of being surrounded by a completely other reality, as different as water is from air, that takes over all of our attention, our whole perceptual apparatus" [11].

1.2 Immersion as a System Property

In work describing experiences in virtual, computer-generated environments delivered through head-mounted displays (HMD), immersion is commonly used to describe a property of a system capable of submerging a subject in a simulated world. In McKenzie's early (1992) and exploratory work, the term immersion is interpreted as a defining feature of virtual reality (VR), encapsulating the quality of experiencing a computer performance from the inside-out, rather than the traditional outside-in, perspective [12]. A more commonly cited definition for immersion as a system property stems from various publications within an extensive area of investigation known as presence research [13, 14, 15, 16, 17]. Presence is commonly defined as the sense of feeling "really there", in a mediated world, or alternatively the sense of being in one location, whilst truly situated in another [18]. Within presence research, this phenomenon is explored during experiences that occur when using various mediating technologies including both: new immersive technologies such as virtual reality and spatial audio; and visual monitors of varying sizes viewed at various angles [6, 4]. In 2003, Slater expressed the following proposition for immersion as a system property:

> "Let's reserve the term "immersion" to stand simply for what the technology delivers from an objective point of view. The more the system delivers displays (in all sensory modalities) and tracking that preserves fidelity in relation to their equivalent real-world sensory modalities, the more that is "immersive."" [14]

In interpreting this extract, one may reflect that immersion is a binary, with something immersive delivering immersion into the digitally mediated world. However, elsewhere in this publication, Slater describes a "level of immersion", which implies immersion is graded. Arguably, this is a confusing point of ambiguity - is immersion graded, or are things immersive or not immersive? In Slater & Wilbur's earlier, 1997 work, Immersion is described as "objective and quantifiable" and therefore graded. Furthermore, in this same publication immersion is also evaluated to be made up of various measurable components [13]. Specifically, these components are considered to be the characteristics of a "computer display" as; inclusive, the extent to which physical reality is shut out; extensive, the range of sensory modalities accommodated; surrounding, the extent to which this virtual reality is panoramic rather than limited to a

narrow field; and vivid, the resolution and fidelity of each sensory modality. Slater & Wilbur, still within the domain of presence research, subsequently also describe immersion as also being graded in regards to the level of matching between sensory modalities, as in the extent to which body movement should be matched by appropriate and low latency changes in the audiovisual display [13].

In Cuadrado et. al's 2018 study, immersion is used to describe the system characteristic which differentiates audiovisual narrative content with stereo and 3D surround sound audio [19]. Within this study, both of these mixes are delivered to participants through headphones, with the latter being created from a 5.1 surround sound mix processed binaurally using specialised software. The influence of the spatial audio mix on learning outcomes and emotional response was analysed for a sample of 330 students. The results of the study show an improvement in listening comprehension and suggest a more intense emotional impact for the 3D surround sound group compared to the stereo group. Inconsistently, within the abstract, the use of voice recordings and electrodermal response is implied for emotion analysis, however, within the body of the paper neither of these modalities are mentioned.

1.3 Immersion as a Subjective Experience

Immersion is also defined as a subjective experience in many cases, despite contradicting work cited in the preceding section. This category of conceptualisations is broadly made up of two aspects: immersion as a perceptual response to spatial, generally surrounding, and/or multi-sensory, qualities of a digitally mediated experience (referred to as *spatial immersion* from this point on); and immersion as a deep state of positive mental absorption in some challenge or narrative within the experience (referred to as *absorption* from this point on). Sometimes, these aspects are presented together in models with authors using terminology (e.g. presence & immersion) in an inconsistent way [5, 7, 20, 21].

Within interpretations of immersion as a subjective experience, the source of this subjectivity is often questioned. Notably, this has led to a concept known as immersive tendencies (IT's), introduced by Witmer & Singer [18], which relates to the idea that some individuals are more likely to experience presence or immersion than others. Moreover, researchers have, expansively, conceived of the idea of influencing factors (IFs) for immersion [5, 21]. Within the construct of IFs, the medium, media and context are also presented as having notable effects on the qualities of a subjects' sense of immersion (a.k.a. their immersive experience).

1.3.1 Immersion as Spatial Immersion

The first identified grouping of interpretations of immersion as a subjective experience, as *spatial immersion*, has been labelled in many alternative ways, notably as sensory, or perceptual, immersion [5, 7, 20].
Arguably, this concept has a clear perceptual pairing with Slater & Wilbur's objective properties of system immersion (inclusive, extensive, surrounding, vivid, matching) [13]. Notably, Lombard & Ditton [22] synthesise an understanding of immersion categorisable as *spatial immersion* from previous literature, defining the term perceptual immersion as being: a user's impression of having their perceptual system submerged [23], objectively measured by counting the number of senses that are provided with input, and by the extent to which the physical environment is shut out [24]. Building upon this description, this heavily cited paper describes, "Not only immersive virtual reality systems but also simulation rides, IMAX theaters, and even standard movie theaters can be said to immerse the senses of media users."[22]. In this work, as well as almost all other cited work in this paragraph, presence is seen as the common outcome of immersion.

Within the context of audio, researchers have explored the perception of spatial aspects of sound, developing methods and tools for self-assessment of what has recently been described as *auditory immersion* [25], or more commonly as quality of spatial audio [26, 27, 28]. Rees-Jones et al. [29], in the context of exploring the influence of audio rendering on the quality of player experience, assesses the quality of spatial audio rendering by asking subjects to assess their sense of localisation accuracy (how easy it is to identify the direction in which a sound source is originating), distance accuracy (perceived distance of sound sources), sense of depth (the perceived front-back definition of the sound scene), sense of width (perceived left-right definition of the sound scene), and envelopment (sense of being surrounded by the sound presented in the scene). Eaton & Lee [25] use a similar set of perceptual factors to define *auditory immersion* in the context of VR, validating the inclusion of each item in their construct by confirming the association of each with the term *auditory immersion* by surveying audio professionals.

In the context of film research, Rooney et al. leverage the conceptualisation of immersion as being "defined in terms of the sensory information the technology provides to the user" [31]. This interpretation of immersion as purely the perception of the technical characteristics of an experience references a response paper to Witmer Singers' [18] previously cited work exploring presence, by Slater [17]. With the aim of exploring the difference between 2D and 3D film viewing experiences, Rooney et al. present a study comparing psychological and physiological responses to short film clips, between 13 and 101 seconds long. Between the two groups, comprising a total of 39 subjects, no

significant difference was found between conditions in terms of self-reported feelings of emotion, skin conductance or skin temperature. However, a significant difference was found in heart rate between conditions with the 3D condition inducing a larger increase in heart rate during emotive periods in the film.

1.3.2 Immersion as Absorption

The second identified grouping of interpretations of immersion as a subjective experience, as *absorption*, is often presented as relating to a sense of emotional, and/or cognitive involvement in some characteristics of the media (e.g narrative, or challenge in interactive experiences) [5, 7, 20, 21]. The term *absorption* has been chosen due to its recurrence within associated literature.

The concept of a shift in attention, where a subjects' focus moves away from their surroundings and towards the mediated experience, is often included within many interpretations of immersion. This shift is often considered to be facilitated through a subjects' motivation, or by technologies which spatially surround - isolating the senses and shutting out the real world. Hence, this provides a link between the concepts labelled here as *spatial immersion* and *absorption*. Specifically, this bridging factor is based on the idea that immersive technologies force our attention towards an experience and away from our real environment due to the masking of one's true environment and submerging of perceptual apparatus [7, 22]. In this section, this shift in attention is considered *absorption*, however, as described above it may arguably also be associated with *spatial immersion*. In the context of audiovisual experiences, Agrawal describes immersion purely relating to *absorption* [21, 30].

> "Immersion, also known as deep mental involvement, can be described as being mentally lost (absorbed) in the experience. Immersion is encountered when the experience is involving and absorbs you mentally by capturing your attention. For example, immersion may be experienced when reading a book, playing video games, watching a movie, etc." [30]

As a method for subjective assessment of immersion, Agrawal presents the approach of providing subjects with this definition and then subsequently asking them to assess their immersion using a graphic line scale, anchored at each end with the word "Not very immersed" and "Immersed", after viewing various audiovisual media clips from films and documentaries. Using this method, a study showed that this form of immersion assessment can produce significant differences in average scores between short 4 - 12 minute clips delivered with a high technical specification (e.g. 4K with Dolby Atmos audio) [30].

In a similar domain, focusing on immersion in film, Bjorner et. al. [31] develop a model for *absorption*, labelled as narrative immersion, which builds upon Brown and Cairns' understanding of the existence of three distinct depths of immersion in video games [32]. Between these three levels, labelled as: engagement, engrossment, and total immersion, the subject incrementally loses perception of their true surroundings and becomes more focused on the film. During total immersion, they feel as if they are truly located within the film and unaware of time passing due to their intense and enjoyable sense of focus. As part of this study, a 28-question, 11-factor self-assessment questionnaire is developed. Subsequently, the questionnaire is deployed in a study which implemented distractions in an attempt to decrease viewers' involvement whilst viewing a 15 minute long 3D animate film. The finding of the study highlighted the large interpersonal differences in how the subjects reacted to the distraction events. The authors also state that some subjects are able to reach full immersion even when external distractions are present, however, this finding should arguably be interpreted cautiously as total immersion is an ambiguous concept which has been measured in this study using a novel questionnaire.

The Film IEQ study [33] represents another conceptualisation of immersion adapted from gaming research. This study aimed to adapt Jenett et.al's Immersive Experience Questionnaire (IEQ) [34], a validated and widely used self-assessment for gaming experiences, into the domain of non-participatory media viewing. In the original IEQ study, items were adapted from Brown & Cairn's grounded theory study [32], as well as a questionnaire for Cognitive Absorption (CA) - defined as "a state of deep involvement with software" [35]. The approach taken by Rigby et al. in developing the Film IEQ was to adapt each question in the IEQ to the context of Film and then deploy the 31-item questionnaire in an exploratory factor analysis study asking the 414 participants to respond based on the last video they watched [33]. Subsequently, the questionnaire was redeployed in a validation study, with the removal of seven items which did not load well with the emergent factors found in the factor analysis process.

The validation study for the IEQ was a repeated-measures study, deploying the well-supported hypothesis that bigger screens lead to a greater sense of presence [36, 37] and attention [38]. Specifically, the hypothesis stated that the same content viewed on a larger screen would deliver a greater score within the relevant factors of the developed questionnaire, as well as a greater overall score. 19 subjects were asked to choose a film on Netflix which they had not seen before, and then watch three 10 minutes clips sequentially on three devices: a phone (4.5"), laptop (13"), and monitor (30") with sound being delivered through headphones for each condition. After each clip, subjects were asked

to fill out the Film IEQ questionnaire. In analysing the results of the study, the increase in overall immersion score, calculated by adding up the response to each item, was deemed statistically significant between the monitor and phone, laptop and phone, but not monitor and laptop. However, the only statistically significant increase in subfactor between devices was found in the comprehension factor, when comparing the monitor to the phone condition.

1.3.3 Immersion as Spatial Immersion & Absorption

In many conceptualisations of immersion, and presence, some interpretation of both *spatial immersion* and *immersion as absorption* are included. Furthermore, in some studies both of these aspects are assessed as subjective measures. Within the domain of presence research, a questionnaire which allows for this, called the ITC Sense of Presence Inventory (ITC-SOPI) exists [39]. Although the outcome of this self-assessment questionnaire is labelled as presence, there are clear similarities between the factors of this questionnaire and those theorised to represent the subjective experience of immersion, notably engagement (relating to *absorption*) and sense of physical space (relating to *spatial immersion*). This questionnaire has been validated and deployed in numerous studies, and aims to allow for the assessment of both interactive and non-interactive experiences. Within the associated 2001 publication outlining the development of this questionnaire, 63 items are initially sourced from uncited "theoretical and empirical papers" [39]. Subsequently, an exploratory factor analysis of 604 respondents was conducted, with subjects self-assessing various digitally mediated experiences. The participants evaluated non-participatory film viewing experiences of various types and lengths (e.g fiction film, non-fiction TV, sports highlights), with a further 49 subjects completing the questionnaire after playing a rudimentary racing game. After analysis, 44 items were retained as they loaded well with the emergent factors.

In a subsequent study, Freeman & Lessiter conducted a study exploring the influence of different surround sound mixes on the perception of an audiovisual car driving simulation, using the questionnaire [40]. Alongside the ITC-SOPI, another unshared 18-item questionnaire was also deployed for assessing items relating to the perceived quality of audio and visual content, based on Gabrielsson & Lindstrom's dimensions of perceived sound quality [41]. Specifically, a designed 5.1 mix was downmixed to stereo, creating two mixes, which were presented with and without a subwoofer delivering the low-frequency components of the mix. Due to the increase in sound pressure level introduced by the addition of the low-frequency loudspeaker, a control stereo mix condition was added, which delivered the same sound pressure level as the 5.1 and 2.1 mixes. Hence, in combination, there were 5 different audio delivery conditions: 2.0 (70dB SPL), 2.0 control (83/84dB SPL), 5.0 (70dB SPL), 5.1 (83/84dB SPL). The study, which included 30 participants, deployed a repeated measures study in which the subjects moved between the different surround sound mixes, filling out the two questionnaires between each trial. It is unclear how often the change in these conditions was made, although it is noted that trial orders were counterbalanced to some extent. Through analysing the results of the study, the authors conclude that there may be no perceived advantage when comparing the 5 channel mixes over 2 channel mixes, and that "the majority of measures – presence and audio quality evaluations – were rated more highly when the presentation included bass"[40]. However, the generalisability of these findings to film should be interpreted cautiously, notably due to the novelty of the driving simulation video used as the sole stimuli.

The final item of literature explored in this section is Zhang's study exploring which aspect of immersion is more immersive: spatial/embodied immersion or emotional/empathetic immersion [20], reflecting a similar taxonomy of immersion used to guide the review of literature in this paper (*spatial immersion* and *absorption*). Definitions are synthesised, and the decision to split the concept into two is validated through citation of 12 existing "attempts to classify immersion" [42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53]. As part of this publication, Zhang also creates a questionnaire by combining items from various validated questionnaires. A mixed-method approach is used with the intent of exploring both the influences of media and medium, between a spatial video (a realistic visualisation of the interior of a building) and an emotional video (an animated love story with elements of humour and surprise) delivered via a tablet with built-in loudspeakers or a TV with basic computer loudspeakers. A 45 participant sample was split into two groups both receiving two 7 minute media stimuli. However, in the analysis section, quantitative analysis is only deployed for assessing the influence of media, and not the influence of medium, on responses to the questionnaire. Zhang cites that the lack of clear difference between responses between devices as validation for this decision, and speculates that this may be due to the fact subjects were allowed to view the tablet from up close. The outcome of this study is described as follows:

> "...emotional immersion is significantly more immersive than spatial immersion in terms of sense of "being there", time perception, realism, sense of engagement, emotional aspects, sensory cues, etc. Spatial immersion is almost as immersive as emotional immersion in terms of attention and image motion, and spatial immersion is more immersive than emotional immersion in terms of spatial dis-orientation"[20].

These findings are based on the specific question items which showed statistically significant differences in mean scores between the two media stimuli, hence further study should be made to support the inferred generalisability of results.

2   SURVEY

To gain an alternative perspective on how immersion is interpreted in the context of Film, a Qualtrics [54] form was distributed via email throughout the University of York, and via social media, with the aim of reaching a broad range of demographics. Specifically, an online survey was circulated consisting of: an information form, confirmation of consent based on a described use of data, a series of demographic questions, and the following two items:

1.) Are you familiar with the concept of immersion, as in being immersed in media not in a liquid?
2.) How would you describe immersion, in the context of Film?

The second clause of the first question was added to improve clarity and to avoid false positive responses from individuals who believed the question referred to the literal form of immersion. Participants were recruited, data gathered, and results analysed in line with the University of York Ethical approval processes (Williams291121). In total, the study was completed in full by 36 participants. The participants in the study were all over 18 and prominently under the age of 40 (75%), with a high proportion of males (64%). Almost all participants spoke English as their native language (89%). 31 (86%) of the subjects were currently in higher education or were university educated.

2.1   Analysis

The immersion description data collected consisted of categorical, yes/no/unsure, answers to the question: "Are you familiar with the concept of immersion, as in being immersed in media not in a liquid?"; and a short written response to the question, "How would you describe immersion, in the context of film?". Out of these participants, 27 of the 36 (75%) respondents indicated they were familiar with the concept of being immersed in media, with 3 (8%) participants stating that they were unsure of the concept, and the remaining 6 (16%) stated they did not understand the concept. The decision was made to include all subjects' responses in the analysis, due to consistency between subjects who assessed their familiarity as unsure or unfamiliar, compared to those who stated they did understand the concept.

To analyse the responses, the text data was first read through in detail and split into meaning units: near-verbatim extracts from the text separated based on the points they communicated, commonly encapsulated within a sentence or phrase. Once this preparation was complete, analysis was made by the first author of this study along with two fellow researchers. Initially, these researchers were asked to familiarise themselves with the data. After this task was completed, a group session was held to inductively generate codes relating to the meaning units. This was facilitated using Trello [55], an online interactive list-making application, with lists being labelled with codes and associated meaning units being placed within all relevant lists.

The researchers aimed to initially create a broad range of codes, often based on common uses of specific words (e.g involved, engaged), or concepts for which descriptions were agreed upon (e.g Transition). After this process was completed, the codes were categorised creating a broader construct of abstraction. It should be noted that throughout the process the original text was referred back to so as to ensure the codes were applied correctly, in the context of the original response, not just the meaning unit. This method was influenced by both Elo [56] & Bengtsson's [57] suggestions for familiarisation with data, coding, creating categories then abstracting the categories into some construct. In Table 1, examples of the coding process are presented. The taxonomy of categories and codes developed, descriptions for each and values representing the number of respondents who were deemed to communicate each idea are presented as follows paragraphs and visualised in Figure 1.

2.1.1   Technology

This code was not placed in a category and contains responses associated with specific technologies, notably virtual reality and augmented reality. 3 subjects (8%) were deemed to communicate this interpretation by the investigators.

2.1.2   Emotion

This code was not placed in a category and contains responses associated with a sense of emotional connection to the narrative or characters. 4 subjects (11%) were deemed to communicate this interpretation by the investigators.

Text Response	Meaning Units	Codes	Categories
"Being completely captivated and taken into the world of the film"	Being completely captivated	Completely	Intensity
		Positive Attention	Attention
	taken into the world of the film	Transition	Placement
"The user engages with the story mentally as if he/she/they would enter that reality and leave the physical realty behind them"	The user engages with the story mentally	Positive Attention	Attention
	enter that reality	Transition	
			Placement
	leave the physical reality behind	Dissociation	
"the stronger the immersion, the more intense these emotions become"	The stronger the immersion	Scaled	Intensity
	Intense emotions	Emotion	Emotion

Table 1: Table showing examples of responses to the question: "How would you describe immersion, in the context of film?", split into meaning units, coded, and then placed into categories.
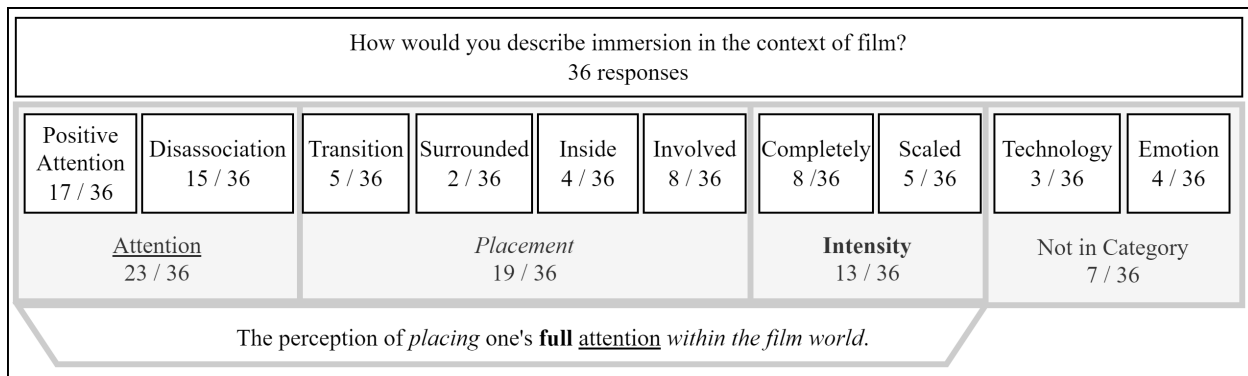

Figure 1: Visualisation of the qualitative content analysis of survey responses. The numbers (e.g. 15/36) denote how many individual participants had a meaning unit from their response placed within each corresponding code or category. The statement in the trapezium is a definition for film immersion synthesised from the three connected category names.

2.1.3 Intensity

This category contains two codes: *Completely* & *Scaled*. These represent responses that infer that immersion is an intense, possibly binary, psychological phenomena, or immersion is scaled and varies in intensity, respectively. 8 subjects (22%) were deemed to consider immersion 'complete', whereas 5 other subjects (14%) were deemed to infer that immersion was 'scaled'. These codes are independent, and hence there was a total of 13 (36%) subjects who were deemed to infer *Intensity* of immersion by the investigators.

2.1.4 Attention

This category contains two codes: *Positive Attention* and *Dissociation*. The first of these codes is associated with responses which used words such as attentive, engaged, engrossed, captivated or focused to describe a positive feeling of attention towards the content. It should be noted that the term *positive* here is used to mean direction of attention towards the media, as well as some generalised association between the terms used and a desirable experience. *Dissociation* relates to meaning units associated with a lack of attention toward one's surroundings and real-world concerns. The *Positive Attention* code was deemed by the investigators to be communicated in 17 (47%) subjects' responses, and *Dissociation* in 15 (42%). In total 23 (64%) of individuals were deemed to conceptualise immersion in a way relating to the *Attention* category.

2.1.5 Placement

This category contains four codes: *Involved*, *Surrounded*, *Inside*, & *Transition*. *Involved* relates to responses which use the term involved or close synonyms relating to feeling part of the film, *Surrounded* relates to being surrounded by the media in your environment, *Inside* relates to feeling as though placed within the media, and *Transition* relates to responses which reflect the idea of being transported or somehow moved into the film world. These were categorised together as they relate to an idea of *Placement*: either being placed within the media, or the media being placed around you. The *Involved* code was associated with 8 responses (22%), *Surrounded* was associated with 2 responses (6%), *Inside* with 4 responses (11%), and *Transition* with 5 responses (14%). In total 19 (53%) of the individuals were deemed to conceptualise immersion in a way relating to *Placement*. Putting the *Involvement* code within this category is based on the interpretation that being 'placed within' is a similar descriptor to 'being involved', within the context of films due to their low agency nature.

3 DISCUSSION

The immersion, in the context of film, descriptions from the sample appear to reflect an interpretation of immersion as *the perception of placing one's full attention within the film world*. This definition is synthesised by combining the labels for the three main categories of description found through the qualitative content analysis: *Placement*, *Attention*, and *Intensity* as shown in Figure 1. Survey analysis also shows that immersion in film is not commonly associated with any specific technologies. Notably, there was mention of cinemas in the survey responses, although VR was mentioned. Although not included in the analysis, it should also be noted that only 2 (6%) respondents used the terms *audio* and *visual*, and sound and vision, respectively. Similarly, this result is deemed to disassociate film immersion from the technical qualities of a system as defined in the literature explored in Section 1.1.

The data also shows emotional or empathetic response is not commonly considered to be a defining aspect of immersion in film within the sample, rather immersion appears to be more commonly interpreted as an enjoyable sense of attention - described as captivation, engrossment or engagement. It may therefore be conceived that subjects feel that immersion is the process of being, or becoming, involved in the film rather than the specific emotional outcomes of experiencing its content. Finally, it may also be inferred that immersion is considered an ambiguous and unclear concept by the sample investigated: 25% of subjects stated that they were 'unfamiliar or unsure' of the concept.

Within the subjects' responses, it appears that the sense of *Placement*: being involved in, surrounded by, within, or transition into, the film world - is part literal, and part figurative. Considering this in the context of the literature, the literal/perceptual aspect of *Placement* may be considered as *spatial immersion* (e.g. I felt surrounded by audio and video), and the more figurative deployment of *Placement* (e.g. I felt transported into the world of the film) as well as the meaning represented by the *Attention* category, may be interpreted as *absorption*. This interpretation of the survey analysis using the same approach as in the literature supports a theoretical framework of immersion as *spatial immersion* & *absorption*.

Within research, the relationship between *spatial immersion* and *absorption* has been explored in a visual sense, with many studies finding that spatial qualities of a screen, notably size, influences aspects of the subjective experience relating to *absorption* [36, 37, 38, 6]. However, in the context of audiovisual content with varying spatial audio qualities, there is little evidence to support that a similar connection exists [6, 58, 19, 40]. Hence, in future work exploring immersion, in the context of films with spatial audio, the suggestion is made to ensure that *spatial immersion* and *immersion as absorption* are clearly distinguished. Considering *spatial immersion*, interpreted in the context of film by the author of this paper as *the sense of being placed within the film through the delivery of perceptually congruent sensory information and masking of external audiovisual information*, a taxonomy of two sub-aspects is conceptualised: *visual immersion* and *auditory immersion*. This broadly relates to the perception of being placed inside the film world through the video and audio content respectively. In section 2.2.1 the term *auditory immersion* [25] is considered effectively synonymous with some interpretations of the more established concept of quality of spatial audio (e.g [29]). To synthesise a combined unambiguous definition contextualised to film, the factors of [25] and [29] are combined with interpretations of film immersion relating to placement within the film drawn from the survey, arriving at the following interpretation: *auditory immersion is the perception of placement within a 3D auditory environment made up of localisable sound sources consistent with the visual film media*. To subjectively assess *auditory immersion*, a process suggested to be necessary for quantifying the difference between spatial deliveries of film soundtracks (e.g stereo, 5.1, 7.1.4, binaural), [25, 29, 27, 28, 59] are all considered viable solutions. Based on this definition for *auditory immersion*, an interpretation of *visual immersion* is presented, tentatively as *the perception of placement within a visual environment made up of imagery consistent with the film's soundtrack*. These two sub-aspects and the two higher level aspects, which make up the model for film immersion, are shown in Figure 2.
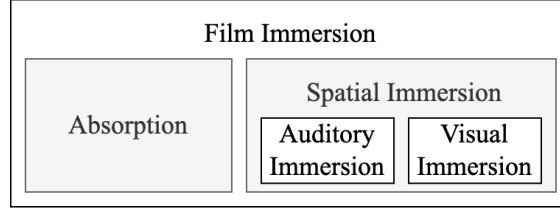
Figure 2: The model for film immersion developed in this paper, made up of two aspects (*absorption & spatial immersion*) and two sub-aspects of spatial immersion (*auditory immersion & visual immersion*).

In this paper the term *absorption* is used to encapsulate a common interpretation of immersion discussed by various authors investigating experiences of non-participatory audiovisual media [20, 21, 31, 30, 33]. Through analysis of literature, combined with analysis of the survey results, we arrive at an interpretation that *absorption* in film may be described as: *an enjoyed sense of attention towards the film and shift of attention away from one's true environment*. Furthermore, we consider that *absorption* may be measurable as a reaction contextualised to the content, alongside or instead of the assessment of a subjects' sense of engagement. For example, it may be measured as the sense of excitement during an action film, or presence in a driving simulation [40]. To measure absorption, using a self assessment questionnaire, the use of the Film IEQ [33] or items associated with the engagement factor of the ITC-SOPI [39] is suggested.

The use of physiological monitoring is also supported for assessing *absorption*, with the concept of producing a biometric storyboard of continuous data, reflecting the changing themes in the film [60]. For example, analysing if a subject's heart rate is accelerated more during a tense section of the film, or if a group of subjects repeatedly looked away from the visual monitor during a slow section. Quantification of these effects may allow for insightful comparison of *absorption*. To allow for these biometric storyboards, or alternative *physiological journeys*, to be comparable between subjects, we consider easily comprehensible short films, or film clips, with clear emotive events to be most appropriate as complex stimuli may have multiple interpretations and hence multiple viable responses to *absorption* in the film narrative.

4   CONCLUSION

In this paper, a review of literature and a survey of 36 subjects leads to the development of understanding for immersion in the context of films with spatial audio. Specifically, the survey analysis revealed three categories of immersion description, which can be combined into a definition for film immersion. Moreover, two of the categories are deemed to closely relate to the two fundamental aspects of immersion interpretation found in the literature *spatial immersion & absorption*. This leads to the development theoretical model, further framing immersion in film as these two interrelated aspects, with definitions and suggestions for assessment being provided for both. This is considered a contribution as it helps bring clarity to a concept which is often regarded as difficult to understand.

To test this model and progress this field of research, the suggestion is made to explore the influence of the auditory aspect of *spatial immersion* (*auditory immersion*) on *absorption*, through subjective and physiological assessment of a film viewing experiences delivered through varying audio technologies. The results of such a study should allow for a better understanding of how our perception of spatial audio can influence or cognition of media content. Hence, moving towards a better understanding of the effect of new immersive audio technologies.

5   Acknowledgements

References

[1] "Dolby atmos - official site." https://www.dolby.com/technologies/dolby-atmos/, 2022. Accessed: 2022-4-28.

[2] "About spatial audio with dolby atmos in apple music." https://support.apple.com/en-gb/HT212182, Mar. 2022. Accessed: 2022-7-1.

[3] R. Lyster, "Immersion pedagogy and implications for language teaching," *Multilingual Matters*, pp. 64–95, 1998.

[4] "immersion, n." https://www.oed.com/view/Entry/91885?redirectedFrom=immersion, 2022. Accessed: 2022-3-14.

[5] A. Perkis, C. Timmerer, and many more, "QUALINET white paper on definitions of immersive media experience (IMEx)," June 2020.

[6] J. J. Cummings and J. N. Bailenson, "How immersive is enough? a Meta-Analysis of the effect of immersive technology on user presence," *Media Psychol.*, vol. 19, pp. 272–309, Apr. 2016.

[7] Nilsson, Nordahl, and Serafin, "Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence," *Hum. Technol. Interdiscip. J. Hum. ICT Environ.*, 2016.

[8] Google, "Google books NGrams viewer: Immersion & immersive." https://books.google.com/ngrams/graph?content=immersion%2C+immersive&year_start=1950&year_end=2019&corpus=26&smoothing=3&direct_url=t1%3B%2Cimmersion%3B%2Cc0%3B.t1%3B%2Cimmersive%3B%2Cc0, Mar. 2022. Accessed: 2022-4-28.

[9] "immersive, adj. : Oxford english dictionary." https://www.oed.com/view/Entry/91887?redirectedFrom=immersive, 2022. Accessed: 2022-3-10.

[10] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *J. Commun.*, vol. 42, pp. 73–93, Dec. 1992.

[11] L. Randolph, "Hamlet on the holodeck: The future of narrative in cyberspace," 1998.

[12] J. McKenzie, "Virtual reality: Performance, immersion, and the thaw," *TDR (1988-)*, vol. 38, no. 4, pp. 83–106, 1994.

[13] M. Slater and S. Wilbur, "A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 6, pp. 603–616, 1997.

[14] M. Slater, "A note on presence terminology," *Presence connect*, vol. 3, no. 3, pp. 1–5, 2003.

[15] M. Slater, M. Usoh, and A. Steed, "Depth of presence in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 3, no. 2, pp. 130–144, 1994.

[16] M. Slater, "Immersion and the illusion of presence in virtual reality," *Br. J. Psychol.*, vol. 109, pp. 431–433, Aug. 2018.

[17] M. Slater, "Measuring presence: A response to the witmer and singer presence questionnaire," *Presence: Teleoperators and Virtual Environments*, vol. 8, pp. 560–565, Oct. 1999.

[18] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[19] F. Cuadrado, I. López-Cobo, T. Mateos, and B. Valverde, "Film sound," in *Audio Engineering Society Convention 144*, 2018.

[20] C. Zhang, A. Perkis, and S. Arndt, "Spatial immersion versus emotional immersion, which is more immersive?," in *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, May 2017.

[21] S. Agrawal, A. Simon, S. Bech, K. Bæntsen, and S. Forchhammer, "Defining immersion: Literature review and implications for research on audiovisual experiences," *JAES*, vol. 68, pp. 404–417, July 2020.

[22] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *J. Comput. Mediat. Commun.*, vol. 3, no. 2, p. JCMC321, 1997.

[23] F. Biocca and B. Delaney, "Immersive virtual reality technology," *Communication in the age of virtual reality*, vol. 15, no. 32, pp. 10–5555, 1995.

[24] T. Kim, "Effects of presence on memory and persuasion," *University of North Carolina, Chapel Hill, NC*, 1996.

[25] C. Eaton, *Quantifying Factors of Auditory Immersion for Virtual Reality*. PhD thesis, University of Huddersfield, Jan. 2020.

[26] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkman, and S. Weinzierl, "A spatial audio quality inventory (SAQI)," *Acta Acustica united with Acustica*, vol. 100, no. 5, pp. 984–994, 2014.

[27] M. Narbutt, J. Skoglund, A. Allen, M. Chinen, D. Barry, and A. Hines, "AMBIQUAL: Towards a quality metric for headphone rendered compressed ambisonic spatial audio," *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, vol. 10, p. 3188, May 2020.

[28] M. Narbutt, A. Allen, J. Skoglund, M. Chinen, and A. Hines, "AMBIQUAL - a full reference objective quality metric for ambisonic spatial audio," in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, May 2018.

[29] J. Rees-Jones and D. T. Murphy, "The impact of multichannel game audio on the quality and enjoyment of player experience," in *Emotion in Video Game Soundtracking* (D. Williams and N. Lee, eds.), pp. 143–163, Cham: Springer International Publishing, 2018.

[30] S. Agrawal, S. Bech, K. Bærentsen, K. De Moor, and S. Forchhammer, "Method for subjective assessment of immersion in audiovisual experiences," *J. Audio Eng. Soc.*, vol. 69, pp. 656–671, Sept. 2021.

[31] T. Bjørner, A. Magnusson, and R. P. Nielsen, "How to describe and measure obstacles of narrative immersion in a film? the wheel of immersion as a framework," *Nordicom Review*, vol. 37, no. 1, pp. 101–117, 2016.

[32] E. Brown and P. Cairns, "A grounded investigation of game immersion," in *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, (New York, NY, USA), pp. 1297–1300, Association for Computing Machinery, Apr. 2004.

[33] J. M. Rigby, D. P. Brumby, S. J. J. Gould, and A. L. Cox, "Development of a questionnaire to measure immersion in video media: The film IEQ," in *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '19, (New York, NY, USA), pp. 35–46, Association for Computing Machinery, June 2019.

[34] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton, "Measuring and defining the experience of immersion in games," *Int. J. Hum. Comput. Stud.*, vol. 66, pp. 641–661, Sept. 2008.

[35] R. Agarwal and E. Karahanna, "Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage," *Miss. Q.*, vol. 24, no. 4, pp. 665–694, 2000.

[36] W. IJsselsteijn, H. d. Ridder, J. Freeman, S. E. Avons, and D. Bouwhuis, "Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence," *Presence*, vol. 10, pp. 298–311, June 2001.

[37] M. Lombard, R. D. Reich, M. E. Grabe, C. C. Bracken, and T. B. Ditton, "Presence and television," *Human Communication Research*, vol. 26, no. 1, pp. 75–98, 2000.

[38] B. Reeves, A. Lang, E. Y. Kim, and D. Tatar, "The effects of screen size and message content on attention and arousal," *Media Psychol.*, vol. 1, pp. 49–67, Mar. 1999.

[39] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff, "A cross-media presence questionnaire: The ITC-Sense of presence inventory," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 282–297, 2001.

[40] J. Freeman and J. Lessiter, "Here, there and everywhere: the effects of multichannel audio on presence," 2001.

[41] A. Gabrielsson, B. Lindström, and G. Elger, *Assessment of perceived sound quality of eighteen high fidelity loudspeakers*. Department of Technical Audiology, Karolinska institutet, 1983.

[42] L. Ermi, "Fundamental components of the gameplay experience: Analysing immersion," 2005.

[43] A. McMahan, "Immersion, engagement, and presence: A method for analyzing 3-D video games," in *The video game theory reader*, pp. 89–108, Routledge, 2013.

[44] L. N. Taylor, *Video Games: Perspective, Point-of-view, and Immersion*. University of Florida, 2002.

[45] W. R. Sherman and A. B. Craig, "Understanding virtual reality—interface, application, and design," *Presence*, vol. 12, pp. 441–442, Aug. 2003.

[46] D. Carr, D. Buckingham, A. Burn, and G. Schott, *Computer Games: Text, Narrative and Play*. Polity, Mar. 2006.

[47] S. Bjrk and J. Holopainen, "Patterns in game design. hingham, MA: Charles river media," 2005.

[48] D. Arsenault, "Dark waters: Spotlight on immersion," in *GAMEON-NA International Conference*, pp. 50–52, 2005.

[49] E. Adams and A. Rollings, *Fundamentals of Game Design (Game Design and Development Series)*. USA: Prentice-Hall, Inc., 2006.

[50] J.-N. Thon, "Immersion revisited: on the value of a contested concept," 2008.

[51] M. Slater, V. Linakis, M. Usoh, and R. Kooper, "Immersion, presence and performance in virtual environments: an experiment with tri-dimensional chess," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '96, (New York, NY, USA), pp. 163–172, Association for Computing Machinery, July 1996.

[52] K. M. Lee, "Presence, explicated," *Commun. Theory*, vol. 14, pp. 27–50, Feb. 2004.

[53] N. Curran, *The psychology of immersion and development of a quantitative measure of immersive response in games*. PhD thesis, University College Cork, 2013.

[54] "Qualtrics XM // the leading experience management software." https://www.qualtrics.com/uk/, Oct. 2015. Accessed: 2022-2-2.

[55] "Manage your team's projects from anywhere." https://trello.com/en, 2022. Accessed: 2022-6-15.

[56] S. Elo and H. Kyngäs, "The qualitative content analysis process," *J. Adv. Nurs.*, vol. 62, pp. 107–115, Apr. 2008.

[57] M. Bengtsson, "How to plan and perform a qualitative study using content analysis," *NursingPlus Open*, vol. 2, pp. 8–14, Jan. 2016.

[58] M. Kerins, "Understanding the impact of surround sound in multimedia," *Psychology of music in multimedia, Oxford University Press, New York*, pp. 365–390, 2013.

[59] J. Berg, *Systematic Evaluation of Perceived Spatial Quality in Surround Sound Systems*. School of Music, Division of Sound Recording, 2002.

[60] P. Mirza-Babaei and L. Nacke, "Introducing the biometric storyboards tool for games user research," in *2014 IEEE Games Media Entertainment*, pp. 1–7, Oct. 2014.