Multimedia Tools and Applications (2024) 83:43377–43417
https://doi.org/10.1007/s11042-023-16918-1

Measuring users' emotional responses in multisensory virtual reality: a systematic literature review

Mariana Magalhães[1] · António Coelho[1] · Miguel Melo[2] · Maximino Bessa[3]

Received: 18 March 2023 / Revised: 14 August 2023 / Accepted: 7 September 2023 /
Published online: 13 October 2023

Abstract
Virtual reality and emotions have become inseparable concepts over the past few years, supported by the increasing number of studies relating them. However, these studies' methodologies are often poorly justified or dependent on the authors' subjective definition of emotion and its classification. Moreover, frequently, these studies only consider two stimuli, specifically audiovisual, despite being known the relevance of including a greater variety of sensory channels to improve the relationship between the individual and the virtual environment. So, to address these gaps, and considering the importance of multisensory stimulation, this paper aims to review the methods and instruments found in the literature regarding the analysis of the users' emotions in virtual reality. Also, we provide an overview of the main limitations of such studies. Little information can be found in the literature regarding the connection between the input stimulus and the users' emotional responses. This corroborates the difficulty in creating and evaluating immersive virtual experiences when stimulating more than two human senses, typically audiovisual. Nevertheless, we address some clues on the impact of visual, auditory, haptic, smell, and taste elements to trigger specific emotions. Also, we address the association between the research area and the method used. Finally, the main gaps and challenges are discussed. We expect that the combination of these results acts as guidelines for designing richer multisensory virtual experiences. Moreover, we intend to contribute to future research on emotions-based immersive virtual reality by providing a review of the most suitable methodologies and instruments for specific contexts.

Keywords Emotions in human-computer interaction · Evaluation/methodology · Multimodal systems · Multisensory stimulation · Systematic literature review · Virtual reality

✉ Mariana Magalhães
mariana.o.magalhaes@inesctec.pt

[1] Department of Informatics Engineering, Faculty of Engineering of the University of Porto/ INESC TEC, Porto, Portugal

[2] Centre for Information Systems and Computer Graphics, INESC TEC, Porto, Portugal

[3] Department of Informatics Engineering, University of Trás-os-Montes and Alto Douro / INESC TEC, Vila Real, Portugal

1 Introduction

Many psychological dimensions, such as the user's emotions or sense of presence [1, 2], have been involved in the evaluation of Virtual Reality (VR) systems. Many studies have proven that immersion associated with VR is an essential feature in generating both sense of presence and emotional states [3–7]. Typically, emotion-based studies engage one or two human senses [8, 9], most commonly vision and auditory [10]. As Gall, Roth, Stauffert, et al. [11] stated: "multisensory feedback provides means to manipulate the strength of illusion". The introduction of multisensory features improves users' affective responses [12–14], sense of presence and realism [15–17], enhancing the fidelity of the world [10], creating a sensorial richer virtual world [10, 18, 19] and an increased emotional connection with it [20], in contrast with unisensory interfaces [12]. All this is possible thanks to specific hardware that helps to isolate the user from the real world, such as head-mounted displays (HMDs) [16], haptic [21], smell [22], and gustatory interfaces [23–25], devices that simulate wind, warmth [13, 26], among others. Unisensory interfaces, in turn, are more likely to provide the undesired sense of hollow, fictitious, and unrealistic experiences [27].

Despite the positive aspects, to date, studies examining immersive VR regarding their ability to induce emotions are still scarce [4, 10, 19, 28, 29], due to the high complexity the system development encompasses [30, 31] and consequently due to the difficulty of evaluating the user experience [32]. The combination of different sensory channels implies that they can no longer be seen as separate [33]. Instead, they interplay, affecting each other, which hinders research in the multisensory VR context [28]. These obstacles are even more present when a large number of senses are explored [9, 31]. Despite studies involving just two stimuli (typically visual and audio) being easy to implement and very common in literature, multisensory virtual environments are an increasingly popular topic and transversal to numerous areas (games, therapy, treatment, training, education, tourism, to name a few) [7, 10, 18, 28, 29]. Moreover, multisensory stimulation in VR has been pointed out as more likely to increase immersion and, consequently, presence [2]. Compared to less immersive environments, immersive environments present greater potential to gather users' different emotional states [34–36] and more emotional arousal [2].

This paper explicitly explores studies that address, either as a primary or secondary objective, the users' emotional responses in multisensory immersive VR, particularly those that engage at least three human senses, typically composed of visual and audio cues, and complemented by other(s). This criterion will prevent obtaining studies on an audiovisual basis only, which are more likely to restrict the results related to the investigation of users' emotional responses in VR. Moreover, one of the most important properties of the emotional aspects in VR is the "multisensory synergetic stimuli", which can be achieved by the use of "a great variety of elements, colors, sounds, objects and smells" [37]. Based on that, we consider that studies with more than two stimuli provide more accurate results in emotional experiments. As far as we know, this investigation represents a novelty, as no other studies can be found reviewing the relationship between the input stimuli and the obtained emotions, considering the research area and the methodology used, as previously discussed. Therefore, this is one of the main research goals we intend to achieve, vital to understand how the qualitative and effective design of multisensory VR experiences can be affected by these factors, as mentioned by Kruijff, Marquardt, Trepkowski, et al. [10]. The obtained results will contribute as guidelines for designing richer multisensory virtual experiences. They also contribute as relevant topics for a research agenda, which should be taken into account by investigators in this field to keep the evolution of studies involving users' psychological dimensions in VR.

1.1 The assessment of human emotional responses

Emotions are complex to understand and explore due to their abstract and subjective nature [38], making their measurement challenging [39]. Although there is still no consensus in the literature on an exact definition for "emotion", Cabanac [40] defines it as "any mental experience with high intensity and high hedonic content (pleasure and displeasure)". Emotions are the basis for understanding the human mind regarding progress and evolution. As they are not voluntary expressions [38], they can provide relevant information regarding one's unconscious state [10, 39]. Involuntary changes in the Autonomic Nervous System regulate one's physiological processes. Results obtained with the measurement of physiological responses can act as an indicator of an emotional variation [39, 41]. They can be provided by monitoring the heart rate (HR) or the heart rate variability (HRV), respiration rate (RR), via electrodermal activity (EDA), also known as galvanic skin response (GSR), which measures the skin conductance level (SCL), or by examining the finger pulse volume (FPV) or the blood pressure (BP), for example. This variation can also be expressed through behavior: verbally, through facial expressions, and body movements [39], to name a few.

Two general models are used to study emotions: the categorical approach and the dispositional theory. In the first one, emotions are represented as labels [42, 43]. In contrast, in the second one, the concepts of valence (pleasure and displeasure) and arousal (high activation and low activation) determine the individuals' emotional experience, which is represented in a two-dimensional model – the circumplex model of affect. So, each emotion can be defined as a combination of these two dimensions or by the position it takes in the model of valence and arousal [44]. More informally and amply, emotions can also be divided into positive, such as happiness, and negative, such as anxiety, fear, sadness, and disgust [41].

Several performance tasks and scenarios in VR can provide insights into the relationship between emotions and cognitive-behavioral responses. Among them, we highlight stress-inducing tasks, cognitive load tasks, attention and vigilance tasks, motor skills and coordination tasks, decision-making and risk assessment tasks, and social interaction tasks. By combining such performance tasks with subjective and objective measures, researchers can comprehensively understand how emotions and stress impact cognitive and behavioral responses in VR [45].

1.1.1 Objective and subjective methods and instruments

The study of emotions can be done consciously, with subjective methods and instruments [29], or unconsciously, resorting to objective methods and instruments. Also, humans' emotional state can be recognized by combining these two methods – a mixed approach [39].

Objective measures must be used if the aim is to measure the most unconscious level of the individual, which can be done by resorting to psychophysiological measures, like EDA, HR, or electroencephalography (EEG) [10], as previously discussed. They allow the collection of physiological and biometric data, which give essential information about how a person feels, even though not conscious of it [39]. Though, they also present some contraindications, mainly because of being obtrusive and noisy [46].

In turn, subjective measures should be used if the objective is to evaluate the emotional experience through an individual's subjective point of view. That includes established user scales, e.g., the "Visual Analogue Scale" (VAS) or the "Self-Assessment Manikin" (SAM), interviews, thinking aloud, and questionnaires (e.g., the "Check-All-That-Apply" (CATA) procedure, the "Positive and Negative Affect Schedule" (PANAS), or the "State-Trait Anxiety Inventory" (STAI)). Such measures can be pictorial, such as SAM, created by Bradley and Lang [47], which presents a representative drawing of the human figure for the inquiries to express their emotions. Regarding interviews, structured, semi-structured or unstructured interviews might be used. Respectively, they follow a predetermined set of questions or a standardized interview protocol (structured), or a combination of a protocol and flexibility for the interviewer to freely adapt the conversation with the participants according to their responses (semi-structured), and, finally, with no predetermined set of questions (unstructured interview). Focus groups, also a type of interview, are a method to obtain subjective reports. It involves a group of participants (typically 6 to 12), who engage in a facilitated discussion led by a moderator. Participants are invited to share their opinions, experiences, and perspectives regarding a specific topic, allowing the interaction and exchange of ideas [48]. Regarding the understanding of human emotions in VR, a phenomenological interview can be beneficial. It is frequently a semi-structured or unstructured interview, as the emphasis is the first-person perspective of a specific experience, considering the individual's subjective meanings, emotions, body sensations, and intentions related to the phenomenon under investigation [49].

Questionnaires can be presented as a list of a limited number of adjectives, requiring respondents to express how they feel at that moment through a scale. Some examples are PANAS, composed of twenty adjectives (ten positives and ten negatives) and a scale from 0-5 [50]; the STAI, which focuses on the level of anxiety the respondents feel, which is expressed in a 0-3 scale [51]; VAS, initially created to assess the patient's pain in clinical research in 1921 by Hayes and Patterson [52], and nowadays used broader to specify the level of agreement in a continuous scale; and CATA questionnaires, which are used to widespread investigate the users' perceptions on a variety of attributes [53]. Given the low complexity of the application [54], subjective measures are frequently used, despite the number of biases [55] and measurement issues [56], for instance, the interference of individual and socio-demographic aspects. As Prescott [57] pointed out, age, education, culture, socio-economic status, and personality might destabilize the results.

2 Methodology

This systematic literature review was inspired by the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" (PRISMA) [58] to minimize bias and maximize its contribution to science.

2.1 Research objectives

We propose the following research objectives:

1. To survey the current methods and instruments found in the literature related to the study of the users' emotions during multisensory VR experiences;

2. To explore the reported limitations regarding the study of users' emotional responses in multisensory VR experiences;
3. To assess the impact of the stimuli on emotional responses, to help understand the relationship between a specific stimulus and the obtained emotional response(s);
4. To investigate whether there is an association between the used method and the research area;
5. To identify the main gaps and challenges in this context.

2.2 Research questions

To meet the above-mentioned research objectives, we propose the following research questions (RQ):

RQ 1. What are the most common methods and instruments used to measure users' emotional responses during multisensory VR experiences?
RQ 2. What are the main limitations regarding the study of users' emotional responses in multisensory VR experiences reported by the author(s), associated to:

RQ 2.1 the used VR equipment?
RQ 2.2 the instruments for emotional responses collection?
RQ 2.3 the experimental design?

RQ 3. Is there any association evidence between the stimulus provided to the users and their emotional response(s)?
RQ 4. Is there any association evidence between the used method and the research area?
RQ 5. What gaps and challenges remain in the literature for future research?

2.3 Search strategy

We extensively searched four electronic data sources (Web of Science, Scopus, ACM Digital Library, and IEEE Xplore). To develop this Systematic Literature Review (SLR), a selection of data sources to perform the search was made, as it would be impracticable to search all the existent data sources. We chose three (Web of Science, Scopus, and ACM Digital Library) of the principal search systems meeting the necessary performance requirements to conduct systematic reviews, according to Gusenbauer and Haddaway [59]. Although not on this list, IEEE Xplore was considered for the search process, as it provides high-quality technical literature in engineering and technology, which is this paper's general field of research. According to its website [60, 61], IEEE Xplore claims to provide the "highest quality technical literature in electrical engineering, computer science, electronics, and related disciplines".

To summarize, two multidisciplinary (Web of Science and Scopus) and two computer science specialized databases (ACM Digital Library and IEEE Xplore) were considered to ensure relevant results. This combination allows obtaining a wide variety of papers pertinent to this paper's subject. All four are comprehensive data sources, comprising peer-reviewed journal articles and conference proceedings, ensuring the source materials' quality.

We considered all the available results (inclusion criteria applied), disregarding the year of publication and the subject area. To restrict the results, two filters were applied: (1) related to language (Portuguese or English writing), and (2) to the document type (conference proceedings paper or journal paper).

The primary search terms consisted of "virtual reality", "emotions", and "stimulus", whether in the title, abstract, or keywords. However, due to the inconsistent terminology, some synonyms or similar terms were used, in some cases with wildcards, as follows:

- To designate "virtual reality", we used the following terms: "virtual reality", "VR", "virtual environment*", "immersive environment*", "immersive technolog*" and "virtual scenario*";
- For “emotions”, we used only: “emotion*”;
- Finally, for "stimulus", we considered the terms "*sensory OR stimul*".

We also used Booleans ("AND" and "OR") in our search, resulting in the following query: ("virtual reality" OR VR OR "virtual environment*" OR "immersive environment*" OR "immersive technolog*" OR "virtual scenario*") AND emotion* AND (*sensory OR stimul*).

We performed the last search on all databases on January 24th, 2023.

2.4 Inclusion criteria

The inclusion criteria were as follows:

1. The paper is published in one of the mentioned data sources;
2. The paper is written in English or Portuguese (the authors’ native language);
3. The publication type is a research article or proceedings paper, published in a refereed journal or conference;
4. The search terms include, whether in the title, abstract, or keywords, the following query: ("virtual reality" OR VR OR "virtual environment*" OR "immersive environment*" OR "immersive technolog*" OR "virtual scenario*") AND emotion* AND (*sensory OR stimul*).

2.5 Exclusion criteria

The exclusion criteria taken into consideration were:

1. The paper’s full text is not available;
2. The paper’s methodology does not include the assessment of the users’ emotional responses in virtual reality environments, either through subjective or objective measures, or a mixed approach;
3. The paper’s methodology does not focus on a fully-immersive system. The criterion adopted was based on Costello [60], who defends a classification of VR systems according to the level of immersion provided, which can be non-immersive (desktop-based VR), semi-immersive (projection systems), or fully-immersive (HMD systems). For the author, VR systems are more immersive the less the user can perceive (see, hear, touch) the outside world [60], which is achieved nowadays by using high-resolution 360-degree vision HMD [62]. Furthermore, as previously discussed, fully immersive systems have greater potential to gather users’ different emotional states [2, 34–36] compared to non-immersive ones.
4. The paper does not consider a multisensory virtual environment (the addition of at least one more sense to the base pair vision and audio).

2.5.1 Criteria used to identify the stimuli provided

The user isolation from the real world is achieved with the help of specialized hardware, such as HMDs [16], haptic [21], smell [22], and gustatory interfaces [23–25]. This section will overview the available hardware regarding the five human senses used to immerse the users in VR. Considering that this work focuses on fully-immersive VR setups, regarding the visual stimulus, the only requirement is the use of advanced visualization systems on the methodology, specifically HMDs, independently of the brand or its specifications. No specific criterion was applied regarding the auditory stimulus. Any audio configuration was considered, for instance, headphones, earphones (integrated or not in the HMD), or external devices. For haptic stimulus, we used the list of haptic interfaces distinguished by Wee, Yap and Lim [63], which consists of "handhelds", "wearables", "encountered-types", "physical props", and "mid-air". By handhelds, the authors consider controllers held by the user, attachments, or add-ons that enhance the haptic feedback of default devices, such as the controllers provided by Oculus Rift, HTC Vive, and Sony Playstation VR. Wearable devices are worn anywhere on the body (fingers, wrists, hands, etc.), for instance, the Emogle [64]. Encountered-types include devices that provide on-demand haptic feedback, typically robotic arms, drones, or specialized devices with an end effector attached. Physical props consist of tangible objects placed in the physical space aligned to the virtual object to deliver haptic feedback. Mid-air devices can provide ultrasonic vibrations through the air to deliver haptic sensations [63, 65]. Besides such classification, whole-body tactile stimulation has also been explored in VR, such as floor vibration [21, 66], which is also considered haptic or tactile feedback. New ways of providing haptic feedback have been developed and implemented recently. These methods include the use of heat to deliver the users the sensation of a warm environment [13], wind [10, 13, 28], pain [67], fire, or a ghostly breeze by means of a fan [19]. Such techniques are frequently considered haptic feedback [67, 68], which can be divided into two categories: "active haptic feedback", when the computer-controlled actuators exert forces on the user, and “passive haptic feedback”, which corresponds to the interaction with tangible objects [69]. Together, active and passive haptic feedback integrate the so-called "Smart Substitutional Reality" (SSR) [19]. This paper will include both active and passive haptics as haptic stimuli.

Finally, regarding smell and taste, no specific criterion was applied. Despite having an essential role in evoking human emotion [27] and enhancing the user experience, especially regarding the perception of realism and sense of presence [15–17], these two senses are much less explored, especially when compared to auditory and visual stimuli [30].

2.6 Data collection process

For data collection, the procedure was conducted using piloted forms. We used the reference manager software Endnote™ 20 to upload all the studies from the four databases (Web of Science, Scopus, ACM Digital Library, and IEEE Xplore). After removing the duplicates, a primary analysis of the paper’s title and abstract was performed to screen the results for full-text analysis, following the PRISMA flow diagram [58]. At this stage, the results were divided according to the exclusion criteria (Subsection 2.5). Results are presented in Table 1. The ones selected for the quantitative analysis were thoroughly read, and the information regarding the proposed objectives and research questions was retrieved and synthesized into Table 1.

Table 1 Number of discarded papers after the title and abstract analysis, based on the exclusion criteria

Exclusion Criteria	Number of discarded papers
1	N= 9
2	N= 441
3	N= 201
4	N= 251
Total	N= 902

2.7 Assessing studies' relevance

After gathering the final results for qualitative analysis (N=37), the studies' relevance was assessed by a scoring process consisting of an individual evaluation of the papers. The scoring system adopted was similar to the approach made previously by some authors [30, 45, 70, 71]. As defined by them, the scoring system ranges from 1 (lowest score) to 3 (highest score) per question. The total score results from the sum of the scores of the created questions. Considering that three questions were developed, the total score ranges between three (lowest relevance) and nine points (highest relevance). Thus, papers with a score of 3 were considered low-relevance papers; papers between 4 and 6 (both included) were addressed as medium-relevance papers; a score equal to or greater than 7 determined high-relevance papers. The first author completed the scoring process for all the studies, which the coauthors double-checked. The papers' scores depended on the consensus between the authors. The three questions for assessing the relevance of the studies were:

1. Is the paper technically sound? For this question, we have created some guidelines to assess the papers' performance based on the work made by Spezi, Wakeling, Pinfield, et al. [72]. The authors define "technical soundness", also named "scientific soundness", as the papers' "methodological precision, coherence, and integrity", which includes the quality of argumentation, the logic of research, and the interpretation of data. Based on this, for the scoring options, we considered 1 point for those papers that present several issues regarding the quality of argumentation, the logic of research, and the interpretation of data; 2 points for the papers that present scarce issues regarding the same topics; and 3 points for the papers that sound technically accurate regarding the same three topics.
2. Is the sample size (number of participants) suitable, according to the recommendations? According to Macefield [73], although the specification of the participant group size remains challenging, the author considers that for comparative studies aiming to obtain valid results, a group size of 8-25 participants is a good baseline. However, increasing the participants' size allows to obtain more statistically significant results. Accordingly, in this literature review, those papers that included less than 8 participants (Sample < 8) were assigned 1 point; 2 points were attributed if the sample was between 8 and 25 participants ($8 \leq$ Sample ≤ 25); 3 points were assigned to studies with a sample greater than 25 (Sample > 25).

3. How relevant is the study for concluding the correlations between the sensory stimulus and users' emotions? This question was assessed based on each paper's conclusions regarding the correlation between the stimulus and the users' emotional responses, as it is one of the main objectives of this literature review. If no correlations could be found ("Inconclusive"), the paper was assigned 1 point – low relevance; if no concrete correlations were established, 2 points were assigned – medium relevance; 3 points were attributed when objective correlations had been found by the authors – high relevance.

It is worth noting that this score, particularly the third point, concerns the studies' relevance only for this paper's purposes. This means that this score should be reviewed for other scientific purposes.

3 Results

3.1 Study selection

The data collection process was conducted in four steps ("Identification", "Screening", "Eligibility", and "Included"), according to the PRISMA flow diagram [58] (Fig. 1), to gather relevant information to answer all the research questions.

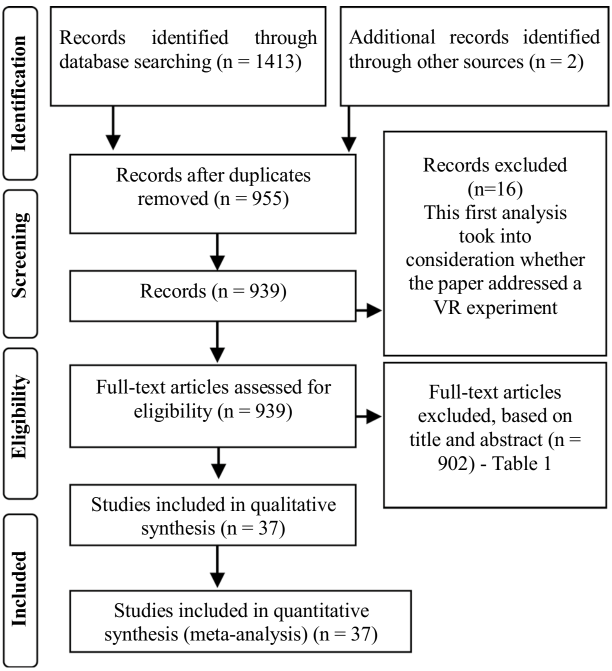

Fig. 1 PRISMA search flow guidelines

Initially, in the "Identification" step, we ran a search in the four electronic database sources (Web of Science, Scopus, ACM Digital Library, and IEEE Xplore) for the search terms, applying only language filters (Portuguese or English) and document type filters (conference proceedings paper or journal paper).

From this initial search, we obtained 1413 results. Additionally, we applied the so-called "snowballing method", whereby we could locate other relevant studies by checking the references in the already selected articles. Two more results were obtained, resulting in a total of 1415 results. The "Screening" stage, the second step of the process, consisted of removing the duplicate results (n=460), resulting in 955 studies. Then, these results were analyzed to check whether the paper addressed a VR experiment (16 results removed). In the third step, "Eligibility", the full text of the remaining papers (n=939) was assessed. In this process, we discarded 902 results, registering the reason according to the exclusion criteria, as previously noted (Table 1). Finally, we obtained 37 results for the qualitative and quantitative analysis, all written in English, which were included in the final stage (“Included”).

3.2 Qualitative analysis

Relevant data were extracted from these results for in-depth review. This information was organized and synthesized into Table 2, created based on the research questions we would like to answer. First, studies were segmented according to the research’s context area. This classification was based on the title, keywords, and abstract analysis. Six categories were created accordingly (“Therapy”, “Food research”, “General VR Research”, “Games”, “Tourism”, and “Marketing”). The third category was developed to comprise the papers that do not fit into any particular category, given their wide range of applications.

For each selected study, the table contains the following variables: 1) identification of the study; 2) research’s context area; 3) the sample size used for emotional responses collection, considering that some studies only investigate it with a sub-sample ("N="); 4) the methods and instruments used in the immersive VR experience, distinguished into “VR System” and “Measurement of the users’ emotions”; 5) the stimuli provided by the immersive VR experience to the user, that can be Visual (V), Auditory (A), Haptic (H), Smell (S) and Taste (T); 6) the “Context and main findings” of the study that might be relevant for conclusions; 7) the relationship found between the stimulus and the obtained users’ emotional responses (if applicable). Whenever the study’s main findings (column 7) were related to the relationship between the stimulus and the users’ emotional responses (column 8), both columns were merged.

Little work has been performed regarding the emotions triggered according to each specific stimulus. It is still challenging to pinpoint the effect of specific cues on the users’ emotional responses, as attested by some authors [10, 28, 33, 74]. Also, considering the emotions’ subjective nature, this relationship is challenging due to the multiple factors that may affect the users’ perception, as addressed by Kruijff, Marquardt, Trepkowski*,* et al. [10]. We can identify this gap as of now, which will be further discussed.

Table 2 Data obtained from all the articles selected for final analysis

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Chen, Dey, Billinghurst, et al. [7]	General VR Research	19	HTC Vive	Mixed (Objective: HR, pupil size; Subjective: SAM, PANAS, interview)	V, A, H	To investigate whether providing HR feedback enhances the participants' experience in VR, this paper compared the design space of different types of multisensory HR representations. The authors reported that the participants preferred having multisensory HR feedback (compared to no-feedback), specifically audio-haptic. Visual feedback was found to be distracting.	Inconclusive.
Kruijff, Marquardt, Trepkowski, et al. [10]	Games	20	Oculus Rift DK1	Subjective: Specifically designed questionnaire	V, A, S	The authors conducted a pilot study to reveal specific correlations between the stimulus and the users' emotional responses in specific situations. Results provided some important clues: *Happiness: normal sound (bee swarm situation); Surprise: tactile back (spider behind back situation); Surprise: back vibration, low-frequency sound (bad weather situation); Happiness: wind, smell (sea view situation); Surprise: low frequency and normal sound, back vibration, seat vibration (zombie swarm situation).*	
Gall, Roth, Stauffert, et al. [11]	General VR Research	21	HTC Vive	Subjective: SAM	V, A, H	This study intends to demonstrate that the illusion of a virtual body might modulate users' emotional responses. Compared to the low embodiment condition, the high embodiment condition intensifies the users' arousal, dominance, and valence.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Kono, Miyaki and Rekimoto [18]	General VR Research	7	Oculus Rift CV1	Subjective: VAS, specifically designed questionnaire + Interview	V, A, H	In this study, the authors aim to understand which design aspects may induce negative emotional feedback in the users. Virtual experiences with more than just visual stimuli *(Tactile and Electrical Muscle Stimulation feedback (EMS)) affect realism and emotional aspects such as fear and pain. EMS can be more effective in inducing pain in first-person experiences.*	
Kruijff, Marquardt, Trepkowski, et al. [28]	Games	20	Oculus Rift DK2	Subjective: Specifically designed questionnaire	V, A, S	This paper explored the challenges and potential methodological directions to assess the effects of multisensory cues to enhance users' engagement in VR. Results highlight the role of presence and interaction in improving users' engagement but do not provide clues on which specific multisensory cues can trigger certain emotions.	Inconclusive.
Dey, Chen, Billinghurst, et al. [29]	General VR Research	19	HTC Vive	Mixed (Objective: HR, GSR; Subjective: SAM, PANAS, interview)	V, A, H	This paper explores the influence of presenting manipulated physiological feedback on users' emotions in VR. *Visual manipulated HR feedback (decreased – "-15%/-30%", increased - "+15%" and accurate) can significantly enhance five emotions (interest, excitement, scariness, nervousness, and fear) in virtual experiences.* Specifically, *interest and excitement are awakened for "-15%", "-30%", and "+15%" HR feedback situations*; *"+15%" was found to increase negative emotions (scariness, nervousness, and fear)*. The accurate HR feedback was proven to create the least effect on any of these emotions.	

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Brengman, Willems and De Gauquier [31]	Marketing	235	Oculus Rift DK1	Subjective: Specifically designed questionnaire	V, A, S	This paper addresses the power of smell and its congruence with auditory cues to enrich VR and to create more engaging experiences in advertisement in terms of attitude toward the ad and the brand, customer delight, purchase intention, sensory user experience, telepresence, and vividness. To do so, 3 scenarios were developed: congruent, incongruent, and no scent. In general, the introduction of smell in VR enriches ad scenarios in such a way that incongruent smells result in a slightly better sensory experience than the total absence of smell.	Inconclusive.
Wu, Weng and Xue [41]	General VR Research	15	Oculus Rift DK2	Mixed (Objective: HR; Subjective: PANAS, Fear Survey Schedule, behavior scale (composed by the reported feelings of "rapid heartbeat", "dodge eyes and don't dare to look", and "the mouth opens out a cry")	V, A, H	This paper aimed to verify the effectiveness of VR in inducing emotional changes and eliciting fear. Visual, auditory, and tactile stimuli are positively related to emotion-induced. Visual stimulation correlated more with the user's emotions than haptic and auditory stimulation. In turn, the haptic stimulus via air injection and a sweep leg device had a more significant influence on the user's emotions than the auditory stimulus when watching terror scenes.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Steinhaeusser, Eckstein and Lugrin [55]	Games	12	Not specified	Subjective: STAI, FRS	V, A, H	The authors explored the relationship between immersion, presence, and negative emotion induction, resorting to 3 different setups: desktop, VR, and SSR (VR + thermal and haptic stimuli). Results demonstrated VR and SSR setups to elicit more fear and sense of presence than desktop.	Inconclusive.
Kong, Sharma, Kanala, et al. [62]	Food Research	67	Oculus Go	Subjective: CATA procedure	V, A, T	Three chocolate types (milk, white and dark) were evaluated under three contextual settings (sensory booths, a 360° VR experience in a pleasant sightseeing tour – positive VR (PVR), and a 360° VR experience in a noisy live music concert – negative VR (NVR)). No significant differences were found between the context type (NVR and PVR) and the chocolate-tasting experience. Regarding the correlations found, *milk chocolate and white chocolate in the PVR were associated with neutral and positive emotions: "peaceful," "pleasant," "good," "satisfied," "glad," "pleased," and "polite"*. Curiously, *in the NVR, milk and white chocolate were also associated with positive emotions: "affectionate," "interested," "happy," "loving," "joyful," and "friendly". Dark chocolate under NVR was related to ardent descriptors: "adventurous," "energetic," "wild," "active," and "enthusiastic". Under PVR condition, dark chocolate was related to negative terms such as "bored," "worried,", "disgusted," and "aggressive".*	

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Karafotias, Korres, Teranishi, et al. [65]	Therapy	50	Oculus Rift DK1	Subjective: VAS + Modified version of Igroup Questionnaire (for valence and arousal)	V, A, H	This study tests whether haptic VR sensations provide a unique opportunity for pain distraction. Results demonstrated that VR task associated with tactile stimulation enhances pain distraction when touch interactions are involved. Their combination also induced positive feelings (valence and arousal).	Inconclusive.
Li, Jung, McKee, et al. [66]	General VR Research	26	HTC Vive Pro Eye	Objective: Pupil size, GSR	V, A, H	Three conditions (walking sound matching vibration (V), mismatching vibration, and no vibration) were developed to assess its impact on users' cybersickness and sense of presence. No differences were found for presence or cybersickness in the three tested conditions. The results were inconclusive for emotional arousal, although participants reported they preferred the "V" condition over the other two conditions.	Inconclusive.
Steinhaeusser and Lugrin [19]	Games	61	HTC Vive	Subjective: Fright Reaction self-reported Survey (FRS), State Anxiety subscale from STAI	V, A, H	A comparison between the effects of game presentation (PC, VR, and SSR) on the perceived sense of presence and induced emotions was performed. Users' sense of presence is higher in VR and SSR than in PC condition. No differences in anxiety, fear levels, or desensitization effects were found in the three conditions.	Inconclusive.
van Veelen, Boonekamp, Schoonderwoerd, et al. [74]	Therapy	3	HTC Vive Pro Eye	Mixed (Objective: RR; Subjective: Interview)	V, A, S, H	Four design aspects of VR (system properties, sensory cues, narrative, and challenge) were explored with veterans with PTSD to understand their effects on presence and emotional responses. The design aspects studied were recognized to maintain the presence levels and influence the subjects' emotional responses.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Torrico, Han, Sharma, et al. [75]	Food Research	53	Oculus Go	Subjective: CATA procedure	V, A, T	In the context of wine tasting, this paper explores the effects of the context and the environment on the consumers' acceptability and emotional responses. Light conditions (bright and dark) in a VR experience influence consumers' emotions which, in turn, can affect their taste perception. *Bright-VR was associated with "free", "glad", "aggressive", and "enthusiastic" emotions. Dark-VR was associated with "nostalgic", and "daring".*	
Chin, Thompson and Ziat [76]	Therapy	5	Oculus Rift CV1	Subjective: Specifically designed questionnaire	V, A, H	This work's primary goal is to determine whether a multimodal experience can change the emotional responses related to injections. Results show that compared to non-immersive look-at and look-away conditions, multisensory VR environments can effectively reduce stress, anxiety, and boredom related to needles through distraction and increasing happiness.	Inconclusive.
Kaminska, Smolka, Zwolinski, et al. [77]	Therapy	28	HTC Vive	Mixed (Objective: GSR, EMG, HR; Subjective: General Mood Scale)	V, A, H	These authors tested whether integrating multisensory VR in Eye Movement Desensitization and Reprocessing (EMDR) – a treatment for PTSD – can be a tool to relieve stress. Results could demonstrate so. In general, EMDR sessions were found to be most effective when using a visual-only method, compared to V+A, V+T, and V+A+T conditions.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Sinesio, Moneta, Porcherot, et al. [78]	Food Research	513	Oculus Rift DK1	Subjective: CATA procedure	V, A, S, T	Three immersive scenarios (projection in flat walls at 180°, VR-3D with Oculus Rift, and Immersive Room (IR) - VR) were developed to compare the effectiveness of consuming a beer in a pub in a real-life situation. Results demonstrated comparable feelings of presence and engagement between the immersive conditions and the real environment experience. All the immersive technologies tested improved consumer engagement.	Inconclusive.
Schweizer, Schmitz, Plempe, et al. [79]	Therapy	40	TriVisio VR Vision (3D)	Subjective: STAI	V, A, S	This paper tested if individual differences in high trait anxiety (HA) modulate affective and cognitive processing of VR analogue trauma. Results demonstrated that individuals with HA showed increased stress and anxiety responses, emotion dysregulation, and intrusive memories upon VR trauma exposure.	Inconclusive.
Worch, Sinesio, Moneta, et al. [80]	Food Research	513	Oculus Rift DK1	Subjective: CATA procedure	V, A, S, T	Five conditions were tested to simulate a pub experience and to understand the consumers' self-reported emotional responses: lab condition, the actual pub, a recreated pub (IR-Room), VR goggles projecting 360° video, and VR combining 3D modelling and 360° videos. VR condition induced negative emotions (disappointed and bored), whereas Pub and IR-Room induced excitement and friendliness. Consumers tended to find the experience more adjusted and consistent with the beer experience in the actual pub or the VR-3D condition. Also, they felt more comfortable and involved in the same two conditions than in the IR-VR 360° or lab conditions.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Kim, Bae and Lee [81]	General VR Research	81	HTC Vive	Subjective: Specifically designed questionnaire	V, A, H	This paper explored whether the tactile perception of soft vs. spiky textures in two scenes (negative and positive) influence emotional responses and the user experience in VR. Users' emotions increased when users watched positive scenes with soft stimuli compared to no-haptic or spiky-tactile conditions. Spiky stimuli did not influence emotions for either negative or positive scenes.	Inconclusive.
Nivedhan, Mielby and Wang [82]	Food Research	50	HTC Vive Pro Eye	Subjective: Specifically designed questionnaire	V, A, T	In the context of tasting cold brew coffee, this paper aimed to explore the association between emotions and the colors and music in VR. Results revealed that background color significantly influenced coffee liking when only color or music was manipulated. When color and music were simultaneously used, there was an overall effect of music valence on coffee sweetness and an interaction effect of color and music on liking.	Inconclusive.
Parsons and Courtney [83]	General VR Research	50	eMagin Z800	Objective: HR, RR, GSR	V, A, H	Two conditions were tested to evaluate the differences in psychophysiological responses associated with the completion of an affective task alone vs. the completion of an affective task that also included a Stroop task. Results showed that the increased cognitive workload was associated with more cognitively challenging Stroop conditions, resulting in increased SCL, HR, and RR response levels.	Inconclusive.
Munyan, Neer, Beidel, et al. [84]	Therapy	60	Oculus Rift DK2	Mixed (Objective: EDA; Subjective: STAI)	V, A, S	This paper explored whether olfactory stimuli could increase presence, physiological arousal, and anxiety in VR. Results showed a positive effect of smell on presence (when this stimulus was removed, the sense of presence decreased). No influence of olfactory stimuli on anxiety or arousal was found.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Saladin, Brady, Graap, et al. [85]	Therapy	11	VFX3D HMD	Objective: GSR + HR	V, A, H	The authors developed a 3D-VR immersive environment for cocaine consumption (with various content scenes), to evaluate its ability to elicit subjective craving and cue reactivity (emotional responding, HR, and SCL) in crack cocaine-dependent individuals. Immersion in VR provided more craving for cocaine-related content scenes than neutral scenes. Stimulus richer VR environment is efficient in eliciting craving and physiologic reactivity.	Inconclusive.
Schweizer, Renner, Sun, et al. [86]	Therapy	66	TriVisio VR Vision (3D)	Mixed (Objective: GSR; HR; Subjective: VAS)	V, A, S	A study was conducted to find the differences in the induced analogue trauma symptoms between multisensory analogue trauma by either VR or Script-Driven Imagery (SDI) versus a neutral condition. VR and SDI induce psychophysiological stress responses, coping behavior, and intrusive memories. Psychophysiological arousal tended to be higher for VR. They both offer a novel way to study trauma exposure.	Inconclusive.
van't Wout, Spofford, Unger, et al. [87]	Therapy	43	eMagin Z800	Objective: EDA	V, A, S, H	In this study, two situations of VR combat scenes were compared to examine the physiological responses of veterans with and without PTSD: combat-related events and non-combat classroom events. Overall, PTSD veterans showed increased emotional arousal compared to those without PTSD. Veterans with PTSD reported larger SCL across VR combat events but not for non-combat events, compared to veterans without PTSD. Nevertheless, veterans with or without PTSD showed a similar reduction of emotional arousal to repeated combat scenes.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Wilson and McGill [88]	Games	16	PlayStation VR (PS VR)	Subjective: STAI + Interview	V, A, H	This paper explored the impact of interactive experiences with violent videogames. Higher values of presence, body ownership, and involvement were found for PlayStation VR condition, considering the survival horror game "Resident Evil 7" compared to the TV version. No differences were reported between the two formats regarding fear. The effects of immersion lasted longer (after leaving the experience) for the PS VR than for the TV condition.	Inconclusive.
Junker, Hutters, Reipur, et al. [89]	Therapy	48	Not Specified	Subjective: Interview	V, A, H	A virtual womb was developed to allow users to view, hear and feel a fetus. Three groups tested the application: Miscellaneous participants found the experiment "calm", "secure", and "comfortable"; Expert participants found it "untraditional", "powerful", and complex"; Pregnant participants found it "meaningful", "relevant", "exciting", "stimulating" and "innovative". Overall, all groups had a positive experience.	Inconclusive.
Salgado, Flynn, Naves, et al. [90]	General VR Research	57	Oculus Rift DK1	Mixed (Objective: HR, EDA; Subjective: SAM)	V, A, H	This paper assessed the quality of experience of an immersive haptic-based VR wheelchair simulator, including exploring users' emotions, cognitive task load, user expectations, and presence. Tests were performed in three configurations: desktop (non-immersive), headset1 (immersive with a high rate of motion acceleration), and headset2 (immersive with a lower rate of motion acceleration). In general, results revealed higher arousal, immersion, and usability scores for both headset groups compared to the desktop configuration.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Haraguchi and Kitazaki [91]	General VR Research	18	Valve Index	Subjective (VAS)	V, A, H	The authors explored the perception of pleasant touch modulated by the velocity of a tactile brushing. Results demonstrated higher levels of pleasure when the tactile and visual velocities were identical.	Inconclusive.
Archer, Bluff, Eddy, et al. [92]	Games	22	PlayStation VR	Mixed (Objective: HR, EDA, body temperature; Subjective: SAM)	V, A, S	This paper examined the influence of simulated odors in VR on the user experience considering a horror game (Resident Evil). In general, the addition of odors significantly affected the participants' psychological and physiological responses. However, only the first exposure to an odor was found to be effective on the user experience. *The presence of unpleasant and intense odors (rotten food, smoke, and a rotting head) increased participants' anxiety.*	
Cornelio, Dawes, Maggioni, et al. [93]	Food Research	32	Oculus Quest 2	Subjective: SAM	V, A, T	The authors explored the effects of some visual aspects of food (coloring (blue and red) and shape (rounded or spiky)) on people's taste and expectations. Results suggested that both color and shape might influence the taste of sweetness. Also, some clues on its influence on peoples' arousal, valence, and dominance ratings were found. *Regarding the neutral samples, a significant effect of blue and red light was found to predict lower valence and dominance levels. Regarding the sweet samples, red lighting predicted reduced valence levels, and blue lighting was found to predict low arousal levels*. The effects of shape were neither significant for valence, arousal, or dominance.	

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Colla, Keast, Mohebbi, et al. [94]	Food Research	80	HTC Vive	Subjective: Affect Grid, Interview	V, A, T	This study aimed to compare the sense of presence, liking, beverage desire, consumption, and choice, considering the taste of 5 beverages, between four conditions: a real-life setting, a sensory laboratory with no distractions (control condition), a sensory laboratory with printed pictures being shown at the participant's eye level (evoked context), a 360° video presented with a VR headset (360VR), and real-life condition. Results show that the 360VR and the evoked contexts provided the users with more sense of presence in a café. Regarding liking measures, 360VR and the control condition showed similar results. Considering choice and desire, no significant differences were found. For the affective mood states, no significant differences were found between the conditions.	Inconclusive.
Tamtama, Santoso, Wang, et al. [95]	Tourism	15	Oculus Quest 2	Mixed (Objective: EDA, HR; Subjective: Interview)	V, A, H	Neutral, cold, and warm thermal conditions were tested in a VR experiment to understand their impact on emotional responses, considering virtual dark tourism context. The stimuli were presented on the participant's left hand's thenar eminence. Some relationships between the stimulus and the users' emotional responses were found. *The warm condition triggered more arousal and valence than cold. The cold was perceived to provide more relaxation and less arousal. Auditory stimuli triggered pity, empathy, sadness, nervousness, fear, self-awareness, and horror.*	
Kampa, Finke, Stalder, et al. [96]	Therapy	100	HTC Vive Pro	Mixed (Objective: HR, SCL, BP, RR; Subjective: VAS)	V, A, H	This study encompasses the validation of a multisensory VR environment to support relaxation in acute stress. Based on the objective and self-reported measures, the authors could prove the efficacy of the proposed IVE.	Inconclusive.

Table 2 (continued)

Study	Area	Methods and Instruments			Stimuli	Context and Main Findings	Relationship between specific stimulus and users' emotional responses
		Sample size (N=)	VR system	Measurement of the users' emotions			
Melo, Coelho, Gonçalves, et al. [97]	Tourism	80	HTC Vive	Subjective: Specifically designed questionnaire	V, A, S, H	Two VR setups (audiovisual and multisensory) were compared to examine their impact and the impact of participants' gender on their sense of presence, satisfaction, emotions, and attitudes. Results showed women scored significantly higher spatial presence in general VR and specifically more involvement and overall presence in the audiovisual condition. Regarding emotion, no statistically significant differences were found between the VR setup and the user's gender.	Inconclusive.

3.3 Studies' relevance

As explained in section 0, to assess the relevance of the studies, three topics were examined (technical soundness, sample size, and general relevance). The mean rating for the fully analyzed papers (N=37) was 6.92, and the modal was 7. According to Fig. 2, the majority of the selected studies (N=26, approximately 70%) are high-relevance papers, considering having a score greater than or equal to 7. The other studies (N=11, approximately 30%) consist of medium-relevance papers. No low-relevance papers were found. These results indicate good methodological robustness.

3.4 Quantitative analysis

3.4.1 Yearly Scientific production

A chart was created to understand the study distribution according to the research area. By looking at Fig. 3, it is possible to notice an increase in the number of studies regarding users' emotional responses in multisensory immersive VR (with at least 3 stimuli provided) starting in 2015, which is in line with conclusions from Brooks, Lopes, Amores, et al. [98]. Since then, every year, at least two studies have been published. This result can be supported by the democratization of VR, mainly for the gaming industry, with the release of the first more efficient and affordable VR headsets in 2014 [99]. We highlight the only study published before 2015, conducted by Saladin, Brady, Graap, *et al.* [85] in 2006, that does not match this reasoning. Also, the years 2019 and 2021 are not in line with the increasing number of results since 2015. Although for 2019 we cannot find any reason, 2021 might have been atypical due to the high restrictions caused by the pandemic of Covid-19 that year and the year before (considering that some of the papers written in 2020 would be published in 2021).

Finally, Fig. 3 also allows concluding that therapy (N=11, approximately 30%) and general VR research (N=10, approximately 27%) are the two main research areas explored over time. Less emphasis was given to food research (N=7), games (N=6), tourism (N=2), and marketing (N=1).

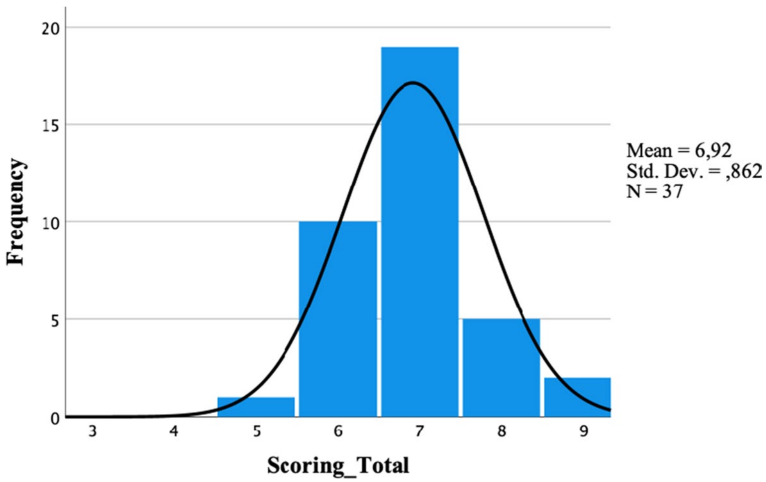

Fig. 2 Relevance assessment (Scores Histogram)

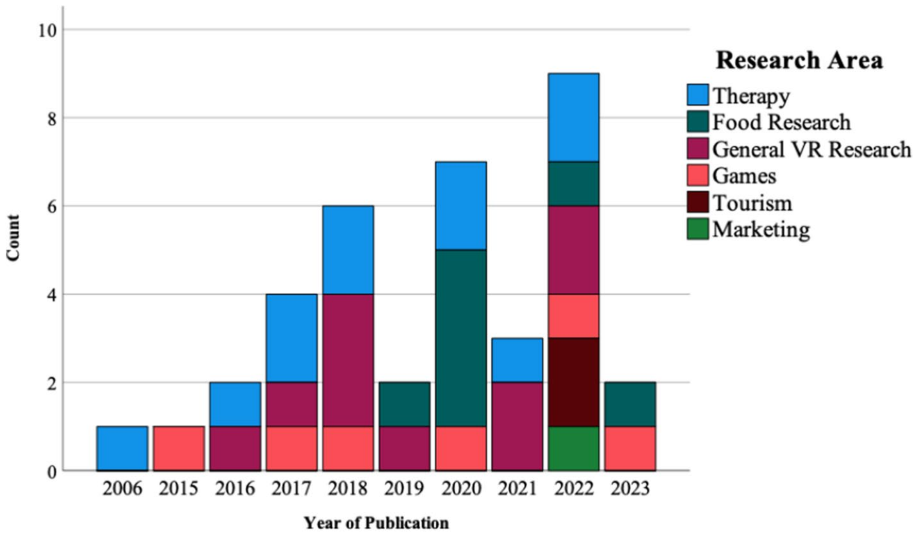

Fig. 3 Distribution of studies according to the research area

3.4.2 Methods and instruments for measuring users' emotional responses

To understand the most common methods and instruments to measure users' emotional responses during multisensory immersive VR experiences, we defined variable sets for subjective and objective measures using SPSS®, to analyze them individually or as a group. Subjective measures were composed of 13 variables, i.e., the analyzed studies used 13 subjective instruments for users' self-report of their emotional responses (Table 3). Objective measures comprised 7 variables, meaning the analyzed studies used 7 objective measures to assess the users' emotional responses (Table 4).

A crosstabulation between the research area and the used methods was built to analyze the most common methods for studying users' emotional responses in immersive multisensory VR (Table 5). This table allows concluding that the majority of the studies (21 studies, equivalent to 56,7%) used a subjective measure, either exclusively (9 studies, corresponding to 24,3%) or as part of a mixed approach (12 studies, corresponding to 32,4%). In turn, objective measures were the least used (4, equivalent to 11%). 16 studies (equivalent to 43,2%) presented objective instruments in their methodologies, either individually (4 studies, corresponding to 10,8%) or as part of a mixed approach (12 studies, corresponding to 32,4%). According to the combined analysis of Tables 3 and 5, the most common method used to measure users' emotional responses in multisensory immersive VR experiences is subjective, independent of the research area, and mainly resorting to specifically designed questionnaires (17,4%), i.e., custom questionnaires developed by the authors, and Interviews (17,4%). With the simultaneous analysis of Tables 4 and 5, we can observe that objective measures are the least used, and the most recurrent instruments were GSR (38,7%) and HR (35,5%), whether measured through heartbeats (HR), its monitoring (HRM), or its variability (HRV).

Table 3 Frequency of the instruments used as subjective measures

Subjective Instruments	Percent	Percent of cases
Specifically designed questionnaire	17,4%	24,2%
Interview	17,4%	24,2%
Self-Assessment Manikin	13%	18,2%
State Anxiety Inventory	10,9%	15,2%
Visual Analog Scale	10,9%	15,2%
Check-All-That-Apply Procedure	8,7%	12,1%
Positive and Negative Affect Schedule	6,5%	9,1%
Fright Reaction self-reported Survey	4,3%	6,1%
Fear Survey Schedule	2,2%	3%
Behavior Scale	2,2%	3%
Modified version of the Igroup Questionnaire	2,2%	3%
General Mood Scale	2,2%	3%
Affect Grid	2,2%	3%

Table 4 Frequency of the instruments used as objective measures

Objective Instruments	N	Percent	Percent of cases
Galvanic Skin Response (or Skin Conductance Level or Electrodermal Activity)	12	38,7%	80%
Heart Rate (or Heart Rate Monitoring or Heart Rate Variability)	11	35,5%	73,3%
Respiration rate	3	9,7%	20%
Pupil size	2	6,5%	13,3%
Electromyography	1	3,2%	6,7%
Body Temperature	1	3,2%	6,7%
Blood Pressure	1	3,2%	6,7%

Table 5 Association between the research area and the used methods for measuring users' emotional responses

Research Area	Used methods for the measurement of users' emotions			
	Subjective	Objective	Mixed approach	Total
Therapy	4	2	5	11
Food Research	7	0	0	7
General VR Research	4	2	4	10
Games	4	0	2	6
Tourism	1	0	1	2
Marketing	1	0	0	1
Total	21	4	12	37

3.4.3 Association between the used method and the research area

Goodman and Kruskal's λ was run to determine whether the used method for measuring the users' emotions (dependent variable) could be better predicted by the research

Table 6 The association between the used method and the research area (Goodman and Kruskal's λ)

		Value	Asymptotic Standard Error	Approximate T	Approximate significance
Lambda	Symmetric	0,095	0,122	0,762	0,446
	Research area Dependent	0,115	0,120	0,915	0,360
	Used methods for the measurement of users' emotions Dependent	0,063	0,182	0,334	0,739
Goodman and Kruskal tau	Research area Dependent	0,075	0,033	.	0,198
	Used methods for the measurement of users' emotions Dependent	0,182	0,055	.	0,217

area (independent variable) – Table 6. Goodman and Kruskal's λ was very close to 0 (p=0,063), meaning there is no association between the variables [100].

3.4.4 Limitations on the study of emotional responses in multisensory VR

To answer research question 2, we analyzed the main limitations pointed out by the author(s) whenever they mentioned them. This data was summarized and classified into Tables 7, 8 and 9, respectively, according to the limitations regarding the used VR equipment (RQ 2.1), the used instruments for emotional responses collection (RQ 2.2), and the experimental design (RQ 2.3). The Discussion (Section 4.2) will further address them in more detail. For each limitation, the studies reporting it are identified. According to the authors' point of view, the context or justification for such limitation is indicated when necessary.

Table 7 The main limitations regarding the used VR equipment

Limitations	Studies
Intrusiveness (N=4)	Sometimes HMDs are uncomfortable and heavy-weight [94]. VR headsets limit the precise assessment of the tasting sample's appearance [75, 78, 94], and impede the user from tasting it without removing the headset; also, the user has to remove the VR headset after each tasting experience to answer a questionnaire [78, 94]; VR headsets might hamper or bias the assessment of behavioral reactions, such as facial expressions or gestures [77].
Social interaction inhibition (N=1)	In some VR studies, social interaction with real people is fundamental to simulating real-life scenarios, like drinking a beer, which in VR is not as real as in actual life [78].

Table 8 The main limitations regarding the used instruments for emotional responses collection

	Limitations	Studies
Considering objective instruments (N=4)	Bias caused by external factors	The test condition, the technological material, or the presentation order [80], might have contributed to obtaining inconsistent emotional responses not related to the tasting experience itself.
	Bias caused by the intrusiveness	EEG setup can be intrusive, although nowadays, there are some more recent low-cost technology, wireless, and wearable EEG devices [77].
	Bias caused by the sensitivity	GSR can be influenced by the room temperature conditions [77]; Physical movements might affect data collection of HR [29] and EDA [55]; "Empatica E4" wearable, used to collect EDA, does not work well consistently throughout the entire time of use (some occasional disconnections and server failures were reported) [55].
Considering subjective instruments (N=2)	The user's self-report of emotions	Especially when it comes to emotional responses, self-reports are subjected to response bias [55, 94], which can be due to the social desirability and ideas of role models [55]; The use of more than one subjective instrument might result in a data mismatch [55].
Considering a mixed approach (N=1)	The mismatch between the results	Using a mixed approach might result in a mismatch between the instruments used. For instance, Salgado, Flynn, Naves, et al. [90] reported conflicting results of the data obtained with the application of the objective instruments HR+EDA and the subjective instrument SAM.

Table 9 The main limitations regarding the experimental design

Limitations	Studies
Many authors pointed out the absence of a mixed approach (N=9)	Several authors reported this limitation [10, 19, 28, 31, 55, 79, 81, 82, 85]. As Steinhaeusser, Eckstein and Lugrin [55] concluded, integrating objective measures (e.g., HR) can mitigate the eventual conflict between subjective data.
The non-use of other complementary measures from the same method used (objective or subjective) (N=2)	The use of other physiological measures, such as GSR to complement HR and pupil size could be helpful [7], as well as the use of neuroimaging to understand the effects of the Stroop Task on specific brain areas, complementing HR, GSR, and RR [83].
To analyze the stimulus and its effects individually should have been considered (N=2)	Dey, Chen, Billinghurst, et al. [29] emphasize the importance of understanding what levels of manipulation are most effective for each key emotion; Chin, Thompson and Ziat [76] consider it beneficial to test each stimulus individually to compare modal vs. multimodal effectiveness.
Arousal ratings might be biased by the lack of images that induce intense emotional reactions (N=1)	Due to ethical reasons, the authors did not use such virtual content [11].
Filling out the questionnaire was done a long time after the experience (N=1)	The authors explain that this gap can bias the results [65].
Incongruency between the stimuli (N=3)	Colla, Keast, Mohebbi, et al. [94] recognize that the incongruency between the visual and the taste stimuli, for instance, may have contributed to inconsistent data. Also, the uncertainty regarding consistency between smell and the audiovisual stimuli was addressed by Brengman, Willems and De Gauquier [31] in terms of the IVE effectiveness, and by van Veelen, Boonekamp, Schoonderwoerd, et al. [74], who stated that the incongruency between the olfactory and tactile stimuli caused a disturbance on the users' experienced presence.
The integration of a large number of sensory stimuli represents a challenging task (N=1)	The more stimuli included in an IVE, the more complex the experience development and user evaluation process becomes [31].

4 Discussion

This work aimed to overview the methods and instruments used to assess users' emotional responses during multisensory virtual experiences exploring the addition of at least one more sense to the base pair vision and audio. A total of 37 articles matched the defined criteria and were thoroughly analyzed. Our main results were distributed along Section 3 (Results), which is the core information for the following discussion and conclusions. This section will be organized into topics that will answer the proposed research questions to simplify.

4.1 Methods and instruments for measuring users' emotions during multisensory VR experiences

To answer RQ 1, an analysis of the most used instruments was done (Section 3.4.2). As revealed, subjective measures are the most common to assess users' emotional responses in multisensory immersive VR experiences, representing 56,7% of the methods used in the analyzed studies. Specifically, researchers tend to assess the subjective emotional responses by resorting to questionnaires (17,4%) and interviewing the participants (17,4%). We believe that one possible reason for this result is related to the more accessible and less expensive implementation process, compared to the objective measures, as previously discussed. Also, when obtaining subjective reports, researchers gain a more comprehensive understanding of the relationship between emotions and performance in VR, which might be another explanation for this result.

Among the objective instruments, GSR remains the most popular in emotional experiments, according to our results, which is also supported by recent statements from Hosany, Martin and Woodside [54]. However, there is still a controversy in the literature regarding its power and accuracy to measure emotional responses, considering, for instance, the reported mismatch between EDA and the subjective method SAM [90]. This is contradicted by the results of Kaminska, Smolka, Zwolinski, *et al.* [77], who have proven SCL to have the highest correlation with the participants' subjective reports regarding the assessment of stress level and mood [77]. Also Felnhofer, Kothgassner, Schmidt, et al. [4] concluded that, due to not discriminating different affective states, SCL might not be the best instrument to measure emotions, which is in line with IJsselsteijn [101], who believes HR may be a better indicator.

Despite our results showing a trend in the use of subjective measures, using a mixed approach for assessing the users' emotional responses in multisensory immersive VR experiences has been recommended by several authors [10, 19, 28, 31, 55, 79, 81, 82, 85]. This idea is also supported by Johnson and Onwuegbuzie [102], who stated that a mixed approach might result in better validity and more robust conclusions [103]. However, our literature review does not allow us to predict a clear advantage of resorting to it, as it is still very incongruent whether its usage represents a benefit or a disadvantage. On the one hand, if the data between the two methods present similar results, the mixed approach will understandably bring more robustness to the results. However, supposing a large discrepancy between objective and subjective data, using a mixed approach is less beneficial than using only one objective or subjective measure. So, a previous study should be done by researchers to define which method and instruments are more advantageous to achieve the investigation goals. Nevertheless, we could not find in any of the studies a discussion regarding the reason why the researchers had resorted to that instrument or measure. That is, the pros and cons are not adequately balanced when choosing the method to assess users' emotional responses. Our conclusion concerning this topic is that using a mixed approach does not certainly bring more reliable results. Instead, researchers must equate the pros and cons of each approach (subjective, objective, or mixed), according to the specific VR experiment. Further discussion on this topic will be found in Conclusions (Section 5).

4.2 Limitations found regarding the study of users' emotional responses in multisensory VR experiences

According to the authors' reports, we concluded that several pointed out the same limitations, although in different research areas and contexts, which we believe can contribute as an alert to ameliorating future research. These results, corresponding to RQ 2, were synthesized into Tables 7, 8 and 9, and will be discussed in Sections 4.2.1 regarding the used VR equipment, Section 4.2.2 regarding the instruments for emotional responses collection, and Section 4.2.3 regarding the experimental design.

4.2.1 Limitations regarding the used VR equipment

The limitations regarding the VR equipment (RQ 2.1) were classified according to its intrusiveness and social interaction inhibition, independent of the used method (subjective or objective). Generically, our results demonstrate that the VR headset is mainly intrusive for tasting experiences. From a researcher's perspective, VR equipment might hamper the assessment of the users' emotional responses, mainly due to covering a significant part of the face, which avoids assessing behavioral reactions, such as facial expressions or gestures [77]. From a user's point of view, the fact that participants cannot precisely assess the appearance of the tasting sample [75, 78, 94] will affect their experience and, consequently, their reported feelings. Considering that the users had to remove the VR headset after each part of the experience to answer the questionnaire was also pointed out as a limitation [78, 94]. We highlight the use of in-VR questionnaires, which has become more relevant in VR experiments, as users can answer any questionnaire without taking out the headset and not breaking their levels of presence [104, 105], e.g., the VR Questionnaire Toolkit from Feick, Kleer, Tang, et al. [106]. It contributes to alleviating the transition from VR to the real world, which, in turn, contributes to maintaining the user's immersion and presence [104, 107], and reducing disorientation [105].

Finally, VR equipment was pointed out by Sinesio, Moneta, Porcherot, et al. [78] to inhibit social interaction, which the authors consider fundamental to simulating real-life scenarios, despite the eventual presence of avatars.

4.2.2 Limitations regarding the used instruments for emotional responses collection

The primary limitations regarding the instruments to collect users' emotional responses (RQ 2.2) were distinguished by their subjectiveness, objectiveness, or the use of a mixed approach. Objective instruments are subjected to bias caused by external factors, and the hardware's intrusiveness and sensitivity. Worch, Sinesio, Moneta, et al. [80] commented on the negative impact of external factors such as the test condition, the technological material, or the presentation order to obtain more accurate results. Another limitation worth noting is how intrusive hardware can be, such as EEG [77], which can be overpassed by resorting to other objective methods or less invasive EEG hardware. Finally, regarding the hardware's sensitivity, the interference of the room temperature on the GSR data collection process was pointed out by Kaminska, Smolka, Zwolinski, *et al.* [77]; physical movements can affect HR [29] and EDA [55], which provides little exploratory interaction between the user and the VR scenario [29]. Specifically, the "Empatica E4" wearable used to collect EDA data was reported not to work trustworthily [55].

Subjective instruments denote a self-report, which is considered a limitation per se. One's perception of the own emotional responses is an arduous task, highly subjected to bias [55, 94], for instance, due to the social desirability and ideas of role models [55]. So, although using more than one subjective measure can be a good method to validate emotional responses, it also can result in a mismatch between the data collected [55].

The same occurs when using a mixed approach, i.e., the mismatch between subjective and objective data, as reported by Salgado, Flynn, Naves, et al. [90].

4.2.3 Limitations regarding the experimental design

The most reported limitation considering the experimental design (RQ 2.3) was the absence of a mixed approach, independently of the use of a subjective or an objective approach [10, 19, 28, 31, 55, 79, 81, 82, 85]. It means that when researchers used an objective approach, a subjective method would be necessary to complement results and vice-versa. Although rare, some researchers suggest using additional instruments of the same method to validate the results [7, 83]. However, as discussed in Section 4.1, our results do not allow us to conclude that a mixed approach will necessarily result in more solid, consistent, or rigorous outcomes, considering the potential risk of inconsistency between the subjective and the objective data.

Concerning the investigation of multisensory stimuli, as we will address later, some authors claimed that the stimulus effects should be explored individually to compare their modal effectiveness *vs.* multimodal [76], and to understand what levels of manipulation are most effective in inducing specific emotions [29].

Although sporadic, it was also pointed out as a limitation the absence of intense emotional images due to ethical reasons, which could have contributed to inducing stronger emotional responses in the user [11]. For instance, this limitation could be reduced by using an affective picture system, such as the International Affective Picture System (IAPS), which allows more control over the emotional responses elicited in the user. Such a system includes a great variety of emotion-eliciting photographs, such as nature, erotic scenes, weapons, animals, etc. [108]. Also, one of the most cited and frequently used [4], the "Velten Mood Induction Procedure" could be a good option, as it combines photo material with music [109].

It was commented that the questionnaires had been filled out long after the experience led the users to report what they remembered rather than what they actually felt [65]. This limitation allows concluding the importance of applying the questionnaires as soon as possible after the experiment and, ideally, to create short VR experiences to avoid a long period between the experiment's beginning and the questionnaire's fill. As previously discussed, this issue could be overpassed by resorting to in-VR questionnaires.

Another considerable limitation reported by some authors [31, 74, 94] regarding the experimental design of multisensory VR systems is the incongruency between the stimuli. This problem has been increasingly highlighted in the literature. Many authors have shown that VR users subjected to incongruent stimuli tend to report several negative feelings, such as a disturbance in presence [74], the sense of disownership [110, 111], the loss of agency, and the sensation of being out of ones' own body [111].

Finally, Brengman, Willems and De Gauquier [31] commented that multisensory VR systems are, in general, hard to implement and evaluate, but even harder the more stimulus they include. Indeed, considering the essential requisite of stimuli congruency, the several obstacles of implementing and assessing certain human senses such as taste and smell, and all other previously discussed limitations, designing multisensory VR systems requires significant effort and expertise. All these limitations contribute to understanding why this topic has been so poorly explored.

4.3 The relationship between the stimulus provided to the users and their emotional response(s)

One of this research's main goals was to find a relationship between the stimulus provided to the users and their emotional response (RQ 3). Unfortunately, only a few studies explored such association, furthermore, generically. As Dey, Chen, Billinghurst, et al. [29] and Chin, Thompson and Ziat [76] noted, stimuli should be explored individually to understand their isolated effects on users' emotions. Regarding the individual effect of visual stimulus, results from Wu, Weng and Xue [41] confirm that its correlation with emotions is greater than haptic or auditory stimuli. In turn, the haptic stimulus had a greater influence on users' emotions than the auditory. Yet, none of these conclusions provide concrete information, such as which emotions are triggered according to each specific stimulus. An attempt to achieve such a clear conclusion was made, for instance, by Torrico, Han, Sharma, et al. [75], who found that bright-VR was associated with "free", "glad", "aggressive", and "enthusiastic" emotions, considering different light conditions in taste perception. Dark-VR was associated with "nostalgic" and "daring". Similarly, Cornelio, Dawes, Maggioni, et al. [93] found a significant effect of blue and red lighting to predict lower valence and dominance when the users tasted neutral samples. When the users tasted sweet samples, red lighting was found to reduce valence, and blue lighting predicted low arousal levels. Dey, Chen, Billinghurst, et al. [29] demonstrated that visually manipulated HR feedback could significantly enhance "interest", "excitement", "scariness", "nervousness", and "fear" in VR.

Regarding the auditory stimuli, Tamtama, Santoso, Wang, et al. [95] demonstrated that the presence of it in a dark tourism context triggered the negative feelings of “pity”, “empathy”, “sadness”, “nervousness”, “fear”, “self-awareness”, and “horror”. Work from Kruijff, Marquardt, Trepkowski, et al. [10] revealed significant correlations between a “bee swarm” (normal sound condition) and “happiness”, while “bad weather situation” (low-frequency sound) and “zombie swarm” (low frequency and normal sound) correlated to “surprise”. No more auditory correlations were found between sensorial stimulus and the users' emotional responses. For smell correlations, the same authors [10] found that the smell provided in the “sea view situation” happiness to the user. Still regarding smell, unpleasant and intense odors demonstrated a positive association with anxiety [92].

Regarding taste, only one study was found relating the tasting sample of milk, white and dark chocolate to several positive and negative emotions [62].

For the haptic stimulus, Kruijff, Marquardt, Trepkowski, et al. [10] found a correlation with “surprise” (in the "spider behind the back" situation). Also, the back vibration during the “bad weather situation” and the “zombie swarm” situation caused “surprise”. Kono, Miyaki and Rekimoto [18] found that the haptic feedback from EMS or Solenoid induces “fear” and “pain”. Considering thermal feedback, warm condition triggered more arousal and valence than cold, which, in turn, was perceived to provide more relaxation and less arousal [95].

To summarize, none of the obtained results allow concluding a pattern on which specific stimulus triggers each emotion, as they all present vague or dispersed results unrelated to each other. This limitation will be further addressed in Sections 4. 5 and 5.

4.4 The association evidence between the used method and the research area

Goodman and Kruskal's λ was run to verify the existence of an association between the used method and the research area (RQ 4). The result showed no association between these two variables. Still, there should be considered the fact that little research has been published so far related to the assessment of users' emotional responses in multisensory immersive VR experiences, which is reflected in the low number of studies analyzed in this literature review.

4.5 Gaps and Challenges

In this paper, one of our goals was to address the main gaps and challenges remaining for future research (RQ 5). First, as previously stated, our results demonstrate a significant gap regarding the study of users' emotions in multisensory immersive VR experiences, specifically in understanding which stimuli evoke certain users' emotional responses, regardless of the research context. As Kruijff, Marquardt, Trepkowski, et al. [10] highlighted, "there is hardly any evidence that supports which combinations of stimuli can trigger specific emotional responses", which our paper corroborates six years later. Indeed it is still impossible to achieve a pattern or a set of guidelines regarding the immersive design aspects within an IVE to stimulate specific emotional responses in a certain context (positive, negative, or neutral). We believe that, in the next few years, with the constant development of technology, specifically VR, a significant number of studies regarding this subject will appear, which is in line with our results regarding the distribution of studies by year of production (Section 3.4.1) and the findings from Brooks, Lopes, Amores, et al. [98]. Thus, it will provide the chance to obtain more concrete results in this field. Although some exploratory experiments have been done to understand the correlation between a specific stimulus and the triggered emotional responses, for instance, according to light conditions [75, 93], auditory stimulus [10], haptic thermal feedback [95], the studies we analyze were, in general, recent, revealing that there is still a long way to go in this regard.

For future research, we recommend comparing our results considering less immersive VR environments, i.e., the inclusion of unisensory and two-stimuli environments, which our work does not encompass. Such investigation could provide more robustness and consistency regarding the association between the input stimuli and the obtained emotional response. Finally, it is also worth noting that other results could have been obtained using a different VR systems classification (e.g., Slater's classification of immersion in terms of sensorimotor contingencies [112]). So, it could be pertinent for future work to consider such dimensions.

5 Conclusions

This paper reviewed the relationship between multisensory information in VR and the users' emotional responses, considering papers exploring at least three different stimuli, i.e., considering the addition of at least one more sense to the base pair vision and audio. Although this exclusion criterion might be seen as reductive, as unimodal and bimodal VR experiences can also provide valid information regarding emotions, this study contributed to understanding the significant lack of knowledge remaining on this subject. It also acts as an alert to the urgency of more research stimulating more than just two human senses. We consider that the existence of so little information in this regard so far derives from the complex task for researchers to simulate and evaluate other than audiovisual stimuli, as previously addressed. Also, lately, the focus of VR research has been the portable mass-market VR, which is not as compatible with extra hardware, and makes it harder to explore other senses, such as smell and taste. Lastly, due to restrictions caused by the COVID-19 pandemic, fewer studies with VR in general than expected were conducted, which indirectly limited this paper's results and conclusions.

Indeed, if the fourth exclusion criterion had not been included, at least 251 more articles would have been added to this study. Although results originated from that investigation could have provided more consistent results, we believe that the future of VR is the multisensory experiences providing as many sensorial inputs as possible. As previously discussed, the more senses stimulated, the more immersive and emotionally powerful the IVE becomes, which justifies the importance of adding more than just two stimuli to a VR experience. The rapid evolution of technology and VR in particular, which can be proven, for instance, by the noticeable increase in publication rates regarding the less explored stimuli so far (smell and taste) [98], also contributes to reinforcing such conviction. When the focus on such studies is more frequent, it will be possible to clearly outline the association between specific stimuli and the triggered emotions.

Finally, this literature review allowed us to conclude, contrary to what some authors defend, that the use of a mixed approach to measuring users' emotional responses in VR does not clearly guarantee more reliable results, mainly considering the potential risk of biased results caused by the mismatch between the objective and subjective data. So, the researcher should ponder in advance the inherent risk of using each approach. Many aspects of the VR system can cause the users to feel more cybersickness. For instance, as enumerated by Bockelman and Lingum [113], technical aspects (e.g., the type of display and its comfort of design), or, regarding the visual experience (e.g., the navigation speed, the scene oscillation, the 3D experience), and the hardware interaction, contribute to cybersickness symptoms. Considering an investigation that recognizes the existence of certain related risks that contribute to a greater feeling of cybersickness (for example, the use of an HMD with low resolution, or hardware interaction delay), and taking into account that the individuals tend to respond according to what they believe researchers expect (one of the behaviors of the so-called "good-subject effect" [114], which, in this case, refers not to feeling cybersickness), then, inevitably, the objective data is likely to contradict the users' subjective responses. That is, objective data is likely to demonstrate a negative emotional experience caused by the negative symptoms of cybersickness, contrary to the user's subjective report, which might reveal a biased, and, consequently, a different result caused by the good-subject effect. These are examples of eventual situations to which only objective data should be applied, as there is a high risk of mismatch between objective and subjective measures. On the other hand, if these conditions are not expected, and it is predictable to have low levels of cybersickness (which can be anticipated, for instance, by resorting to pre-tests), a mixed approach should be considered. Finally, supposing the VR experience per se is complex and intrusive, for example, due to the equipment it requires the participant to use (e.g., haptic, smell, or taste interfaces), it might be beneficial not to resort to objective methods to assess the users' emotional responses, as they will cause even more intrusiveness and, consequently, affect the collection of objective data.

Acknowledgments This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme and by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project POCI-01-0145-FEDER-030740. The work of Mariana Magalhães was also supported by the Portuguese Foundation for Science and Technology (FCT) under Doctoral Grant PD/BD/150491/2019.

Funding Open access funding provided by FCT|FCCN (b-on).

Data availability All data generated or analyzed during this study are included in this published article.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval The authors assure that this manuscript is a completely original work and that it has not been published elsewhere.

References

1. Baños RM, Botella C, Rubio I, Quero S, Garcia-Palacios A, Alcaniz M (2008) Presence and emotions in virtual environments: the influence of stereoscopy. Cyberpsychol Behav 11(1):1–8
2. Riva G, Mantovani F, Capideville CS, Preziosa A, Morganti F, Villani D, Gaggioli A, Botella C, Alcañiz M (2007) Affective interactions using virtual reality: the link between presence and emotions. CyberPsychol Behav 10(1):45–56
3. Brivio E, Serino S, Negro Cousa E, Zini A, Riva G, De Leo G (2020) Virtual reality and 360° panorama technology: a media comparison to study changes in sense of presence, anxiety, and positive emotions. Virtual Reality 25(2):303–311
4. Felnhofer A, Kothgassner OD, Schmidt M, Heinzle A-K, Beutl L, Hlavacs H, Kryspin-Exner I (2015) Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. Int J Human-Comput Stud 82:48–56
5. Bindman SW, Castaneda LM, Scanlon M, Cechony A (2018) Am I a Bunny? The impact of high and low immersion platforms and viewers' perceptions of role on presence, narrative engagement, and empathy during an animated 360° video. In: Proceedings of the 2018 CHI conference on human factors in computing systems. Montreal, pp 1–11
6. Mania K, Chalmers A (2001) The effects of levels of immersion on memory and presence in virtual environments: A reality centered approach. CyberPsychol Behav 4(2):247–264
7. Chen H, Dey A, Billinghurst M, Lindeman RW (2017) Exploring the design space for multi-sensory heart rate feedback in immersive virtual reality. In: Proceedings of the 29th Australian Conference on Computer-Human Interaction. Brisbane, pp 108–116
8. Raheel A, Majid M, Anwar SM (2021) DEAR-MULSEMEDIA: Dataset for emotion analysis and recognition in response to multiple sensorial media. Inform Fusion 65:37–49
9. Masood N, Farooq H (2018) Multimodal paradigm for emotion recognition based on EEG signals. In: Human-computer interaction. Theories, methods, and human issues: 20th International Conference, Las Vegas, pp 419–428

10. Kruijff E, Marquardt A, Trepkowski C, Schild J, Hinkenjann A (2017) Designed emotions: challenges and potential methodologies for improving multisensory cues to enhance user engagement in immersive systems. Vis Comput 33(4):471–488
11. Gall D, Roth D, Stauffert JP, Zarges J, Latoschik ME (2021) Embodiment in virtual reality intensifies emotional responses to virtual stimuli. Front Psychol 12:1–11
12. Mishra A, Shukla A, Rana NP, Dwivedi YK (2021) From "touch" to a "multisensory" experience: The impact of technology interface and product type on consumer responses. Psychol Market 38(3):385–396
13. Hülsmann F, Fröhlich J, Mattar N, Wachsmuth I (2014) Wind and warmth in virtual reality: Implementation and evaluation. In: Proceedings of the 2014 Virtual Reality International Conference, Laval, pp 1–8
14. Hoppu U, Puputti S, Mattila S, Puurtinen M, Sandell M (2020) Food Consumption and emotions at a salad lunch buffet in a multisensory environment. Foods 9(10):1–11
15. Dinh HQ, Walker N, Song C, Kobayashi A, Hodges LF (1999) Evaluating the importance of multisensory input on memory and the sense of presence in virtual environments. In: Proceedings IEEE Virtual Reality, Houston, pp 222–228
16. Gutiérrez MAA, Vexo F, Thalmann D (2008) Stepping into virtual reality. Springer, pp 1–7
17. Cheok AD, Karunanayaka K (2018) Electric taste. In: Vanderdonckt J, Tan D (eds) Virtual taste and smell technologies for multisensory internet and virtual reality, Springer, pp 49–67
18. Kono M, Miyaki T, Rekimoto J (2018) In-pulse: Inducing fear and pain in virtual experiences. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, pp 1–5
19. Steinhaeusser SC, Lugrin B (2020) Horror laboratory and forest cabin - A horror game series for desktop computer, virtual reality, and smart substitutional reality. In Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY ’20). Association for Computing Machinery, New York, NY, USA, pp 80–85
20. Hoyer WD, Kroschke M, Schmitt B, Kraume K, Shankar V (2020) Transforming the Customer Experience Through New Technologies. J Interact Mark 51:57–71
21. Feng M, Lindeman RW, Abdel-Moati H, Lindeman JC (2015) Haptic ChairIO: A system to study the effect of wind and floor vibration feedback on spatial orientation in VEs. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI), Arles, pp 149–150
22. Yanagida Y, Kawato S, Noma H, Tetsutani N, Tomono A (2003) A nose-tracked, personal olfactory display. In: International Conference on Computer Graphics and Interactive Techniques. ACM SIGGRAPH 2003 Sketches & Applications, p 1
23. Narumi T, Nishizaka S, Kajinami T, Tanikawa T, Hirose M (2011) Augmented reality flavors: gustatory display based on edible marker and cross-modal interaction. In: Proceedings of the SIGCHI conference on human factors in computing systems. Vancouver, pp 93–102
24. Ranasinghe N, Cheok A, Nakatsu R, Do EY-L (2013) Simulating the sensation of taste for immersive experiences. In: Proceedings of the 2013 ACM international workshop on Immersive media experiences, pp 29–34
25. Ranasinghe N, Nguyen TNT, Liangkun Y, Lin L-Y, Tolley D, Do EY-L (2017) Vocktail: A virtual cocktail for pairing digital taste, smell, and color sensations. In: Proceedings of the 25th ACM International Conference on Multimedia - MM ’17, New York, pp 1139–1147
26. Brooks J, Nagels S, Lopes P (2020) Trigeminal-based temperature illusions. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI ’20). Association for Computing Machinery, New York, pp 1–12
27. Maggioni E, Cobden R, Obrist M (2019) OWidgets: A toolkit to enable smell-based experience design. Int J Human-Comput Stud 130:248–260
28. Kruijff E, Marquardt A, Trepkowski C, Schild J, Hinkenjann A (2015) Enhancing User engagement in immersive games through multisensory cues. In: 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games), vol 2015, Skovde, pp 1–8
29. Dey A, Chen H, Billinghurst M, Lindeman RW (2018) Effects of manipulating physiological feedback in immersive virtual environments. In: Proceedings of the 2018 annual symposium on computer-human interaction in play, Melbourne, pp 101–111
30. Melo M, Goncalves G, Monteiro P, Coelho H, Vasconcelos-Raposo J, Bessa M (2020) Do multisensory stimuli benefit the virtual reality experience? A systematic review. IEEE Trans Vis Comput Graph 28(2):1428–1442
31. Brengman M, Willems K, De Gauquier L (2022) Customer engagement in multi-sensory virtual reality advertising: the effect of sound and scent congruence. Front Psychol 13:1–20

32. Wang G, Suh A (2021) A literature review on a neuro-psychological approach to immersive technology research. In: Schmorrow DD, Fidopiastis CM (eds) Augmented cognition. Lecture notes in computer science, vol 12776. Springer, pp 97–115
33. Csikszentmihalyi M (1990) Flow: The psychology of optimal experience. Harper & Row New York, New York
34. Botella C, Rey A, Perpiñá C, Baños R, Alcañiz Raya M, Garcia-Palacios A, Villa H, Alozano J (1999) Differences on Presence and Reality Judgment Using a High Impact Workstation and a PC Workstation. Cyberpsychol Behav 2:49–52
35. Chirico A, Cipresso P, Yaden DB, Biassoni F, Riva G, Gaggioli A (2017) Effectiveness of Immersive Videos in Inducing Awe An Experimental Study. Sci Rep 7(1):1218
36. Marin-Morales J, Higuera-Trujillo JL, Greco A, Guixeres J, Llinares C, Scilingo EP, Alcaniz M, Valenza G (2018) Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors. Sci Rep 8(1):13657
37. Marcolin F, Wally Scurati G, Ulrich L, Nonis F, Vezzetti E, Dozio N, Ferrise F, Stork A, Basole RC (2021) Affective Virtual Reality: How to Design Artificial Experiences Impacting Human Emotions. IEEE Comput Graph Appl 41(6):171–178
38. Patrono L, Podo L, Rametta P (2019) An innovative face emotion recognition-based platform by using a mobile device as a virtual tour. J Commun Softw Syst 15(2):1–9
39. Ferreira HA, Saraiva M (2019) Subjective and objective measures. In: Ayanoğlu H, Duarte E (eds) Emotional design in human-robot interaction: Theory, methods and applications, vol 337. Springer, Cham, pp 143–161
40. Cabanac M (2002) What is emotion? Behav Proc 60(2):69–83
41. Wu D, Weng D, Xue S (2016) Virtual reality system as an affective medium to induce specific emotion: A validation study. Electron Imaging. 28:1–6
42. Ekman P (1992) An argument for basic emotions. Cogn Emot 6(3–4):169–200
43. Calvo RA, Mac Kim S (2013) Emotions in Text: Dimensional and Categorical Models. Comput Intell 29(3):527–543
44. Russell JA (1980) A circumplex model of affect. J Personal Soc Psychol 39(6):1161–1178
45. Feng Z, González VA, Amor R, Lovreglio R, Cabrera-Guerrero G (2018) Immersive virtual reality serious games for evacuation training and research: A systematic literature review. Comput Educ 127:252–266
46. Samara A, Galway L, Bond R, Wang H (2017) Affective state detection via facial expression analysis within a human–computer interaction context. J Ambient Intell Human Comput 10(6):2175–2184
47. Bradley MM, Lang PJ (1994) Measuring emotion: the self-assessment manikin and the semantic differential. J Beh Ther Exp Psychiatry 25(1):49–59
48. Williamson K (2018) Questionnaires, individual interviews and focus group interviews. In: Williamson K, Johanson G (eds) Research methods: Information, systems, and contexts, vol 2, San Diego, pp 379–403
49. Gomes WB (1997) A Entrevista Fenomenológica e o Estudo da Experiência Consciente. Psicologia USP 8(2):305–336
50. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: the PANAS scales. J Personal Soc Psychol 54(6):1063
51. Spielberger C, Gorsuch R, Lushene R (1970) Manual for the State-Trait Inventory STAI (Form Y). Mind Garden, Palo Alto
52. Hayes MHS, Patterson DG (1921) Experimental development of the graphic rating method. Psychol Bullet 18:98–99
53. Meyners M, Castura JC, Carr BT (2013) Existing and new approaches for the analysis of CATA data. Food Qual Prefer 30(2):309–319
54. Hosany S, Martin D, Woodside AG (2021) Emotions in tourism: Theoretical designs, measurements, analytics, and interpretations. J Travel Res 60(7):1391–1407
55. Steinhaeusser SC, Eckstein B, Lugrin B (2023) A multi-method approach to compare presence, fear induction and desensitization in survival horror games within the reality-virtuality-continuum. Entertain Comput 100539(45):1–13
56. Meiselman HL (2017) Emotion measurement: Theoretically pure or practical? Food Qual Prefer 62:374–375
57. Prescott J (2017) Some considerations in the measurement of emotions in sensory and consumer research. Food Qual Prefer 62:360–368
58. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 62(10):1006–1012

59. Gusenbauer M, Haddaway NR (2020) Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Res Synth Methods 11(2):181–217
60. Costello P (1997) Health and safety issues associated with virtual reality: a review of current literature, LE11 INL edn. Loughborough University Department of Human Sciences Advanced VR Research Centre, Leicestershire
61. General Information and content coverage (2023) https://ieeexplore.ieee.org/Xplorehelp/administrators-andlibrarians/administrator-faqs
62. Kong Y, Sharma C, Kanala M, Thakur M, Li L, Xu D, Harrison R, Torrico DD (2020) Virtual reality and immersive environments on sensory perception of chocolate products: A preliminary study. Foods 9(4):1–19
63. Wee C, Yap KM, Lim WN (2021) Haptic Interfaces for Virtual Reality: Challenges and Research Directions. IEEE Access 9:112145–112162
64. Kang J, Lee J, Jin S (2018) Personal Sensory VR interface Utilizing Wearable Technology. In: International Conference on Information and Communication Technology Convergence (ICTC). Jeju, Korea, pp 546–548
65. Karafotias G, Korres G, Teranishi A, Park W, Eid M (2018) Mid-air tactile stimulation for pain distraction. IEEE Trans Haptics 11(2):185–191
66. Li RC, Jung S, McKee RD, Whitton MC, Lindeman RW (2021) Introduce floor vibration to virtual reality. In: Proceedings of the 2021 ACM Symposium on Spatial User Interaction (SUI '21). Association for Computing Machinery, New York, Article 21, pp 1–2
67. Singhal Y, Wang H, Gil H, Kim JR (2021) Mid-air thermo-tactile feedback using ultrasound haptic display. In: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology (VRST '21). Association for Computing Machinery, New York, Article 28, pp 1–11
68. Peiris RL, Feng Y-L, Chan L, Minamizawa K (2019) Thermalbracelet: Exploring thermal haptic feedback around the wrist. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, Paper 170, pp 1–11
69. Nilsson N, Zenner A, Simeone A, Johnsen K, Sandor C, Billinghurst M (2021) Propping Up Virtual Reality With Haptic Proxies. IEEE Comput Graph Appl 41:104–112
70. Gonçalves G, Coelho H, Monteiro P, Melo M, Bessa M (2022) Systematic review of comparative studies of the impact of realism in immersive virtual experiences. ACM Comput Surv 55(6):1–36
71. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM (2012) A systematic literature review of empirical evidence on computer games and serious games. Comput Educ 59(2):661–686
72. Spezi V, Wakeling S, Pinfield S, Fry J, Creaser C, Willett P (2018) "Let the community decide"? The vision and reality of soundness-only peer review in open-access mega-journals. J Doc 74(1):137–161
73. Macefield R (2009) How to specify the participant group size for usability studies: A practitioner's guide. J Usability Stud 5:34–45
74. van Veelen N, Boonekamp RC, Schoonderwoerd TAJ, van Emmerik ML, Nijdam MJ, Bruinsma B, Geuze E, Jones C, Vermetten E (2022) Tailored immersion: implementing personalized components into virtual reality for veterans with post-traumatic stress disorder. Front Virtual Reality 2:1–9
75. Torrico DD, Han Y, Sharma C, Fuentes S, Viejo CG, Dunshea FR (2020) Effects of context and virtual reality environments on the wine tasting experience, acceptability, and emotional responses of consumers. Foods 9(2):1–17
76. Chin K, Thompson M, Ziat M (2021) The effect of multimodal virtual reality experience on the emotional responses related to injections. In: VISIGRAPP 2021 - Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, vol 2, pp 128–134
77. Kaminska D, Smolka K, Zwolinski G, Wiak S, Merecz-Kot D, Anbarjafari G (2020) Stress Reduction Using Bilateral Stimulation in Virtual Reality. IEEE Access 8:200351–200366
78. Sinesio F, Moneta E, Porcherot C, Abbà S, Dreyfuss L, Guillamet K, Bruyninckx S, Laporte C, Henneberg S, McEwan JA (2019) Do immersive techniques help to capture consumer reality? Food Qual Prefer 77:123–134
79. Schweizer T, Schmitz J, Plempe L, Sun D, Becker-Asano C, Leonhart R, Tuschen-Caffier B (2017) The impact of pre existing anxiety on affective and cognitive processing of a Virtual Reality analogue trauma. PLoS ONE 12(12):1–19

80. Worch T, Sinesio F, Moneta E, Abbà S, Dreyfuss L, McEwan JA, Porcherot-Lassallette C (2020) Influence of different test conditions on the emotional responses elicited by beers. Food Qual Prefer 83:1–8
81. Kim A, Bae H, Lee K (2019) Effects of tactile perception on emotion and immersion to film viewing in a virtual environment. In: Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology (VRST '19). Association for Computing Machinery, New York, Article 66, pp 1–3
82. Nivedhan A, Mielby LA, Wang QJ (2020) The Influence of Emotion-Oriented Extrinsic Visual and Auditory Cues on Coffee Perception: A Virtual Reality Experiment. Companion Publication of the 2020 International Conference on Multimodal Interaction. Virtual Event, Netherlands, pp 301–306
83. Parsons TD, Courtney CG (2018) Interactions between Threat and Executive Control in a Virtual Reality Stroop Task. IEEE Trans Affect Comput 9(1):66–75
84. Munyan BG, Neer SM, Beidel DC, Jentsch F (2016) Olfactory stimuli increase presence in virtual environments. PLoS ONE 11(6):1–19
85. Saladin ME, Brady KT, Graap K, Rothbaum BO (2006) A preliminary report on the use of virtual reality technology to elicit craving and cue reactivity in cocaine dependent individuals. Addict Behav 31(10):1881–1894
86. Schweizer T, Renner F, Sun DL, Kleim B, Holmes EA, Tuschen-Caffier B (2018) Psychophysiological reactivity, coping behaviour and intrusive memories upon multisensory Virtual Reality and Script-Driven Imagery analogue trauma: A randomised controlled crossover study. J Anxiety Disord 59:42–52
87. van't Wout M, Spofford CM, Unger WS, Sevin EB, Shea MT (2017) Skin Conductance Reactivity to Standardized Virtual Reality Combat Scenes in Veterans with PTSD. Appl Psychol Biofeedback 42(3):209–221
88. Wilson G, McGill M (2018) Violent video games in virtual reality: Re-evaluating the impact and rating of interactive experiences. In: Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '18). Association for Computing Machinery, New York, pp 535–548
89. Junker A, Hutters C, Reipur D, Embøl L, Nilsson NC, Nordahl R (2020) Virtual womb: Experiencing human sensory development from a fetal point-of-view in virtual reality. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20). Association for Computing Machinery, New York, Article 85, pp 1–8
90. Salgado DP, Flynn R, Naves ELM, Murray N (2022) A questionnaire-based and physiology-inspired quality of experience evaluation of an immersive multisensory wheelchair simulator. In: Proceedings of the 13th ACM Multimedia Systems Conference (MMSys '22). Association for Computing Machinery, New York, pp 1–11
91. Haraguchi G, Kitazaki M (2022) Virtual touch modulates perception of pleasant touch. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops. VRW, pp 782–783
92. Archer NS, Bluff A, Eddy A, Nikhil CK, Hazell N, Frank D, Johnston A (2022) Odour enhances the sense of presence in a virtual reality environment. PLoS ONE 17(3):1–20
93. Cornelio P, Dawes C, Maggioni E, Bernardo F, Schwalk M, Mai MCL, Pawlizak S, Zhang JX, Nelles G, Krasteva N, Obrist M (2022) Virtually tasty: An investigation of the effect of ambient lighting and 3D-shaped taste stimuli on taste perception In virtual reality. Int J Gastron Food Sci 30:1–11
94. Colla K, Keast R, Mohebbi M, Russell CG, Liem DG (2023) Testing the validity of immersive eating environments against laboratory and real life settings. Food Qual Prefer 103:1–11
95. Tamtama GIW, Santoso HB, Wang JC, Windasari NA (2022) Aw… The Museum is so "Dark": The effect of thermal stimuli for virtual reality experience and emotion. In: 2022 Seventh International Conference on Informatics and Computing (ICIC), pp 1–7
96. Kampa M, Finke J, Stalder T, Bucher L, Klapperich H, Mertl F, Zimmer C, Geiger C, Hassenzahl M, Klucken T (2022) Facilitating relaxation and stress reduction in healthy participants through a virtual reality intervention: study protocol for a non noninferiority randomized controlled trial. Trials 23(1):1–12
97. Melo M, Coelho H, Gonçalves G, Losada N, Jorge F, Teixeira MS, Bessa M (2022) Immersive multisensory virtual reality technologies for virtual tourism: A study of the user's sense of presence, satisfaction, emotions, and attitudes. Multimedia Syst 28(3):1027–1037
98. Brooks J, Lopes P, Amores J, Maggioni E, Matsukura H, Obrist M, Lalintha Peiris R, Ranasinghe N (2021) Smell, taste, and temperature interfaces. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, Article 76, pp 1–6

99. Elmqaddem N (2019) Augmented reality and virtual reality in education. Myth or reality? International Journal of Emerging Technologies in Learning (iJET) 14(03):234–242
100. Goodman LA, Kruskal WH (1954) Measures of Association for Cross Classifications. J Am Stat Assoc 49(268):732–764
101. IJsselsteijn WA (2004) Presence in depth. Technische Universiteit Eindhoven, Eindhoven, p 2004
102. Johnson RB, Onwuegbuzie AJ (2004) Mixed Methods Research: A Research Paradigm Whose Time Has Come. Educ Res 33(7):14–26
103. Bazeley P (2011) Integrative Analysis Strategies for Mixed Data Sources. Am Behav Sci 56(6):814–828
104. Safikhani S, Holly M, Kainz A, Pirker J (2021) The influence of in-VR questionnaire design on the user experience. In: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology (VRST '21). Association for Computing Machinery, New York, Article 12, pp 1–8
105. Schwind V, Knierim P, Haas N, Henze N (2019) Using presence questionnaires in virtual reality. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, Paper 360, pp 1–12
106. Feick M, Kleer N, Tang A, Krüger A (2020) The virtual reality questionnaire toolkit. In: Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology. Virtual Event, USA, pp 68–69
107. Tamaki R, Nakajima T (2021) Shoot down drones with your answer, an integration of a questionnaire into a VR experience. In: Proceedings of the 2021 ACM Symposium on Spatial User Interaction (SUI '21). Association for Computing Machinery, New York, Article 24, pp 1–2
108. Verschuere B, Crombez G, Koster E (2001) The international affective picture system. Psycholo Belgica 41:205–217
109. Kenealy PM (1986) The velten mood induction procedure: A methodological review. Motiv Emot 10(4):315–335
110. van der Waal NE, van Bokhorst JAW, van der Laan LN (2022) Identifying emotions toward an overweight avatar in Virtual Reality: The moderating effects of visuotactile stimulation and drive for thinness. Front Virtual Reality 3:1–16
111. Valzolgher C, Mazzurega M, Zampini M, Pavani F (2018) Incongruent multisensory stimuli alter bodily self-consciousness: Evidence from a first-person perspective experience. Acta Psychol 191:261–270
112. Slater M (2009) 'Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. Phil Trans R Soc B Biol Sci 364:3549–3557
113. Bockelman P, Lingum D (2017) Factors of cybersickness. In: Stephanidis C (ed) HCI International 2017 – Posters' Extended Abstracts. HCI 2017, vol 714. Communications in Computer and Information Science, pp 3–8
114. Nichols AL, Maner JK (2008) The good-subject effect: Investigating participant demand characteristics. J Gen Psychol 135(2):151–166