Synergia: A Multimodal Virtual Reality System for Creative Expression and Positive Change Through Cognitive Flow

Oana Camelia Burca, Maros Pekarik, and Brian Bemman(✉)

Aalborg University, 9000 Aalborg, Denmark
maros.pekarik@pm.me, bb@create.aau.dk

Abstract. In recent years, virtual reality (VR) technologies for positive change have emerged as a way to combat various physical and mental issues, such as anxiety and depression, which have been linked to an increased use of digital technologies. Moreover, a state of *cognitive flow* in VR has been shown to have positive effects on human well-being. When designing for such VR technologies that can support positive change, feedback through aural and visual stimuli as well as interaction through movement have been suggested. However, evaluations of cognitive flow when using such technologies and with those designed for creative and artistic expression are lacking. In this paper, we present the multimodal VR system, *Synergia*, which encourages creative and artistic expression through bodily movement that is used to generate aural and visual feedback. In particular, participants' experiences with Synergia were evaluated using the Flow State Scale (FSS) in conjunction with semi-structured interviews. Our results indicate that Synergia shows promising potential for inducing aspects of cognitive flow related to increased concentration and an autotelic experience. Furthermore, these findings highlight the importance of multi-modalities to the flow experience but suggest also that visual-sound mappings may present a problem when designing for similar systems in the future.

Keywords: Interactive art · Positive technology · Virtual reality · Cognitive flow · Cross-modality · Movement

1 Introduction

As technology has become an increasingly more integral part of what makes a person's life both productive and fulfilling, a growing interest from researchers in human-computer interaction (HCI) and artists alike in designing technologies for improving mental and physical well-being has emerged [16]. Studies across a variety of disciplines have shown that virtual reality (VR), in particular, can have a positive impact on human flourishing, such as improving mental wellness [30], supporting mindfulness [32], improving physical and psycho-social well-being [18]

A. Brooks et al. (Eds.): ArtsIT 2020, LNICST 367, pp. 271–290, 2021.
https://doi.org/10.1007/978-3-030-73426-8_16

as well as inducing a state of cognitive flow – a mental state largely characterized by an increased focus desirable in many physical and creative activities [5,6]. Designing effective VR technologies for achieving these ends, however, remains challenging. Moreover, evaluations of existing systems designed for both cognitive flow and creative pursuits (as opposed to e.g., computer gaming) are limited [27].

In this paper, we present the multimodal VR system, Synergia, which was designed to allow for creative expression and to elicit positive mental change in users through cognitive flow. Concretely, a user's experience with Synergia involves 'painting' with free-flowing particles in a virtual environment and creating sound using the movement of their arms. We evaluate how well our system brings about cognitive flow using a standard questionnaire known as the Flow State Scale (FSS) with two experiments intended to (1) establish which modalities of the system – sound, visuals, or a combination of both – contribute most to cognitive flow and (2) determine for the result from (1) which of two possible design elements – a visual response to sound or not – factor most significantly. In Sect. 2, we provide an overview of cognitive flow, strategies for designing immersive, interactive technologies for positive change, as well as those that have been employed in mapping the modalities of sound and visuals in multimodal systems. In Sect. 3, we explain the conceptual design of Synergia, how this design was improved through usability testing, and provide insight into the implementation of the system. In Sect. 4, we present the two aforementioned experiments and discuss our findings as they relate to cognitive flow and the design of our system. In Sect. 5, we conclude with a discussion of possible future work.

2 Related Work

In this section, we discuss the nine dimensions which characterize cognitive flow and their importance to creative expression and practice. We conclude with an overview of design considerations for immersive, interactive technologies for positive change. In particular, we discuss such systems for VR and cognitive flow as well as various mapping strategies that have been employed in designing multimodal systems.

2.1 Cognitive Flow

Cognitive flow is a multidimensional mental state largely characterized by the feeling of being fully immersed in and focused on an activity which is intrinsically rewarding [9]. Situations in which a person may experience being in a state of cognitive flow include being particularly 'locked in' during a physical activity, sport, computer game, or meditation [7,9]. The flow experience has often been described as 'optimal' as one feels an accompanying profound sense of enjoyment which has the potential for long-lasting effects on a person's well-being [9]. The markers of cognitive flow have been categorized along nine dimensions in [9] and three categories in [7], shown in Table 1.

Table 1. The nine dimensions of cognitive flow and their respective explanations divided into the three categories, "Antecedents", "Characteristics", and "Consequences", of being in a state of flow, as described in [7,9].

Category		Flow Dimension	Explanation
Antecedents	1.	A challenging activity that requires skill	Accomplishing a physical or mental activity while being challenged enough by its requirements
	2.	Clear goals	The objectives of the activity are clear
	3.	Immediate feedback	Straightforward and immediate feedback allows for understanding how well one is doing
Characteristics	4.	Concentration on the task at hand	Worries and concerns from everyday life disappear from consciousness
	5.	The merging of action and awareness	Attention is completely absorbed by the activity which becomes spontaneous, almost automatic
	6.	A sense of control	Taking control over a novel, unpredictable situation without conscious control
Consequences	7.	Loss of self-consciousness	The concern for the self disappears allowing for immersion and unity with the environment
	8.	Transformation of time	Perception of time is altered – time seems to pass slower or faster
	9.	Autotelic experience	The activity is an end in itself, self-contained and intrinsically rewarding

As shown in Table 1, the nine dimensions of cognitive flow according to [7,9] consist of three antecedents, three characteristics, and three consequences. These antecedents involve the difficulty of an activity being in alignment with a person's skill level, clear goals for completing this activity, and immediate feedback on this person's progress. The characteristics of cognitive flow are full concentration, a merging of action and awareness, and an increased sense of control. The consequences of being in a state of cognitive flow are a loss of self-consciousness, a transformation of time, and a sense that this activity is intrinsically rewarding.

Importantly, cognitive flow has been noted as being a significant facet of creativity and artistic expression [6]. Creativity has been defined as an "interaction among aptitude, process, and environment by which an individual produces a perceptible product that is both novel and useful as defined within a social construct" [29, p. 90]. In [6], semi-structured interviews with artists demonstrated that the creative process often depends in part on the activation of flow states, with artists intentionally striving to foster the conditions and preconditions of flow. When asked to describe their creative process, respondents explicitly mentioned flow and associated this process with being in a meditative state or having a profound sense of concentration and calmness. Further analysis of their responses revealed significant overlap between the creative process and the nine dimensions of cognitive flow such as having a sense of purpose in creating an artwork which serves to clarify one's goals, receiving immediate feedback as this artwork takes shape, requiring a balance between challenge and skill when working with their chosen medium, as well as being an intrinsically rewarding pursuit, among others [6].

Measuring Cognitive Flow. Assessing whether or not a person is in a state of flow is not a straightforward process due to the complexity and subjective nature of such an experience. Nonetheless, several methods for measuring cognitive flow in individuals have been suggested, such as interviews, questionnaires, experience sampling (e.g., through diary studies), and various physiological measures (e.g., heart rate and eye tracking) [25–27,33]. Interviews were particularly useful to researchers in the early stages of developing theories around the construct of cognitive flow as these can often provide a rich understanding of a person's subjective experience [24]. However, questionnaires presently comprise the bulk of methods employed in research on this phenomenon – whether it be cognitive flow in physical activity [15], media use [35], work settings [31], computer games [8], or task absorption [20]. Perhaps the most widely used of these questionnaires is the Flow State Scale (FSS) which was designed to measure the extent to which a person experiences a state of cognitive flow during physical activity shortly after having completed it [15]. The FSS questionnaire consists of 36 Likert items using a 5-point rating scale (ranging from 1 being 'strongly disagree' to 5 being 'strongly agree') grouped into sets of four items according to the nine dimensions of cognitive flow described in [9] and shown previously in Table 1. Examples of these items include "I felt I was competent enough to meet the high demands of the situation", "I performed automatically", "I knew what I wanted to achieve", "I found the experience extremely rewarding", and "Time seemed to alter (either slowed down or sped up)". Due to the difficulty of measuring cognitive flow, it has been suggested that combining multiple methods consisting of both qualitative and quantitative data can provide a more valid measure and understanding for researchers [15,33]. Further still, in many situations such as user testing of various systems designed for cognitive flow, large numbers of participants are not always possible. Even with a sufficient number of participants, multiple measures and lengthy questionnaires are not often practical in terms of time. For these reasons, other approaches such as surveys have been employed [12], or such systems are not fully evaluated at all [27].

2.2 Designing Immersive Interactive Technologies for Positive Change

Artists and researchers, working in HCI-related endeavours and who are interested in developing immersive, interactive systems that have a positive impact on the well-being of the people who use them, are tasked with determining how exactly to design such technologies. This design process can often be quite complex and involve, for example, guiding theories of interaction from the fields of psychology or persuasive technology as well as methods intended to prioritize the user's needs. Moreover, with complex multimodal systems such as those found, for example, in VR or when creating new musical interfaces, knowledge concern-

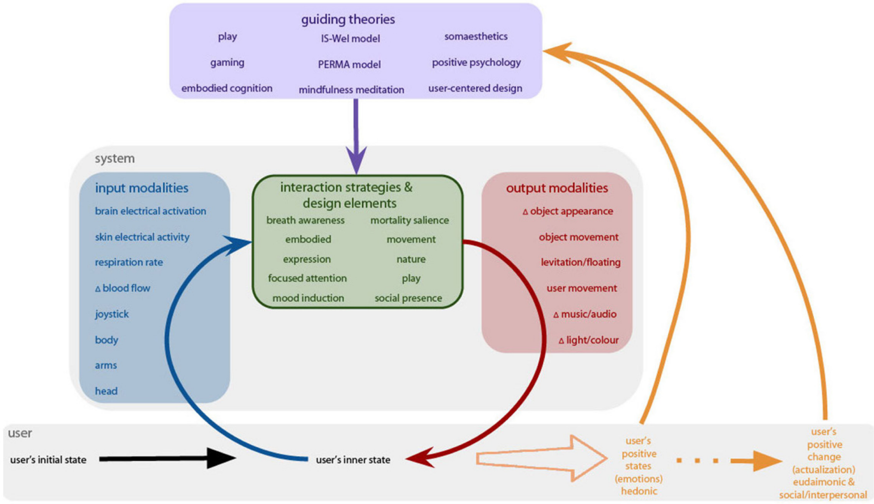


Fig. 1. A framework for designing immersive, interactive systems which elicit positive states and produce experiences which support positive change in users (as taken from [16]).

ing appropriate mapping strategies of these various modalities and their effects is crucial.

With these points in mind, in work by [16], the authors constructed a suggested framework shown in Fig. 1, based on observations from existing literature and systems of the various design elements, interaction strategies, input-output modalities, and the positive effects these had on users, for designing immersive, interactive technologies that promote positive states. The researcher begins using this framework in Fig. 1 by identifying a certain positive change for the user that he or she wishes to design for. This change is grounded in one or more possible theories or models (purple box at top), which, in turn, influence the choice of design elements and interaction strategies (green box at middle). These elements and strategies include, for example, breath awareness, expression, and play, among others, and collectively, these form a feedback loop in connection with the input and output modalities (blue box at left and red box at right, respectively). When the user interacts with the system, input data is collected through various possible modalities such as physiological measures, bodily movement, or the use of controllers, and transformed by the system into one or more possible output modalities such as object movement or changes in sound or visuals. These output modalities are intended to influence the user's inner state which, in turn, affects his or her interaction with the system, and thereby leading to the intended positive change [16].

As noted in [16], there is a relatively low number of such technologies for positive change that use controllers as input modalities – further stating that "traditional controllers do not map well to eliciting positive states" and that their use might "lead to a break in presence, immersion and flow ... distracting from the goal of eliciting positive states" [16, p. 13]. Of all the systems the authors surveyed, output modalities which produced changes to an object's appearance, music/audio, and light/color were found to be the most common. Furthermore, the interaction strategies and design elements of play and movement were also shown to have an important role in inducing positive change and led to a sense of curiosity, imagination, and embodiment for users which gave them the opportunity to interact with the system in a natural way.

As a framework for the design of future immersive, interactive technologies such as VR, the authors note how, in particular, mappings between physical and virtual movements can facilitate users' immersion in their experience [16]. According to the authors, such immersion is facilitated by natural control, mimicry of movement, proprioceptive feedback, and some physical challenge. Moreover, they suggest that any kind of movement, whether it be user movement or object movement, is correlated with positive states such as calmness, clarity, and focus while sensory changes can enhance relaxation, enjoyment, balance, and harmony for users [16]. Perhaps most importantly, the authors address the need for further investigation and empirical evidence for determining how exactly immersive, interactive technologies can elicit positive states and support positive change.

Cognitive Flow in Creative Applications of Virtual Reality. Cognitive flow and its related constructs of presence and immersion have been extensively studied subjects in computer gaming and VR largely also in the context of gaming [5,22,36]. Far less research, however, has explored cognitive flow in the context of creative and artistic applications of VR – despite their strong connection – and of these efforts, many lack formal evaluations [27]. Nonetheless, researchers have noted several criteria for designing explicitly for cognitive flow in virtual environments with creative applications in mind, such as those used in the creation of *Flow Zone* [27], a cross-modal VR system for music creation and visual exploration through movement. Moreover, these design criteria overlap significantly with several of the dimensions of cognitive flow (shown in Table 1). According to [27], such virtual environments must (1) maintain a challenge-skill balance, (2) present clear goals, (3) facilitate concentration on the present moment, (4) encourage strong interactions for inducing a sense of control, and (5) offer a space for inward motivation. Moreover, the rich environments provided by VR and cross-modal stimuli, deep embodiment through music and movement, and the intrinsically rewarding experience of creative expression are further noted design elements for cognitive flow [27].

2.3 Mapping Strategies for Sound, Visuals, and Movement in Interactive Systems

Simply knowing which of the possible input and output modalities (as shown e.g., in Fig. 1) one wishes to use in the design of a system is insufficient, as how these modalities can be mapped to one another can vary significantly. Moreover, the choice of one particular mapping over another affects the expressiveness afforded to the user and ultimately, the overall effectiveness of the system. There are several variations of possible mappings whether one is working with mapping sound to movement when creating, for example, new musical interfaces, or mapping visuals to movement as is often required in interactive art installations.

In addition to a many-to-one or *convergent* mapping, in which several controls affect a single parameter, and its inverse – a *divergent* one-to-many mapping, other variations such as one-to-one and the more complex many-to-many are possible [13]. It has been suggested in [13], however, that when working with sound and movement, mappings which are not one-to-one are generally more engaging and less frustrating for users. Such strategies encourage a certain level of effort in which "the output sound energy should be in some way proportional to the amount of movement, momentum, or acceleration" [11, p. 157]. On the other hand, having too many mappings can unnecessarily increase the complexity of interaction in a way which is confusing for users, so a proper balance must be found.

When tasked with determining effective mappings of visuals to movement, artists who design interactive art installations have employed a number of compelling strategies with various goals in mind e.g., amplifying the expressiveness of gestures [3], enhancing the 'liveness' of performative movement [4], establishing a sense of presence [28], extending body awareness [19], and stimulating creative expression [2]. Many of the same considerations when mapping sound to movement, are relevant to mapping visuals to movement as well as to visuals to sound, such as the visual perceptual correlates to one's movements (e.g., what constitutes a 'calm' looking visual in response to 'calm' movements) and the phenomenological effects of cross-modal stimuli.

3 Design of Synergia: A VR System for Creative Expression and Positive Change

In this section, we provide an overview of how Synergia was designed according to an adaptation of the aforementioned framework for designing immersive, interactive technologies for positive change [16] (shown in Fig. 1) that includes design elements from [27] for inducing cognitive flow in virtual environments (discussed in Sect. 2.2). Next, we discuss how this framework informed the mapping of sound and visual components found in Synergia to arm movements and provide the results of our usability testing of this design. We conclude with an overview of the implementation details of the final design of the system.

Figure 2 shows the framework from [16], adapted to include the design elements suggested in [27], used in guiding the design elements, interaction strategies, and input/output modalities in Synergia. As noted in Fig. 2, Synergia is a multimodal VR system which makes use of the arms as input modalities (blue box at left) that, through creative and expressive movements (green box at middle), produce output modalities of changes in sound, light appearance, and

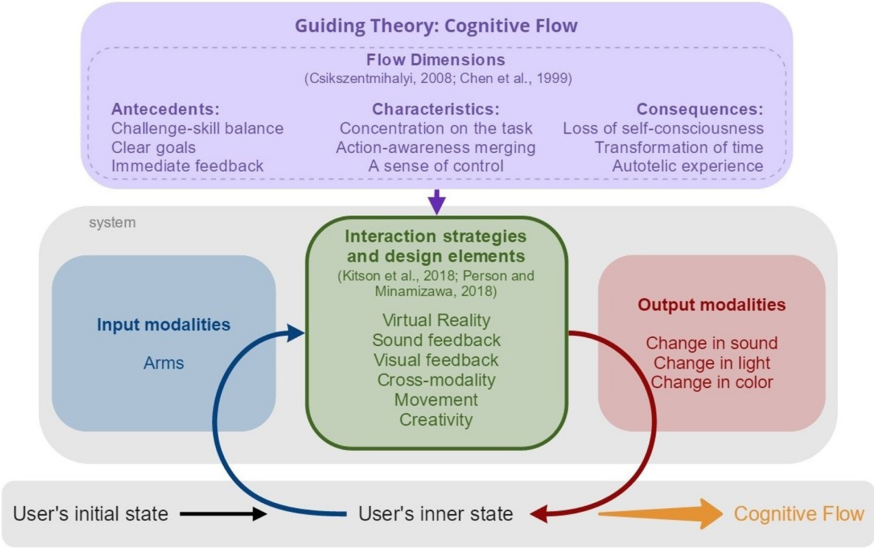


Fig. 2. Proposed design framework for Synergia – an adaptation of a framework in [16] for designing immersive, interactive technologies for positive change which includes design elements suggested in [27] for inducing cognitive flow in virtual environments. (Color figure online)

color (red box at right). These output modalities, in turn, serve as aural and visual feedback for the user. In line with the original framework [16], this aural and visual feedback can be used to stimulate concentration and focused attention, inform the users about their performance, facilitate emotional expression, and support feelings of contentment and harmony. Where our proposed design framework in Fig. 2 differs considerably from the suggestions presented in [16] is in its guiding theory (purple box at top). With Synergia, we have chosen to use the dimensions of cognitive flow [9] as well as its preconditions [17,27] as criteria for guiding the design of the system's sound and visual feedback (discussed further in Sect. 3.1). Furthermore, our choice to use VR as an interaction strategy and design element (green box at middle) was motivated by its potential to provide several preconditions of cognitive flow, as discussed in Sect. 2.2, through an immersive, rich environment as well as to possibly facilitate further dimensions of flow such as an increase in concentration and a transformation of time. Moreover, a cross-modal interaction between sound and visuals (green box at middle) was incorporated as a means for increasing the complexity of the VR environment and in doing so, potentially provide a more appropriate challenge-skill balance for the user.

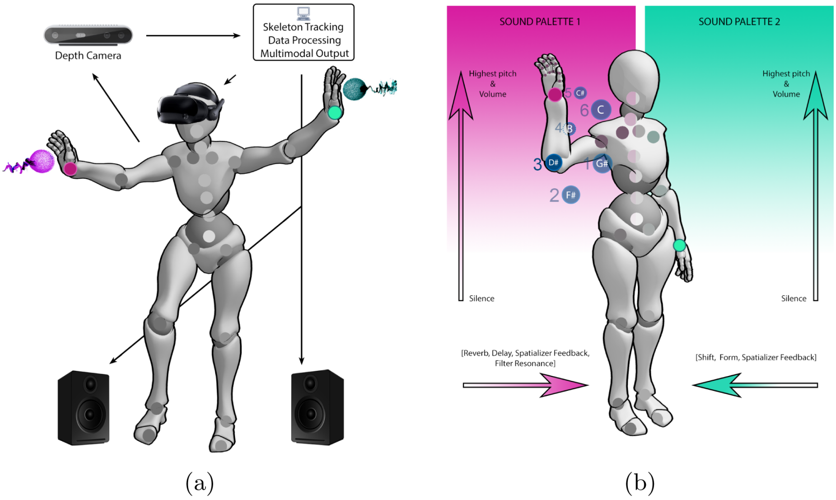


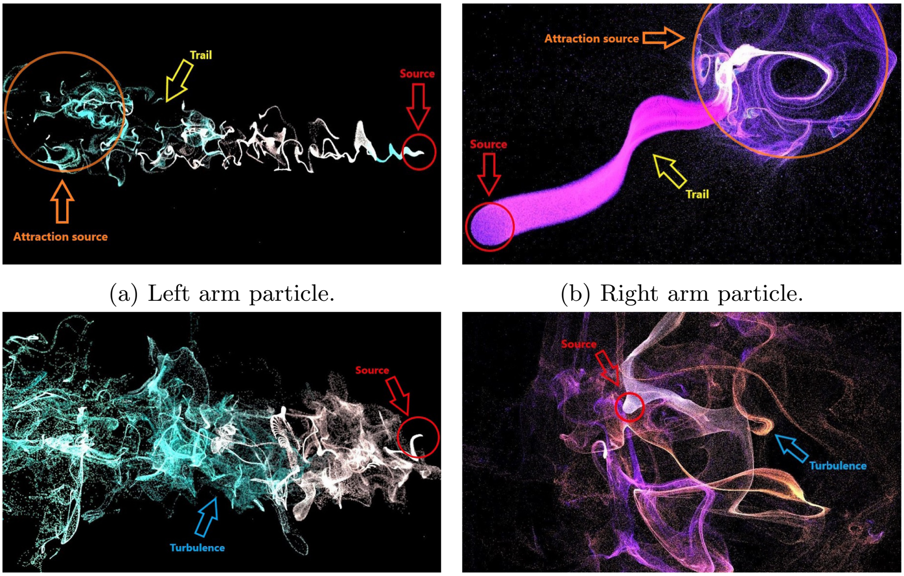
Fig. 3. Conceptual design of the Synergia system in (a) and its mapping of various parameters of sound to movements of the arms in (b).

3.1 Mapping Sound and Visuals to Arm Movements

Synergia was designed to allow users the ability to creatively 'paint' with free-flowing particles in a virtual environment and create sound using the movement of their arms. This sound, in turn, affects the visual appearance of the particles and adds to the complexity of the interaction in a way which we suggest will enrich the experience. The conceptual design of the system and mapping of sound to arm movements is illustrated in Fig. 3(a) and (b), respectively.

As indicated in Fig. 3(b), the right and left arms of the user control separate sound palettes, chosen so as to be distinguishable but consonant to the ear. The pitch of each sound palette is generated through a one-to-one mapping with one of six possible positions of the user's elbows relative to his or her torso (as indicated by the six blue dots circling the figure's shoulder). A divergent, one-to-many mapping utilized the position of the wrists to control several additional parameters of sound such as reverb, and delay, among others. As a means for facilitating an effort-based mapping strategy (discussed in Sect. 2.3), high and low positions of the wrists were further mapped to the amplitude of each sound palette (as indicated by the vertical arrows).

Detailed images of the left and right hand particle systems controlled by the movements of a user's arms (as depicted in Fig. 3(a)), are shown in Fig. 4(a) and (b), respectively. Each particle system consists of a source, responsible for emitting the actual particles (circled in red), a trail (circled in yellow) that slowly disappears over the lifetime of the particles, and an associated attraction source (circled in orange) which invisibly encapsulates the body of the user and provides a direction in which the particle's trail can float towards. The left and right hand particle systems differ in their size, visual appearance, and gradient of



(a) Left arm particle.

(b) Right arm particle.

(c) Left arm particle with sound-induced turbulence.

(d) Right arm particle with sound-induced turbulence.

Fig. 4. Two example particles created with Synergia using a person's moving left hand in (a) and moving right hand in (b), shown in gradients of turquoise and pink, respectively. These same particles are shown in (c) and (d) with a turbulence effect after having been perturbed by certain frequencies of sound produced by movement. (Color figure online)

color so that they are easily distinguished in the same way as their corresponding sounds described above. With the preconditions of flow discussed in Sect. 2.2 in mind, we introduced an additional element of complexity to the interaction by making the visuals reactive to certain frequencies of sound. As shown in Fig. 4(c) and (d), the particles can be perturbed with a turbulence effect in which their respective color gradients are changed and their visual appearances are made more disparate.

Prototype Testing. Usability tests were conducted with nine participants who were asked to complete a set of tasks using their right and left hands separately while interacting with three versions – sound alone, visuals alone, and sound and visuals – of a prototype design of Synergia. Participants were asked to think aloud as they attempted each task and upon completion of all tasks, we conducted a brief semi-structured interview for further clarification on their experience when interacting with the system. The goal of these tests was to assess (1) how aware participants were of what they could and could not control, (2) how intuitive the various mappings (i.e., sound-movement, visuals-movement, and sound-visuals) were, and (3) how immediate the feedback in the form of visuals and sound

Table 2. Usability tasks participants were asked to complete while interacting with three versions – sound alone, visuals alone, and sound with visuals – of Synergia. Note that each task for a given version was completed for each arm except for the version with both sound and visuals.

	Version	Arms	Tasks
1.	Sound	Both	1. Produce a high pitched sound 2. Choose an interesting sound and repeat it
2.	Visuals	Both	1. Draw a rectangle 2. Bring the particles closer to the body
3.	Sound and Visuals	Single	1. Right arm only: Create a 'smoky' visual effect 2. Left arm only: Create a 'chaotic' visual effect

was. Table 2 shows the set of usability tasks carried out in the design testing of Synergia.

The participants' successful completion (or not) of the usability tasks, their thoughts while attempting them, and their feedback expressed during the following interviews demonstrated that (1) the sound-movement mapping could be made more intuitive, and (2) the visual-movement mapping lacked full control. With respect to (1), we reduced the number of one-to-many sound parameters manipulated by the movements of each arm and incorporated a velocity component to the movements such that more energy and effort would be required to produce sound. The problem in (2) appeared to be caused by the inherent latency of the system, but several participants noted that not being fully in control proved to be a motivating challenge so we opted not to make any adjustments. Another participant expressed that the attraction speed of the particles was too fast, so we connected their speed to the overall amplitude of sound so that the attraction occurred more slowly. While all participants were able to successfully complete the tasks associated with the sound-visual mapping, they expressed that they found it challenging to determine what aspects of their actions contributed to their experience. For example, some participants correctly attributed the turbulence effect of the visuals to sound, while others misattributed this effect to rapid arm movements or nothing at all. Based on the challenging nature of this mapping and the importance to cognitive flow of having such an antecedent, we decided against making any adjustments to the sound-visual mapping.

3.2 Implementation Overview

The hardware used in Synergia is a single Intel RealSense D415 depth camera [14], an HTC VIVE head-mounted-display (HMD), and a set of floor-mounted Magnat speakers. The software for running Synergia consists of a real-time, machine learning skeleton tracking SDK from Cubemos [10], a 3D virtual environment created in Unity [34], sound processing through Ableton Live [1], and

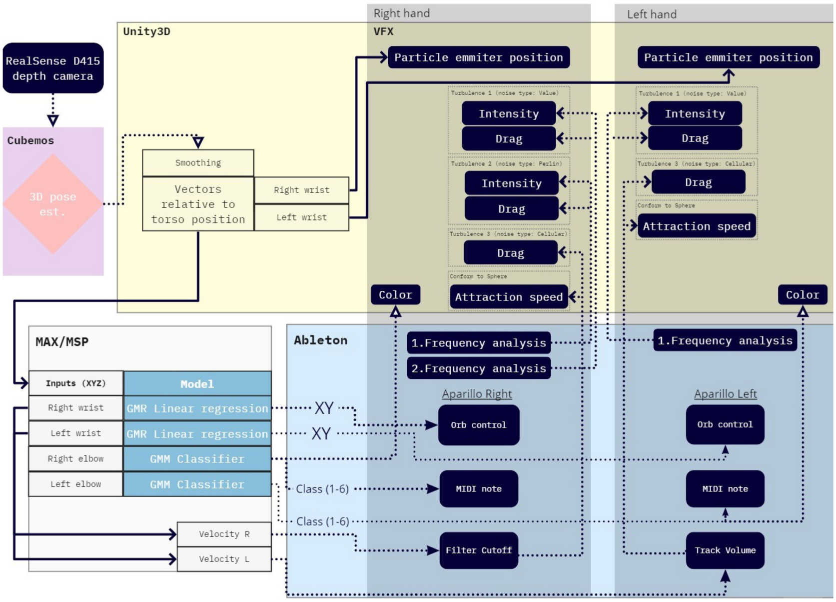


Fig. 5. Overview of the implementation details of the Synergia system using Cubemos, Unity3D, Max/MSP, and Ableton Live.

a Max/MSP [21] patch available in the MuBu toolbox [23] for interactive sound and motion applications using machine learning. All communication between programs is handled through a network using the Open Sound Control (OSC) protocol. Figure 5 shows the overall technical diagram of the final design of the Synergia system[1].

As shown in the upper left corner of Fig. 5, Synergia first uses Cubemos to detect the coordinates in 3D space corresponding to the 18 skeletal joints of a user. These coordinates and corresponding confidence levels for each joint are then sent to Unity (yellow box) where the positions of the wrist and elbow joints for each arm are computed relative to the coordinates of the torso. Subsequently, these positions are used to generate locations in virtual space for the sources that produce the particles which have been set to emanate at a slight offset from the left and right wrist joints using Unity VFX (grey overlays). Determining the positions of the joints in this way addresses the problem of individuals with different body sizes by ensuring that their arm movements are independent from their body position relative to the location of the depth camera. The positions of these joints are then sent from Unity to a Max/MSP patch (grey box at lower left) which uses MuBu to handle their respective mappings to the sound palettes. The actual sound generation is carried out in Ableton Live (blue box)

[1] The complete code for the Synergia system can be found in the following repository: https://github.com/marospekarik/synergia.

and the mapping of the particles to color and any turbulence effects is handled by Unity VFX. In VFX (grey overlay on yellow box), the 'color' parameter is assigned a 'sample gradient' operator that returns a color value depending on the particular pitch produced by the locations of the user's elbows while the 'attraction speed' parameter is connected to the amplitude of the sound. Two parameters, 'intensity' and 'drag', are then used by the 'Turbulence 1' block to modify the appearance of the particles for both hands based on their respective pitch frequencies. These same parameters in the 'Turbulence 2' block further modify the appearance of the particles for the right hand based on the range of frequencies present. Finally, the 'drag' parameter in the 'Turbulence 3' block is modified by the amplitude of the sounds for both hands.

4 Evaluation

In this section, we discuss two experiments carried out with Synergia. The first of these experiments was intended to discover which of the three design versions (discussed in Sect. 3.1) of our system – sound alone, visuals alone, and both —- contributed most to cognitive flow. The second of these experiments was carried out in order to establish for the version of our system which most contributed to cognitive flow, as determined by the first experiment, which of two design elements had the greatest influence along the nine dimensions of cognitive flow. In analyzing the participants' ratings from the FSS questionnaire, we have elected to consider the data as interval (as opposed to ordinal), as has been commonly done in previous work [7,25]. In our case, this data were the mean participant ratings in the aforementioned two experiments both across the entire FSS questionnaire and along the nine dimensions of cognitive flow. All participants were informed of what actions would take place during both experiments as well as how their data would be used. Consent was obtained in accordance with the participating university's ethical guidelines for conducting experiments with human participants.

4.1 Experiment 1

In our first experiment, we constructed a repeated measures design in which each participant interacted with Synergia in three conditions: (1) sound alone, (2) visuals alone, and (3) both sound and visuals. Conditions were completely counter-balanced so that any observed order effects were minimized. Following each respective condition, participants were asked to complete the Flow State Scale (FSS) and after having completed the third questionnaire, participants took part in a follow-up semi-structured interview designed to collect qualitative data concerning the nature of their experience with the system and any perceived positive change to their physical or mental well-being.

Participants. We collected data from seven, volunteer participants (six male and one female) of largely university students with an average age of 25.3 ± 3.8

years. Of these participants, three claimed to have 10 or more years of experience in art, while three others noted less than 10 years of experience and one stated no such experience at all. Four of the participants stated having more than 10 years of experience with music while the remaining three stated having no such experience. Furthermore, six participants stated having previous experience with VR with only one having no such prior experience.

Procedure. Participants were asked one at a time to enter the room in which the Synergia system had been set up and two experimenters were present. They were then asked to respond to a few questions concerning their age, gender, and prior experience with art, music, and VR. Afterwards, they were informed that they would be experiencing a VR system a total of three separate times and that with each time, their arms should be used to interact with this system. Participants were fitted with the VR head-mounted display (HMD) and situated in a standing position at the center of an approximately 2.5 by 2.5 m square area facing the depth-camera. Each condition was stopped by an experimenter after 10 min and the participant was given 5 min of resting time without the HMD followed by time to complete the FSS questionnaire. Following the third condition and the completion of the third FSS questionnaire, participants were asked to take part in an interview with both experimenters, where further information was gathered regarding their experience such as "Which experience had a stronger impact on you and why?", "How natural was it for you to control the environment through movement?", and "Were you able to anticipate what would happen next in response to your actions?". Following the completion of this interview, the test was concluded and participants were thanked for their time.

Results and Discussion. The mean ratings from each of the three conditions met the assumptions for a repeated measures ANOVA. However, the mean rating of cognitive flow proved statistically insignificantly different across the three conditions ($F(1.15,6.93) = 1.315$, $p > 0.2$, ges $= 0.113$). It is worth noting that treating the actual participant ratings as ordinal data (rather than considering the mean rankings as interval data) in which each individual item on the FSS is a separate observation, the differences between the three conditions proves significant when using a Friedman ranked sum test ($F_r = 23.2$, $p < 0.001$), where the assumptions for this test have similarly been met. Nonetheless, some interesting trends can be observed in the mean ratings. Figure 6 shows box plots of the mean ratings from the FSS for our participants interacting with Synergia across the three conditions in (a) and along the nine dimensions of cognitive flow for these same conditions in (b).

As shown in Fig. 6(a), the mean rating (noted with a diamond symbol) in the condition with both sound and visuals was highest of all, followed by the second condition with only visuals, and lastly, the third condition with only sound. When looking at Fig. 6(b), one will note this trend persists across a majority of the dimensions of cognitive flow (i.e., CHAL, GOAL, FDBK, ACT, TRAN, and

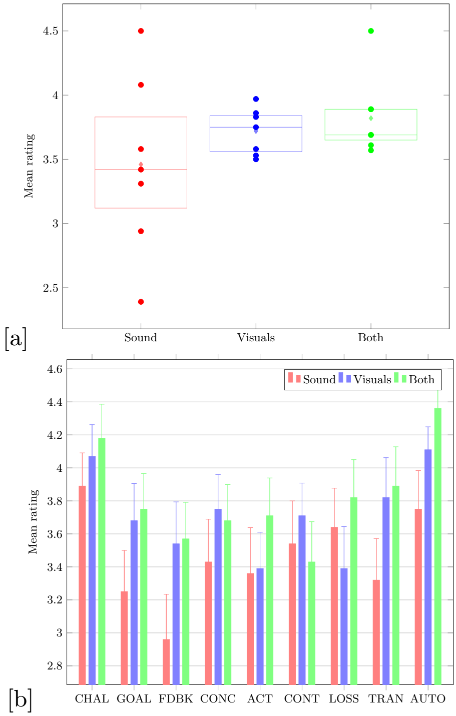


Fig. 6. Mean participant ratings from the Flow State Scale (FSS) questionnaire after interacting with Synergia across each of the three conditions, sound alone, visuals alone, and sound and visuals, in (a) and along the nine dimensions of cognitive flow for these same conditions in (b).

AUTO) except for those concerning concentration, a sense of control, and loss of self (i.e., CONC, CONT, and LOSS), however, these differences in mean ratings between conditions within all dimensions proved insignificant ($p > 0.1$). Were we to again consider the actual ratings (as opposed to their means), the difference between conditions observed in the dimensions concerning clear feedback and a sense of a transformation of time (i.e., FDBK and TRAN) proved significant ($p < 0.05$).

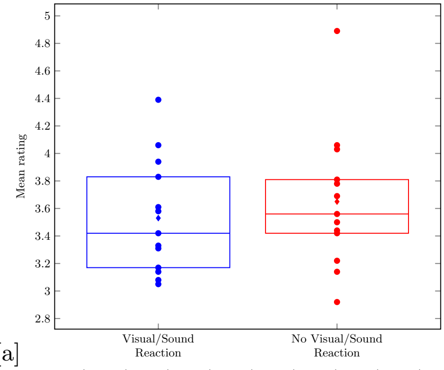
The fact that for the majority of dimensions, the sound and visual condition resulted in an overall greater mean rating than the others suggests that having both modalities is an important component of what may constitute a flow experience with Synergia. Concerning the dimensions of concentration and a sense of control, the visual only condition was rated higher than the other two conditions likely because users are in general more acutely attuned to visual stimuli rather than sound (as indicated by the lowest overall rating for feedback in the sound only condition) and the visual particles were designed to have fewer adjustable parameters in comparison to the sound. Moreover, the mapping of movement-generated sound to affect the visual appearance of particles possibly proved difficult for participants trying to understand the effects of their actions which might explain the comparative lack of perceived control in the sound and visual condition. This lack of control in the third condition likely factored into the greater perceived challenge-skill balance as well. However, participants found the perceived goals and feedback more or less equally as clear in the second and third conditions, possibly indicating that any perceived lack of control in the third condition was not overly negative. More promising still is that participants found the sound and visual condition more intrinsically rewarding (i.e., AUTO).

4.2 Experiment 2

Our second experiment was carried out in response to the finding from our first experiment that the sound and visual version of Synergia appeared to contribute most to cognitive flow but that some aspect of its design elements seemed to result in a reduced sense of concentration and control for these participants when compared to the other versions (i.e., sound alone or visuals alone). Based on the qualitative feedback received during the interviews, the nature of how these visuals reacted to the sound appeared responsible for this response from these participants. Moreover, it is likely that participant fatigue was a negative factor in the design of the first experiment. For these reasons, in our second experiment, we adopted a paired and completely counter-balanced two samples design in which participants interacted with a sound and visual version of Synergia (as was done in the third condition of experiment 1) where the appearance of the visuals was changed in response to the sound produced by the participants' arm movements (as shown e.g., by the turbulence effect added to the particles in Fig. 4(c) and (d)), and a version where the appearance of visuals did not react to the sound produced by the participants' arm movements (as shown e.g., in Fig. 4(a) and (b)).

Participants. We collected data from 13, volunteer participants (seven male and six female) with an average age of 28.4 ± 4.7 years. Of these participants, four claimed to have 10 or more years of experience in art, while four others noted less than 10 years of experience and five stated no such experience at all. Five of the participants stated having more than 10 years of experience with music while two others noted less than 10 years of experience with music and the remaining six stated having no such experience. Furthermore, six participants stated having previous experience with VR while seven others reported having no such prior experience.

Procedure. The procedure for the second experiment followed that of the first experiment with only minor adjustments made. First, participants were asked to fill out the FSS questionnaire only twice – once after each of their two respective conditions. Second, participants were given up to 20 min to interact with the system instead of 10 min. Finally, the area in which participants were free to explore in the physical space was blocked from view by the experimenters so that the users might feel less inhibited in their interactions with the system.



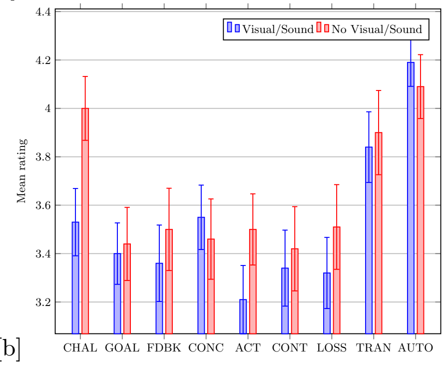


Fig. 7. Mean participant ratings from the Flow State Scale (FSS) questionnaire after interacting with a sound and visual version of Synergia across each of the two conditions, visuals reacting to sound and visuals not reacting to sound in (a) and along the nine dimensions of cognitive flow for these same conditions in (b).

Results and Discussion. The mean ratings from each of the two conditions met the assumptions for a two-tailed dependent (paired) samples Student's t-test. However, the mean rating of cognitive flow proved statistically insignificantly different across the two conditions ($t = -1.4842$, $p > 0.1$). It is worth noting that, as with the first experiment, treating the participant ratings here as ordinal data in which each individual item on the FSS is a separate observation, the differences between the two conditions proves significant when using a Wilcoxon signed-rank test ($W = 10670$, $p < 0.05$), where the assumptions for this test have also been met. Nonetheless, as with the first experiment, some interesting trends can be observed in the mean ratings.

As shown in Fig. 7(a), the mean rating in the condition with visual particles reacting to sound is marginally lower than in the condition with no such reaction. When looking at Fig. 7(b), this difference holds across most of the dimensions of cognitive flow (i.e., CHAL, GOAL, FDBK, ACT, CONT, LOSS, and TRAN) with the notable exceptions being concentration and an autotelic experience (i.e., CONC and AUTO). Of the nine dimensions, however, only the difference in mean ratings for a challenge-skill balance proved significant ($p < 0.05$) with all other dimensions proving insignificantly different ($p > 0.5$). Were we to again consider the actual ratings (as opposed to their means), the difference between conditions observed in the dimensions concerning both a challenge-skill balance and merging of action and awareness (i.e., CHAL and ACT) proved significant ($p < 0.05$).

Recall that our motivation for having visuals which react to sound was to increase the perceived complexity of the experience in compliance with the dimension of cognitive flow concerning a challenge-skill balance. Surprisingly, however, our results indicate that the mean rating in this dimension (i.e., CHAL) is significantly lower in the first condition with visuals reacting to sound than in the second condition without this reaction. If we look then to the dimension of increased control (i.e., CONT), the lower mean rating in the first condition when compared to the second condition suggests perhaps that participants found the added complexity prohibitively large, resulting in a perceived challenge that exceeded their perceived skill (rather than the reverse). We might interpret also that the lower mean ratings in the first condition when compared to the second condition for the dimensions concerning immediate feedback and merging of action and awareness (i.e., FDBK and ACT), indicate for the participants a sense of confusion or frustration with respect to this added complexity. Regardless, our finding with respect to the dimension of autotelic experience (i.e., AUTO), suggests that interestingly, participants still found their experience with visuals reacting to sound more intrinsically rewarding than not having this reaction.

5 Conclusion and Future Work

In this paper, we presented a multimodal VR system called Synergia for creative expression and positive change which allows users to 'paint' with particles in a virtual environment and create sound through movements of the arms. We evaluated this system through two experiments which demonstrated that (1) the use of both sound and visuals contributes most to cognitive flow over sound or visuals alone, and (2) when using sound and visuals, a cross-modal design element, which uses sound to further modify the appearance of visuals produced by arm movements, factored negatively into the overall contribution of cognitive flow. In particular, this cross-modal design element resulted in rather lower levels of perceived challenge-skill balance and feedback but contributed to an increased sense of concentration and greater autotelic experience for our participants. In future work, it would be necessary to look further into what exact parameters of the sound and visuals in Synergia contribute to the users' experiences along the nine dimensions of cognitive flow. Furthermore, it would be worthwhile to consider additional measures of cognitive flow beyond the FSS. With the work presented here, we hope that we have provided some ways forward for artists and researchers in HCI interested in positive uses of technology and looking to design VR systems for creative expression and cognitive flow.

References

1. Ableton: Ableton (2020). https://www.ableton.com/
2. Akten, M.: Body paint (2009). http://www.memo.tv/works/bodypaint/
3. Alaoui, S.F., Bevilacqua, F., Pascual, B.B., Jacquemin, C.: Dance interaction with physical model visuals based on movement qualities. Arts Technol. **6**(4) (2013)
4. Almena, M.: Transcendence: can live performance art in combination with interactive technology induce altered states of consciousness. Electronic Visualisation and the Arts (EVA) (2018). https://doi.org/10.14236/ewic/EVA2018.38
5. Bian, Y., et al.: A framework for physiological indicators of flow in VR games: construction and preliminary evaluation. Pers. Ubiquit. Comput. **20**(5), 1–12 (2016)
6. Chemi, T.: The experience of flow in artistic creation. In: Harmat, L., Ørsted Andersen, F., Ullén, F., Wright, J., Sadlo, G. (eds.) Flow Experience, pp. 37–50. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-28634-1_3
7. Chen, H., Wigand, R., Nilan, M.: Optimal experience of web activities. Comput. Hum. Behav. **15**, 585–608 (1999). https://doi.org/10.1016/S0747-5632(99)00038-2
8. Choi, D., Kim, J.: Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. Cyberpsychol. Behav. **7**(1), 11–24 (2004). The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society
9. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. HarperCollins (2008)
10. Cubemos: Skeleton Tracking SDK for Intel RealSense Depth Cameras (2020). https://www.intelrealsense.com/skeleton-tracking/
11. Dahlstedt, P., Dahlstedt, A.S.: Otokin: mapping for sound space exploration through dance improvisation. In: Queiroz, M., Sedó, A.X. (eds.) Proceedings of the International Conference on New Interfaces for Musical Expression, Porto Alegre, Brazil, pp. 156–161. UFRGS, June 2019. http://www.nime.org/proceedings/2019/nime2019_paper031.pdf
12. Hassan, L., Jylhä, H., Sjöblom, M., Hamari, J.: Flow in VR: a study on the relationships between preconditions, experience and continued use. In: Proceedings of the 53rd Hawaii International Conference on System Sciences, pp. 1196–1205 (2020). https://doi.org/10.24251/HICSS.2020.149
13. Hunt, A., Kirk, R.: Mapping strategies for musical performance. Trends in Gestural Control of Music (2000)
14. Intel: Intel RealSense Depth Camera D415 (2020). https://www.intelrealsense.com/depth-camera-d415/
15. Jackson, S.A., Marsh, H.W.: Development and validation of a scale to measure optimal experience: the flow state scale. J. Sport Exerc. Psychol. **18**, 17–35 (1996)
16. Kitson, A., Prpa, M., Riecke, B.E.: Immersive interactive technologies for positive change: a scoping review and design considerations. Front. Psychol. **9**(1354) (2018). https://doi.org/10.3389/fpsyg.2018.01354
17. Kotler, S.: The Rise of Superman: Decoding the Science of Ultimate Human Performance. Amazon Publishing (2014)
18. Lee, L.N., Kim, M.J., Hwang, W.J.: Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. Appl. Sci. **9**(17) (2019)
19. Legarnisson, E.: (un)balance (2018). http://www.interactivearchitecture.org/lab-projects/unbalance
20. Martin, A.J., Jackson, S.A.: Brief approaches to assessing task absorption and enhanced subjective experience: examining ‘short’ and ‘core’ flow in diverse performance domains. Motiv. Emot. **32**(3), 141–157 (2008)

21. Max: Max (2020). https://cycling74.com/
22. Michailidis, L., Balaguer-Ballesterand, E., He, X.: Flow and immersion in video games: the aftermath of a conceptual challenge. Front. Psychol. **9**, 16–82 (2018). https://doi.org/10.3389/fpsyg.2018.01682
23. MuBu: Mubu for Max (2008). http://ismm.ircam.fr/mubu/
24. Nakamura, J., Csikszentmihalyi, M.: The concept of flow. Flow and the Foundations of Positive Psychology, pp. 239–263. Springer, Dordrecht (2014). https://doi.org/10.1007/978-94-017-9088-8_16
25. Nijs, L., Coussement, P., Moens, B., Amelinck, D., Lesaffre, M., Leman, M.: Interacting with the music paint machine: relating the constructs of flow experience and presence. Interact. Comput. **24**(4), 237–250 (2012)
26. Novak, T.P., Hoffman, D.L.: Measuring the flow experience among web users. Technical report, Vanderbilt University (1997)
27. Person, T., Minamizawa, K.: Flow zone: inducing flow to improve subjective wellbeing by creating a cross-modal music creation experience in virtual reality. In: SIGGRAPH Asia 2018 Virtual & Augmented Reality (2018)
28. Plant, N.: Sentient flux (2015). http://nicolaplant.co.uk/sentientflux.html
29. Plucker, J.A., Beghetto, R.A., Dow, G.T.: Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. Educ. Psychol. **39**(2), 83–96 (2004). https://doi.org/10.1207/s15326985ep3902_1
30. Roche, K., Liu, S., Siegel, S.: The effects of virtual reality on mental wellness: a literature review. Ment Health Fam Med. **14**, 811–818 (2019)
31. Schaufeli, W., Bakker, A.: Utrecht work engagement scale: preliminary manual. Occupational Health Psychology Unit, Utrecht (2003)
32. Seabrook, E., et al.: Understanding how virtual reality can support mindfulnesspractice: mixed methods study. J. Med. Internet Res. **22**(3) (2020)
33. Tian, Y., Bian, Y., Han, P., Wang, P., Gao, F., Chen, Y.: Physiological signal analysis for evaluating flow during playing of computer games of varying difficulty. Front. Psychol. **8**, 1–10 (2017)
34. Unity: Unity 3D (2020). https://www.unity.com
35. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoper. Virtual Environ. **7**, 225–240 (1998)
36. Yao, S., Kim, G.: The effects of immersion in a virtual reality game: presence and physical activity. In: Fang, X. (ed.) HCII 2019. LNCS, vol. 11595, pp. 234–242. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22602-2_18