



HAL
open science

Is it Real? Measuring the Effect of Resolution, Latency, Frame rate and Jitter on the Presence of Virtual Entities

Thibault Louis, Jocelyne Troccaz, Amélie Rochet-Capellan, François Bérard

► To cite this version:

Thibault Louis, Jocelyne Troccaz, Amélie Rochet-Capellan, François Bérard. Is it Real? Measuring the Effect of Resolution, Latency, Frame rate and Jitter on the Presence of Virtual Entities. ACM ISS 2019 - Conference on Interactive Surfaces and Spaces, Nov 2019, Séoul, South Korea. 10.1145/3343055.3359710 . hal-02323605

HAL Id: hal-02323605

<https://hal.science/hal-02323605>

Submitted on 2 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Is it Real?

Measuring the Effect of Resolution, Latency, Frame rate and Jitter on the Presence of Virtual Entities

Thibault Louis, Jocelyne Troccaz, Amélie Rochet-Capellan, François Bérard

University Grenoble Alpes, CNRS, LIG, TIMC,
Gipsa-Lab Grenoble, France

(thibault.louis,jocelyne.troccaz,francois.berard)@univ-
grenoble-alpes.fr, amelie.rochet-capellan@gipsa-
lab.grenoble-inp.fr

ABSTRACT

The feeling of presence of virtual entities is an important objective in virtual reality, teleconferencing, augmented reality, exposure therapy and video games. Presence creates emotional involvement and supports intuitive and efficient interactions. As a feeling, presence is mostly measured via subjective questionnaire, but its validity is disputed. We introduce a new method to measure the contribution of several technical parameters toward presence. Its robustness stems from asking participant to rank contrasts rather than asking absolute values, and from the statistical analysis of repeated answers. We implemented this method in a user study where virtual entities were created with a handheld perspective corrected display. We evaluated the impact on two virtual entities' presence of four important parameters of digital visual stimuli: resolution, latency, frame rate and jitter. Results suggest that jitter and frame rate are critical for presence but not latency, and resolution depends on the explored entity.

Author Keywords

Presence; HPCD; User Study; Spatial Augmented Reality.

INTRODUCTION

Presence is the sensation felt when exposed to immersive realities using high fidelity devices. It is sought in many application fields like Virtual Reality (VR), teleconferencing, Augmented Reality (AR), exposure therapy and video games as it allows intuitive and efficient interaction and emotional involvement. In this paper, we refer to reality (either physical or virtual) as a complete autonomous environment governed by physical laws, as well as all the objects, living and social creatures that evolves and interact with one another in the environment. We use the term entities to regroup all kind of actors evolving in a reality (i.e. objects, creatures or peoples). As human beings, we evolve in the physical reality. In order to interact with other physical entities, our brain analyzes the data perceived by our sensory system using previous knowledge of the reality's physical laws. Our lifetime interacting experiences allow us to apprehend new entities that we never saw before. From the texture, reflections and shape of an object, we can deduce its matter, weight and purpose.

More than fifty years ago [31], research on virtual reality began to address the creation of other realities with their own entities, rules and interaction metaphors. Within VRs, people can interact with entities, stories and environments that do not or cannot exist in the physical reality. With high fidelity VR systems that reproduces the laws, visual stimuli and interactive potentialities of the

physical reality, users can be immersed into other realities and interact with virtual entities intuitively, benefiting of their own lifetime experiences. When most of the physical reality is adequate, however, it may seem wasteful to reproduce it in a virtual reality only to add a non-physical entity. A more sensible approach is to create an augmented reality where only the virtual entity is immersed in the physical world. Both VRs and ARs share the same objective of offering intuitive interaction with the transfer of skills and knowledge acquired in the physical reality.

Ultimately, with either high fidelity VR or AR, users will not be able to tell if the entity they interact with is physical or virtual. When this happens, we qualify the virtual entity as entirely present for the user. The more present an entity is, the more difficult it is for the user to identify whether the entity is a virtual or a physical one. Various definitions of presence are found in the literature, each one being specific to a given context of use.

Presence is considered beneficial and sought after in several domains from video games to professional immersive training through curing phobia with exposure therapy. Interacting with a very present object is usually intuitive; which reduces learning time and enhances training acceptability. When using virtual simulations to train for physical skills (e.g. flight simulators), it is crucial that the virtual reproduction of the physical reality is as faithful as possible, i.e. very present, in order to maximize the transfer of learning from virtual to physical. In VR games, players want to be immersed into the game's world; which requires a strong presence of its entities.

Rendering a scene with a high level of presence require to produce credible sensory feedbacks coupled with natural interaction. To date, hardware capable of such feedbacks are emerging but expensive, and are still struggling with technical barriers. As presence is a subjective feeling, researchers struggle to find a direct way to measure it, and many studies rely on questionnaires to do so. Empirical works that study the impact of devices input and output performances on presence are scarce and hard to compare given the lack of objective measurement.

Enhancing an entity's presence can be achieved by acting at several abstraction levels, from devices input/output performances to the scenario of use. In this study, we focus on the lowest abstraction level: the device technical performances. A pixelated rendering, for instance, should yield less presence than a high-resolution rendering that saturate human's visual acuity. All technical parameters do not have the same impact on presence, and literature is lacking in experimental work on the topic.

In this paper, we propose a new method to measure the contribution of various parameters towards the perceived presence of a scene. Our method relies on a statistical repetition of comparisons between 2 conditions rather than on a direct questionnaire on the feeling of presence. We argue that the method yields a more objective measurement of presence than questionnaires. Our method might enhance comparison of future studies and help to understand the factors impacting the presence of virtual scenes.

The method was implemented in a user study that questions the impact of 4 major parameters influencing the visual perception of a scene: the display's pixel Resolution, the end-to-end Latency, the display Frame rate and the motion capture Jitter (RLFJ). The interactive system tested was a handheld perspective corrected display (HPCD). This study provided first insights about the impact of these parameters on presence.

RELATED WORK

We organize this review of the related work around four topics: the various definitions of the concept of presence, its applications, the means to increase presence, and how to measure it.

Literature definitions of presence

Virtual Reality

Presence was first introduced in the VR domain as a psychological sense of “being there” in the virtual environment (Slater et al. [29]). It described the human psychological state when exposed to an immersive VR system. Immersive VR systems aim at entirely replacing users’ physical environment with a virtual environment. To do so, VR systems must, for as many of users’ sensory receptors as possible, either provide high-resolution stimuli that saturate the receptors, or silent the receptors.

Remote collaboration

In teleconferencing and telecooperation, the term telepresence refer to the feeling of being physically there in a remote workspace [23]. In this case, users want to feel as if they were present in a remote physical reality. As the focus is essentially on collaboration, the environment becomes superfluous: only the copresence of users with their remote coworkers is essential [21]. Here, the focus is on presence as a feeling of relation to social individuals, not to an entire reality. The realism of the coworkers’ avatars has been shown to positively impact the feeling of copresence as well as the understanding of non-verbal communication. Bailenson et al. tested participants feeling of copresence, verbal and nonverbal disclosure, and emotion recognition of their interlocutor in a teleconferencing communication depending on the realism of the interlocutor avatar [2]. They tested three conditions: no avatar, an impersonal polygonal avatar that reproduces the interlocutor facial expressions, and a videoconference (i.e. a high quality photo-realistic avatar). They reported a negative impact of the polygonal avatar on copresence and a reduction of nonverbal disclosure with both avatar representations.

Augmented Reality

AR transports virtual objects into the user physical reality, in opposition to VR that aims to transport the user out of the physical reality. One important advantage of AR is that applications do not have to reproduce the surrounding environment. However, the realism of the virtual object has to be very high for the illusion that it actually exists in the physical environment to work. To date, realism is limited by several technical issues listed and categorized by Kruijff et al. [13], including: occlusion handling, place stability, lightning and shadowing, accommodation and limited field of view. Chen et al. extended the notion of presence defined in VR to apply it in AR. They considered the presence of the virtual entities in the physical reality instead of the user’s presence in the virtual reality [8]. They measured AR presence by adapting the common Witmer and Singer’s Presence Questionnaire [33] to the AR context. Here, the concept of presence relates to the feeling, for users, that the virtual object is with them in their physical reality.

Applications of presence

HCI performances

With a high level of presence, a virtual entity is perceived as a physical entity, favoring the transfer of skills that were learnt in the physical reality to the virtual environment. Indeed, Bérard et al. used the observation of a transfer of skill between physical and virtual realities as a definition of a “natural” human-computer user interaction [6]. This transfer of skill makes the interaction more intuitive

(i.e. reducing learning time) and more efficient. Jacob et al. proposed a framework to regroup recent interaction systems that tend to reduce the gap between a user's intention and the means to execute the corresponding interaction by getting closer to user's habits [12]. They called this framework Reality-Based Interaction, and claim that such emerging systems get their inspirations from the physical reality.

Emotional responses

Enhancing presence fosters user's involvement with a virtual experience; which, in turn, can lead to similar emotional re-actions that they would have in a physical context. In a 2001 review, Schuemie et al. pointed at several studies in VR that show a correlation between self-reported presence and emotional responses such as fear of heights and public speaking anxiety [22]. More recently, Diemer et al. highlighted a relation between presence and arousal emotions such as fear and excitement [11]: they reviewing more than 30 studies of the past twenty years; which consistently reported correlations between presence and emotional experience. This literature shows that high presence in immersive VR environment tends to increase emotional responses.

Generating emotions from VR experience has an important health application: exposure therapy to cure phobias. Parsons et al. reported a meta-review of numerous VR exposure therapy studies [18]. They quoted that in vivo exposure therapy has shown greater efficacy than imaginary exposure, such as hypnosis therapy, in the treatment of specific phobias. This meta-analysis suggests that VR exposure therapy is effective in reducing patient anxiety. Price et al. analyzed clinical VR exposure therapy followed by 36 patients suffering from agoraphobia or fear of flying [20]. They found correlations between reported presence and phobic anxiety. They conclude that feeling present in the VR environment seems necessary for the success of the exposure therapy but may not be sufficient. Those findings suggests that patients experiencing a high level of presence during VR exposure therapies are able to feel physically exposed to their phobia, which is a crucial factor for the therapy to succeed. Exposure therapy was also implemented in AR in cases related to the phobia of small animals. According to Wrzesien et al. the proximity of the patient with the stimuli, and especially the hands, is of prime importance in the therapy [34]. Creating a motion-tracked avatar of the user's hands that is realistic enough to allow great presence is a very difficult challenge in VR. The AR approach eliminates this difficulty, as users perceive their actual arms and hands. In their study, Wrzesien et al. tested the response of 12 patients suffering from small animal phobia to either an AR or an in vivo exposure therapy. Results indicated comparable reduction of phobia anxiety for patients of both groups after the treatment, supporting the efficiency of Augmented Reality as a treatment.

Learning environment

Chen et al. studied the effect of presence experienced in two kind of AR-based learning environment (head-mounted display versus webcam display) [8]. In a lesson introducing the concept of protein structure, they tested a class of 58 university students enrolled in organic chemistry class in a between-subjects design. Results suggested that students that experienced a high level of presence endured lower cognitive load when interacting with the AR system. Furthermore, the authors reported a correlation between presence and involvement in the course.

Video games

Traditional HCI performance metrics do not apply to video games as players look for different rewards: traditional HCI usually aim at improving task accomplishment performance while games aim at creating emotions and challenges for users. The flow-channel model is frequently used to evaluate

the quality of video games. This concept was first introduced by Csikszentmihalyi in 1975 [9]. An optimal state of flow occurs in a situation where skills and challenges are balanced. The theory provides a 70-items questionnaire to analyze players' subjective feedbacks. Takatalo et al. applied the model to evaluate two popular first-person shooter video games based on the answer from more than a hundred expert gamers. They supported the importance of presence as a prerequisite that fosters an optimal state of flow [32].

How to achieve presence

High quality sensory feedbacks

In the VR research field, presence is the human response to an immersive virtual environment: the more the system is immersive, the more its users feel presence. Rendering high quality sensory feedbacks provides better immersion; which, in turn, is likely to enhance presence. However, the quality of a sensory feedback depends on many factors that are not comparable. It is thus difficult to predict which factor improvement will have the largest effect on immersion, and thus presence. The visual feedback quality, for example, depends on many factors as listed by Slater [25]: display resolution, field of view, frame rate, implementation of stereo, mesh resolution of the displayed entity, system latency and motion tracking precision and stability. In addition, Slater mentioned that the consistency between the various sensory feedbacks is mandatory to allow presence. For instance presence is hindered if a big metallic cube has a perfect visual feedback but users feel no weight for the cube. This, incidentally, provides evidence that presence is task-dependent: if this same cube were only to be observed at a distance, its presence would not be reduced by its lack of weight. Aiming to perfect presence for complex scenes may require very high quality visual feedbacks, imperceptible latency, full body interaction with haptic feedbacks, and so on. All these challenges have a cost and are constrained by the available technology. In practice, it is important to inform about the effect on presence of each of these parameters in order to choose the best trade-off.

Preeminence of the visual feedback

Cummings et al. reported a large meta-analysis on the effect of various parameters on presence. [10]. To perform the analysis, they normalized the effect size for all the parameters tested in the reviewed studies. The main parameters were the update rate, the "tracking level" defined as the number of tracked degrees of freedom, the field of view, the image quality, the stereoscopy, and the use of sound. They reported a strong effect on presence of the update rate and the tracking level. The lack of sound and "image quality" was found to have a lower impact on presence. While this survey is quite extensive, its results are based on disparate studies involving tasks and experimental contexts that differ across studies. As such, it is not as detailed as in a single controlled user study focused on comparing the effect of each parameter. The low impact of sound is consistent with the visual capture effect: a tendency of the human brain to trust visual information when multimodal information conflicts [16]. A well-known example of visual capture is the ventriloquism effect: it produces the illusion to hear the dummy's voice because its mouth is moving, although the sound actually comes from the ventriloquist. The visual capture effect stresses the importance of the visual feedback and motivated our choice of tested parameters in the present study.

Plausibility

At a higher level of abstraction, one must consider the plausibility of the scenario presented to the user. Slater defined the plausibility illusion as "the illusion that what is apparently happening is really happening" [27]. In VR, the situation in which the users are immersed must be credible for them to feel totally present. I.e. immersing participants into a room that looks alike the physical room that

are in is more likely to favor great presence than immersing them into a wild jungle scene. In AR, a virtual dragon is less likely to be perceived as present than a virtual butterfly. When interacting with virtual humans, the way the virtual entity will look at the user and answer questions is crucial to avoid breaking the presence. In addition, according to Diemer et al., VR scenarios that engage emotions tend to increase presence [11].

HPCD: a promising approach

Introduced by Stavness et al. [30] and recently improved by Berard et al. [5], the Handheld Perspective Corrected Display (HPCD) approach seems to be particularly suitable to create a strong feeling of presence. HPCDs render 3D content with a perspective projection that depends on the viewpoint of a user. This can create the illusion that the virtual entity is inside the manipulable display, as if the user was holding a physical transparent object with a physical entity inside. HPCDs are closer to AR than to VR in the sense that they create the illusion of presence of a virtual entity in the surrounding of the user rather than immersing the user in a virtual environment. They may not be called AR displays, however, in the sense that the virtual entity appears to reside in the display rather than in the general surrounding of the user. HPCD present several benefits to foster presence: as in AR, users are not cut from their physical environment; which does not have to be reproduced digitally. The display itself has a credible meaning in the physical reality as a transparent container. The display being physical, it has natural passive haptic feedback making users understand that they cannot reach inside to touch the virtual entities. Finally, as long as the weight of the display matches that of the virtual entity, the sense of touch is credibly rendered without any implementation effort.

In summary, the HPCD approach solves important issues of consistency between sensory feedbacks with no particular effort while avoiding the VR struggle to create a faithful avatar of users' body. Displaying virtual entities with a high level of presence seems feasible, and mainly depends on the quality of the visual stimuli created by the device. This motivated the choice of this approach in the current study.

Measuring presence

In a survey of VR literature related to presence, Schuemie et al. analyzed the various methods used to measure presence [22]. In an overwhelming majority of studies, presence was evaluated with subjective questionnaires. As presence is a subjective notion by definition, a self-evaluation makes sense. Echoing the diversity of definitions, a lot of questionnaires have been used to measure presence. The most commonly used are the Slater–Usuh–Steed Questionnaire [28], the Witmer and Singer's Presence Questionnaire [33], and the ITC Sense of Presence Inventory [15].

Objective measurement methods for presence were only explored in a few studies. Slater et al. proposed to measure how the users prioritize cues when virtual and real cues are conflicting [29]. They implemented this method by immersing participants in a virtual replication of the experimental room where a radio was present. When immersed, the physical radio was turned on and secretly moved while the virtual radio stayed still. Participants were then asked to point at the radio. The results suggested a correlation between self-reported presence using the Slater–Usuh–Steed questionnaire and the objective measurement. Another approach is to record skin conductance as several studies noticed a correlation between presence and skin conductance when dealing with arousal emotions (like phobias) [11]. However, skin conductance is known to be a direct measure of stress, not presence. Hence it can only be applied as an indirect metric for presence in specific contexts where an induced stress depends on a high level of presence, and such method is quite intrusive.

Given the subjective nature of presence and the lack of a reliable objective metric for presence, the consensus is to use self-reported questionnaire. A recent survey provided an evidence of this consensus: Cummings et al. reviewed more than 80 studies that analyzed the effect of immersive technology on presence [10]. In a concern for results comparisons, they chose to reject studies that did not evaluate presence through questionnaires.

However, self-reported presence does not achieve unanimity. Slater advanced an important argument against the validity of questionnaires on presence. In everyday life, presence is a fact. It is where we are, and we do not question whereas our surrounding environment is coherent in response to our actions as it indubitably is. As we never get to evaluate presence, it is hard to believe that we could grade the level of presence of any virtual situation in an absolute but yet undefined scale [26]. In their recent survey of presence, Skarbez et al. [24] investigated methods to measure presence that did not implicate subjective questionnaires that “add complexity to presence research”. They presented “behavioral methods” which analyze how participant behave when confronted to a given scenario. As objective measures, these methods address the major shortcomings of questionnaires. However, they usually require additional specific elements to trigger the evaluated behavior that depends on the scenario. As such, they are unsuited for generalization.

Skarbez et al. also presented “psychophysical methods” as methods meant to evaluate the relative importance of different aspects of a virtual experience on presence. Those novel methods have been used in VR to evaluate the impact of some immersion and scenario coherence parameters on participant’s self-reported presence. Bergstrom et al. implemented such a method to evaluate the contribution of four parameters of a virtual string quartet performance on the reported plausibility [3]. After experiencing the performance with optimal parameters, participants were exposed to a low level of fidelity and were asked to change parameters to maximize plausibility while minimizing the number of parameter change. Participants enhanced the gaze and the environmental sound first, suggesting the importance of these two features to enhance the performance plausibility. Psychophysical methods seems promising for studying different factors of virtual experience and their relative effects on presence.

A NEW METHOD TO EVALUATE PRESENCE

On the one hand, self-reported questionnaire are commonly used but the difficulty to rate presence on an absolute scale yields large inter-participant variability that hides small effects. Even for a single participant, it is difficult to estimate the absolute distance between two experiences on the scale of presence. In addition, questions are usually asked only once in subjective questionnaire. This does not support the computation of basic statistics (i.e. means, spread) that improve the precision of the measure and inform about its robustness. On the other hand, a direct measurement of the presence felt by participants, such as using bio-signals, seems illusory, as presence is a subjective feeling of high level of abstraction. We introduce a new approach founded on the following observation: while grading presence on an absolute scale is difficult, ranking two immediate experiences in terms of the presence felt is much easier. As a parallel, recognizing the pitch of a single note requires a rare absolute ear, but given two music notes most people are able to identify the higher one.

Our approach has been inspired by the just noticeable difference (JND) protocol used in recent HCI research to measure the smallest perceivable system’s latency [17]. The JND protocol consists in presenting pairs of stimulus and asking participants to rank them, for example by choosing the best one, or the lowest one. For each correct answer, the difference between the stimuli is slightly reduced. It is greatly enlarged in case of wrong answer. This occurs when the two stimuli are too closed and the participant cannot perceive the difference. Changes in the stimuli and order of

presentation are hidden from participants so that their answers are not influenced by the evolution of the parameters. This protocol robustly converges to the smallest perceivable difference for the participant. Ng et al. always used the lowest possible latency, i.e. 1ms, as one of the stimulus in the pair. Assuming that 1ms was below the absolute threshold of perception for all participants, they changed the purpose of the JND protocol from measuring a smallest perceivable difference between two latencies to measuring an absolute threshold of perception. The JND protocol can be seen as a way to make an objective measure of a subjective question, e.g. "what is the smallest system latency that you can perceive?".

The JND protocol as implemented by Ng et al. is not directly applicable to our aim. We are interested by the relative effect of various heterogeneous parameters on presence, but the JND is used to accurately find a threshold on a single parameter. Another main obstacle is that a system that creates the perfect presence does not exist. The physical world is not a candidate as its parameters such as latency or resolution cannot be reduced in a way that is unnoticeable to participants. However, we keep the principles of a repetition of many comparisons between pairs of stimulus; which parameters are unknown to the participants.

The main principles of our approach is thus to expose participants to two variant of a virtual experience, that we call representations. We ask them which of the two representations generates the strongest feeling of presence. To facilitate such comparison, participants can switch at will between the two representations before giving their answer. A single comparison only yields a binary measure. However, we arrange for a comparison to be quick: in the order of 10s. This way, we can ask for many comparisons where we control which parameter is altered in each representation, and to which level, yielding a finer measure of the relative effect of pairs of parameters. The differences between representations are not explained so that participants can only rely on their feelings to make the choice. In addition, the same comparison is requested several times during a participant's experimental session. This improves the precision of the measure by allowing the computation of a mean. It also provides an estimate of the robustness of the measure by exposing the variability of a participant's answers to the same comparisons.

Every representation corresponds to one of the experimental condition: each representation is rendered by optimizing every parameter except for one that is altered in a negative way. Experimental conditions differ by the level of alteration for one particular parameter, such as latency or resolution, or by the parameters themselves. The approach allow testing very different parameters, such as interaction possibilities, or the use of a stereoscopic rendering, as long as participant can quickly and repeatedly switch between representations. In a concern for generalization, this method is repeated on various virtual scenes.

Similarly to the psychophysical methods, our method relies on the participants' sense of presence to evaluate the influence of several parameters. But unlike methods such as in Bergstrom et al. [3], participants are not familiarized with the parameters and they cannot manipulate them freely. By hiding the information and the control on the parameters, we insure that participant are not influenced by their own expectations about the parameters.

DESIGN OF THE USER STUDY

Selecting a definition of presence

The goal of our study is to measure the effect of the main technical parameters of a visual stimulus on presence. However, a universal definition of presence is hard to find in the literature, as the various studies tend to subdivide the concept to fit their specific needs: physical presence, social

presence, co-presence, and telepresence. All these definitions, however, describe a psychological sense of sharing a reality with one or several entities. Lee [14] proposed a definition of presence that unifies the VR and AR approaches: they define presence as “a psychological state in which virtual (para-authentic or artificial) objects/social actors/selves are experienced as actual objects in either sensory or nonsensory ways”. We use this definition, and use the term “entity” to denote all possible actors from unanimated objects to the self. We also acknowledge that presence is a continuous metric rather than a binary value, as evaluated in many studies [29, 20, 8], and that presence depends on the task (e.g. a weightless metallic virtual object may seem present, as long as users don’t interact with it). In summary, our study aims at measuring presence defined as “a continuous metric for the psychological state corresponding to the degree to which virtual entities are experienced as actual entities in either sensory or nonsensory ways for a specific task”.

Acknowledging the lack of consensus on the definition of presence and its associated concepts, we note that our experiment can also be seen as measuring the plausibility of the virtual entity.

Choice of parameters

In order to isolate the effect of technical parameters, we implemented an HPCD technique where parameter’s values could be independently altered from a reference optimal value. The optimal value corresponds to our best implementation effort. We chose to implement an HPCD device, as it allowed us to surpass commercialized Head-Mounted Display (HMD) hardware in term of jitter and resolution, and to equalize the bests HMD in term of latency.

According to Cummings et al. [10], the most studied and impactful technical parameters when dealing with presence are display resolution, frame rate, tracking jitter, the field of view, and stereoscopy. We altered this set of parameters with the following rationales. We created virtual entities that appeared to reside inside a Handheld Perspective Corrected Display (HPCD); hence the field of view was not relevant and was left out of this study. We found in a pilot study that the absence of stereoscopy when using a HPCD would break the feeling of presence in a strong way. We thus considered that stereoscopy is an essential requirement for presence and removed this parameter from the study. The effect of the system’s latency on presence has not been studied in depth, however several studies have shown its strong negative impact on the quality of the interaction [7, 17, 19]. We suspected a similar negative impact on presence and chose to include it in the study. We thus tested the effect on presence of display resolution, latency, frame rate and tracking jitter (RLFJ).

Alteration levels

All parameters were tested at two or more levels in the study. For each parameter, one level was our best effort, i.e. the best level we could achieve with our experimental system. We reproduced the HPCD system described in Berard et al. [5]: for all representations, the entity was displayed in a 30cm diameter polystyrene sphere used as a wireless HPCD. Motion capture was implemented with an Optitrack optical tracking system (8 Prime 13 cameras at 240Hz). We recorded the reported position of a static marker to evaluate tracking jitter and observed a Gaussian white noise. Images were created by a Barco F-50 projector (2560x1600@120Hz pixels). Stereoscopy was implemented using shutter glasses and time multi-plexing the left and right images. We evaluated the system’s end-to-end latency using the approach from Cattani et al. [7]. This system achieved the following best-effort levels for the 4 parameters: resolution at 90dpi in the focus area (resolution lowers on the side of the projection on the sphere), latency at 27ms, frequency at 60Hz per eye, and jitter at 0.022mm of standard error.

To determine relevant alteration levels according to the state of the art, we first diagnose popular VR and AR devices for their RLFJ values. We evaluated two VR HMD (HTC Vive and Vive pro), two AR HMDs (Metavision Meta 2, Microsoft HoloLens) and an AR tablet (Tango).

Software alteration was chosen as follow.

- **Resolution.** The majority of AR ready tablets, as well as the HoloLens, have a resolution in the order of 100dpi; which is superior to the 90dpi of our HPCD implementation. Most current HMDs have a much lower resolution, in the order of 50dpi, although high resolution VR HMDs are becoming available. We chose a first alteration level at 50 dpi. We also tested a second alteration level, recognizing that it is common practice to lower the rendering resolution in software to keep an acceptable frame rate with virtual scenes of high complexity. We chose 35pdi ($\sqrt{2} * 50 \text{ dpi} = 35 \text{ dpi}$) as the resolution reduction needed to halve the number of pixels from the first alteration level.
- **Latency.** The end-to-end latency of systems used to create virtual entities is generally not published. Sometimes, a soft-ware prediction is used to compensate for part of the latency. However, predictions have the drawback of increasing the jitter in position. We chose to avoid predictions in this study to concentrate on the fundamental impact of latency and jitter. Nevertheless, the prediction used in the HoloLens prevented us to estimate its baseline latency. We measured a latency of 28ms for the vive pro (nearly identical to our optimal value), and around 60ms for the Meta 2. Ng et al. report that the latency of commercial touch devices is in the range 50-200ms [17]. We chose to use a single alteration level for latency at 60ms.
- **Frame rate.** Frame rates are easy to evaluate, as they are negotiated with the graphic card; which report it to the system. We observed a native frame rate (i.e. optimal) of 60Hz for the HoloLens, 72Hz for the Meta 2, 90Hz for the Vive and Vive pro. The frame rate of most tablets is 60Hz. Current hardware thus affords 60Hz stereoscopic display, but rendering complex scenes often result in frame rate drop. We chose a single frame rate altered level at 30Hz to illustrate rendering limitations. This frame rate level also represents what time multiplexed stereo would provide on 60Hz tablet computers.
- **Jitter.** We measured a standard error of the jitter at 0.13mm, 0.20mm and 1.7mm for the Vive pro, the HoloLens, and the Meta 2, respectively. We first intended to use these as 3 alteration levels from our baseline (0.022mm). However, we determined in pilot studies that 0.13mm was unnoticeable and that 1.7mm was conversely breaking presence when used with the HPCD. We chose not to include 0.13mm, to halve 1.7mm, and to include a value in-between the two remaining alteration levels. We thus tested 0.2mm, 0.5mm and 0.8mm.

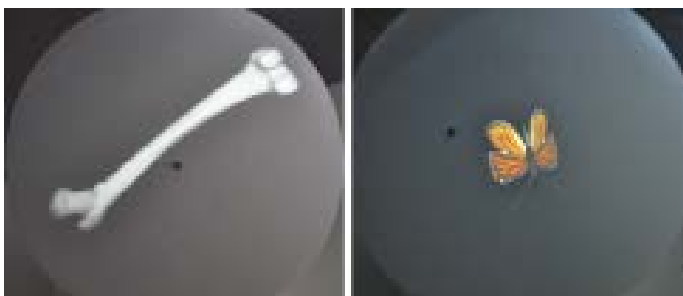


Figure 1. The two entities displayed. On the left, the static bone. On the right, the animated butterfly while it is flying inside the sphere. For illustration purpose, stereo has been disabled in this picture.

Entities and task

In a concern for generalization, we chose two entities with different complexity in their behavior and interaction possibilities. The first entity was a femur. Participants could explore the bone by rotating and moving the sphere but the bone was static in the sphere coordinate system. The second one was a butterfly that acted like it was engaged in the sphere: it could walk on the surface of the sphere or fly inside it. When on the surface, participants could force it to fly if they shook the sphere. The animation was randomized to appear realistic. Both entities are illustrated in figure 1. The task that we asked participant to perform was to freely explore the presented entities by means of changing their viewpoint and manipulating the sphere (including shaking it).

Protocol

We implemented the method introduced above. At the beginning of the session, we explained the definition of presence to participants. Every representation was associated with a single level of alteration of a single RLFJ parameter. In other words, each representation was our best effort, except for one parameter that was altered. The table 1 summarize all the representations used in the experiment. Each representation is named after its altered parameter and level. It should be noted that the best effort was not presented as a representation.

The set of comparison presented to participant was made of all possible pairs of representations, except that we did not present pairs of representations of the same parameter at different levels. The same comparison was presented 8 times to each participant (4 times per entity). This resulted in $17 \times 8 = 136$ comparisons per participant. As the experimental session was quite long, in the order to 90 minutes, we split it on two days to limit a tiredness effects. Comparisons were presented in a randomized order in a total of 8 blocks (4 per day). The entity displayed was the same within a block and was switched between each block. Participants could take a short break between each block.

At the end of the first day of experiment, participants filled a short questionnaire with three questions designed to highlight the subjective differences between the two entities presented. For the two entities, they had to report: the variability of presence levels experienced (on a straight line without scale), how easy it was for them to answer the comparisons, and how much they enjoyed to explore the scene (on a 5 level scale for the two last questions).

representation:	resolution (dpi)	latency (ms)	frame rate (Hz/eye)	jitter (mm)
Resolution 50	50	27	60	0.02
Resolution 35	35	27	60	0.02
Latency 60	90	60	60	0.02
Frame rate 30	90	27	30	0.02
Jitter 0.2	90	27	60	0.2
Jitter 0.5	90	27	60	0.5
Jitter 0.8	90	27	60	0.8

Table 1. The value of each RLFJ parameter for every compared representation. Values in blue are the optimal values of the given parameter.

Participants

We welcomed 16 participants: 8 women, 28.2 mean age (23 to 37). All had correct stereoscopic vision (tested with the Stereo Optical RANDDOT stereopsis test), and none was afraid of butterflies. Figure 2 shows a participant exploring two representations: one for each entity. While exploring, she can tap her right foot at any time to switch between the two representations. She identifies the one generating the most presence by tapping with her left foot while it is active.

RESULTS

Methodology

We recorded participants' preference in terms of presence felt when presented a comparison of two representations. To compute mean and confidence intervals, we associated numerical values to participants' answers. Given a comparison A vs. B, we associate a value of 1 whenever A has been preferred, and a value of -1 whenever B has been preferred.



Figure 2. A participant exploring a representation of the bone entity (left) and of the butterfly entity (right). Daylight was covered during the experiment to maximize the projector contrast.

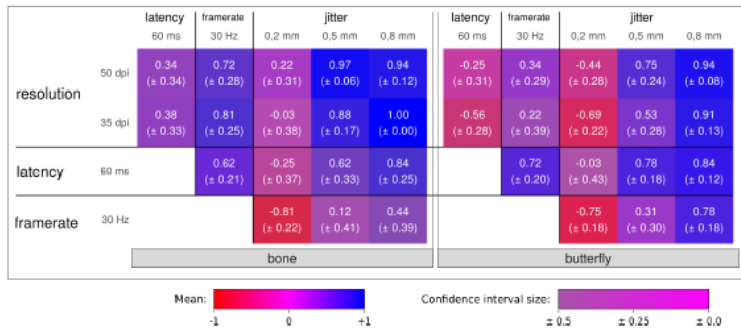


Figure 3. Mean user preferences per comparison, with 95% confidence intervals. A value of 1 means that the condition defined by the cell's row was always preferred to the cell's column. -1 means the opposite. Left: bone entity, right: butterfly entity. Hue encodes the mean; saturation encodes the size of the confidence interval.

In Figure 3, Figure 4, and Figure 5, each cell represents a A vs. B comparison where A and B are defined by a cell's row title and column title, respectively. In a concern for readability we did not present redundant information: resolution conditions are only presented on the left and jitter conditions only on top.

We acknowledge the current concerns on the use of statistical significance categorization in various research fields [1, 4]. Hence, we report the exact values of the probability of the NULL hypothesis and we avoid the use of the significance vocabulary.

We computed the probability of results being explained by random sampling and measurement errors (p-value of the NULL hypothesis) with logistic regressions for every comparison. We used R glmer, anova and glht methods from packages lme4 and multcomp.

Effect of the entity

We performed an analysis of deviance, with single-step correction, to test the effect of the entity on participants' preferences. Results are shown in Figure 4. They indicate that participants' feeling of presence differed depending on the entity mostly when resolution was involved, although it also

differed in the frame rate 30Hz vs. jitter 0.8mm comparison. Conversely, results appear to be similar with both entities when resolution is not involved.

		latency 60 ms	framerate 30 Hz	jitter		
				0,2 mm	0,5 mm	0,8 mm
resolution	50 dpi	p= 6e-05	p= 0.005	p= 8e-05	p= 0.002	p= 1.000
	35 dpi	p= 8e-10	p= 9e-08	p= 1e-05	p= 0.003	p= 0.023
latency	60 ms		p= 0.470	p= 0.087	p= 0.111	p= 0.998
framerate	30 Hz			p= 0.528	p= 0.160	p= 8e-04

Figure 4. Effect of the entity: probability of the NULL hypothesis for the difference of mean preferences between entities, for each comparison.

		latency 60 ms	framerate 30 Hz	jitter							
				0,2 mm	0,5 mm	0,8 mm	latency 60 ms	framerate 30 Hz	0,2 mm	jitter 0,5 mm	0,8 mm
resolution	50 dpi	p= 0.043	p= 6e-05	p= 0.117	p= 0.006	p= 0.006	p= 0.110	p= 0.038	p= 0.003	p= 0.031	p= 0.006
	35 dpi	p= 0.030	p= 5e-04	p= 0.607	p= 1e-05	p< 1e-15	p= 0.001	p= 0.236	p= 4e-05	p= 0.003	p= 0.025
latency	60 ms		p= 3e-06	p= 0.089	p= 0.005	p< 1e-15		p= 3e-06	p= 0.430	p= 6e-04	p< 1e-15
framerate	30 Hz			p= 1e-05	p= 0.264	p= 0.054			p= 1e-05	p= 0.047	p= 0.004
						bone	butterfly				

Figure 5. Robustness of the comparisons: probabilities of the NULL hypothesis (p-values) for the preferences reported in Figure 3.

Participants' preferences

For each participant, we computed the mean answer for the 4 repetitions per entity of each comparison. We then computed the mean and standard deviation across participants for each comparison, per entity. This is shown on Figure 3. We computed the p-value for each comparison by testing the linear hypothesis of the logistic regressions. The p-values are shown in Figure 5.

An alteration that is often preferred indicates that the parameter has not a strong negative impact on presence. However, this does not mean that this parameter has a positive impact on presence as presence is resulting from the combination of all the parameters. Conversely, an alteration that is rarely preferred indicates that the parameter has a strong negative impact on presence. We thus focus in the latter.

Jitter. We observed that jitter 0.8mm was never preferred, which indicate that jitter was a critical parameter for presence and that a high jitter has a strong negative impact. This effect is robust (with

low p-values) and the effect size is large, with mean preferences against this condition at 0.78 or more, except when compared with frame rate 30Hz with the bone.

Even with the lesser alteration at 0.5mm, jitter was rated more detrimental than resolution and latency for both entities. The effect sizes observed are lower than with the highest jitter level but the effect remains robust. This level of jitter, however, appears to have an effect that is close to the effect of the 30Hz frame rate: in these comparisons, preferences are closer to 0, answers are more variable, and the p-value indicates a fragile effect (e.g. $p=0.26$ for the bone).

The smallest jitter level that we tested (0.2mm) is far less detrimental for presence. It was preferred to the frame rate alteration for both entities and to the resolution alterations for the butterfly. The other comparisons were less clear; which indicate an effect similar to those of the resolution alteration for the bone and the latency alteration for both entities.

Frame rate. Frame rate 30Hz is more detrimental than latency 60ms for both entities, and more detrimental than the two resolution alterations for the bone. Frame rate 30Hz appears to be more detrimental than resolution 50dpi with the butterfly entity, even though this last result has a smaller effect size and is less robust (mean preference of 0.34, $p=0.038$). Finally, frame rate 30Hz seems equivalent to resolution 35dpi with the butterfly entity (mean preference of 0.22, $p=0.236$).

Latency. Latency is less detrimental with the butterfly than both resolution alterations, even if the results are not so clear with the Resolution 50dpi (mean preference of -0.25, $p=0.11$). With the bone entity, latency appears to be slightly more detrimental than resolution (mean preference values of 0.34 and 0.38 for the 50dpi and the 35dpi condition respectively, $p<0.043$).

Effect size - task duration correlation

For both entities, we computed the Spearman correlation factor between the effect size for every comparison and the mean time spent by the participants to answer to the comparison. We obtain a correlation factor of -0.78 for the bone and of -0.12 for the butterfly. We assumed that large effect size corresponded to an easy comparison, i.e. a comparison where the best representation for presence was obvious. Hence, we observed that time spent in each comparison was highly related to the difficulty of the task with the bone entity but not with the butterfly.

Subjective results

Participants expressed a range of presence that delimited all representations of the same entity that they explored, i.e. from the less present representation to the most present one. The range was expressed on continuous scale with no graduation from no presence to perfect presence. The results are reported in figure 6.

The range expressed by participant is quite large: over 50% of the entire scale. This indicates that the loss of presence in the worst conditions was clearly felt by the participants. For both entities, results are extremely similar.

For the two questions: 'I liked to explore the scene' (exploration enjoyment) and 'In average, I could easily identify the most present representation within each comparison' (ease for identification), participants answered on 5-levels scale from 1 (not at all) to 5 (totally). Results are illustrated on Figure 7. Participants slightly preferred to explore the scene with the butterfly but they found the choice task easier with the bone.

DISCUSSION

Effect of the entity

We observed that results are similar with both entities except when resolution is involved. The two studied entities are quite different: the bone is a simple non-textured mesh whereas the butterfly is made of highly detailed mesh with colorful textured wings. This could explain why alterations of resolutions were more detrimental to presence in the case of the butterfly. This result reveals how the nature of the entity may put specific technical requirements to create a good presence. In our study, the two entities did not have a tightly coupled interaction with the participant, which may explain why the alteration of latency did not have a strong detrimental effect on presence. A virtual pen, however, may have a stronger requirement in terms of latency.

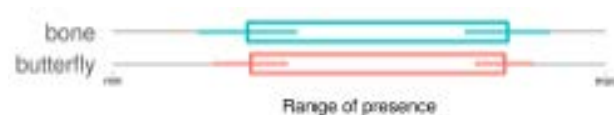


Figure 6. Self reported range of presence of all explored representations, per entity. The left and right sides of the boxes are the mean of the min-imum and maximum limits of the range, respectively. 95% confidence intervals are represented by the colored segments.



Figure 7. Mean subjective enjoyment and task difficulty, per entity, with 95% confidence intervals.

General analysis of the comparisons

Jitter 0.8mm was the worst condition for presence, whereas Jitter 0.2mm was one of the conditions with the weaker effect, alongside Latency 60ms. Frame rate 30Hz and jitter 0.5mm belongs to the same class in their impact on presence with both entities, along with the 35dpi resolution condition with the butterfly.

Figure 8 summarizes the results of the comparisons. Alter-ations are either separated per entity or regrouped when the results do not depend on the entity. Rectangles regroup conditions in same class.

Performance trade-offs

Resolution vs. frame rate

Displaying complex scenes can reduce the frame rate when the computing time for each frame is greater than the refresh period. Our study indicates that reducing the display resolution to keep the optimal frame rate is a good strategy in terms of presence, although caution is required when displaying highly detailed entities. Still, with the highly detailed butterfly, going from 90dpi to 50dpi was less detrimental to presence than going from 60Hz to 30Hz. This could be explained by degree of physicality of the two phenomena. There are many physical explanations for seeing things in low details: air mist, dirty glasses or window, eye moisture, etc. This is not the case for jerky motion.

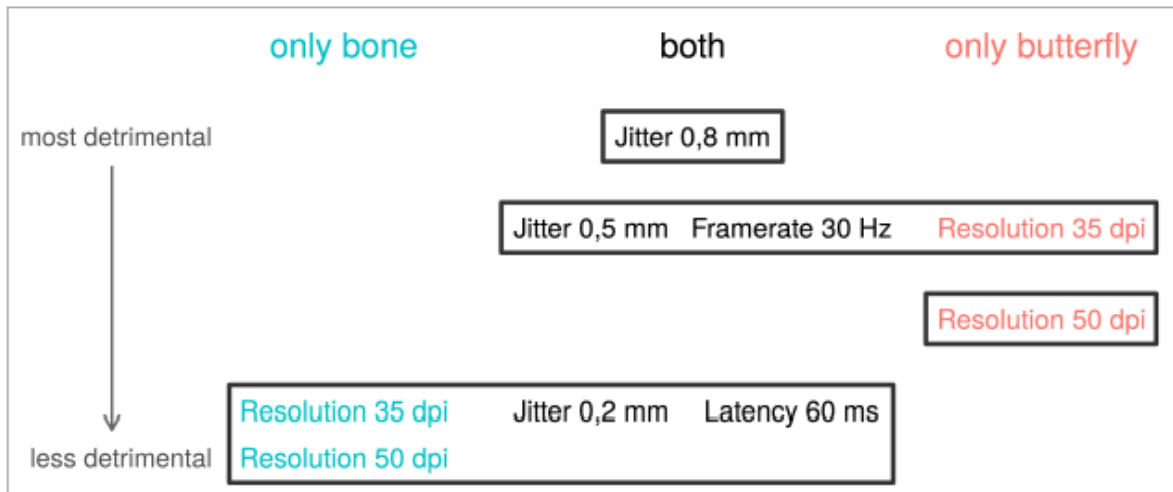


Figure 8. Alterations ordered by the strength of their detrimental effect on presence, from the most detrimental (top) to the less detrimental (bottom). The rectangles group alterations that are in the same class. The effect of the resolution alterations depends on the entity.

Latency vs. jitter

The study revealed that tracking jitter could have a large detrimental effect on presence. One way to reduce jitter is to smooth the tracking data over several capture frames at the cost of additional latency. Our study indicates that this could be a good strategy in terms of presence, as the latency alteration had a smaller negative effect to presence than the two largest jitter alteration, and a similar effect to the smallest jitter alteration. In our experiment, participants explored the virtual entities residing inside a sphere. The vibration of the entities due to jitter did not seem to have a physical interpretation; while the lag of the motion of the entities created the illusion that the entities were immersed in a viscous fluid. However, this trade-off may be task dependent. It may only be beneficial when there is no tightly coupled interaction with the entities, as the opposite trade-off is more suitable in terms of pointing performance [19]. In addition, we only tested a latency alteration from 27ms to 60ms. It would be interesting to investigate how far the trade-off holds when going beyond 60ms.

Informal feedback from participants

Some participants reported that the butterfly appeared more present than the bone because it moved by itself. They took some times to play with it and explore its movements, like they would have done with real butterflies. Three participants also gave it a name. These comments partly explain the weaker correlation between the effect size and the answering time: even when a comparison was easy to rank, some participants did not answer immediately as they were compelled to play with the butterfly. However, some participants were disturbed by the wings animation: they found that it was not plausible, reducing the presence felt. One participant found the bone extremely convincing as a 3D-printed plastic bone, resulting in total presence with the best representations. Two participants reported that latency, especially with the bone, produced an immersion effect, as if the sphere was full of formalin. This is another illustration on how the human brain manages to produce a rational explanation to a strange stimulus, as with the ventriloquism effect. Another participant reported that the “vibrations” of the bone when it was supposed to be static (i.e. with high jitter) was strongly breaking the presence because static objects that lay in stable equilibrium never move in “real life”. Here, the lack of a credible physical explanation appears to have broken the presence.

Several participants reported that the movements of the butterfly made it hard to compare the representations, whereas they could easily reproduce a similar viewpoint in the two representations with the static bone. This probably explains why the task was rated more difficult with the butterfly than with the bone. Finally, the fact that participants also liked to explore the static bone without getting bored (enjoyment score superior to 3.5 over 5) is in line with a high involvement in the task when presence is concerned, as previous work suggested [8].

Presence comparison method

We observed that the presence comparison method that we introduced had a high level of acceptance: all participants declared that they clearly understood the task. In the subjective questionnaire, the question about the ease to answer the task had a score of 3.2 or more (for a maximum of 5) for both entities. The method yielded several large effect sizes (e.g. score of 0.81 for jitter 0.2mm when opposed to frame rate 30Hz) that appear to be robust ($p=1e-05$ for this comparison for both entities).

Although we used the method in a study focused on a small set of technical parameters, we think that its utility is quite general. For example, the method could be used to thoroughly test a single parameter with many alteration levels; which would bring it closer to the JND protocol. Also, it could be used to test the effect on presence of non-technical parameters such as the complexity of the task or the plausibility of the entity.

Limits of the study

As in any user study, participants' time and endurance are bounded. We thus had to limit the number of tested conditions. The results should be interpreted with these limits in mind. In particular, we only tested each parameter individually. Yet, some parameters may interact with others. In addition, for now, we have confidence in the results only for one particular task, two entities, and one particular interactive device. Further studies are required to test for the generality of these results. More conventional VR and AR HMDs could be used, although a lesser baseline may prevent a direct comparison of the results with the spherical HPCD. In addition, it would be particularly interesting to test other tasks involving more tightly coupled interaction between the participants and the entity, in order to interrogate the latency-jitter trade-off. Finally, more complex entities should be tested, such as human avatars, although the main bottlenecks on presence may stem from the quality of the human model rather than the technical parameters of the visual stimulus.

This first study based on the proposed method shows the viability of the method. However, further studies on the effect of these parameters on presence will be required to test the coherence of the results and thus to test the external validity of the method.

CONCLUSION

Recent improvements in AR and VR hardware have resulted in virtual experiences that come close to creating the perfect subjective feeling of presence. However, the final step may require a methodological analysis of the factors coming into play to create this feeling. One way to foster this methodological analysis is to unify the concept of presence across the various research fields involved, so that studies' results can be more easily consolidated. The method that we introduced is not tied to any particular system, interactive setup, or research field. The user study implementing this method serves two purposes. It provides evidences of the viability and utility of the method. It also provides a first measurement of the relative effect on presence of 4 important parameters of

the virtual visual stimulus in the context of a spherical Handheld Perspective Corrected Display. These measurements inform about important design guidelines: avoiding a perceivable motion jitter and a loss in frame rate were the most important requirement in our setup. Reducing the resolution does not reduce the feeling of presence for entities that do not have fine details. And contrary to our expectations, a perceivable latency of the system did not have a strong negative effect on presence.

ACKNOWLEDGEMENTS

This work was supported by the An@tomy2020 project ANR-16-CE38-0011.

REFERENCES

- [1] Valentin Amrhein, Sander Greenland, and Blake McShane. 2019. Scientists rise up against statistical significance. *Nature* 567 (2019), 305–307. <https://www.nature.com/articles/d41586-019-00857-9>
- [2] Jeremy N. Bailenson, Nick Yee, Dan Merget, and Ralph Schroeder. 2006. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 359–372.
- [3] Ilias Bergström, Sérgio Azevedo, Panos Papiotis, Nuno Saldanha, and Mel Slater. 2017. The plausibility of a string quartet performance in virtual reality. *IEEE transactions on visualization and computer graphics* 23, 4 (2017), 1352–1359.
- [4] Lonni Besançon and Pierre Dragicevic. 2019. The Continued Prevalence of Dichotomous Inferences at CHI. In *ACM Conference on Human Factors in Computing Systems (CHI) Extended Abstracts*. ACM, alt14:1–alt14:11. DOI : <http://dx.doi.org/10.1145/3290607.3310432>
- [5] François Bérard and Thibault Louis. 2017. The Object Inside: Assessing 3D Examination with a Spherical Handheld Perspective-Corrected Display. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*.
- [6] François Bérard and Amélie Rochet-Capellan. 2015. The Transfer of Learning as HCI Similarity: Towards an Objective Assessment of the Sensory-Motor Basis of Naturalness. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, Seoul, Republic of Korea, 1315–1324. DOI : <http://dx.doi.org/10.1145/2702123.2702359>
- [7] Elie Cattan, Amélie Rochet-Capellan, Pascal Perrier, and François Bérard. 2015. Reducing latency with a continuous prediction: Effects on users' performance in direct-touch target acquisitions. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM, 205–214.
- [8] Yu-Chien Chen, Sheng-Jo Wang, and Ya-Ling Chiang. 2009. Exploring the Effect of Presence in an AR-based Learning Environment. In *13th Global Chinese Conference on Computers in Education, Taipei*. Citeseer.
- [9] Mihaly Csikszentmihalyi. 1975. *Beyond boredom and anxiety: Experiencing flow in work and play*. Vol. 721. Jossey-Bass San Francisco.
- [10] James J. Cummings and Jeremy N. Bailenson. 2016. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology* 19, 2 (2016), 272–309.
- [11] Julia Diemer, Georg W. Alpers, Henrik M. Peperkorn, Youssef Shiban, and Andreas Mühlberger. 2015. The

- impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology* 6 (2015), 26.
- [12] Robert JK Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 201–210.
- [13] Ernst Kruijff, J. Edward Swan, and Steven Feiner. 2010. Perceptual issues in augmented reality revisited. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*. IEEE, 3–12.
- [14] Kwan Min Lee. 2004. Presence, explicated. *Communication theory* 14, 1 (2004), 27–50.
- [15] Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. 2001. A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 282–297.
- [16] Stefan Mateeff, Joachim Hohnsbein, and Thomas Noack. 1985. Dynamic visual capture: apparent auditory motion induced by a moving visual target. *Perception* 14, 6 (1985), 721–727.
- [17] Albert Ng, Julian Lepinski, Daniel Wigdor, Steven Sanders, and Paul Dietz. 2012. Designing for low-latency direct-touch input. In *ACM Symposium on User Interface Software and Technology (UIST)*. ACM, 453–464. DOI :
<http://dx.doi.org/10.1145/2380116.2380174>
- [18] 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.
- [19] Andriy Pavlovych and Wolfgang Stuerzlinger. 2009. The tradeoff between spatial jitter and latency in pointing tasks. In *ACM SIGCHI symposium on Engineering Interactive Computing Systems (EICS)*. ACM, 187–196. <http://doi.acm.org/10.1145/1570433.1570469>
- [20] Matthew Price and Page Anderson. 2007. The role of presence in virtual reality exposure therapy. *Journal of anxiety disorders* 21, 5 (2007), 742–751.
- [21] Ralph Schroeder. 2002. Social interaction in virtual environments: Key issues, common themes, and a framework for research. In *The social life of avatars*. Springer, 1–18.
- [22] Martijn J. Schuemie, Peter Van Der Straaten, Merel Krijn, and Charles APG Van Der Mast. 2001. Research on presence in virtual reality: A survey. *CyberPsychology & Behavior* 4, 2 (2001), 183–201.
- [23] Thomas B Sheridan. 1992. Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments* 1, 1 (1992), 120–126.
- [24] Richard Skarbez, Frederick P Brooks Jr, and Mary C Whitton. 2018. A survey of presence and related concepts. *ACM Computing Surveys (CSUR)* 50, 6 (2018), 96.
- [25] Mel Slater. 2003. A note on presence terminology. *Presence connect* 3, 3 (2003), 1–5.
- [26] Mel Slater. 2004. How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence* 13, 4 (2004), 484–493.
- [27] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557.
- [28] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144.
- [29] Mel Slater, Martin Usoh, and Anthony Steed. 1995. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)* 2, 3 (1995), 201–219.

- [30] Ian Stavness, Billy Lam, and Sidney Fels. 2010. pCube: a perspective-corrected handheld cubic display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1381–1390. <http://dl.acm.org/citation.cfm?id=1753535>
- [31] Ivan E. Sutherland. 1965. The Ultimate Display. In *Proceedings of the IFIP Congress*. 506–508.
- [32] Jari Takatalo, Jukka Häkkinen, Jyrki Kaistinen, and Göte Nyman. 2010. Presence, Involvement, and Flow in Digital Games. *Evaluating user experience in games* (2010), 23–46.
- [33] Bob G. Witmer and Michael J. Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments* 7, 3 (1998), 225–240.
- [34] Maja Wrzesien, Jean-Marie Burkhardt, Mariano Alcañiz, and Cristina Botella. 2011. How technology influences the therapeutic process: a comparative field evaluation of augmented reality and in vivo exposure therapy for phobia of small animals. In *IFIP Conference on Human-Computer Interaction*. Springer, 523–540.