Agency and the Elicitation of Sadness in Virtual Reality

Literature Review

Julia Kinghorn
Department of Computer Science
University of Cape Town
Cape Town, South Africa
KNGJUL010@myuct.ac.za

ABSTRACT

To successfully create a Virtual Reality Environment to elicit and measure sadness, several elements must be considered. The level of presence in a Virtual Environment is important in eliciting emotions, and an immersive environment with several elicitation techniques, such as visual, auditory, and sensory prompts, is required. Measuring sadness can be achieved most thoroughly with a triangulated approach: both physiological and qualitative data must be collected. This literature review discusses the elicitation, design, and measuring methods to be considered.

KEYWORDS

Virtual Reality, Sadness, Presence, Immersion, Emotion Elicitation

1 Introduction

Emotions often dictate human behaviour, and it is important to understand how to elicit them in a controlled environment so that their effects on individuals can be studied further [14]. Virtual Reality (VR) is an application that allows for real-time interaction with synthetic stimuli, where users can be immersed in a computer-generated 3D environment [12]. It has been used increasingly in Psychology because of its ability to transport patients beyond spatial and resource constraints, and because it allows individuals a safe space to interact with their emotions. Many argue that VR's degree of immersion creates a more reliable elicitation of emotion than less immersive displays such as film and photographs. It is also argued that, because researchers have full control over a VE and can modify conditions when necessary, VR is useful as a relatively safe and accessible environment in which to elicit sadness [62, 59, 75].

Human experience of sadness is ubiquitous, but our understanding of it is limited. While it can be recognised with well-developed self-report techniques, sadness is often conflated with other emotions, and its psychophysiological measurement has returned inconsistent results [18]. Given its key role in facilitating social attachment and empathy, finding a reliable elicitation of pure sadness would be invaluable for a range of psychological applications.

So far, one of the ways VR has been effective in emotion elicitation has been in VR Exposure Therapy (VRET). VRET involves exposing individuals with phobias to their triggers in a virtual environment (VE) in the hopes of desensitisation [23], and has inspired further research.
This paper will analyse existing research on how to design an effective VE, and how to elicit and measure sadness. It will then discuss how these areas of study can be, and have been, combined to elicit and reliably measure sadness in VR.

2 What Makes VR Effective

There are numerous factors that go into creating an effective, or successful, VE. A large consideration is the environment's level of immersion and presence. So far, the relationship between effective VR and immersion and presence has been studied most in VRET, which, although different in its aims to emotion elicitation in isolation, provides a good understanding of how presence is achieved.

2.1 Immersion and Presence

There is much debate around the difference between immersion and presence, so this paper uses the most clearly defined and useful explanations. Immersion refers to what technology objectively delivers. The greater the number of sensory displays afforded by a technology, the greater its immersion. Presence then becomes the human reaction to immersion. Presence is the extent to which one feels that they are in a specific "place". It is determined by the perceptual processing of immersive technology's stimulated sensory data [34, 48, 49, 73]. Different people may experience different levels of presence in the same immersive technology [51].

It is widely believed that increased immersion leads to an increase in presence [34, 60], and this has been validated by many studies assessing different degrees of immersion. For example, the emotional reactions to a neutral VR scene were larger when a high-quality Head Mounted Display (HMD) was used, in comparison with a medium-quality HMD [8, 34]. Similar results were found in a comparison between an HMD and computer monitor [2, 34], as well as in a comparison between a Cave Automatic Virtual Environment (CAVE), where users are placed in a cube shaped room in which the walls, floor, and ceiling are projection screens, and an HMD [34, 42, 54]. In a study to test whether multimodal perceptual cues affect presence, it was determined that a combination of visual and tactile cues lead to greater presence, confirming an association between immersion and presence. All considered, the effect of immersion on presence is well documented [32, 34, 76].

Ijsselsteijn [2003] and Wiederhold and Wiederhold [1999] argue that the effectiveness of a VE depends on how much presence the immersed user experiences, which has been found to affect the success of VRET [5, 70, 77]. For example, Robillard [2003] found a significant correlation between presence and anxiety [17, 34]. In a study of height phobias, presence was deliberately manipulated, and anxiety levels were found to be much higher for higher degrees of presence [43, 70]. The correlation between presence and phobic emotions is bidirectional [70], and it is assumed that this correlation extends to presence and eliciting general emotions in VR.

Stanney and Sadowski [2002] suggest that seven dimensions of a VE affect presence: the realism of the VE, the social interactions included in the VE, how easy these interactions are, the user's perception of control, the duration of exposure to the VE, relationships with other VEs, and characteristics of the VR system itself. Note that, of these dimensions, the user's perception of control has been researched least, and so it is an important consideration when attempting to increase our understanding of presence and elicitation of sadness [36, 70].

From these observations, it can be argued that there are two ways to achieve presence. Firstly, one could construct a system with fidelity that is high enough to be indistinguishable from reality (i.e., perfect immersion). This is aspirational and unlikely to be achieved. Alternatively, one could evaluate the human perceptual system to determine what convinces us of reality. If evaluated successfully, one could match a system's displays and interactivity to those required by human perceptual and motor systems. Here, the level of immersion does not necessarily have to be high to achieve presence. Realistically, presence in VR is a result of how immersion and the human perceptual and motor systems interact [51].

2.2 Measuring Immersion and Presence

Slater suggests that a good indication of presence is when people respond to a VE in the same way as they would to a similar real-life environment [51]. He also argues that presence cannot be scientifically determined post-experience. Instead, he has proposed a measurement coined the Virtual Presence Counter (VPC), which has not yet been fully validated [50, 52].

A presence measurement, known as the Brief Measure of Presence, has been validated by Diemer et al. [2015], Bouchard [2010], and Bouchard et al. [2008] [34, 69, 70]. It simply asks participants the question: *On a scale of 0 to 100, how much did you have the impression of really being there in the virtual environment?* Bouchard [2010] also included a second measure of presence in their study, however, under this paper's definitions, "The Presence Questionnaire" (TPQ) [6, 70] would be considered a measure of immersion. It examines how the hardware and software of a project induces immersion using a 7-point Likert scale [36, 70], by measuring the following dimensions: realism, possibility of action, quality of the interface, possibility of examination, self-evaluation of performance, and auditory and haptic possibility of touching certain objects [70].

Additionally, Riva [2007] and Freedman [2005] used two questionnaires to evaluate presence [19, 32, 34]. The UCL Presence Questionnaire (UCL-PQ) is a subjective measure of presence taken after VR immersion that also uses a 7-point Likert scale rating for three presence-related questions [50]. The Independent Television Company Sense of Presence Inventory (ITC-SOPI) [28, 32] is a longer questionnaire that measures a participant's feelings post-experience, as well as their feelings during the experience. Riva [2007] assumes that this questionnaire measures several dimensions of presence, including one's sense of physical space, engagement, ecological validity, and negative feelings [19]. ITC-SOPI was similarly used by Gilpin et al. [18].

In choosing between these measures, one should note the argument that presence cannot be evaluated post-immersion, and one should thus consider using a during-immersion measure, or, for the sake of completeness, a during-immersion and post-immersion measure.

2.3 Other Considerations

Although hardware limitations have lessened over time, they must be considered when creating an effective VE. Their limitations restrict the complexity of VEs in terms of the number of objects in a VE, its resolution, and its level of interactivity. Some research suggests that visual complexity, when added strategically, is beneficial for eliciting emotions. For example, an animated audience induces greater anxiety than a static one in VRET for public speaking [59, 66], and it has been proven that the way that spiders move is the most fear-inducing element of their appearance [20, 59]. This suggests that high quality graphics may not be necessary for non-critical exposure elements, as long as those that provoke the emotion are perceived as realistic, for example, a spider's movements [59].

When human characters appear in a VE, it is important to consider the "uncanny valley effect" [35, 59]. People usually respond to human characters with comfort and familiarity, especially if the human simulation is highly realistic. However, the uncanny valley effect comes into play when a human simulation looks human-enough, but is not perfectly human. At this point, measured emotion changes severely, and could even be replaced with revulsion. Because of this, it is suggested that virtual human characters who have a prominent role in a VE should be made either semi-realistic, or photorealistic, and nothing in between [59].

Finally, the interactivity of the system must be considered. Movement in a VE strongly contributes to cybersickness symptoms [41, 59], which means that designers should consider if it is necessary for participants to move around. Haptic feedback is also a consideration, as it may lead to greater effect from a VE. This can be seen in a VRET study of arachnophobia, where a lack of tactile simulation made treatment less successful. This could be remedied by haptic feedback in the form of vibrations found in many consumer VR motion controllers, but this would need further research.

3 How to Elicit Sadness

Two factors are considered here, beginning with the method of elicitation. Siedlecka and Denson [2018] determined that a single method does not work for every emotion, and that the most effective elicitation method for emotions, in general, is using visual stimuli [14]. Unsurprisingly, the most notable classical methods of eliciting emotions are emotionally charged pictures and films [45, 64]. However, a reliable method of sadness elicitation in a controlled setting has not been thoroughly determined [86]. Visual stimuli, music and autobiographical recall have been suggested as the most effective methods, but Siedlecka and Denson [2018] suggest that a combination of methods may be more effective than testing them individually [14]. VR allows for such a combination, while ensuring that users feel present, so intuitively, it seems like a valid option [18].

The second factor is the content of an elicitation method, which significantly affects its success. Lauwerijssen [2008] argues that a specific relationship between situation, appraisal, and attribution must be achieved in order to elicit sadness [39]. Appraisal refers to one's perception of a situation, and attribution refers to the sense of responsibility one feels for a situation [7]. If one were to appraise a situation as a loss of some kind, one would likely be sad. Usually, when one feels that another person is responsible for an unfavourable situation, anger is induced. When no one is perceived as responsible, and one feels unable to cope alone, sadness is most likely [7, 39]. An alternative hypothesis to this relationship suggests that cultural normalities determine which emotions are elicited in certain circumstances [39, 68]. Studies suggest that the audience and situation determine how people feel that it is appropriate to respond [7, 39]. In practice, sadness seems to be caused mostly by events in relationships, especially their ending (often by death). Body- and mind-centred experiences, usually related to illness and bad news, and being unable to achieve a goal, also induce sadness to a great extent [9, 39]. These findings suggest that an effective method of eliciting sadness would involve a combination of stimuli, including visual elements. This combination should create a sense of (culturally appropriate) loss for participants, in such a way that no one is perceived as directly at fault.

4 How to Measure Sadness

Like most emotions, sadness can be recognised both physiologically and via qualitative self-report. Many studies have focused on one of the two methods, rather than combining them. Here, physiological and qualitative methods are discussed and compared in an attempt to find the most effective measuring device.

4.1 Physiologically

Sadness presents itself physiologically and behaviourally [37, 15, 39, 57], and the collection of physiological data is evident in many studies [64]. However, as highlighted by Gilpin et al. [2021], physiological indications of sadness change depending on the elicitation method used [18]. In 2010, Kreibig did a meta-analysis of psychophysiological measures, and concluded that the variability in findings could be because subtypes of sadness exist. The subtypes identified were 'deactivating' sadness, in which decreased heart rate (HR), respiration rate (RR) and parasympathetic activation is recorded [85], and 'activating sadness', in which increased HR, RR and parasympathetic withdrawal is observed [64, 82, 86]. Activating sadness is related to imminent, but not inevitable, loss where an individual has agency over a situation. Deactivating sadness is more likely found in situations where a loss is inevitable or has already happened [18].

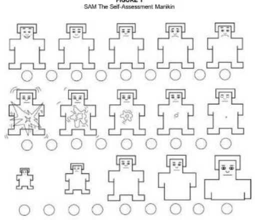
Following this classification, a study by Gilpin et al. [2021] used the Biopac MP160 system to measure HR, RR, and respiratory sinus arrhythmia (RSA), which is a good indication of parasympathetic activity. They found that eliciting sadness in VR resulted in an activating pattern of sadness, whereas exposure to film clips caused a deactivating sadness pattern [18]. This was despite the elicitation content remaining constant in both conditions. This suggests that the sadness induced in a VE is felt to be more imminent, as if the participant is really experiencing a situation, and has agency over it. This further suggests that a participant's sense of agency may determine the sadness subtype elicited. In other words, greater agency leads to an activating pattern of response. These results are supported by a study to test the response to sadness-inducing film clips, in which Kreibig et al. [2007] used autonomic, self-report, and behavioural measures of sadness to study its effects [39, 71]. Their results indicate that the resultant sadness response was deactivating, as the amplitude of ear pulse, RR, HR, and skin temperature decreased, and diastolic blood pressure and skin conductance response (SCR) increased. However, it must be noted that only eight participants were involved in this study, and thus the results are not necessarily generalisable [39].

4.2 Qualitatively

Emotions can be categorised according to several dimensions [64]. These include valence, which refers to positive or negative feelings, arousal, which indicates the strength of a feeling [62], certainty about future outcomes, approach, avoidance, and one's perception of control (e.g. [26, 29, 30, 55, 62, 64]). Models of emotion have been proposed from these dimensions, most notably the valence-arousal model, which categorises emotions as linear functions of valence and arousal [26, 62]. Sadness is traditionally associated with low arousal and negative valence [62]. However, through the physiological methods discussed above, Gilpin et al. [2021] have found that its arousal level changes depending on the elicitation technique employed [18]. This means that sadness could be low or high arousal, depending on its context, and that discrete qualitative measures of sadness may not be as useful as other literature suggests. A third axis, dominance, is often included in this model to indicate how much control one feels they have over an emotion [61, 62]. These models are used in the collection of qualitative data about emotion, and usually rely on self-reports.

The Positive and Negative Affect Schedule (PANAS) [7] is one of the most popular methods of measuring emotions induced by situational manipulation [10]. It involves rating a list of twenty adjectives, ten describing positive emotions, and ten describing negative emotions, on a five-point scale. It has been used by Freeman [2005], Riva [2007], and an expanded version of it, PANAS-X [13, 64], was used alongside physiological measures by Susindar et al. [19, 32, 34, 64]. Another common self-report measure is the 9-point Self-Assessment Manikin scale (SAM). It has been used to measure emotional response to advertising [31], and in a study concluding that VR could be used to research emotions and emotion elicitation, given how similar its effects are to reality [62]. SAM is a subjective assessment that captures the intensity of emotion by visualising an emotion's valence, arousal, and dominance [44, 62], as seen in Figure 1. It was proposed as an alternative to more inefficient verbal self-report measures [58, 31]. However, as recognised by Rivu et al. [2021], it is suggested that objective physiological measurements be used alongside SAM to garner more substantial results.

FIGURE 1
SAM The Self-Assessment Manikin



Several other measures have been used and considered. Gilpin et al. [2021] recorded a Primary emotion Rating (PER), by asking participants what primary emotion they felt after being exposed to the sadness-inducing stimuli [18]. Rivu et al. [2021] used the Ten Item Personality Measure (TIPI) [62, 65] to measure participant's personalities, to determine if they had any effect on how emotions are induced. In a study to test the impact of inducing emotions on the feeling of presence, Bouchard et al. [2010] measured emotions using The Brief Mood Introspection Scale (BMIS) [25, 70]. It divides emotions into two sub-scales, one measuring the intensity of positive emotions, the other measuring the intensity of negative emotions [70]. The State Trait Anxiety Inventory (STAI) has also been used to measure levels of anxiety [19]. The Visual Analogue Scale (VAS), a variation of Gross's measure [19, 27], assesses how participants feel with reference to 7 visual analogue scales measuring emotions. This scale has been used by Riva [2007] and Freeman [2005] [19, 32, 34].

Perhaps the most interesting qualitative measure is the Discrete Emotions Questionnaire (DEQ). This novel instrument was designed to detect eight distinct emotions: anger, anxiety, desire, disgust, fear, happiness, relaxation, and sadness. It has been validated with four studies that directly compared it to PANAS, and was found to be more sensitive at detecting self-reported emotions [10]. It was designed in response to the failings of measurement tools, most notably PANAS. When PANAS is used to measure situational manipulations of emotion, null effects are often reported. Since emotions are only probabilistically linked to their measurement, it seems inappropriate to assert that a null effect from PANAS is definitive evidence that emotion was not involved in a situation of interest. The DEQ aims to identify emotion that would otherwise be discounted by null effects on measures of self-report [10].

The DEQ approaches emotions discretely rather than dimensionally, as certain behaviours, subjective feelings, eliciting situations, and appraisals cannot be measured dimensionally. It is a simple questionnaire that is quick to fill out, and was built on a foundation of words used by the everyday person when describing their emotions, and clearly understood by English speakers. This allows it to be more effective as a self-report measure.

5 How to Make VR Effective in Eliciting Sadness

It has been proven that VR is an effective elicitor of emotion, and that sadness is no exception. Rivu et al. [2021] conducted a user study with 39 participants to determine how emotion elicitation methods of image, video, audio, and autobiographical memory recall perform in VR compared to the real world [62]. They recorded SAM values for valence, arousal, and dominance every 30 seconds to track changes in emotion over time. When compared to the real world, sadness in VR was found to have no statistically significant difference in valence, arousal, and dominance. According to the median values however, arousal was slightly higher than in the real world, and dominance slightly lower. This makes sense given Gilpin et al. [2021]'s findings, which suggest that the type of sadness identified in VR is an activating pattern of sadness.

Because it has been proven that VR can elicit sadness, it is important to discuss how best to use VR to do so going forward, by evaluating existing work.

5.1 Evaluating Immersion and Presence

The relationship between presence, immersion, and emotions in VEs is heavily debated, but many regard presence, which can be measured using the techniques outlined above, as necessary for real emotions to be elicited by a VE [34, 46, 72].

It has been reported that increasing immersion (and thus presence) increases emotion elicitation when compared to less immersive systems [8, 34, 42, 74]. The correlation between emotion and presence has been consistently described in literature, especially in literature about VRET [17, 19, 24, 34, 46, 47, 69]. Others have argued that immersion has a non-linear relationship with emotion [32, 34, 78]. They claim that the only reason that this relationship has been reported is because of the arousal levels of emotions that are being studied [34, 74]. Generally, strongly arousing emotions (like anxiety and fear) are more intense in more immersive VEs [34, 42]. Less arousing emotions, like relaxation and happiness are not as influenced by the immersive quality of a VE [32, 34, 60]. Freeman [2005] has thus theorised that presence and emotion are only related for arousing stimuli, and that, for content to be arousing, it must be perceived as deeply significant and personally relevant [32, 34]. Regardless of whether presence is correlated to all, or only arousing emotions, this definition of arousing stimuli can be applied to sadness in VR. In fact, Baños et al. tested the effects of sadness on presence, and found a correlation between the two [60, 34]. Gilpin et al. extended these findings by identifying that a subtype of sadness, activating sadness, is more likely induced when presence is high. Thus, this paper assumes that presence (and thus immersion) has an effect on eliciting sadness. Because of this apparent relationship, it is generally agreed that an objective measure of presence and immersion would be useful in determining whether a VE will be an effective affective medium [64].

Another consideration when looking at presence is the method of control that the user has, or their perceived agency over a situation. To investigate this, Osking and Doucette [2019] compared using voice control in interactive VR narratives with more traditional methods of "point and click" interfaces [83]. They argue that older selection systems, like "point and click" create a jarring interruption in one's sense of presence, especially when they interrupt a narrative that one is trying to follow. Instead, a voice control dialogue system allows participants to remain grounded in their illusion of presence. The results of their study support this hypothesis.

5.2 Evaluating User Experience and Other Design Considerations

Beyond the idea of presence, user experience in VR is also an important consideration. Sutcliffe et al. [3] developed and tested a heuristic evaluation for assessing VE user interfaces based on Nielsen's usability heuristics [1994] and a previous study on VE-specific principles by Sutcliffe and Kaur [2000] [79]. The evaluation consists of twelve heuristics focusing on usability and presence, and when used correctly is generally agreed to be reliable and informative.

According to Somarathna's [2021] summary of research into emotion elicitation [63], there are various considerations when creating a reliable model for understanding emotion. A benchmark dataset would be useful for making consistent comparisons across different studies of emotion in VR, because it would allow for a more universal model of how to form emotion. Few such databases exist in the literature, limiting the development of new research. For efficiency, emotion-sensing interfaces could also be incorporated into VR headsets to measure physiological signals and facial expressions. The effect of haptic feedback on emotion must be further studied, with more intensive data collection strategies, as more indications of emotion would be present. Appraisal theory [1, 63] and the Component Process Model [38, 63] reconcile well with current neuroscience findings on emotion [16, 63]. It would be useful to include these models, as well as other topical theories of emotion like the theory of constructed emotion [40, 63], in affective media research [63].

5.3 Experiment Design in Previous work

Much research has been done on VR's ability to induce emotions, but very few include sadness, and even fewer focus exclusively on sadness. Therefore, it is useful to examine previous work on eliciting emotions in general in VR, to gain insight into what the best way forward is. A summary of the main studies presented here can be found in Appendix A.

Rivu et al.'s [2021] study began with filling out the TIPI, and then briefing participants about the elicitation materials and the SAM scale. It was not confirmed whether participants had interacted with VR before, which could severely influence their experience and thus the study's results. Regardless, the study included showing them how to use the VR equipment and fill in the SAM scale. At the end of the study, participants were interviewed about how their experiences differed in the real would versus VR. The SAM results were compared to the participants' VR experience, age, gender, and TIPI results. In their sample, no statistically significant difference was identified between the SAM results and other factors. This is an interesting finding, and suggests that affecting research studies need not worry about external factors like personality, although Rivu et al. suggest investigating this result with a larger sample.

Bouchard et al. [2010] recruited adult participants with no VR experience [70]. Participants were checked for exclusion criteria via a telephone pre-screening. This included the likelihood of epilepsy, simulator sickness, cardiovascular problems, high blood pressure, and diabetes. Each participant was made aware of the nature of the experiment, and had to fill out an "Immersive Tendencies Questionnaire" (ITQ) [6, 70] and "Simulator Sickness Questionnaire" (SSQ) to assess and control their predisposition to presence and simulator sickness. From here, participants were immersed in an unrelated VE to familiarise themselves with VR. After seven minutes immersed in the testing environment, the participant had to respond to several questionnaires. Participants could only leave the lab after fifteen minutes in case of significant simulator sickness. This study did well in looking after its participants' best interests, and controlling various factors that may have influenced their results, like VR experience health concerns.

Riva's [2007] experiment to evaluate the elicitation of emotions in VEs with different levels of presence involved 61 undergraduate psychology students [19]. As a prerequisite, those who scored eighteen or higher on the Beck Depression Inventory (BDI) questionnaire could not partake. The BDI has become the most widely used assessment of the main cognitive aspects of depression. Those with a history of certain injuries and emotional problems were also excluded [19]. Again, it was unclear if participants had previous experience with VR, but each was briefed on using VR equipment and how to explore in a VE, before experiencing all three VEs (anxious, relaxing, and neutral parks) for three minutes in a random order. They were expected to use two self-report measures. At the beginning and end of immersion, emotion was measured using a 10-point scale (e.g., "To what extent do you feel anxious at this moment?"). Presence was also measured on a 10-point scale, but in response to two questions: (1) *Do you feel you are here, in this park?* (2) *Do you feel this park is real, is it a place you are visiting?*. After each immersion, the participants filled out the PANAS, VAS, STAI, UCL and ITC-SOPI questionnaires to keep track of their emotions and sense of presence. After all participants had experienced all environments, a debriefing session occurred [19]. The qualitative measuring techniques in this study are impressive, as they thoroughly captured the states of participants throughout the study. However, the efficacy of qualitative measurements alone is debatable, and physiological measures would add valuable depth to the meaning of the results. Furthermore, by putting each participant in all three environments consecutively, the emotions elicited in one environment may have affected the emotions elicited in others.

ANOVAs, correlation analysis, and multiple regressions were used to evaluate how well the VEs elicited their related emotions, and to analyse the link between presence and emotions [19]. The statistical power of these results is low because a small sample size was used. Having only tested two emotional states (anxiety and relaxation), Riva [2007]'s study also required more research to verify if it is possible to induce others, and if the link between emotions and presence remains the same [19].

Susindar et el. [64] used only eleven adult participants, who all reported less than ten hours of previous use of VR equipment in a pre-study questionnaire. Participants were informed that they were taking part in a study to test the influence of VR on decision-making tasks, and were not informed of the nature of the material that would be presented to them in VR or on the desktop. At the end of the study, the participants were informed that the materials presented to them aimed to elicit an emotional reaction, and were asked to assess the material in an informal interview. Responses to a Simulator Sickness Questionnaire (SSQ) were also collected to check if participants had experienced any simulator sickness [64]. The use of a blind experiment design is interesting, as it seems unfair that the elicitation material was not disclosed. However, no complaints were reported, so it is uncertain whether this posed a real issue to participants.

The investigation of different techniques in generating sadness by Gilpin et al. [2021] had sixty participants, with each allocated to one of three elicitation methods [18]. While they had fairly good coverage, their study was a between-subject study. This meant that differences between their participants, rather than differences between their methods, may have influenced the results. Like Bouchard et al. [2010] and Riva [2007], they used several screening measures when selecting participants. These included criteria such as fluency in English, no history of psychiatric or neurological conditions, and no current use of psychiatric/chronic medication. They also noted the influence depression, anxiety, and PTSD may have on physiological readings [84, 88], and the distressing effect their study may have on depressed individuals [81], so excluded them from the study. Their pre-screening involved a score of less than 21 on the BDI-II [81], the next iteration of the BDI used by Riva [2007]. They also used the STAI to evaluate general anxiety, and the 4-item primary Care Post-Traumatic Stress disorder screen (PC-PTSD) to ensure that no PTSD symptoms would be triggered by the experiment. To confirm Rivu et al. 's [2021] findings, they found no significant differences as a result of external factors including age, sex, home language, BDI, STAI, and PC-PTSD-4 scores. They also made use of manipulation checks to ensure that sadness was the primary emotion elicited in all three conditions, which was not found to be an issue for film and VR. Although the results of the study showed that VR elicited the highest level of presence, a limitation of their study is that, to ensure that all conditions remained consistent, participants were seated during all three conditions. This mitigated possible effects of cybersickness, but potentially harmed the immersive and interactive impact of the VR experience.

6 Discussion

The literature discussed above provides much to reflect on. Importantly, most studies focused on more than one emotion at a time. Sadness is an interacting emotion, meaning that it is influenced by multiple factors, and often occurs together with other emotions [39]. Consequently, it can be easily conflated with other emotions. As such, studying sadness in isolation is most likely to yield the purest elicitation of sadness. Beyond Gilpin et al. [2021], Rivu et al. [2021] have come closest to discovering VR's ability to elicit sadness, but their results may be distorted as all of their participants experienced four different emotions [62].

Another observation of note is that measurements of parasympathetic activation have been used most when measuring sadness elicitation [71], making them easiest to compare against previous work. However, measuring sympathetic activation would strengthen the understanding of what kind of sadness is induced, and so should be considered moving forward [18]. A measure of how long elicited emotions last could also strengthen results [64].

In a review of the use of VR in exposure therapy, Lidner [2017] determined that larger study samples, and a more precise recording of treatment effects in general would be necessary moving forward [59, 67, 64]. This can be seen in most of the studies reviewed here. For example, Susindar needed a larger sample size to properly assess responses to the PANAS-X with statistical analysis, in order to validate their results [64]. It also seems important to include an objective measure to assess presence and immersion when aiming to conduct an emotion elicitation study, as noted by Susindar et al. [2019].

Obscuring the aim of emotion induction from participants could also be a vital element that was missed by some of the above studies. This is important, as some techniques produce warped results when the desired emotion is made clear [11]. This is well-documented as the Hawthorne effect, which suggests that users behave differently when they know that they are being studied, especially for some kind of response [87]. There is a potential moral dilemma to keeping studies blind, as participants may feel that they have been lied to or harmed without consent. However, by having a clear set of exclusion criteria to ensure minimal harm, and by ensuring that general care and a debriefing is involved in the experiment, this dilemma can be largely avoided. Thus, it seems most appropriate to keep emotion elicitation undisclosed [14].

In general, some opportunities may have been missed when it comes to elicitation content. It has been widely agreed that including a narrative in a VR scenario is likely to enhance its emotional experience, as it gives participants aims and a clear path to follow while immersing them deeper into the experience [2, 4, 21, 34, 69]. Short stories have been emphasised in particular, as they allow participants to form personal and emotional connections to the narrative [80]. Moreover, a greater emotional response is recorded when personally relevant stories are recorded, and narratives around loss cause the highest sadness ratings [18, 85].

7 Conclusions

Considering the above analysis, there are several requirements when creating a successful VE for eliciting and measuring sadness. First, when looking to create a successful VE in general, a high level of presence should be prioritised. Hardware implications, interactivity, and human depiction should also be carefully considered. To successfully elicit sadness, a combination of elicitation methods should be used, while remaining careful to elicit sadness in isolation. It is important to have a clear exclusion criterion so that vulnerable, or otherwise biased, participants are not included in the study. The most effective way to measure sadness requires triangulation. First, physiological data should be collected, either through self-reports or more thorough measurements of HR, RR, RSA etc. Second, qualitative self-report measurement should be obtained. According to the research reviewed here, the DEQ seems to be a promising qualitative measurement, but study limitations and practicality should be considered when deciding which measurement to use. In terms of testing a VEs effectiveness in eliciting sadness, presence should be measured before using the VE for elicitation, and user experience should also be considered.

REFERENCES

[1] Agnes Moors, Phoebe C. Ellsworth, Klaus R. Scherer, and Nico H. Frijda. 2013. Appraisal Theories of Emotion: State of the Art and Future Development. Emotion Review, vol. 5, pp. 119-124, 03/20 2013.

[2] Alessandra Gorini, Claret S. Capideville, Gianluca De Leo, Fabrizia Mantovani, and Giuseppe Riva. 2011. The role of immersion and narrative in mediated presence: The Virtual Hospital experience. Cyberpsychology, Behavior, and Social Networking 14, 3 (2011), 99–105. DOI:http://dx.doi.org/10.1089/cyber.2010.0100

[3] Alistair Sutcliffe and Brian Gault. 2004. Heuristic evaluation of virtual reality applications. *Interacting with Computers* 16, 4 (2004), 831–849. DOI:http://dx.doi.org/10.1016/j.intcom.2004.05.001

[4] Andreas Mühlberger, Roland Neumann, Ljubica Lozo, Mathias Müller, and Marc Hettinger. 2012. Bottom-up and top-down influences of beliefs on emotional responses: fear of heights in a virtual environment. Stud. Health Technol. Inform. 181, 133–137.

[5] B. K. Wiederhold, and M. D. Wiederhold, M. D. (1999). Clinical observations during virtual reality therapy for specific phobias. Cyberpsychology and Behavior, 2(2), 161-168.

[6] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: a Presence Questionnaire. Presence, 7(3), 225-240.

[7] C. Barr-Zisowitz. 2000. "Sadness"- Is there such a thing? In M. Lewis, & J.M. Haviland- Jones, (Eds.), Handbook of Emotions (pp. 607-622). New York: The Guilford Press.

[8] C. Botella et al. 1999. Differences on presence and reality judgment using a high impact workstation and a PC workstation. CyberPsychology & Behavior 2, 1 (1999), 49–52. DOI:http://dx.doi.org/10.1089/cpb.1999.2.49

[9] C. Z. Stearns. 1993. Sadness. In M. Lewis, & J.M. Haviland (Eds.), Handbook of Emotions (pp. 547-562). New York: The Guilford Press

[10] Cindy Harmon-Jones, Brock Bastian, and Eddie Harmon-Jones. 2016. The Discrete Emotions Questionnaire: A new tool for measuring state self-reported emotions. PLOS ONE 11, 8 (2016). DOI:http://dx.doi.org/10.1371/journal.pone.0159915

[11] Daniel Västfjäll. 2001. Emotion induction through music: A review of the musical mood induction procedure. Musicae Scientiae 5, 1_suppl (2001), 173–211. DOI:http://dx.doi.org/10.1177/10298649020050s107

[12] David R. Pratt, Michael Zyda, and Kristen Kelleher. 1995. Virtual Reality: In the Mind of the Beholder: Computer 28, 7 (July 1995), 17–19

[13] David Watson and Lee Anna Clark. 1999. The PANAS-X: Manual for the positive and negative affect schedule expanded form.

[14] Ewa Siedlecka and Thomas F. Denson. 2018. Experimental methods for inducing basic emotions: A qualitative review. Emotion Review 11, 1 (2018), 87–97. DOI:http://dx.doi.org/10.1177/1754073917749016

[15] Galen V. Bodenhausen, Lori A. Sheppard, and Geoffrey P. Kramer. (1994). Negative affect and social judgment: the differential impact of anger and sadness. European Journal of Social Psychology, 24, 45-62.

[16] Gelareh Mohammadi and Patrik Vuilleumier. A Multi-Componential Approach to Emotion Recognition and the Effect of Personality, IEEE Transactions on Affective Computing, pp. 1-1, 2020

[17] Geneviève Robillard, Stéphane Bouchard, Thomas Fournier, and Patrice Renaud. 2003. Anxiety and presence during VR immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. CyberPsychology & Behavior 6, 5 (2003), 467–476. DOI:http://dx.doi.org/10.1089/109493103769710497

[18] Gina Gilpin, James Gain, and Gosia Lipinska. 2021. The physiological signature of sadness: A comparison between text, film and virtual reality. Brain and Cognition 152 (2021), 105734. DOI:http://dx.doi.org/10.1016/j.bandc.2021.105734

[19] Giuseppe Riva et al. 2007. Affective interactions using virtual reality: The link between presence and emotions. CyberPsychology & Behavior 10, 1 (2007), 45–56. DOI:http://dx.doi.org/10.1089/cpb.2006.9993

[20] Graham C.L. Davey. 1991. Characteristics of individuals with fear of spiders. Anxiety Research 4, 4 (1991), 299–314. DOI:http://dx.doi.org/10.1080/08917779208248798

[21] Henrik M. Peperkorn and Andreas Mühlberger. 2013. The impact of different perceptual cues on fear and presence in virtual reality. Stud. Health Technol. Inform. 191, 75–79.

[22] I. Pavlidis et al. 2016. Dissecting driver behaviors under cognitive, emotional, sensorimotor, and mixed stressors. Scientific Reports 6, 1 (2016). DOI:http://dx.doi.org/10.1038/srep25651

[23] Ivan Alsina-Jurnet, Cristina Carvallo-Beciu, and José Gutiérrez-Maldonado. 2007. Validity of virtual reality as a method of exposure in the treatment of test anxiety. Behavior Research Methods 39, 4 (2007), 844–851. DOI:http://dx.doi.org/10.3758/bf03192977

[24] Ivan Alsina-Jurnet, José Gutiérrez-Maldonado, and María-Virgina Rangel-Gómez. 2011. The role of presence in the level of anxiety experienced in clinical virtual environments. Computers in Human Behavior 27, 1 (2011), 504–512. DOI:http://dx.doi.org/10.1016/j.chb.2010.09.018
[25] J. D. Mayer and Y. N. Gaschke. 1988. The experience and meta-experience of mood. Journal of Personality and Social Psychology, 55(1), 102-111.
[26] James A. Russell. 1980. A circumplex model of affect. Journal of personality and social psychology 39, 6 (1980), 1161.
[27] James J. Gross and Robert W. Levenson. 1995. Emotion elicitation using films. Cognition & Emotion 9:87–108.
[28] Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. 2001. A Cross-Media Presence Questionnaire: The ITC Sense of Presence Inventory. Presence: Teleoperators and Virtual Environments, 10(3), 282-297
[29] Jennifer S. Lerner and Dachter Keltner. 2001. Fear, anger, and risk. Journal of personality and social psychology, 81(1), 146.
[30] Jennifer S. Lerner, Ye Li, Piercarlo Valdesolo, and Karim S. Kassam. 2015. Emotion and decision making. Annual review of psychology, 66.
[31] Jon D. Morris. 1995. Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. Journal of advertising research, 35(6), pp.63-68
[32] Jonathan Freeman, Jane Lessiter, Katherine Pugh, and Ed Keogh. 2005. When presence and emotion are related, and when they are not. In Proceedings of the Conference at Presence 2005, London. Available at: http://www.temple.edu/ispr/prev_conferences/ proceedings/2005/freeman,%20lessiter,%20pugh,%20keogh.pdf
[33] Jonathan Rottenberg, Rebecca D. Ray, and James J. Gross. 2007. Emotion elicitation using films. In J. A. Coan & J. J. B. Allen (Eds.), Handbook of emotion elicitation and assessment (pp. 9–28). Oxford University Press
[34] Julia Diemer, Georg W. Alpers, Henrik M. Peperkorn, Youssef Shiban, and Andreas Muhlberger. 2015. The impact of perception and presence on emotional reactions: A review of research in virtual reality. Frontiers in Psychology 6 (2015). DOI:http://dx.doi.org/10.3389/fpsyg.2015.00026
[35] Jun'ichiro Seyama and Ruth S. Nagayama. 2007. The uncanny valley: Effect of realism on the impression of artificial human faces. Presence: Teleoperators and Virtual Environments 16, 4 (2007), 337–351. DOI:http://dx.doi.org/10.1162/pres.16.4.337
[36] K. M. Stanney and W. Sadowski. 2002. Presence in virtual environments. In K. M. Stanney (Ed.), Handbook of virtual Environments. Design, Implementation, and applications (p.791-806). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
[37] Klause R. Scherer and H. G. Wallbott. 1994. Evidence for universality and cultural variation of differential emotion response patterning. Journal of Personality and Social Psychology, 66, 310-328.
[38] Klaus R. Scherer. 2009. The dynamic architecture of emotion: Evidence for the component process model. Cognition and Emotion, vol. 23, no. 7, pp. 1307-1351, 2009/11/01 2009.
[39] Krista Lauwerijssen. 2008. Sadness. Bachelor's Thesis. Tilburg University, Tilburg, Netherlands.
[40] Lisa Feldman Barrett. 2006. Solving the emotion paradox: categorization and the experience of emotion, (in eng), Pers Soc Psychol Rev, vol. 10, no. 1, pp. 20-46, 2006.
[41] Lisa Rebenitsch and Charles Owen. 2016. Review on cybersickness in applications and visual displays. Virtual Reality, 20, 101–125. http://dx.doi.org/10.1007/s10055-016-0285-9
[42] M. Carmen Juan and David Pérez. 2009. Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or cave. Presence: Teleoperators and Virtual Environments 18, 3 (2009), 232–248. DOI:http://dx.doi.org/10.1162/pres.18.3.232
[43] M. Michaud, Stéphane Bouchard, Stéphanie Dumoulin, and X. W. Zhong. 2004. Manipulating presence and its impact on anxiety. Presented at the Cybertherapy Conference 2004, San Diego, January 10-12 2004
[44] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry 25, 1 (1994), 49–59. DOI:http://dx.doi.org/10.1016/0005-7916(94)90063-9
[45] Margaret M. Bradley and Peter J. Lang. 2007. The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan & J. J. B. Allen (Eds.), Handbook of emotion elicitation and assessment (pp. 29–46). Oxford University Press.
[46] Matthew Price and Page Anderson. 2007. The role of presence in virtual reality exposure therapy. Journal of Anxiety Disorders 21, 5 (2007), 742–751. DOI:http://dx.doi.org/10.1016/j.janxdis.2006.11.002
[47] Matthew Price, Natasha Mehta, Erin B. Tone, and Page L. Anderson. 2011. Does engagement with exposure yield better outcomes? components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia. Journal of Anxiety Disorders 25, 6 (2011), 763–770. DOI:http://dx.doi.org/10.1016/j.janxdis.2011.03.004
[48] Mel Slater and Sylvia Wilbur. 1997. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. Presence: Teleoperators and Virtual Environments 6, 6 (1997), 603–616. DOI:http://dx.doi.org/10.1162/pres.1997.6.6.603
[49] Mel Slater. 1999. Measuring presence: A response to the Witmer and singer presence questionnaire. Presence: Teleoperators and Virtual Environments 8, 5 (1999), 560–565. DOI:http://dx.doi.org/10.1162/105474699566477
[50] Mel Slater and Anthony Steed. 2000. A virtual presence counter. Presence: Teleoperators and Virtual Environments 9, 5 (2000), 413–434. DOI:http://dx.doi.org/10.1162/105474600566925
[51] Mel Slater. 2003. A Note on Presence Terminology.
[52] Mel Slater. 2004. How colorful was your day? why questionnaires cannot assess presence in virtual environments. Presence: Teleoperators and Virtual Environments 13, 4 (2004), 484–493. DOI:http://dx.doi.org/10.1162/1054746041944849
[53] Mel Slater. 2004. Presence and emotions. CyberPsychology & Behavior 7, 1 (2004), 121–121. DOI:http://dx.doi.org/10.1089/109493104322820200
[54] Merel Krijn, Paul M.G. Emmelkamp, Roeline Biemond, Claudius de Wilde de Ligny, Martijn J. Schuemie, and Charles A.P.G. van der Mast. 2004. Treatment of acrophobia in virtual reality: The role of immersion and presence. Behaviour Research and Therapy 42, 2 (2004), 229–239. DOI:http://dx.doi.org/10.1016/s0005-7967(03)00139-6
[55] Norbert Schwarz and Gerald L. Clore. 1983. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. Journal of Personality and Social Psychology 45, 3 (1983), 513–523. DOI:http://dx.doi.org/10.1037/0022-3514.45.3.513
[56] Peter Gathmann, Leo Linzmayer, and Josef Grünberger. 1985. Activation, deactivation, depression: Psychosomatic in-patient therapy. Psychiatry the State of the Art (1985), 449–452. DOI:http://dx.doi.org/10.1007/978-1-4684-4697-5_75
[57] Peter J. Freed and John J. Mann. 2007. Sadness and loss: toward a neurobiopsychosocial model. American Journal of Psychiatry, 164, 28-34.
[58] Peter J. Lang. 1985. The cognitive psychophysiology of emotion: Fear and anxiety. In A. H. Tuma & J. D. Maser (Eds.), Anxiety and the anxiety disorders (pp. 131–170). Lawrence Erlbaum Associates, Inc.
[59] Philip Lindner et al. 2017. Creating state of the art, next-generation virtual reality exposure therapies for anxiety disorders using consumer hardware platforms: Design considerations and Future Directions. Cognitive Behaviour Therapy 46, 5 (2017), 404–420. DOI:http://dx.doi.org/10.1080/16506073.2017.1280843
[60] R.M. Baños, E. Etchemendy, D. Castilla, A. García-Palacios, S. Quero, and C. Botella. 2012. Positive mood induction procedures for virtual environments designed for elderly people. Interacting with Computers 24, 3 (2012), 131–138. DOI:http://dx.doi.org/10.1016/j.intcom.2012.04.002
[61] R.W. Picard, E. Vyzas, and J. Healey. 2001. Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence 23, 10 (2001), 1175–1191. DOI:http://dx.doi.org/10.1109/34.954607
[62] Radiah Rivu, Ruoyu Jiang, Ville Mäkelä, Mariam Hassib, and Florian Alt. 2021. Emotion elicitation techniques in virtual reality. Human-Computer Interaction – INTERACT 2021 (2021), 93–114. DOI:http://dx.doi.org/10.1007/978-3-030-85623-6_8
[63] Rukshani Somarathna, Tomasz Bednarz, and Gelareh Mohammadi. 2021. Virtual Reality for Emotion Elicitation - A Review.
[64] Sahinya Susindar, Mahnoosh Sadeghi, Lea Huntington, Andrew Singer, and Thomas K. Ferris. 2019. The feeling is real: Emotion elicitation in virtual reality. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 63, 1 (2019), 252–256. DOI:http://dx.doi.org/10.1177/1071181319631509
[65] Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann Jr. 2003. A very brief measure of the Big-Five personality domains. Journal of Research in personality 37, 6 (2003), 504–528 (Gosling 2003)
[66] Sandra Poeschl and Nicola Doering. 2014. Virtual speech anxiety training — effects of simulation fidelity on User Experience. 2014 IEEE Virtual Reality (VR) (2014). DOI:http://dx.doi.org/10.1109/vr.2014.6802073
[67] Sarah Page and Matthew Coxon. 2016. Virtual reality exposure therapy for anxiety disorders: Small samples and no controls? Frontiers in Psychology 7 (2016). DOI:http://dx.doi.org/10.3389/fpsyg.2016.00326
[68] Stephen D. White. 1998. The politics of anger. In B. Rosenwein (Ed.), Anger's past (pp. 127-152). Ithaca, NY: Cornell University Press.
[69] Stéphane Bouchard, Julie St-Jacques, Geneviève Robillard, and Patrice Renaud. 2008. Anxiety increases the feeling of presence in virtual reality. Presence: Teleoperators and Virtual Environments 17, 4 (2008), 376–391. DOI:http://dx.doi.org/10.1162/pres.17.4.376
[70] Stéphane Bouchard. 2010. Emotions and the emotional valence afforded by the virtual environment. Virtual Reality (2010). DOI:http://dx.doi.org/10.5772/12957
[71] Sylvia D. Kreibig, Frank H Wilhelm, Walton T. Roth, and James J Gross. 2007. Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. Psychophysiology, 44.
[72] Thomas D. Parsons and Albert A. Rizzo. 2008. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis.

Journal of Behavior Therapy and Experimental Psychiatry 39, 3 (2008), 250–261. DOI:http://dx.doi.org/10.1016/j.jbtep.2007.07.007
[73] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. Presence: Teleoperators and Virtual Environments 10, 3 (2001), 266–281. DOI:http://dx.doi.org/10.1162/105474601300343603
[74] Valentijn T. Visch, Ed S. Tan, and Dylan Molenaar. 2010. The emotional and cognitive effect of immersion in film viewing. Cognition & Emotion 24, 8 (2010), 1439–1445. DOI:http://dx.doi.org/10.1080/02699930903498186
[75] Ville Mäkelä et al. 2020. Virtual field studies. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (2020). DOI:http://dx.doi.org/10.1145/3313831.3376796
[76] Wijnand IJsselsteijn, Huib de Ridder, Jonathan Freeman, S.E. Avons, and Don Bouwhuis. 2001. Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. Presence: Teleoperators and Virtual Environments 10, 3 (2001), 298–311. DOI:http://dx.doi.org/10.1162/105474601300343621
[77] Wijnand Ijsselsteijn, and Giuseppe Riva. 2003. Being there: The experience of presence in mediated environments. In G. Riva, F. Davide, & W. A. Ijsselsteijn (Ed.), Being There. Concepts, Effects and Measurements of User Presence in Synthetic Environments (p. 3- 17).Washington, DC: IOS Press
[78] Yun Ling, Willem-Paul Brinkman, Harold T. Nefs, Chao Qu, and Ingrid Heynderickx. 2012. Effects of stereoscopic viewing on presence, anxiety, and Cybersickness in a virtual reality environment for public speaking. Presence: Teleoperators and Virtual Environments 21, 3 (2012), 254–267. DOI:http://dx.doi.org/10.1162/pres_a_00111
[79] A.G. Sutcliffe and K.Deol Kaur. 2000. Evaluating the usability of virtual reality user interfaces. Behaviour & Information Technology 19, 6 (2000), 415–426. DOI:http://dx.doi.org/10.1080/014492900750052679
[80] Barbra Zupan and Duncan R. Babbage. 2016. Film clips and narrative text as subjective emotion elicitation techniques. The Journal of Social Psychology 157, 2 (2016), 194–210. DOI:http://dx.doi.org/10.1080/00224545.2016.1208138
[81] George M. Slavich, Scott M. Monroe, and Ian H. Gotlib. 2011. Early parental loss and depression history: Associations with recent life stress in major depressive disorder. Journal of Psychiatric Research 45, 9 (2011), 1146–1152. DOI:http://dx.doi.org/10.1016/j.jpsychires.2011.03.004
[82] Gerardo Herrera, Rita Jordan, and Lucí Vera. 2006. Agency and presence: A common dependence on subjectivity? Presence: Teleoperators and Virtual Environments 15, 5 (2006), 539–552. DOI:http://dx.doi.org/10.1162/pres.15.5.539
[83] Hunter Osking and John A. Doucette. 2019. Enhancing emotional effectiveness of virtual-reality experiences with voice control interfaces. Communications in Computer and Information Science (2019), 199–209. DOI:http://dx.doi.org/10.1007/978-3-030-23089-0_15
[84] Hsiao-Pei Lin, Hung-Yu Lin, Wei-Lun Lin, and Andrew Chih-Wei Huang. 2011. Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: Sympathetic-parasympathetic hypothesis of stress and depression. Journal of Clinical Psychology 67, 10 (2011), 1080–1091. DOI:http://dx.doi.org/10.1002/jclp.20833
[85] Mariko Shirai and Naoto Suzuki. 2017. Is sadness only one emotion? psychological and physiological responses to sadness induced by two different situations: “loss of someone” and “failure to achieve a goal.” Frontiers in Psychology 8 (2017). DOI:http://dx.doi.org/10.3389/fpsyg.2017.00288
[86] Sylvia D. Kreibig. 2010. Autonomic nervous system activity in emotion: A Review. Biological Psychology 84, 3 (2010), 394–421. DOI:http://dx.doi.org/10.1016/j.biopsycho.2010.03.010
[87] Berkely Rice. 1982. The Hawthorne defect: Persistence of a flawed theory. Psychology Today, 16(2), 70-74
[88] Szymon Fedor, Peggy Chau, Nicolina Bruno, Rosalind W. Picard, and Joan Camprodon. 2016. Can we predict depression from the asymmetry of electrodermal activity? Iproceedings 2, 1 (2016). DOI:http://dx.doi.org/10.2196/iproc.6117

Appendix A

Author	Relevant research goal	Emotions measured	Most important findings	Sample size	Immersion/presence measurements	Emotion measurements	Emotion elicitation disclosed	Exclusion criteria	General Comments
Bouchard [70]	Determine emotional valence afforded by VR	• Joy • Sadness	Emotions have an impact on presence	59	• TPQ (ITQ and SSQ used for control)	• BMIS	Yes	• Adults • No VR experience • Unlikely epilepsy, simulator sickness, cardiovascular problems, high blood pressure, diabetes	
Freeman [34]	Determine relationship between presence and emotion	• Happiness • Anger • Disgust • Relaxation • Fear • Sadness • Surprise	Presence and emotion are related only for arousing stimuli	30	• UCL-PQ • ITC-SOPI	• PANAS • VAS	No	• No prescribed medication (except oral contraceptives) • no diagnosed emotional/ psychological disorder • No therapy/counselling • Normal vision • fluent in English	Not much critical reflection on their design process
Harmon-Jones [10]	Create effective emotion measuring device (4 iterations, measured with MTurk)	• Anger, • Anxiety, • Desire, • Disgust, • Fear, • Happiness, • Relaxation, • Sadness	DEQ is sensitive to emotions	1531	N/A	• DEQ • PANAS	No	• Adults	This needs to be used and verified by different studies
Riva [19]	Evaluate elicitation of emotions with different presence levels	• Anxiety • Relaxation	VR parks can elicit anxiety and relaxation, emotion-presence relationship	61	• UCL-PQ • ITC-SOPI • 10-pt. scale during immersion	• PANAS • VAS • STAI • 10-pt. scale during immersion	No	• <18 on BDI • certain injuries and emotional problems	No physiological measures used; small statistical power because small sample size; cannot extend presence relationship to other emotions
Rivu et al. [62]	Evaluate emotion elicitation in VR compared to real life	• Happiness • Sadness • Anger • Excitement	VR's emotion elicitation is similar to reality; personalities do not affect emotions	39	N/A	• SAM • TIPI (control)	Yes	N/A	Physiological measurements should be used; because multiple emotions were focused on, we cannot determine the validity of this for sadness
Susindar [64]	Compare VR and desktop in inducing fear/anger	• Fear • Anger	Target emotions influence decision-making behavior more in VR	11	N/A (SSQ used for control)	• PANAS-X • Informal post-review • Physiological	No	• <10 hours VR experience	a measure to evaluate how long the elicited emotion persisted, and to measure presence could be incorporated
Gilpin et al. [18]	Comparing the types of sadness induced by different elicitation techniques	• Sadness	The type of sadness induced by VR is different to the type of sadness induced with other elicitation techniques	60	ICT-SOPI	• PER • Physiological measures: HR, RR, RSA	Unclear	• STAI • <21 on BDI-II • PC-PTSD	Need more physiological measures