PhD-FSTM-2025-95
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 23/10/2025 in Esch-sur-Alzette

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN INFORMATIQUE

by

Stéven PICARD
Born on 27 December 1997 in Fougères, (France)

EXTENDED REALITY AND TIME PERCEPTION: AN INTEGRATED APPROACH AND EXPERIMENTAL FRAMEWORK

Dissertation defence committee

Dr Jean BOTEV, dissertation supervisor
Research Scientist, Université du Luxembourg

Dr Knut DREWING
Professor, Justus-Liebig-University

Dr Maud MARCHAL
Professor, Univ. Rennes, INSA/INRIA

Dr Steffen ROTHKUGEL, Chairman
Associate Professor, Université du Luxembourg

Dr Denis ZAMPUNIERIS
Professor, Université du Luxembourg

Affidavit / Statement of originality

I declare that this thesis:

- is the result of my own work. Any contribution from any other party, and any use of generative artificial intelligence technologies have been duly cited and acknowledged;
- is not substantially the same as any other that I have submitted, and;
- is not being concurrently submitted for a degree, diploma or other qualification at the University of Luxembourg or any other University or similar institution except as specified in the text.

With my approval I furthermore confirm the following:

- I have adhered to the rules set out in the University of Luxembourg's Code of Conduct and the Doctoral Education Agreement (DEA)[1], in particular with regard to Research Integrity.
- I have documented all methods, data, and processes truthfully and fully.
- I have mentioned all the significant contributors to the work.
- I am aware that the work may be screened electronically for originality.

I acknowledge that if any issues are raised regarding good research practices based on the review of the thesis, the examination may be postponed pending the outcome of any investigation of such issues. If a degree was conferred, any such subsequently discovered issues may result in the cancellation of the degree.

Approved on 2025-07-22

[1] If applicable (DEA is compulsory since August 2020)

Abstract

Experiencing the passage of time is an inherent aspect of human life, and therefore, of the experience in interactive applications. Time perception is a growing research subject with numerous research questions such as how to model the human processing of time, what affects it, how to modulate it, and what can be the benefits of modulating it. This research is embedded in the EU-funded *ChronoPilot* project, which aims to modulate time perception to the advantage of its user, in order to *induce flow* or *reduce stress.* With the recent growth of mediated reality technologies, in particular with the rising popularity of *Virtual Reality* (VR) and *Augmented Reality* (AR) dedicated *Head Mounted Displays* (HMDs), interactive computer applications can now deliver multi-sensory stimuli in immersive experiences; which offers unique possibilities to modulate time perception. However, integrating time modulation as part of the user experience in immersive applications comes with numerous challenges in integration, such as the selection of stimuli, the intent of the time modulation, or even the technical integration. In this thesis, as part of the *ChronoPilot* research project and based on my research from the past years, this thesis will try to provide guidelines and tools to help integrate time perception modulation within *eXtended Reality* (XR) applications. More specifically, the focus lies on the use of *rhythmic stimuli* in time modulation, on the *implications of XR* in a time modulation context, and *specifications of a Unity framework* dedicated to *(Multi-)User studies.* Rhythmic stimuli were approached first as a way to synchronize multi-sensory modulation; however, our results suggest that a multi-sensory rhythmic stimulus have fundamentally different effects on time perception than an equivalent single-sensory stimulus. Furthermore, the rhythmic stimuli tested depend heavily on contextual factors, such as *fatigue* and *task familiarity.* Regarding the implication of using XR, it was observed that using XR in a time perception modulation context may lead to differences both due to the technology being used or the design and implementation of in-environment interactions. The main contribution of the thesis is the development of the *XR MUSE* framework, which is built with the aim of allowing extended reproducibility and cross-experiments integrations. The developed technology has been of great help in developing interactive applications

implementing intentional time perception modulation in non-timing tasks and is ready to be used for any interactive application prototyping.

Acknowledgements

Before beginning this thesis, I would like to express my gratitude to Jean Botev for his careful guidance as a supervisor throughout this work. I am also deeply indebted to Sahar Niknam and Ningyuan Sun, my colleagues, with whom I collaborated at the University of Luxembourg. I also express my appreciation to Knut Drewing, Steffen Rothkugel, Maud Marchal, and Denis Zampunieris, the jury members of this thesis. I would also like to thank all the other ChronoPilot colleagues, especially my close collaborators Eirini Balta, Efthymia Lamprou, Argiro Vatakis, Bora Çelebi, Álvaro Garrido Pérez, Amrapali Pednekar, and Pieter Simoens for the pleasure of working with them on various occasions. I am also grateful to the European Union for funding my research through the ChronoPilot project (Horizon 2020 Research and Innovation Program, Grant Agreement No. 964464). And finally, I thank my friends and family for their support.

Contents

List of Abbreviations

API	Application Programming Interface
AR	Augmented Reality
BPM	Beats Per Minute
DFS	Dispositional Flow Scale
EDA	ElectroDermal Activity
EEG	ElectroEncephaloGraphy
FSS	Flow Short Scale
HR	Heart Rate
HRV-BF	Heart Rate Variability Bio Feedback
LBSD	Laser Beam Scanning Display
OLED	Organic Light-Emitting Diode
PU	Probabilistic Uncertainty
PPG	PhotoPlethysmoGram
RL	Real-Life
SCR	Skin Conductance Response
TOE	Time-Order Error
TU	Temporal Unpredictability
VR	Virtual Reality
XR	eXtended Reality

List of Figures

List of Tables

Chapter 1

Objectives

A watched pot never boils. This simple idiom is an efficient way to understand the importance of our relation as humans to the passage of time. Everyone experienced moments that flew by when having fun, as well as moments that felt like eternity due to boredom. Time always passes, no matter the activity; however, our sensitivity to its passage is varying.

If we have a sense of time, one could eventually ask themselves the questions: *Can we change this sensitivity to our advantage? Can we make boring moments fly by? Can we make fun moments last longer?*

This concept is the core idea behind the *ChronoPilot* project; a project funded by the European Union's *Horizon 2020 research and innovation programme*. *ChronoPilot* as a whole aims to use external stimuli; through audio, visuals, and haptics; in order to modulate one's time experience to their advantage, namely in order to induce *flow* and *reduce stress* (see Fig. 1.1). However, from that simple idea emerge multidisciplinary needs, which gather experts in psychology, robotics, decision-making, haptics, and finally, mediated reality. As such, *ChronoPilot* proposes *design principles*, assuming that the addition of stimuli aimed at modulating one's time perception could be beneficial in the overall experience. In this thesis, we will refer to *ChronoPilot* as both the *interdisciplinary project* and its *design principles*, which are built on the belief that an interactive application could lead its user to *flow* and *lesser stress* through *subjective time modulations.*

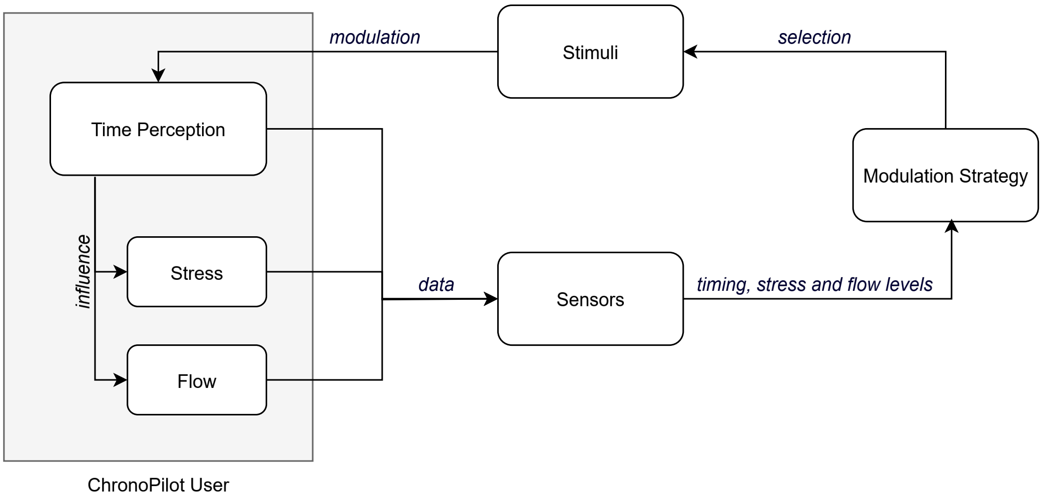


Figure 1.1 Representation of the desired loop in *ChronoPilot*

The work presented in this thesis is deeply tied to this project and its associated design principles. More specifically, we will give answers to the *ChronoPilot* project as designers of interactive applications. Interactive applications are a core need for testing audio-visual time modulation approaches (the stimuli themselves) and time modulation effects in various situations. With results on time perception and rhythmic stimuli, re-usable environments and framework, as well as results in *eXtended Reality* (XR) studies; we will discuss about the delicate usage of multi-sensory modulation, technical consideration of interactive user studies, and question to what extent the *ChronoPilot design principles* should be applied both on top of existing applications or within the design of new ones.

Chapter 2

Background

2.1 Time Perception(s)

Time perception is, while being a self-explanatory concept, a topic full of nuances and with multiple definitions. In this section, we will try to approach a definition of time perception by bringing in elements that are then being put in relation to their usefulness in the *ChronoPilot* project.

2.1.1 The Internal Clock Model

The use of a time model definition is not necessary to be able to use time perception. For example, as shown by Roseboom et al. [1]: in the context of watching videos in different scenarios, a classification network accurately predicted time perception using changes in perceptual content and visual spatial attention (more specifically, gaze position). However, while the use of a model is not necessary for effectively using time perception, this will help us narrow down the definition of time perception, as well as to get some key ideas on how to affect it without looking at the stimuli in detail. Due to its popularity, we will use the *clock model.*

Definition and usual associated mechanisms

This intrinsic *internal clock model*, is a representation which implies that one's body or brain has a dedicated time perception system [2], often accompanied by the idea of an oscillator or pacemaker-counter mechanism. This model suggests that our brain keeps track of time by producing and counting its own time units. Then, as the time perception is dictated by this counting mechanism, the said perception may be altered by either skipping counts or changing the production rate (the *clock speed*) of these time units (see Fig. 2.1). *Skips* and *clock speed* changes may be influenced by external stimuli unrelated to time, the use of

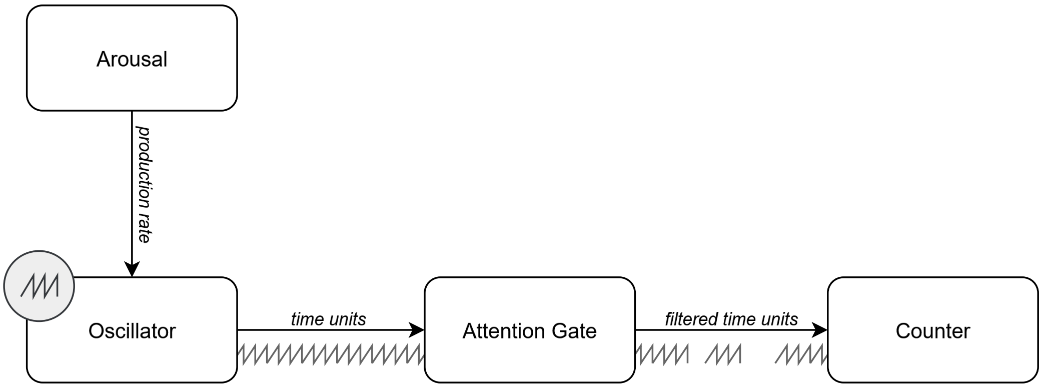


Figure 2.1 Visualization of an internal clock model

external stimuli is the premise of how the ChronoPilot project intends to affect time perception [3].

In these models, it is believed that stimuli can have this effect as they act on attention or attentional resources, which are channels that allow the internal clock to account for the time units produced [4, 5]. As such, internal time changes are associated with two sources in these models: arousal and attention. The tick production rate, or the *clock speed*, would be influenced by the *arousal level*, leading to an extended time perception. Meanwhile, attention would play the role of a gate between the produced ticks and the accumulator. So, the less attention dedicated to time, the more time would feel compressed, as some of these ticks would not be counted [6, 7, 8].

Alternative time models

Although we now have a general representation, the internal clock and pacemaker models have not yet been proven by neural evidence and are still somewhat controversial. For example, this model was often opposed to another *memory storage* model, which correlates time estimations with the amount of concurrent information that a person has to process [9].

It is thus important to remember that the internal clock model is not an absolute and that the model also has different variants. A common variant is to consider multiple clocks; for instance, the literature review by Gorea shows that the same physical duration can be judged differently depending on the reference clock [5]. Another example in favor of this variant is how Fereday and Buehner used temporal binding and causality in time perception. By incorporating an event (such as button presses) and having participants estimate its timing in scenarios involving both causal and non-causal stimuli (stimuli unrelated to any event), they observed no indication of a general slowdown in clock speed on the event itself. This, considering a clock model, indicates either multiple clock mechanisms or that temporal binding has mechanisms separated from the general clock [10].

Continuing on the topic of variants in the clock model, the use of a specific clock model could depend on the considered time length. As shown by Cheng and Creel [11] with the opposition of *interval models* (similar to pacemaker-accumulator) and *entrainment models* (internal self-sustaining oscillator trainable through external stimuli), the former being more suited for longer durations, and the latter for shorter ones.

Implications of the clock model in ChronoPilot

Ending this discussion on the internal clock, it is important to remember that the model in itself is not important in using time perception, but helps in representing time perception while cueing on the use of attention on arousal. A key aspect of ChronoPilot could be to find out how to detect these attention and arousal changes, as well as how to change it. However, these objectives in themselves are not sufficient motivation for the techniques we explored in the ChronoPilot project. And it also does not answer how we express or represent time scales, and as shown by Cheng and Creel's work, may vary depending on the time scale used.

2.1.2 Time Scales

As mentioned earlier, time perception can vary depending on the time scale we consider. We can even find uncommon types of time representation depending on the individual's culture. For example, in competitive fighting video games, players may use time representations based on the game's animations in frames (usually in steps of 1/60 of a second) [12].

Sub-second, Supra-second, and Above

In the literature, we observe different processes depending on the time scale of the events to which participants are subjected. Which vary from sub-second [13, 4] through supra-second [14, 15, 16], to even much longer durations such as multiple days or months long time scales [17]. A good example of this distinction can be found in Droit-Volet et al.'s work, who found that emotional arousal stimuli have varying effects depending on their length, as time distortions decrease with increased durations of the stimuli exposure (varying from 2 to 6 minutes) [7].

Suited scales and resulting biases considering timing-critical modulation

This raises the question for ChronoPilot on what time scales we should focus on, and thus on what stimuli can be used.

One of the claims of the project is to use time modulators in an adaptive way according to the needs of the user and its environment. The needs of a user may be on critical, timely precise events, so sub-seconds to a few seconds long stimuli makes sense for us to use. However, minutes-long stimuli would not suit as it goes against the active modulation aspect of the project.

As such, while we do have an interest in the effects of time modulators in minutes-long scenarios and the general impression they give on a user's experience, the modulators themselves will not reach minute-long scales. As a result, the project and this thesis are biased towards smaller time scales views.

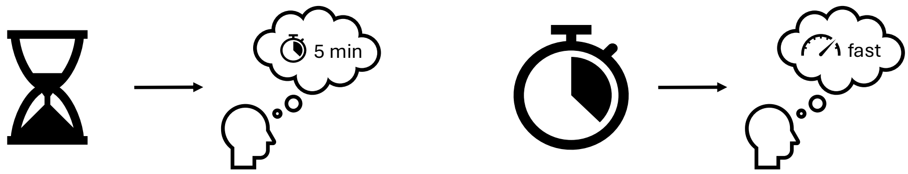
2.1.3 Time Measurements

Subjective time studies may be complex as diverse timing tasks imply diverse processing mechanisms. It is especially true when we compare *time estimations* (such as asking to give an estimation of the duration of an event in seconds) with the feeling of *time passage* (such as an assessment of time passing slow or fast).

As an example, when boredom-prone people are doing a boring task, they will feel a longer subjective time passage compare to non-boredom-prone people, but no difference between the 2 groups is observed on time estimation [18]. Another similar observation can be made on depressed subjects, who feel that time passes slowly but tends to underestimate timing tasks [6].

Another observation of this difference can be seen in a study about flow states induced by the game *Thumper*, these flow states were associated with faster passage of time while not leading to time estimation errors [19]. For a connection with the clock model, it is believed that the time passage is closely related to attention, meanwhile, estimations of durations are linked to memory [6, 20].

As such, we consider time perception through two global subjective measures: *time estimation* and *time passage* (see Fig.2.2). These measures are tied to completely different timing processes. The use of modulators also comes into play depending on whether we want the participants to forget about the passage of time, or to make errors when asked about the time passing.



(a) Time estimation: quantitative evaluation **(b)** Time passage: qualitative evaluation

Figure 2.2 Representation of subjective time perception measures

2.2 Mental States Related to Time Perception

The main goal of using the *ChronoPilot design principles* in an interactive application is to help the users to improve their mental state. Whether through the induction of flow or through the reduction of stress. Both of these states are more popular in scientific literature and present identifiable ties to time perception.

2.2.1 Flow

What is flow?

Since the ChronoPilot project aims to induce flow to its users, we first need to understand what flow is and why it is a desirable mental state. Flow is a specific mental state defined by the psychologist Csikzentmihalyi, it is a state of full attention on a task characterized by nine dimensions: challenge-skill balance, action-awareness merging, clear goals, unambiguous feedback, concentration on task, sense of control, loss of self-consciousness, time transformation and autotelic experience

The currently popular way of measuring flow is through two questionnaires, either the *Flow Short Scale-2* (FSS-2, used for a particular event) or the *Dispositional Flow Scale-2* (DFS-2, used to evaluate flow within a physical activity). When presenting these scales [21], Jackson and Eklund highlights Csikzentmihalyi's focus on the challenge-skill ratio to measure flow, and how previous versions of the scale (FSS and DFS) were unable to represent specific dimensions of flow, especially how the older version had the time transformation dimension showing no relationship with the global flow factor. However, while these new scales reflect it better, through validation studies within that same paper, Jackson and Eklund highlight that "although the higher-order factor loadings of time transformation for the FSS-2 still remain relatively weak" and recommend using factor-level scores rather than the single global flow score[21].

In the context of ChronoPilot, this shows that although flow has a connection to time perception, it is a research subject of itself. Hence, its distinct section within this thesis.

Examples of flow states

Considering how flow is more represented than pure time perception in the literature, and how it is also an objective of ChronoPilot to approach flow through time perception modulation stimuli. We will look into some examples of how flow happens and its effect on user experience.

The first example is a study about telepresence and social media, with the hypothesis that telepresence, here defined as a sense of "being-there" in relation to a medium, precedes flow state. In the context of social media, they found out that flow was influenced by a positive effect of telepresence on enjoyment, concentration, challenge and curiosity; flow then influence the presence of time distortions [22]. For another example involving social media context, the concentration and time distortion components of flow but not the enjoyment were observed to be affected by working as a dyad compared to working alone in virtual worlds (within the social game Second Life)[23]. These social media contexts, on top of showing the importance of telepresence and multi-user aspect of an interactive experience in influencing flow states, also show how flow needs a compatible environment or setting to be induced.

While telepresence favoring flow is a great thing, especially considering the importance of mediated reality in this thesis, it is not a requirement. As seen in a comparison of flow states between *Virtual Reality* (VR) and non-VR setups for play sessions of the rhythmic game *Thumper*, both setups were leading to a flow state despite a higher level of immersion in VR [19].

Continuing on the topic of flow in VR, Reid [24] investigated the relations between flow and playfulness. Here, playfulness is defined as the combination of intrinsic motivation (individual's motive in the activity itself), internal control (the individual is in charge of their actions or has some control over the activity), and freedom to suspend reality (how real the playful activity is for the individual). In their study, Reid describes a model that states that playfulness and flow interact with each other to then influence competence, creativity, and user satisfaction, while both are affected by volitional control (ability to anticipate, choose, experience, and interpret behavior) which itself is affected by self-efficacy. This model was then put to the test through a series of experiments that yielded positive results, but with a limited number of participants.

These two VR-related examples show the importance of the activity itself before the environment in leading to flow. One main idea we can extract from it for ChronoPilot is how flow induction can be context dependent. And thus, one of the research questions of the ChronoPilot project could be about how can we adjust time perception towards flow state, while accounting for the user's needs.

Or in other words, what environmental changes can be used to trigger a chain reaction toward a change in perception of time that precedes a flow state?

2.2.2 Stress

Stress is an essential aspect of the ChronoPilot project, as one of the main objectives is to distract the user away from stress. Stress in itself is not necessarily the perception of time, but one of the factors that affects it. For a first illustration of how time perception and stress are related, but not the same, a bi-directional relationship between distress and time perception in the context of stressful waiting time in very long (multiple days to a month) time scale was reported from Rankin, Sweeny, and Xu. However, the less subjective measures of health and sleep included in the third study of their article (sleep duration and functional limitation due to physical health) were not associated with time perception [17].

Stress and user environment

In the literature, stress research is often associated with the *Attention Restoration Theory* [25], this theory defined by Kaplan states that:

- *Directed attention* (or *voluntary attention*) requires efforts through inhibition of distractions and is susceptible to fatigue, "prolonged mental effort leads to directed attention fatigue"

- The importance of directed attention is conveyed through seven points (selection, inhibition and affect, fragility, perception, thought, action, feeling)

- *Involuntary attention* (or *fascination*) is one of the component required for attention restoration, other components are : being away (conceptually, not physically), the environment must have an extent (be rich and coherent enough, as an endless stream of fascinating stimuli would not be restorative), compatibility between the environment and what one is trying or wants to do

In the same paper, Kaplan explores the idea of how nature environments seem to fit these requirements, going through empirical findings, they suggest that factors leading to stress are either harm (resource inadequacy), direct perceptual pattern or signal (via appraisal, intuition, or through gradual depletion of resources). A challenge is to find whether the correlation between resource decline and stress response is a causal relationship (in either direction) or if there is a confounding of these variables due to factors such as aversive stimulus.

Another example of the usage of nature environments to reduce stress according to the *Attention Restoration Theory* can be found in Beverly et al.'s study [26].

In this study, it was observed that using 360-degree VR cinematic videos of the natural environment managed to significantly reduce perceived stress among a diverse population of frontline healthcare workers with COVID-19.

Since these first two examples, as well as the next ones, are theorizing about the use of natural environments, this allows us to raise questions about the importance of the environment of an experiment when we want to help a participant deal with stress and thus, by proxy, when affecting their time perception.

If in the context of the *Attention Restoration Theory*, most studies would compare a natural environment to an urban one, Anderson et al. instead compared a natural VR environment with a neutral baseline environment [27]. In the said study, VR natural scenes produced improved relaxation compared to the control scene after a stressful event (a math task), which gives credibility to the *Attention Restoration Theory* and its possible usage in VR.

On the same topic, Reese, Stahlberg, and Menzel investigated whether there was a difference in stress reduction between physical urban nature experiences and VR nature experiences [28]. Overall, no significant differences were observed between the two conditions, but the authors mentioned the presence of a larger effect size over time in the real condition.

Keeping in mind the previous examples, we can start thinking that the importance of the environment in relation to stress also translates to virtual environments. As such, in ChronoPilot studies, we need to take into account the environment itself in the design in order to create stress situations through it, but also consider it as a possible stimulus to investigate.

Stress and physiological feedbacks

An advantageous aspect of considering stress is how physiological measures of stress are relatively accessible and can be used both to assess stress levels in real time and to help a participant reduce those. As a first example, using external sensors and a commercial VR headset [29], Cho et al. managed to build a stress-level classification algorithm, with an accuracy over 95%, based purely on physiological inputs. More specifically, using *Heart-Rate Variability*, *Skin Conductance* and *Skin Temperature*. These kinds of measure are easily available through affordable wearable devices, such as the Emotibit[1].

In addition to stress measurement, a use case of physiological signals close to what we aim to do in ChronoPilot is the *Heart Rate Variability Bio Feedback* (HRV-BF), which is a strategy that aims to reduce stress by providing feedback (such as visual representation) of the heartbeat of the participants, a proof of concept of a VR implementation (with VR having the advantage of suppressing

[1] `https://www.emotibit.com/`

distraction) has been demonstrated by Rockstroh, Blum, and Göritz and generated positive results [30].

Continuing on the topic of HRV-BF, in addition to stress, it can also be used to improve performance in specific activities that require concentration, such as darts. As observed by Hiraishi, doing pre-performance routines and throwing the darts in relation to HRV-BF resulted in better performances compared to both metronome-based routines or the absence of routine [31].

If the use of biofeedback is common in stress management techniques, it is not limited to it. For example, VR-based training using *ElectroEncephaloGraphy* (EEG) biofeedback has been used to improve the attention of patients with an *attention-deficit/hyperactivity disorder* [32]. Therefore, even without considering time perception, stress management techniques such as HRV-BF show properties similar to what we are aiming with ChronoPilot. Through the application of pre-selected stimuli in order to improve the experience of a user.

2.3 Approaches to Affect Time Perception

Considering the goals of this thesis, and of the ChronoPilot project, which involve active time modulation, an overview of what kind of modulation can be found in the literature is necessary. In this section, we will separate approaches using the following arbitrary paradigms: *mental processes*, *physical processes*, and *content-based* approaches.

2.3.1 Time Perception and Mental Processes

In this subsection, we will review some of the mental processes involved in time perception. If we go back to the clock model we discussed earlier, counting mechanisms and an attentional gate system were proposed. So, a way to affect time perception is to tackle into these systems directly.

Attention

Attention is believed to act as a gate to the counting mechanism in the clock model. And there are some documented time illusions that involve a shift in attention.

A common attention-based time illusion can be found with the *oddball paradigm*; when the apparition of a low probability event will cause an attentional orienting causing a perceived expansion of time. This mechanism can be exploited through a series of stimuli, when one stimulus will be different from the standard and thus seen as an oddball [33] or by introducing a violation of expectations within a scene (see Fig. 2.3). Tachmatzidou and Vatakis goes further by showing the effects of such violations in both the case of semantic (likelihood of an object to be in a scene) and syntactic (position of an object within a scene) violations [34].

Another temporal sense mechanic related to attention seems to be temporal precision, which means being able to judge different stimuli in their timely aspects. Li et al. in their study brings the topic of alertness and its relation to attention, as well as how it can increase arousal and preparation towards a specific task. If warning cues were already associated with improving visual perception, this study shows significant results that they can also improve temporal precision and this, regardless of the sensory channel used for the warnings [35].

Attention shifting has also been used to reduce the feeling of time passage dragging during a boring task. Xu and David observed that instead of time passage dragging, it would fly by when the boring task is replaced or multitasks with an entertaining task; however, they did not notice an effect on time duration estimations [20]. In a study from Wöllner and Hammerschmidt, mixing music-induced arousal and varying cognitive tasks, they found specific effects from their different tasks on timing experience. Time passage would be felt as faster during math tasks, tapping to the rhythm of the music would make the durations feel shorter, and compressed time estimations and passage were observed when tapping tasks involve tapping to longer rhythmic elements (half notes compared to eighth) [36].

This effect of diverting attention, and ultimately cognitive load, is critical when dealing with time modulation as it implies that the relation of a participant to the activity is a huge confounding effect.

(a) Standard: chef uses a whisk　**(b)** Oddball: chef uses a wrench

Figure 2.3 Oddball example via tools used by a pastry chef

Playing with expectations

If we consider time perception a mental process, then it is possible to change said perception by changing the information that is being presented to the user.

For example, a simple psychological study by Makwana and Srinivasan highlighted the effect of *expected outcomes* on time perception, with time feeling as expended in the expected outcome situations compared to unexpected ones. This was tested by asking the participant what color they wanted to see, then showing either the intended color or another one, the participants were then asked to judge the length of the visual stimuli. If the time distortion effect of the expected outcome is supported by the results, among the duration tested, stimuli of 250 and 500 milliseconds were time distorted by the intention but stimuli of 1000 milliseconds were not [8]. Makwana and Srinivasan consider a pacemaker-accumulator model and suggest that intention acts at the switch level and does not affect the clock speed directly.

Another trick to mislead one's time perception tested by Thönes et al. is the manipulation of time announcements, observing that giving accelerated clock announcements would positively affect task performance [37]. This trick seems consistent with another flow induction observed by Christandl, Mierke, and Peifer about manipulating subjective time progression. Their main idea was to first make the participant perform a task for 10 minutes, then tell them that they were doing it for 5 (*time drags* situation) or 20 (*time flies* situation) minutes; in similar subsequent tasks, participants in the *time flies* situation would perform better and experience higher levels of flow [38]. This means that the recalled flow of an activity can be affected post hoc and influences flow levels in similar activities.

From these examples, we could easily assume that for a project like ChronoPilot, we may be able to achieve our intended effect of inducting flow by selecting what kind of information we communicate to the users. However, this goes against some of the objectives of the project itself, which is more of an assistance to existing problems rather than shaping the problem, a task, to itself.

While these techniques would not be used in ChronoPilot as a part of its active modulation, they may be used to induce a different base time perception situation to later influence with other modulation techniques. However, expectations are not only limited to the time announced. But can also come from expectations of stimuli. For example, in the case of anxiety and anticipation, when there is a chance of having aversive stimuli, Harjunen, Spapé, and Ravaja observed that participants tend to overestimate the time passing [39].

Individual characteristics and time perception

Time perception can vary between people in the same situations. Among individuals, one characteristic that can modulate someone's time perception is *boredom-proneness*, it was observed by Watt that boredom-prone people who had to do boring tasks would experience longer subjective time passage than non-boredom-prone people [18].

Another aspect seems to be the sense of body. By interviewing meditators, Droit-Volet and Dambrun observed that although the passage of time and time estimations are not necessarily tied, these two notions seem to overlap within longer durations; meditation induces *timelessness* and tends to accelerate the passage of time. That feeling of timelessness seems to correlate with the compression of time and is related to the perception of the body in space or the sense of self [40]. However, the authors point out that the effect of meditation on time perception might be due to its slow and motionless aspects, the main point made in that study is about how there seems to be a link between sense of body and sense of time.

2.3.2 Time Perception and Physical Processes

Looking back at the definition of the clock model, in its design, it incorporates elements such as arousal as reasons for the speed of the theoretical pacemaker. We have also seen earlier that states related to time perception, such as stress, can be measured through related physical measures. This raises the question of whether we can directly influence body processes, such as heart rate or temperature, to affect time perception.

Body temperature

One of the stimuli believed to be related to arousal and to affect time perception is temperature. As demonstrated through the review of Wearden and Penton-Voak [41], which theorizes on "a parametric relation relation between the rate of subjective time and body temperature", noting as well that arousal is a proxy on the time experience changes. High body temperature would lead to an increase in arousal, which would then be the cause of time dilation, and the opposite effect on lower temperatures. For a more involved use of body temperature, Kingma et al., through an experiment involving changes in gastrointestinal temperature by using bath conditions, they observed changes in timing accuracy and reaction time depending on the said temperature [42].

Heart rate

If one could think that changing the heart rate could influence time processing, especially considering the previous link with it and stress, it may not be this obvious. As Dormal et al. tried to investigate the link between heart rate (as a proxy of arousal) and time processing, they found that the participant's overestimation of time was not due to physiological arousal, but rather to a distortion of memorized standard durations. According to them, in their study, arousal alone is not enough to explain their supra-second time errors [16].

On a relatively similar topic, exercising has also been associated with arousal and time perception by Lambourne. They noted that their subjects tend to perceive the time interval slower during exercise, but also nothing significant in working memory tasks or episodic timing tasks. Those observations seem to hint that exercising has a similar effect on the internal clock to other manipulations involving arousal [43].

As a result, while we should consider heart rate and arousal as a core aspect of time perception. We cannot just choose to target changes in that specific physiological value, otherwise the changes while exercising in the above text would be directly explainable by heart rate.

Motor action, biological motion

Another aspect that affects time perception is how a user is required to move their body in relation to the content they experience. Interestingly, by recording movement of participants at different speeds, then displaying those movements alongside other participants' movements, Allingham, Hammerschmidt, and Wöllner observed after asking said participants to identify themselves and estimate the duration of the recordings that the faster the movement, the longer it is judged but this was only observed for continuous movement[44]. What was intriguing in this study was that the biasing effect was stronger when watching the movements of someone else [44].

On a more mechanical level, the effect of self-biological motion was also observed on the movement of the hand toward or away from the body. In a study, Tomassini and Morrone asked participants to move their right hand to the right or to the left. While preparing to move, time compression was observed when moving to the left (towards the body), while time dilation was observed when preparing to move to the right (away from the body); the authors suggested that attention, spatial coding, and remapping are unlikely explanations and consider a potential involvement of brain oscillatory dynamics [13].

A set of experiments from Hagura et al. investigates the effect of motor preparation and physical action on time perception. Their results suggest that visual stimuli are judged as longer as participants prepare for action and that this effect is not related to attention; this time dilation seems to be stronger the more prepared is the participant (in these specific experiments, when knowing the direction a target is coming from) [45]. From these experiments, it was also observed that the perceived rate of visual events is subjectively slowed down by motor preparation and that objective performance in processing visual events while in motor preparation is enhanced [45].

It seems that passive motor action (forced by a robot arm) also introduces time distortions, as observed by Kock et al. in the context of supra-second audio event ratings; Kock et al. suggests that motor action or movement time distortions are influenced by viscosity perception mechanisms rather than decision making [15].

Chronostatis and saccades

Chronostasis, also known as *stopped-clock illusion*, is a phenomenon that causes one to feel a second longer than others when looking at the clock. This usually occurs due to an abrupt *saccade* eye movement [46, 47].

Yarrow et al. [46] have conducted investigations on chronostasis and have tried to answer whether the illusion comes from eye movement or a change in visual attention. The results suggest that the extended duration depends mainly on the time required to do the eye movement and that the use of unpredictable movements on the target removes the illusion by breaking the spatial continuity [46].

However, chronostasis has also been observed with non-visual stimuli, more specifically auditory [48] and tactile [49] stimuli. This allows us to think about the intrinsic properties of the stimuli we present to a user and how their body performs to process the said stimuli. An interesting aspect of the saccade is that it can be both visual and non-visual. It is advantageous in the case of ChronoPilot, as these stimuli may work regardless of context if they act on saccades.

2.3.3 Time Perception and Content-Based Approaches

If we can try to map time perception manipulation to pure physical and mental inductions, the notion of time passing by is also a critical aspect of media, games, work and other content-based products. As the work presented in this thesis is related to implementing time perception modulation through XR applications, we will review some of the techniques that act directly on the content presented to a user.

Motion

Representation of motion has evidence to affect the time experience, and that is appearing both in the case of *self-motion* (or *vection*) and *external motion*. In the case of self-motion, Ihaya, Seno, and Yamada observed that vection appears to induce a change in time perception and arousal (from pupil dilation data); however, here the authors considered that *mental tempo* reflects time perception and used the former to evaluate the latter[50].

Even if not mentioned as vection by Weber, Weibel, and Mast, the effect of passive self-motion illusion in VR on time perception, by being driven in a virtual car, seems to non-linearly correlate with the velocity of the self-motion (with longer time judgment correlating to higher velocities). However, the authors mention that it might only be due to the rate of changes in visual effect rather than the velocity itself [51].

As for examples of external motion that affects time perception. In the context of time perception and loading screens, having a faster loading animation (such as a rotating circle animation) seems to yield to a more compressed time perception compared to slower ones [52].

Another important aspect of external motion is how different types of motion appear to affect perceptual duration. A study by Fornaciai, Arrighi, and Burr advances the idea that the repetition of fast translation motion leads to a compressed perception of time [53]. However, they did not observe this effect for motions that are either radial or circular.

Colors

Different colors seem to have different effects on time perception. One simple example can be found in the work of Shi and Huang comparing the effect on time perception between red and blue stimuli, to which the results suggest that red has a significant effect on time perception, but differently depending on the scenarios (here, online dating and job interview scenarios)[14].

On a similar topic of web downloads, according to Gorn et al.'s study testing different fake download web pages that would differ in their colors, color seems to influence feelings of relaxation which then influences perceived quickness of downloads; but the effect of color is attenuated with repetitions. All different components of color (*Hue*, *Chroma* and *Value*) has an effect on perceived quickness, however, *Value* is the most important component in the results of this study [54].

Hue in relation to time perception also received some attention from Thönes et al. with a study comparing *red* versus *blue*. In their results, hue has a clear effect on time perception, but even though red induces more arousal, blue leads to temporal overestimation. Considering the results and the protocol (short time scale), the effect on time perception is not due to arousal or attention, discussion of the authors suggests a possible explanation from eyes' cones or circadian system, but these are simply assumptions [55].

Continuing on the topic of hue, its effect was also observed on *Gansfield stimulation*. In Kübel, Fiedler, and Wittmann's study it was observed that among meditators, color choice during Gansfield stimulation (between red or green) had a lot of impact on time distortions differences (but had a lesser effect on non-meditators). Aside from color, *Ganzfield stimulation* itself also had an effect on time perception [56].

Emotions and mood

Let's now look at the effect of emotions and mood on time perception. Emotional stimuli take various forms, such as music, movies, pictures, etc. Here, we will arbitrarily consider that *emotion* refers to feelings *in the moment*, which could be induced by stimuli, while *mood* refers to a *per individual state* which we have no control over.

A very detailed review of the literature on both mood and emotional stimuli was conducted by Droit-Volet et al., considering an internal clock model with a clock speed. When it comes to depression, time perception is accelerated (which is supposed to be due to less attention and working memory due to the intrusion of negative thoughts), but the sensation of time passing is stretched[6]. More generally, a correlation was also observed between sad state and time compression, high-arousal emotional stimuli, and fear-conditioning situations have a clear effect on time perception, another indirect link between mood and time perception might be present through the effect of mood on emotional stimuli[6]. Droit-Volet et al. in their review of the literature also highlight how the effect of emotional stimuli depends on the context of the time judgment asked and the meaning of the said stimulus. Interestingly but not necessarily related to emotion, in the same review they also mention how showing action-related content, such as people running, can alter duration judgment probably due to reactivation of memory of action dynamics [6].

In the case of music, it was observed by Droit-Volet et al. that different musical pieces will affect time perception mainly through emotional valence and tempo, in a way that is dissociated from arousal, since the orchestration of the same score would lead to different arousal but similar time judgments [57]. Music in general can also be a way to improve the likability and immersion of an activity, which can then lead to altered time states, as observed by Sanders and Cairns in adding background music to video games [58].

It was also observed that high arousal emotional stimuli seem to have a similar effect on time distortions in different periods of time and stimuli length (from a few seconds to several minutes)[7]. This is especially useful in the use cases of ChronoPilot assuming the time scales the project considers.

Environment

If we have seen earlier (2.2.2) that natural environments were used in stress management techniques, a same effect was observed on time perception. Davydenko and Peetz showcased results where walks in nature environment would feel longer than counterparts in an urban environment, indicating that it is the environment that affects time perception in this case. Similarly to studies related to stress, they also observed that walking in nature reduced stress and improved awareness, but they also note that these variables are not correlated with time overestimation [59].

Rhythm

Rhythmic stimuli are a good contender for time perception alterations for practical reasons alone: they are easy to vary (tempo, patterns), they can be applied to different modalities (audio, visual, tactile). Rhythmic stimuli have also been observed to influence users in multiple ways. A first influence on gait can be observed, in an attempt to help in cases of Parkinson's disease, Leow, Parrott, and Grahn observed positive improvements when instructing patients to walk synchronously with rhythmic stimuli [60]. Another effect on improved gait velocity and cadence was observed in older adults by Wittwer, Webster, and Hill [61].

Synchronization with music instructions in a walking task can cause slower, shorter, and variable strides, as observed in one of Leow, Waclawik, and Grahn's experiments [62]. If these results are not too conclusive for the effect of synchronization on gait, it might suggest that the instructions itself require attention resources. In the same study, Leow, Waclawik, and Grahn conducted a second similar experiment but added a cognitive task to reduce the attention allocated to music, which did not affect synchronization but did affect gait itself by making it slower.

Another influence on the *fluency of performance* aspect of flow was also observed in tapping tasks of medium to high rhythmic complexity [63]. Still on the topic of task performance, exergames (exercising through game) seems to positively benefit from the presence of music, the synchronization of music with a cycling task yielded better performances on the task in Yin et al.'s study [64]. If the results of the synchronized music were significantly positive, the presence of music alone was also able to improve the performance of the tasks. Rhythmic auditory stimuli seem to also help in timing and spatial movement tasks through motor performance in the case that vision is unavailable [65].

Synchrony, defined as similar timing between people (which requires either a prediction from one of the persons involved in the synchrony or a mean to synchronize), seems to increase the likability between people when interpersonal. This was observed by Hove and Risen with a series of experiments that involved tapping in rhythm to a metronome, resulting in a significant increase in likability towards the experimenter when the tapping was synchronized with those of the experimenter, but not when it was only synchronized with the metronome (without the experimenter doing the tapping). Stupacher, Wood, and Witte also observed that the synchronicity of the participants with music (but not with a simple metronome) increased the sympathy [67]. As such, since rhythmic stimuli can help in aspects of task performance, health improvement techniques, and sympathy in case of multi-user experiments; it is a perfect candidate for the ChronoPilot project. Its effects on time perceptions are also not to be underestimated.

As mentioned earlier, the appealing aspect of rhythmic stimuli is the ease with which they can be modulated. The speed or tempo of rhythmic stimuli is a simple value change, and tempo changes have evidence of affecting the time experience. That was, for instance, observed on instrumental versions of Disco songs, where a faster tempo caused longer durations in reproduction timing tasks without affecting estimation tasks [68]. The authors also noted that between two musical pieces, a difference of at least 20 *Beats Per Minute* (BPM) was needed to observe a stretched time experience.

In the study mentioned earlier from Droit-Volet et al. on musical pieces and their effect on time experience, they also noted that faster tempos would correlate with longer subjective time. However, the rhythmic effect of a musical piece on time perception appears to be less when the same piece also induces emotional valence [57].

A common misconception is that rhythm would be associated with sound; however, rhythmic patterns are not strictly limited to it. There were even some beliefs that audio was dominant in visuals in relation to temporal judgment of rhythmic events [69]. However, later studies reconsidered it and the rhythmic understanding may instead be related to the amount of information carried by a stimulus. For example, with how Wang, Wöllner, and Shi observed a tendency of participants to prioritize the tempo of the more complex Point-Light-Display dance motions over the concurrent more simplistic audio stimuli[70]. We can also see the use of tactile rhythmic stimuli in timing, with effects such as the improvement in the judgment of *just noticeable differences* [71].

Taking into account the aforementioned studies, we can therefore identify some properties that could differentiate two rhythmic stimuli from each other (see Fig. 2.4). First with the *tempo* (or *speed*) of the stimulus, to which a value in BPM can be assigned (e.g., differences between disco songs[57]), then with the *rhythmic pattern* (e.g., metronome leading to different sympathy than musical pieces [67]), and with how the stimulus is *represented* (or *performed*) to the participant (e.g., rhythmic patterns carried by audio compared to visuals[69, 70]).

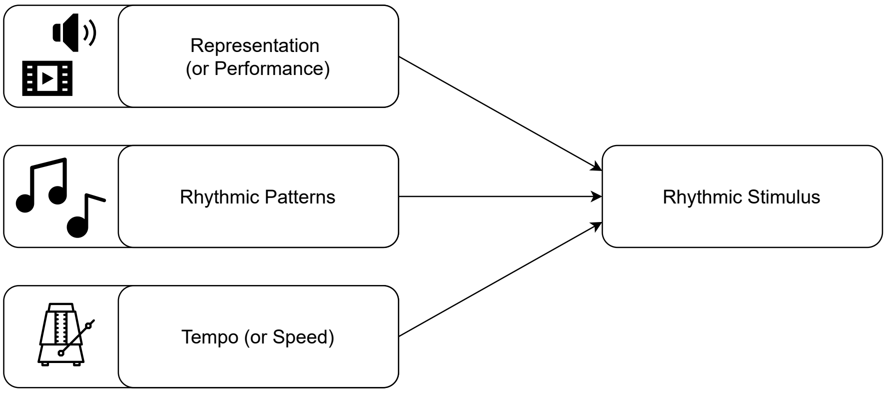


Figure 2.4 Different parameters of a rhythmic stimulus

2.4 Approaches to Measure Time Perception

If the research presented in this thesis has a focus on the modulation of time perception, this effect of the modulation on time perception requires appropriate methods to be measured. There are two main ways to assess a subjective aspect of a user's experience such as time perception, either through user assessments or relying on approximations through proxy data.

2.4.1 Measurement Through User Assessment

In this subsection, we will focus on common methods to measure retrospective and prospective time perception, which differs from other forms of time perception such as circadian rhythm. The methods come from chapters 2-6 of the book *Timing and Time Perception: Procedures, Measures, & Applications* [72] which can be considered as the main reference for time researchers. It is important to note that here measurements only consider *time estimation* (how much time does the participant think has passed) as the *feeling of time passage* (do participants believe that time was passing quickly or slowly) is a purely subjective assessment done through post-trial questionnaires.

Uses of known time units

Two main categories of time estimations can be observed: estimations using known units and estimations using duration comparisons. The former assumes that the participant will make an estimation either of a past event using a verbal estimation (for instance, saying or writing that an event lasted 10 seconds) or through a time production task using these units (such as asking a participant to press a button for 5 seconds, see Fig. 2.5). The use of time units has the benefit of being straightforward to implement; however, the measures tend to have high variability due to an inaccurate representation of a second across participants.

Uses of time comparisons

Time comparisons involve comparing two durations; therefore, intervals are judged relative to another. Although much more precise than using time units, this method has limited use cases. There are two main ways to use comparisons:

- Bisection tasks: participants are asked if a duration is longer than the other
- Reproduction task: participants attempt to produce a presently shown duration

An example of both tasks can be seen in Fig. 2.5. The former often relies on the comparison of a standard duration with another that is affected by a stimulus. It allows building a per-parameter *psychometric function*, allowing a more precise analysis on the effect of a stimulus on time estimations. However, these tasks require a large number of trials with a design centered around said bisection, which can be impractical.

Reproduction tasks involve asking the participant to produce a duration previously experienced. Although precise, these measures may be an index of how consistent a participant is throughout the trials rather than a representation of time estimation changes, since the internal clock speed is already affected when presenting the target duration.

Challenges of implementation

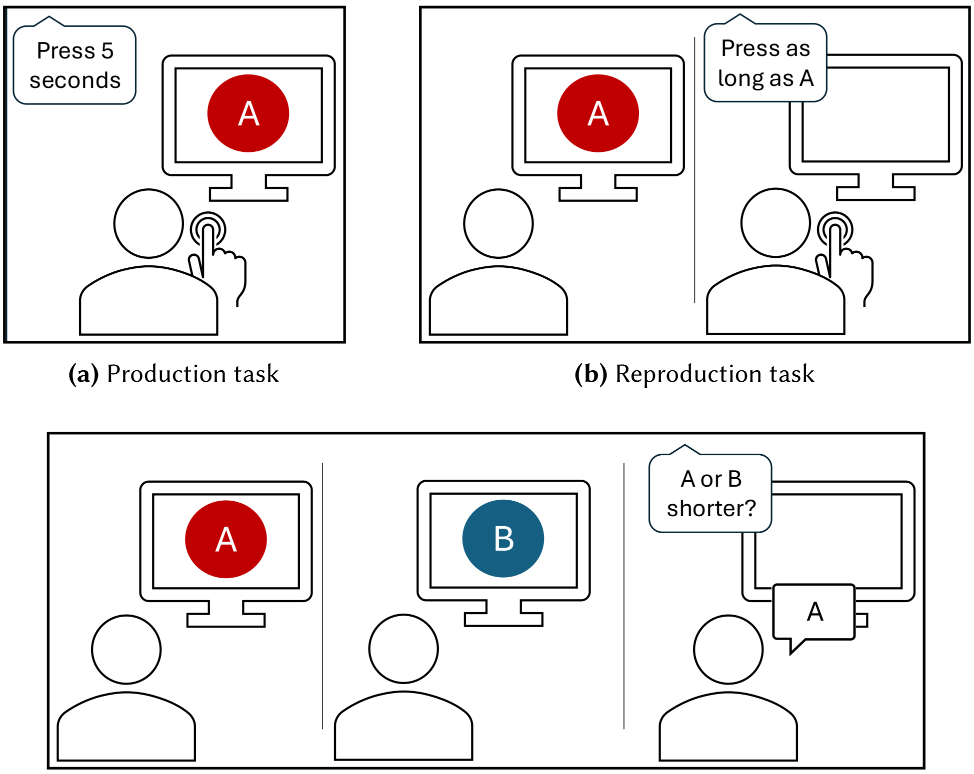
The use of user assessment faces multiple challenges in the design of studies. The precision of the measurement comes at the cost of how practical the task can be. While tasks based on time units are simple to introduce, they still involve interruptions of a participant's experience. Time units tasks are also unfit for short duration tasks, and due to their need for knowledge of conventional units, are not suited to all populations such as young kids.

Reproduction and bisection tasks are even more prominent in changing the design of an experiment, as the compared timings need to be presented in succession, which heavily limits their use cases outside of stimuli studies.

Among perceptual errors, a common error is the *Time-Order Error* (TOE) that occurs on the judgment (of any characteristic, not necessarily duration) of stimuli depending on the order in which they are presented. Relations between TOEs and the estimation of the duration of visual stimuli seem to appear, more especially for saccade-inducing stimuli [73]. Considering the high need of repetitions when using a timing tasks such as bisection, it is also an effect to account for in data analysis.

Evaluating time perception can be difficult in these cases, as ideally experimenters want a report on the participant's experience rather than an evaluation of their performance in how good they are at timing. With the risk of participant using timing strategies on the timing tasks, such as counting, experimenter may need to hide the timing goal of the study and implement a concurrent (non-timing) secondary task.

It could be argued that these timing task methods were also designed specifically for more fundamental research studies; aiming to evaluate the effect of stimuli (or other parameters) specifically on time perception, rather than methods that can measure time perception among other measures. Introducing these methods as more than a post-trial questionnaire already states that the focus of the experiment (or an interaction within the experiment) is on time experience, which biases the uses of results using these methods on the greater context of general experience in interactive applications.



(a) Production task

(b) Reproduction task

(c) Bisection task

Figure 2.5 Example of timing tasks to evaluate a stimulus A

2.4.2 Measurement Through Proxy Data

While the methods extracted from *Timing and Time Perception: Procedures, Measures, & Applications* [72] offer reliable ways to compare the time estimations of a participant, they come with an impractical aspect, mainly due to the requirements of user inputs and for experiment design to be focused around time perception. As such, there has been a growing interest in evaluating time perception from proxy data, mainly physiological signals, to approximate a participant's time perception.

A first strategy can be to evaluate for states related to time perception, such as stress. We have seen earlier the relation between *stress and physiological feedbacks* in Section 2.2.2 which are possible to measure with affordable and non-intrusive sensors.

Emotions are also a good candidate for estimating time perception with wearable sensors. As Orlandic, Valdes, and Atienza managed to successfully build a machine learning model that predicts the perception of the passage of time based on wearable sensors and using data from fear and sadness situations (which leads to slower time passage) as well as mentally taxing tasks (leading to faster time passage). The classifier obtained a 77.1% F-1 score in "distinguishing time passing fast rather than slow" [74], however, such a classifier might instead predict situations that lead to a faster or slower time perception rather than predicting the time perception itself.

If a classification is possible, the emotional response also appears to be correlated with pupil dilation likely due to a raise in arousal, as noted in Chen et al., while providing new methods to normalize pupil dilation values depending on screen brightness [75].

Finally, the use of brain signals, such as EEG, shows promising results in predicting time estimations [76, 77] even in VR [78]. However, while promising, these studies required performing timing tasks while wearing intrusive hardware in the form of an EEG cap, limiting the participant from having head movements. Although not suitable for larger-scale applications with affordable hardware, these results could be especially useful in the future if VR devices start implementing EEG hardware, as seen in OpenBCI's Galea headset[2].

Using proxy data for time perception evaluation can definitely be a way forward both in research and in commercial interactive applications alike; however, it is necessary to keep in mind that as time perception is a multi-facet perception with no definitive (non-proxy) evaluation, proxy data will remain an approximation and may be biased toward *time estimations* without accounting the *feeling of time passage.*

[2] https://shop.openbci.com/products/galea

2.5 Time Perception and XR

In the context of this thesis, XR stands for *eXtended Reality*, which encompasses *computer-mediated realities* such as *Virtual Reality* (VR) and *Augmented Reality* (AR). As such, when we talk about XR environments, we talk about environments that are controlled by a computer and represented to the user through means such as (but not limited to) *Head-Mounted Displays* (HMDs), which actively alters the perception of its users.

In recent years, mediated (or extended) reality technologies have been met with increasing interest for research and commercial use. The technology became more accessible through the releases of both dedicated hardware (e.g., HTC VIVE, Meta Quest, Microsoft Hololens, Apple Vision Pro; HMDs used in this thesis are shown in Fig. 2.6) and software standards such as OpenXR[3]. As such, XR environments can be seen as a new part of some populations' daily life, and thus a subject to research in itself. But its nature as computer-mediated realities also makes it an opportunity for (real-life) research, as it allows for greater scenario control and reproducibility for experimenters.

In the case of the ChronoPilot project, XR is the perfect testbed to evaluate implementations of the *ChronoPilot design principles*. In the following section, we will first see how XR differs from other computer applications due to *telepresence*, and then explore multiple ways in how mediated technologies can affect time perception. In the case of both ChronoPilot and this thesis, we are using XR specifically through HMDs; however, at the moment, use of HMDs is much more documented for VR than it is for AR (where using the camera feed of smartphones is more common). As a result, the following section is biased towards VR studies.

(a) VIVE Pro Eye [4] **(b)** Microsoft Hololens 2 [5] **(c)** VIVE XR Elite [6]

Figure 2.6 The XR HMDs used in this thesis

[3]https://www.khronos.org/openxr/
[4]https://www.vive.com/sea/product/vive-pro-eye/overview/
[5]https://www.microsoft.com/p/hololens-2-industrial-edition/8mqn5pzp01x5
[6]https://www.vive.com/us/newsroom/2023-01-05/

2.5.1 Immersion and Telepresence

Immersion is one of the most used arguments in favor of the use of XR technology, however, it is often misconceived as *telepresence*. While telepresence is a feeling of being present in a different environment than the real one, immersion is a state of deep mental involvement which causes a dissociation from the awareness of the real physical world.

It is not surprising to find relations between telepresence and immersion. Considering a VR flow-based context, Mütterlein modeled the following relationships between interactivity, telepresence, immersion, and satisfaction: interactivity enhances telepresence and immersion, telepresence enhances immersion, and finally, immersion improves satisfaction[79]. However, if this representation was supported by their study results, the effect of immersion on satisfaction is low, so other non-investigated factors might affect user satisfaction in VR.

Using VR means that we have the choice to use an avatar or not, and the presence of an avatar might affect time perception. This has been investigated by Unruh et al. in a waiting scenario in front of a mirror under three conditions: full body avatar in VR, no avatar in VR, and real-life scenario [80]. The only differences observed were between the avatar and the no avatar conditions where that avatar presence leads to, in a retrospective paradigm, significant faster passage but similar time estimations.

In 2019, Kim and Ko investigated what they call flow experiences and VR in the context of sports viewing. But in their study [81], flow is used interchangeably with time distortions. They varied situations using 3 aspects: flat screen or VR, sport involvement (how engaged the viewer is with the sport), and rivalry of the teams involved. They noted that VR had a significant effect on telepresence, which by proxy affects time distortions, since flow was affected by both sport involvement and telepresence. But, with a reduced impact of telepresence if there already was involvement from the viewer. Kim and Ko also suggests that telepresence either is a component or an antecedent of flow experiences.

2.5.2 Virtual Reality Paradigm on Time Perception

While the use of mediated reality technologies has some obvious advantages for user studies, the technology itself induces some biases in time perception. A first example can be found in how the same game is more prone to underestimate time in VR compared to a flat screen [82].

Another time perception difference while being bored or while waiting was observed between real-life and a VR counterpart; as Igarzábal et al. reported in their study that compared the self-reported waiting time in the virtual waiting room compared to a real-life one in the retrospective paradigm [83].

However, Bruder and Steinicke reported that walking in VR does not appear to significantly affect time perception, although there is a tendency to overestimate time while walking in VR compared to non-VR, likely due to discomfort[84]. As such, technology itself is not sufficient to create the perception shift, and it can be a shift in relation to the content or activity of the participant. This thought can also be put in relation to Mallam, Ernstsen, and Nazir's study comparing estimation times between a VR and desktop equivalent in tasks lasting from 30 seconds to 5 minutes, where in situations of overestimation in both cases, the VR situation tends to be overestimated even more [85].

We can thus hypothesize that VR may have an amplification effect on the direction of the timing experience. An amplification that may be related to other aspects of experience, such as the better performance observed in the video game *Thumper* [19]. Although an effect of technological immersion is observed, these kinds of amplification effects do not invalidate the use of mediated technology for time perception studies.

On another note, in relation to the effect of emotional content and virtual reality, no significant differences were observed when emotional content stays the same [86]. We can then expect that, between VR and other setups such as AR, if we use emotional content, only the content itself will, regardless of the setup, influence time perception.

2.5.3 Virtual Environments and Time Perception

One of the key aspects of XR is how it changes the perception of our environment. We may display fully virtual environments to the users, but it still goes through a screen and thus, we could ask if the virtual representation of an environment has the same properties as its real counterparts.

In previous chapters, we mentioned the effect of natural and urban environments. The relation between crowded environment and time perception was also studied in VR. Shimokawa and Sugimori replicated a real-life experiment in VR about subjective time judgment, in an environment varying in being crowded/uncrowded and natural/urban using VR videos, the results were similar to the real-life experiment with a correlation between uncrowded and non-urban situations with time dilation [87]. This suggests that changes in virtual environments through VR can be a valid way to shift time perception.

Environments tend to have temporal cues or *zeitgebers* such as clocks or the movement of the sun. Schatzschneider, Bruder, and Steinicke tested the effect of the movement of the Sun (as a zeitgeber) in a virtual environment on time perception with and without cognitive loads; the results are that the absence or presence of sun movement and the presence of a task would affect time perception, but the speed of the zeitgeber did not affect time perception, if its presence affected time perception, it did not affect task performance [88]. We could then assume that subtle changes in environment, by introducing and removing zeitbegers or other background elements, would affect the time experience in a non-intrusive way.

Although it is not specifically about time perception, but more about flow or just concentration in general, Ruvimova et al. went as far as suggesting the use of virtual reality environment in open offices to induce better flow or concentration, approaching the likes of a closed office. Their results yielded some positive results on flow for one of their environment, despite some notes that their VR hardware was not suited for long work sessions due to weight and eye strain [89].

2.5.4 VR and Motion

VR is a powerful tool in order to test one's relation to motion, Verde et al. used VR to investigate the effect on time perception from self- and world-movement. It appears from their study that only apparent movement of a world (visual) movement causes perceptual time distortions (in the form of time dilation); situations where people are walking but where objects in the virtual world are either not moving or moving in the participant's walking direction do not lead to time distortions [90].

An interesting idea from Bansal, Weech, and Barnett-Cowan was to apply the *sensorimotor contingency theory* to the perception of motion and time. Their main idea was to couple the speed and duration of visual events with the body movement of participants in VR, the results of their study suggest that after the experiment participants underestimated time while moving and overestimated it while static[91].

2.5.5 XR and Synaesthetic Designs

Synesthesia is the phenomenon of a sense (audition, vision, feeling the temperature) being affected by a seemingly unrelated one, such as "hearing pictures". Synesthesia can happen naturally, but XR technologies allow one to create *artificial synesthesia.*

By *synaesthetic designs*, we talk about interactive applications where stimuli are associated either with events or with other stimuli. For example, Wright and Kefalidou designed an experiment involving a virtual maze (as a flat-screen desktop game) with a *synaesthetic-oriented design*, which means that some elements and stimuli of the *user experience* were fused (sound cues accompanying visual cues, colors accompanying the solution). In this experiment, immersion was observed to be affected by user experience but not task solving and navigation or *change blindness*;in terms of user experience, color signals and landmarks were judged positively, while disorientation (from the maze) was judged negatively [92].

As body temperature was observed to influence time perception [41], Erickson et al. tried to synaesthetically induce a feeling of temperature change and see if it has an effect on time perception. Using AR, if showing visual effects that evoke temperature (such as ice or fire) on the participant's hand influenced their perception of temperature, no effect was found on a prospective time production task[93].

In the context of VR, Rogers et al. conducted a study [94] about how the presence of music might alter time perception in a VR game. The study consisted of collecting retrospective time estimations of a VR bow-and-arrow tower defense game under two conditions, one without music and the other using music (which was determined with a prior online survey in order to use music with sufficient arousal and valence). The results imply that time passed faster when music was present; interestingly, the faster time passage did not seem to correlate with a change in immersion. In the same paper, Rogers et al. conducted an exploratory case study comparing sound effects and adaptive music to signal game mechanics in VR, which approaches synaesthetic designs. If the results are limited to the exploratory nature of the case study and the sole presence of qualitative data, these results highlight the potential for future studies, as it was observed that music significantly affected the experience of the player and increased tension was experienced under adaptive conditions.

Artificial VR synesthesia has also been made use of by Outram et al. to showcase full-body vibro-tactile suits. By linking audio with visuals and vibrations, they increased immersion and enjoyment of their participants [95]. Artificial VR synesthesia was also used by Reif and Alhalabi to mitigate attention and immersion in the case of pain therapy [96].

The design of ChronoPilot is, by its concept of tying time perception to multisensory stimuli, synaesthetic. We can expect, depending on the use case, some of the benefits from synaesthetically designed applications seen above. The use of authoritative strategies to synchronize stimuli, such as synchronized rhythmic patterns, also goes in these directions.

2.6 Interactive Applications and Behavioral Studies

Time perception studies may rely on multiple media formats, such as videos [26, 1], music [57, 68] or computer games [19, 23]. As such and just like any other kind of behavioral study, using custom-tailored interactive applications is relatively common.

However, making these custom-tailored applications requires a lot of skill and work, this is even more prominent when it comes to VR or AR applications. In addition to basic programming skills, making these applications may require knowledge in interactive 2D or 3D development, network programming, user interface design, data logging and processing, etc. Since this thesis focuses on the design and implementation of behavioral studies such as XR time perception studies, it is necessary to review the technical challenges and needs related to the technical work of such studies.

In the case of making interactive 3D or 2D experiences, several game engines are of popular choice (Unity, Unreal Engine, Godot Engine), simplifying the required implementation work by alleviating the need to build rendering and audio pipelines, providing physics, and handling player controls. These engines also offer the possibility to import numerous plug-ins or assets, either first-party or third-party, which leads to new markets such as the *Unity Asset Store*[4]). In the case of XR development, the Unity engine is popular in behavioral studies due to its easy integration through Unity-compatible toolkits provided either by Unity themselves (XR Interaction Toolkit, AR Foundation), by other industry's giants like Microsoft's MRTK, or by headset makers such as Meta, Vive or Pico.

However, be it for desktop and XR interactive applications alike, while these engines and integrations do facilitate the development work, a considerable workload is still needed. This is why in recent years we can also see a trend of research-focused extensions being developed. Indeed, we can find examples such as the *vexptoolbox* [97], the *BiomotionLab Toolkit for Unity Experiments* (bmlTUX) [98] or the *Unity Experiment Framework* (UXF) [99], which allow to manage some common requirements of behavioral studies such as data collection or trial generations. The most advanced one being the *Unified Suite for Experiments* (USE) [100], with technically complex features to integrate, such as trial recording and replaying with millisecond precision or *Arduino IDE* integration.

However, these tools do not provide the environments nor the integration of in-environment interactions, this is why we can also see the emergence of scientific articles presenting non-generalist virtual environments. For XR related environments, examples can be found in Outram et al.'s environment for a vibrotactile suit [95], or Landeck et al.'s Unreal Engine environment to test the effect of zeitgebers on time perception [101]. The extensions and environments mentioned above are geared towards the integration of experimental design [98, 98, 99, 100] or specific scenarios [95, 101].

Simplification of interaction implementation in a more all-purpose way, especially in the case of XR interactive applications, can be found in community-

[4] `https://assetstore.unity.com/`

focused commercial games such as VRChat[5] and Rec Room[6], by providing tools and *Application Programming Interface* (API) to create multi-user XR experiences for their platform. Although behavioral studies are possible using these platforms [102], the restrictive nature of these commercial applications can be a problem in terms of practicability (due to the dependence of an external platform and the API eventually deprecating an existing implementation through updates) and data privacy.

Some open-source projects like BasisVR[7] aim to provide similar experiences and associated APIs while having the option to self-host, but are still in development and not production-ready as of now.

Although behavioral studies heavily benefit from recent technological progress, we also observe a growing technical debt for the implementation of experiments, especially compared to commercial products.

2.7 Research Questions

In this chapter, we have over-viewed several subtopics related to the pluri-disciplinary research topic that is time perception.

We first have seen what time perception is and its implications for other mental states (Sections 2.1.1, 2.2.1), followed by a glimpse of the ways researchers have tried to influence it (Section 2.3). Then, after looking deeper into the methodology of the time perception measurements (Section 2.1.3) we went through topics more specific to this thesis with the introduction of XR in timing research (Section 2.5 and the challenges that come with the technical implementation of experiments (Section 2.6).

In the next chapter, we will present research productions that only consider a few specific stimuli; however, the variety of approaches to affect time perception presented in the current chapter comes not only as opportunities (such as additional levers we could use to influence the time experience), but also, since intrinsic aspects of an activity (fatigue it induces, attention needed, presence of movements, colors, etc.) could act on the time experience, as threats of confounding effects.

The results of the research presented in this thesis, bounded by the ChronoPilot project's *design principles*, mainly involve the implementation of timing research through XR experiments, which introduces both technical implementation needs and changes in the representation of stimuli and time experience. As such, three

[5] `https://hello.vrchat.com/`

[6] `https://recroom.com/`

[7] `https://github.com/BasisVR/basis`

main questions may emerge from considering both this *Background* chapter and the following *Results* chapter:

- How can newly tested stimuli, in our case rhythmic stimuli, be integrated as part of the time perception within an experience (such as research experiments, commercial applications, etc.), especially when there are potential interactions with unknown intrinsic aspects of an experience in time perception?
- What changes from the introduction of XR in the timing experience? How should the design of time modulation be changed according to the use of XR technology?
- How to implement such XR experiences?

The aforementioned questions are all elements of the key research question of this entire thesis: *To what extent and with which methods can active time modulation be used in interactive applications to improve flow and reduce stress?*

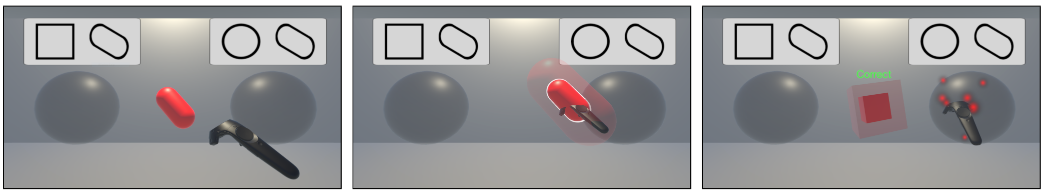
Chapter 3

Results

In this chapter we will review the different studies that have been published as part of this thesis. Following a cumulative thesis format, works that lead to a publication have been contextualized and are also fully available in the Appendix. This chapter also has unpublished works; two are unpublished results with associated sections that are closer to a full article; finally, some ongoing collaborative works are also briefly over-viewed.

3.1 Time Perception and Rhythmic Stimuli

The following three papers have as a main topic of interest the use of audio-visual rhythmic patterns in a shape sorting game. The first paper corresponds to a study conducted through a browser version displayed on a normal screen. The second and third papers refer to a study using the same game but in VR.



(a) **(b)** **(c)**

Figure 3.1 A sorting sequence in the VR experiment (from Appendix B)

3.1.1 Rhythmic Stimuli Effects on Subjective Time Perception in Immersive Virtual Environments (MMVE, ACM MMsys 2022)

This conference paper focuses on the use of stimuli following a metronome-like rhythm and their effect on both time perception and task performance.

Motivated by the goals of ChronoPilot, which are to use time perception as a tool of user experience, we thus wanted to study these effects on a task that is not related to time perception itself.

To this end, we designed a sorting task of primitive 3D shapes, requiring the participant to move said shapes using a computer mouse and drag it to one of the two possible choices. Time perception will then be assessed through a post-trial questionnaire, due to the multifaceted aspect of time experience (Section 2.1.3), both *time estimation* and *time passage* are investigated. Since the task may be exhausting due to its repetitive, cognitive, and physical nature, we also asked about the participant's fatigue over that questionnaire.

As mentioned earlier in the *Rhythm* section of 2.3.3, rhythm is a relatively simple way to both synchronize (as rhythm can be used with different modalities such as audio, visuals or haptic [69, 70, 71]) and modulate stimuli (using different tempo speeds [68]). As such, the study is also designed around multi-modal stimuli; here, we explored combinations of the presence or absence of auditory clicks and visual pulses. Since we want to ensure that the participants are exposed to the stimuli, the visual pulses are on the objects they sort, while the auditory clicks use global and non-spatial audio.

The stimuli were voluntarily kept simple and unrelated to the task for allowing later uses in other ChronoPilot-related studies, to avoid biases on performance metrics due to cueing such as observed in synaesthetic designs (Paragraph *XR and Synaesthetic designs* of Section 2.5.5), and to keep the stimuli rhythm-based rather than emotional (Paragraph *Emotion* of Section 2.3.3) or environmental (Paragraph *Environment* of Section 2.3.3). The results of this study show a varying effect of tempo on time perception depending on the use of single or synchronized stimuli; the presence of single stimuli greatly affects time estimation but with no specific error direction, while tempo affects the error direction for synchronized stimuli. In terms of time judgment, it seems to depend on the tempo for both single and synchronized stimuli. In this study, it was also observed that the presence of visual stimuli, probably due to the task challenging shape recognition, negatively impacted task performance.

This article was published in 2022 as part of the *IMmersive Mixed and Virtual Environment Systems* (MMVE) workshop of the *13th ACM Multimedia Systems Conference* (MMSys 2022). I am the main author of the article's manuscript, with my contributions being the designer and developer of the interactive application, collecting and analyzing the data of the study, and writing of the article.

The full paper can be found in Appendix A

3.1.2 Rhythmic Stimuli and Time Experience in Virtual Reality (EuroXR 2023)

This conference paper is a follow-up study of the previous paper with the main goal being to transpose the previous experiment to VR.

Although VR may introduce some time perception biases depending on the content displayed, the direction of the time experience (under-estimating or over-estimating) from similar content between VR and flat screen or real-life counterparts tends to be similar, as seen in Section 2.5.2.

A limitation of the previous study also came from its use of crowd-sourcing to recruit participants, leading to inconsistent screen setups between them; this would be resolved in this study, as it was hosted in a dedicated lab and also allows us to run more trial rounds per participant.

The virtual environment has been made compatible with VR, replicating the controls using motion controllers. Although participants were allowed to take breaks at any desired moment, this aspect, coupled with the larger amount of trial per participant, changes how we had to approach the data analysis. The added motion and the change in the amount of trials both contribute to inducing extra physical fatigue over time in the participant, as well as changes in physiological aspects such as body temperature or heart rate, which also affect the time experience (Section 2.3.2). Having more trials also means that participants will be more trained and thus more proficient at the task in later trials, potentially affecting the challenge-skill balance which defines flow (Section 2.2.1).

As such, this study aims to explore the effect in VR of rhythmic stimuli, either or both audio and visual, on time experience and task proficiency. More specifically, the study examines the effects of stimuli based on the context of user fatigue and trial repetition.

This VR experiment led to similar but more detailed results compared to its desktop counterpart, with effects of the stimuli on time experience and task performance being dependent on user fatigue and trial repetition. For instance, tempo affected time judgment only during low fatigue, while the type of stimuli was significantly affecting it during the middle of the experiment, where participants would be familiar with the task while keeping low fatigue. Meanwhile, some effects would evolve throughout the session, such as the presence of auditory stimuli improving the performance in earlier trials, while visual stimuli would decrease performance in later ones.

This article was published in 2023 as part of the *20th EuroXR International Conference* (EuroXR 2023) proceedings. I am the main author of the article's manuscript, with my contributions being the designer and developer of the interactive application, collecting and analyzing the data of the study, and writing of the article.

The full paper can be found in Appendix B

3.1.3 Psychophysiology of Rhythmic Stimuli and Time Experience in Virtual Reality (Computer & Graphics)

This journal paper is a revised version of the previous conference paper, published in 2024 as part of *Computer & Graphics Special Section on EuroXR 2023 Best Papers* and includes additional results on physiological data that were collected during the experiment.

As the VR experiment was conducted in person, this also gave the opportunity to use devices to collect physiological data on each trial. The HMD had embedded eye-tracking compatibility, while a non-intrusive wristband collected heart-rate and skin temperature measures.

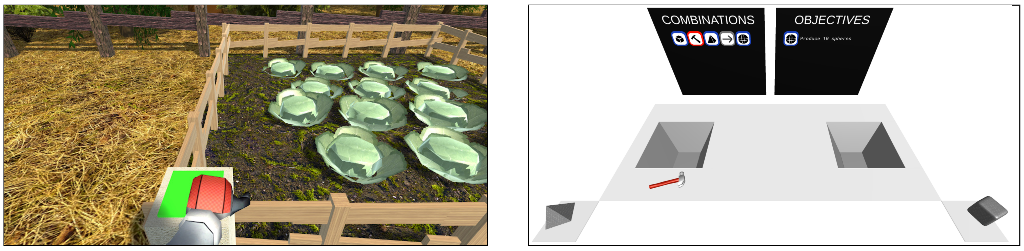
All three physical processes appear to be related to time perception in the literature (Section 2.3.2), especially in the case of eye movement, as it may be indicative of saccades; pupil dilation also has examples of being tied to arousal [75]. Exploring how physiological values correlate with aspects of time perception may give us access to additional proxy data (Section 2.4.2) to use throughout the ChronoPilot project.

The resulting analysis shows promising correlations of time experience with eye-tracking, temperature, and heart rate. However, due to the exhausting and repetitive nature of the experiment, these observations could also be due to a confounding effect of fatigue.

The full paper can be found in Appendix C

3.2 Re-Usable Environments and Framework

The two following papers describe custom virtual environments built as re-usable baselines for later experiments. The first is a poster paper focusing on a VR farming environment, while the second showcases on a multi-user AR industrial one. If the environments differ both visually and in the interactions perceptible by their user, these projects share similar goals in facilitating the technical implementation of user experiments, by relying on the same back-end allowing customization components architecture loading in real-time.



(a) VR environment (from Appendix D) **(b)** AR environment (from Appendix E)

Figure 3.2 The two custom environments

3.2.1 A Dynamic and Scriptable Environment and Framework for Stimulus-Based Cognitive Research in Virtual Reality (IEEE ISMAR2023)

This poster paper explores the features of a custom-made Unity environment and framework tailored for time perception user studies in VR. Although the making of this environment was started due to the obligations of the ChronoPilot project, as a part of the project's deliverables was to test modulation strategies under a virtual precision farming scenario, we also identified several technical integrations that could benefit interactive user studies beyond the specific needs around ChronoPilot.

As such, on the one hand, this paper presents interactive environments with a specific scenario similarly to the environment of Outram et al. for a vibro-tactile suit [95], or the environment of Landeck et al. for studying zeitgebers on time perception [101]. More specifically, our environment allows immersive plant management tasks which can induce varying levels of cognitive load, stress, and engagement through two task types: plant production and plant identification.

However, on the other hand, it also presents a framework integrated in that same environment which offers more generic tools similar to frameworks such as the *BiomotionLab Toolkit for Unity Experiments* (bmlTUX) [98] or the *Unity Exper-*

iment Framework (UXF) [99]. This framework notably allows one to externally define the content and events of the environment through external files, which allows a form of scripting, it also allows to externally control the environment through network messages.

This environment and its associated framework was used as a base for the experiment presented in Section 3.3.2.

This article was published in 2023 as part of the *2023 IEEE International Symposium on Mixed and Augmented Reality* (ISMAR 2023) proceedings. I am the main author of the article's manuscript, with my contributions being the main designer and developer of the core framework and example tasks, as well as writing of the article.

The full paper can be found in Appendix D

3.2.2 XR MUSE: An Open-Source Unity Framework for Extended Reality-Based Networked Multi-User Studies (MDPI Virtual Worlds)

In the associated journal article, the *XR MUSE* Unity framework is presented. Similarly to the environment presented in the previous publication, *XR MUSE* is a project that has been carried out to answer a ChronoPilot deliverable, but has led to a product that can benefit research in a broader scope. *XR MUSE* offers an evolution of the framework present in the previous paper, which aims to provide a more generic approach that completely detaches the framework from the tasks and environment.

Shared collaborative spaces are common situations in daily life, making those research scenarios valuable to investigate. For example, it was observed that working as dyads could improve flow even within virtual environments [23].

As previously observed in Section 2.6, interactive applications for research often use game engines, such as Unity, which is the one used in *XR MUSE*. Fortunately, several extensions are available to implement networking within Unity, such as Photon[1], Mirror[2] or Unity Netcode[3]; however, while these libraries ease the implementation by removing the need to handle lower-level protocols, using those still requires challenging per-application integration.

Several research frameworks, usually built on solutions such as Unity Netcode or Photon, offer multi-user integrations for various XR devices such as VR headsets, AR glasses, and tablets [103, 104, 105]; while these frameworks have associated publications in the scientific literature, they remain as internal prod-

[1] `https://www.photonengine.com`
[2] `https://www.github.com/mirrornetworking/mirror`
[3] `https://docs-multiplayer.unity3d.com`

ucts with no available code base. We also noted earlier in Section 2.6 that, while there are some existing commercial platforms that allow the creation of shared experiences, they remain unpractical for research.

For these reasons, *XR MUSE* offers an open-source framework for multi-user interactive XR experiences. Achieving its goals through three independently usable components: an example scenario and environment, networking integration, operations. The example environment, designed around a workbench setting, allows for natural collaboration-based interactions. The networking integration implements *Photon PUN2* to offer higher-level coding tools such as a *Double Transform Design* and a *Data Registration* system. Finally, the scene description operates similarly to the farming environment 3.2.1 by loading objects and scripts through external files.

The example environment and the *XR MUSE* framework were used as a base for the experiments presented in Section 3.3.4.

This article was published in 2024 as part of the *MDPI Virtual Worlds* journal. I am the main author of the article's manuscript, with my contributions being as the main designer and developer of the core framework, the designer and co-developer of the example tasks, and writing of the article.

The full paper can be found in Appendix E

3.3 XR and Time Perception

The following works explores the effects of XR content on time perception through the use of meaningful modulation. The first focuses on user engagement and environmental dynamics, where users had to handle a memorization task in an immersive VR environment. The second is adapted from a work in progress publication; the presented experiment attempts to use pre-selected time modulators in order to help participants in VR farming situations with varying levels of stress. The third describes unpublished results in a series of experiments comparing, within comparable test environments, timing tasks across VR, AR, and *Real-Life* (RL). Finally, the last section presents ongoing work, while no data analysis is currently available for these ongoing experiments, this serves as a showcase of an implementation of the framework presented in Section 3.2.2.

3.3.1 Some Times Fly: The Effects of Engagement and Environmental Dynamics on Time Perception in Virtual Reality (ACM VRST 2024)

This conference article explores the effect of user engagement and environmental dynamics on time perception in VR. Although it was noted in the *Background* chapter that VR could induce either an underestimation bias or an error amplification bias (see Section 2.5.2), the goal of this paper is to study how the integration (or implementation choices) of a virtual environment could also affect the time experience.

To this end, a custom environment with a time estimation task and a memory side-task was used. The environment can vary in environmental dynamics through the removal (*Static* environment) or addition (*Dynamic* environment) of audio and background (visual) elements, as well as in user engagement through the use of either in-environment interactions and full vision (*Active* environment) or with on-screen overlays and limited field of view (*Passive* environment).

Each participant had to do a memory task on displayed numbers (either in the environment or on overlay depending on the conditions) through all possible combinations of the previously mentioned variables, followed by a time estimation task.

The results indicate significant effects of both variables on time estimation; in fact, the highest difference could be observed between the *Dynamic-Active* condition and the *Static-Passive* condition, which strengthens the idea that the implementation of the VR perspective is crucial in interpreting results in time perception user studies.

This article was published in as part of *ACM Symposium on Virtual Reality*

Software and Technology (VRST 2024) proceedings. I am a co-author of the article's manuscript, with my contributions being helping on data analysis, aspects of technical implementation, and revision of the article.

The full paper can be found in Appendix F

3.3.2 Shaping Time Perception Under Ambiguity: Multi-Sensory Interventions in a VR Precision Farming Scenario

These unpublished results investigate the use of pre-defined time modulation stimuli in response to situations with varying levels of expected stress within an immersive VR farming scenario. Due to the unpublished nature of these results, the following section will be divided into distinct titled paragraphs. The presented results are largely based on the analysis made by Eirini Balta and Efthymia Lamprou.

The content in this section is adapted from a current work-in-progress article. I am a co-author of the article's manuscript, with my contributions mostly being the environment design and implementation, help in the experiment design, and writing of the *VR Environment* section of the article.

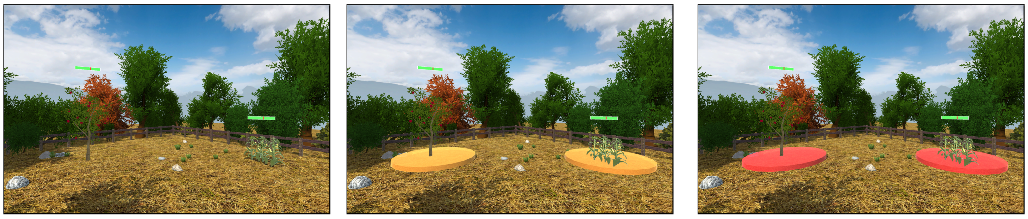
Rationale and environment design

Stress in fear-provoking situations has varying effects on time perception, such as slow time perception when using various aversive stimuli (pictures [106], electric shocks [107], or looming stimulation [108]) or faster time perception when faced with an imminent deadline [109].

Last decade, Lake and LaBar proposed a classification of ambiguous fear-provoking threats with two dimensions involving different cognitive processes: *temporal unpredictability* (not knowing *when* the threat will appear) and *probabilistic uncertainty* (not knowing *if* a threat will appear) [110]. This *ambiguity* aspect of a threat could also be associated to the use of *expectations* as an approach to modulate time perception (Section 2.3.1). These situations of *ambiguity* are perfect candidates to try and apply time perception modulators according to the *ChronoPilot design principles* (see Chapter 1), with the aim of reducing the negative effects of the resulting stressful situations.

To this end, the framework and example VR environment presented in Section 3.2.1 was used with newly implemented tasks. The participants went through trials that involved preparation and reaction to pest attacks on plants. Stress scenarios of varying *Ambiguity* were introduced through the display of *danger zones* (see Fig. 3.3), indicating that for a trial, either no attack would happen

(*Safe* condition), an attack could occur after 2.5 minutes (*Probabilistic Uncertainty* (PU) situation), or that an attack will occur but at an unknown time (*Temporal Unpredictability* (TU)).



(a) *Safe* condition **(b)** Uncertain attack (*PU*) **(c)** Certain attack (*TU*)

Figure 3.3 Example scenes and conditions with associated dangers

Using immersive controls, participants could choose where to place a drone ready to spray pesticides according to the *danger zones* displayed. Then, once pests appear (see Fig. 3.4), they could choose either to deploy the drone and use safe pesticides (with the drawback that the drone travel time could eventually be too long if it was deployed far from the actual attack) or use an emergency pesticide (with the drawback of actively harming the plant in the process).

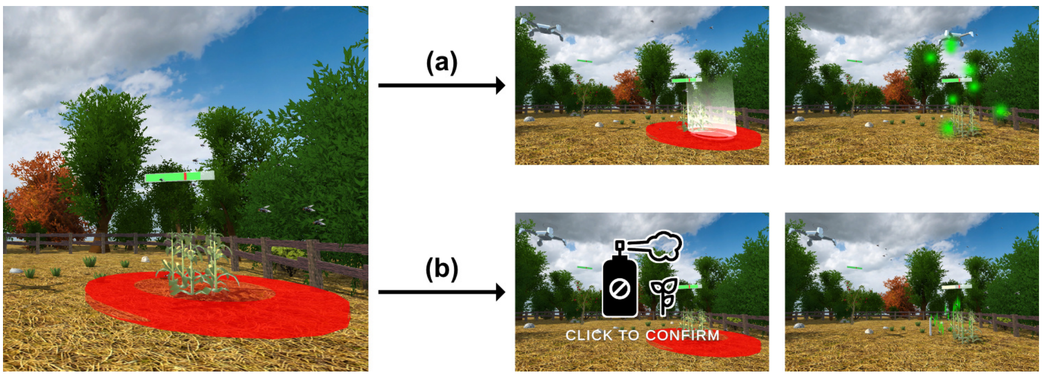


Figure 3.4 During an attack, participants can either send a drone or use emergency pesticides

For each of the *Ambiguity* situations, an association of specific auditory and haptic stimuli was designed. Therefore, in this experiment, we tested whether the presence or absence of these targeted stimuli (the *Modulation*) effectively affect the user's experience both in timing and user experience.

Safe trials were *modulated* using *audio-tactile rhythmic click trains* (through a haptic vest and speakers of the VR HMD) at a high BPM (300) in order to combat expected *Boredom* effects (see the Paragraph on *Rhythm* of Section 2.3.3). The

PU trials used *Modulation* through *soothing nature sounds* (from the *International Affective Digital Sounds Extended* database [111]) as according to the effects of nature environments on both *reducing stress* and *expanding the time experience* (see Sections 2.2.2,2.3.3). Finally, *TU* trials made use of random haptic patterns as a way to create *novelty-inducing tactile feedback*, due to the effects on increased information processing and emotional regulation of such stimuli [112], we expect this *Modulation* to redirect the participants to the present moment.

The timing experience was assessed through timing production tasks at the beginning and end of each trial (see Fig. 3.5), while different aspects of the user experience were assessed through an in-environment post-trial questionnaire and with wearable sensors.

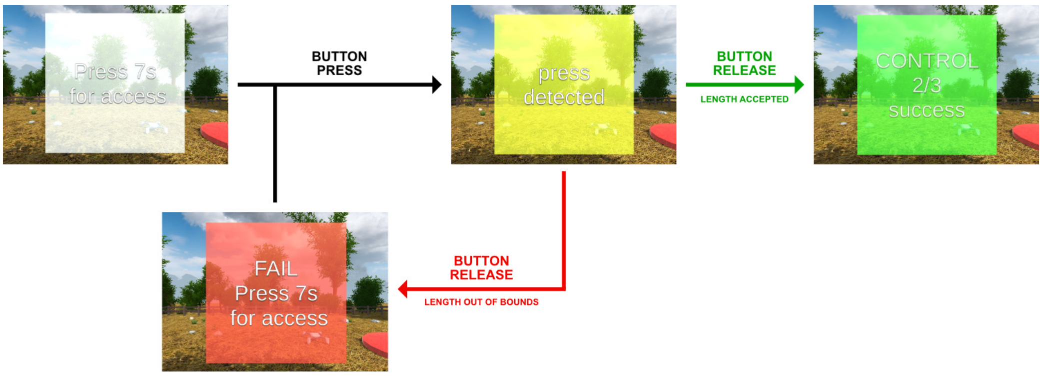


Figure 3.5 Timing task sequence and prompts

Methods

The participants went through trials with varying *Ambiguity* denoted as *Safe*, *PU*, or *TU*. All trials had a sequence using the following structure (see Fig. 3.6): an *Initial Wait* ranging between 30 to 60 seconds, a *Time Production (Pre Attack)* involving 3 timing tasks which give access to drone control, a *Drone Placement* time of 60 seconds (randomized between 30 to 60 for *TU*), an *Attack* phase, a *Time Production (Post Attack)* involving 3 timing tasks which gives access to the post-trial *Questionnaire*, and finally the post-trial *Questionnaire*.

In the case of the *Safe* trials, and the *PU* trials with no pests attacking the plants; a *Pseudo-Attack* phase was still present in the form of a waiting time ranging between 10 to 15 seconds.

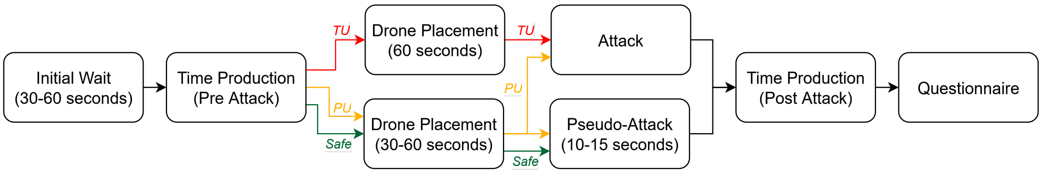


Figure 3.6 Trial sequences with branching paths depending on the *Ambiguity*

Another main parameter of this experiment was that half of the trials had stimuli, or *Modulation*, which were introduced during the *Attack* phase (or during the *Pseudo-Attack* phase for trials with no pests attacking the plants).

Therefore, 28 participants participated in 12 trials (2 repetitions per condition, 4 minutes average per trial with breaks allowed between trials) that are combinations of the following parameters:

- **Ambiguity:** Safe, PU, TU
- **Modulation:** YES, NO

Some variables of the data (timing tasks and physiological data) were repeatedly measured at different moments within a trial, which we categorize as these 3 *Phases*:

- **Pre Attack:** during the *Time Production (Pre Attack)* for timing tasks; continuous measures from the *Initial Wait* until the end of the *Drone Placement* for physiological data
- **Attack:** continuous measures during either the *Attack* or *Pseudo-Attack* for physiological data
- **Post Attack:** during the *Time Production (Post Attack)* for both timing tasks; continuous measures during the *Time Production (Post Attack)* for physiological data

As in this section, the results of this experiment are separated into titled paragraphs; variables of the study will be introduced in their associated paragraph.

Results: questionnaire

Participants had to answer a questionnaire presented as a post-work report after each trial. In this analysis, we focus on the answers to the following questions, each giving a per-trial variable in the form of a 7-point Liker scale:

- **Difficulty:** Was the task easy? (0 = hard, 7 = easy)

- **Relaxation:** Did you feel relaxed during the task? (0 = tense, 7 = relaxed)
- **Boredom:** Was the task engaging? (0 = boring, 7 = engaging)

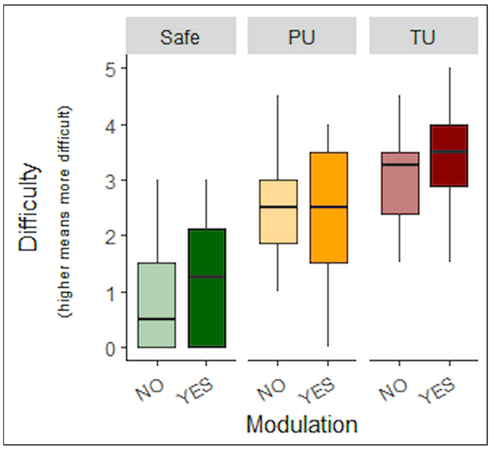
A visual representation of the spread of the answers can be seen in Fig. 3.7. For each questionnaire variable, an ANOVA was performed on *Ambiguity*, *Modulation* and *Ambiguity-Modulation* classes; these ANOVAs were then followed up with post-hoc pairwise comparisons with Bonferroni corrections in the case of a significant ANOVA p-value.

Regarding *Difficulty*, a main effect of *Ambiguity* ($F = 38.31, p < 0.001$) was observed; post-hoc analysis indicates that trials in the *Safe* condition were judged as easier than both *PU* ($t = 8.56, p < 0.001$) or TU ($t = 10.7, p < 0.001$), and that the *PU* condition was judged as easier than *TU*($t = 3.81, p = 0.001$). However, no interactions between *Difficulty* and either *Modulation* ($F = 2.58, p = 0.11$) or *Ambiguity-Modulation* ($F = 1.38, p = 0.25$) could be observed. These results on the *Difficulty* were expected in our design since participants had nothing to do in the *Safe* condition, and *PU* conditions involve an attack in only half the trials whereas they always happen in the *TU* trials.

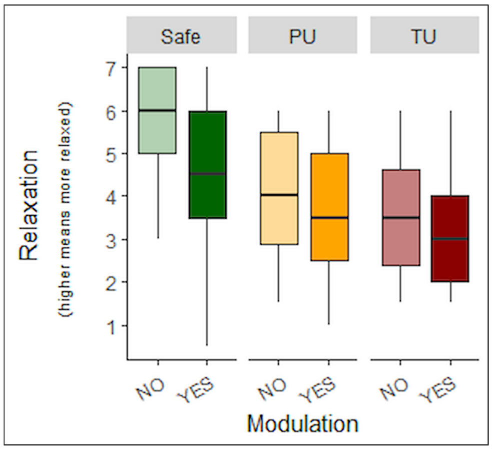
For *Relaxation*, we observed significant effects of *Ambiguity* ($F = 21.39$, $p < 0.001$), *Modulation* ($F = 10.4$, $p = 0.0013$), and *Ambiguity-Modulation* ($F = 4.51, p = 0.011$). Post-hoc analysis on *Ambiguity* indicates that *Safe* trials were perceived as more relaxing than both *PU* ($t = 5.82$, $p < 0.001$) and *TU* ($t = 7.8$, $p < 0.001$), and that *PU* trials were perceived as more relaxing than *TU* ($t = 3.47$, $p = 0.001$). Post-hoc analysis on *Modulation* reveals that trials under *Modulation-NO* were judged as more relaxing than those under *Modulation-YES* ($t = 3.79, p < 0.001$). These results hint that our design on the *Ambiguity* manages to induce varying levels of stress from the experiment design; however, contrary to the goals of this experiment, *Modulation* in trials did not lead to a more relaxing experience.

Regarding *Boredom*, we observed significant effects of *Ambiguity* ($F = 38.89$, $p < 0.001$) and *Modulation* ($F = 6.88$, $p = 0.009$) but not of *Ambiguity-Modulation* ($F = 2.98, p = 0.051$). Post-hoc analysis on *Ambiguity* indicates that *Safe* trials were perceived as less interesting than both *PU* ($t(= 9.58, p < 0.001$) and *TU* ($t = 10.9$, $p < 0.001$), and that *PU* trials were perceived as less interesting than *TU* ($t = 5.15, p < 0.001$). Post-hoc analysis on *Modulation* reveals that trials under *Modulation-NO* were judged as less relaxing than those under *Modulation-YES* ($t = 3.12, p = 0.002$). These results on the *Ambiguity* indicate that the design of the experiment also managed to create varying levels of interest; the results on *Modulation* show a positive global effect of introducing stimuli into the environment; however, with the absence of significant differences between *Ambiguity-Modulation* conditions, it might be due to a global novelty effect from the introduction of a stimulus rather than the intrinsic properties of the stimulus chosen for each condition.

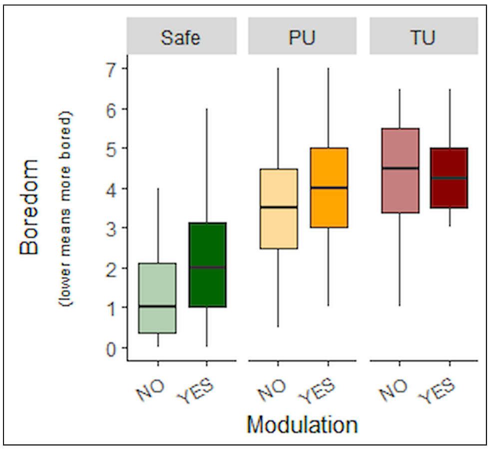
In general, these results hint towards an alignment between our goals in the design of our trials on the *Ambiguity*, as well as interaction effects between *Modulation* and both *Relaxation* and *Boredom*. However, while globally the effect of the *Modulation* aligns with the goals of the *ChronoPilot* project on reducing *Boredom*, *Modulation* goes against these goals with a negative effect on perceived *Relaxation*.



(a) *Difficulty*



(b) *Relaxation*



(c) *Boredom*

Figure 3.7 Responses per *Ambiguity* and *Modulation* conditions for post-trial questions on *Difficulty*, *Relaxation*, and *Boredom* (illustrations courtesy of E.Balta)

Results: physiological

Participants were equipped with an Emotibit wearable device at all times during the experiment. More specifically, the 3-wavelength *PhotoPlethysmoGram* (PPG) sensor was used to monitor changes in blood volume at a sampling rate of 25 Hertz, while the *ElectroDermal Activity* (EDA) sensor was used to monitor skin conductance at a sampling rate of 15 Hertz. These values allow us to derive the following variables:

- ***Skin Conductance Response Rate* (SCR Rate):** amount of changes in the phasic component of the electrodermal conductance signal per second, derived from EDA data

- ***Heart Rate* (HR):** speed in *Beats Per Minute* (BPM) of the cardiac activity, derived from PPG data

In the case of the *SCR Rate*, we could get a measure for each of the *Phases (Pre Attack, Attack, Post Attack)* of the trials (see Fig. 3.8).

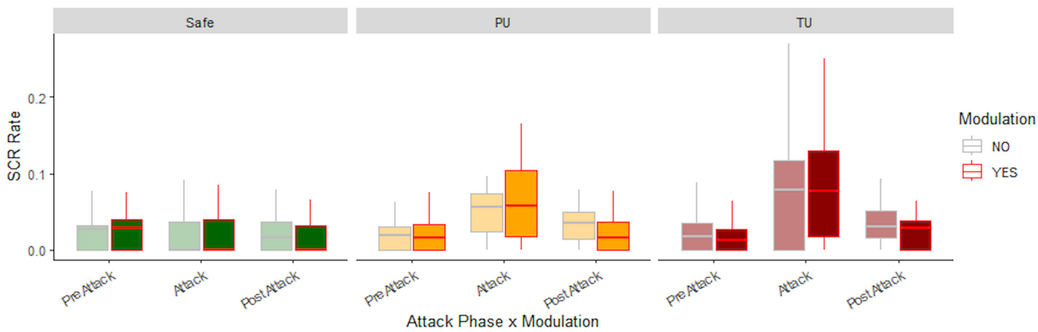


Figure 3.8 *SCR Rate* during the different experimental *Phases (Pre Attack, Attack, Post Attack)*, across all *(Ambiguity-Modulation)* conditions (illustration courtesy of E.Balta)

Using Friedman and Conover follow-up tests, it was found that the *SCR Rate* was different between the *Pre Attack* and *Attack* both in *PU* and *TU* trials, regardless of the presence of the *Modulation* (see Table 3.1).

	$\chi^2(2)$	**Friedman p-value**	**Conover p-value**
PU-NO	8.49	0.0143	0.012
PU-YES	11.9	0.0027	0.004
TU-NO	6.4	0.041	0.005
TU-YES	27.0	<0.001	<0.001

Table 3.1 Friedman tests with significant p-values on *Ambiguity-Modulation* conditions and p-values of the followup Conover test between *Pre Attack* and *Attack*

However, the follow-up Conover test yielded significative differences between the *Attack* and *Post Attack* only in the *Ambiguity-Modulation* situations involving the presence of *Modulation-YES* in both *PU* trials (*PU-YES*, *Attack-Post Attack*: $p = 0.005$) and *TU* trials (*TU-YES*, *Attack-Post Attack*: $p < 0.001$).

These results on the *SCR Rate* suggest that the *Attack* in both *PU* and *TU* situations is able to raise the *SCR Rate* (and thus, the stress) regardless of the *Modulation* compared to the *Pre Attack*. The specific differences observed only during the presence of *Modulation* between the *Attack* and *Post Attack* suggest that the introduction of stimuli during the *Attack* either (or both) raised the stress while the attack was happening, or allowed stress to drop at a faster rate during the *Post Attack*.

Regarding *HR*, we could obtain a single measure for each of the trials (instead one per *Phase* of each trial); this is due to the way PPG data must be derived from multiple measurements. No effects were observed in the ANOVAs between the *HR* (BPM) and both *Ambiguity* alone or *Ambiguity-Modulation* combined factors. Although a significant difference can be observed in the *t-test* between global *Modulation* conditions ($t = 2.21$, $p = 0.03$), the *HR* values are within 1 BPM of difference, which is not conclusive. These results on *HR* are likely due to an incompatibility of the experiment with the measurement, which requires minimal movements and longer exposure to the sensor to detect changes, while our trial design generates short-lived changes depending on the *Attack Phase*.

Results: timing tasks

In each trial, the participants had to perform a *Time Production* block of 3 randomized production timing tasks (with targets of 3, 5, and 7 seconds) in 2 *Phases* of the experiment: the *Pre Attack* and the *Post Attack*. Therefore, we can assess the following variables for each involved *Phase* and for each trial:

- **Task Accuracy(i):** with i as the position of the task within the block (1, 2, or 3), the $\frac{production}{target}$ ratio of a singular timing task during a *Time Production* block

- **Production Accuracy:** the average of *TaskAccuracy* within a *Time Production* block

A visualization of the *Production Accuracy* across all *Phase-Ambiguity-Modulation* conditions can be seen in Fig. 3.9.

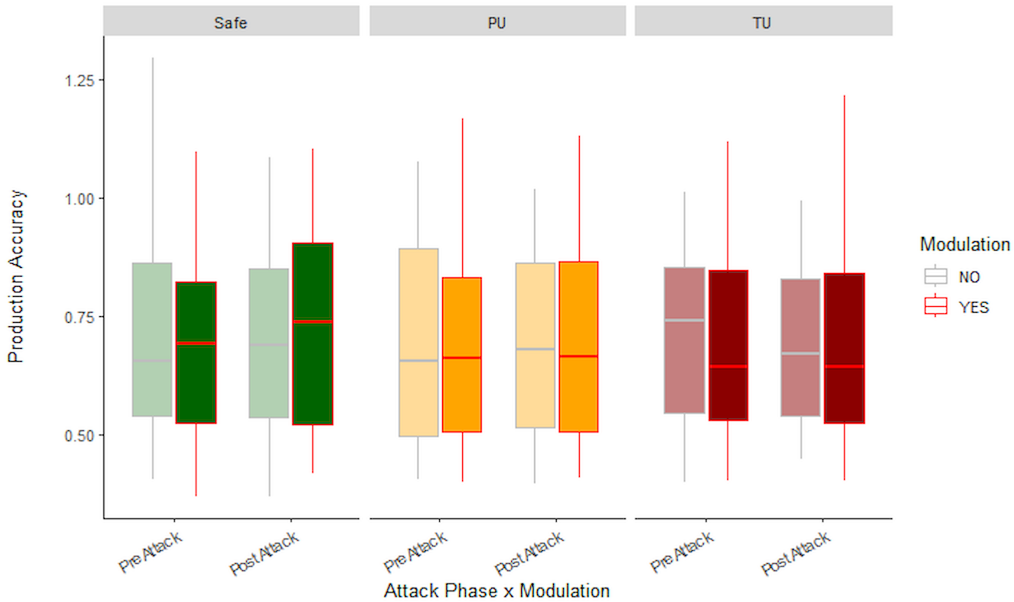


Figure 3.9 *Production Accuracy* per *Phase-Ambiguity-Modulation* conditions (illustration courtesy of E.Balta)

A first $3 \times 2 \times 2$ repeated measures ANOVA (*Ambiguity*, *Modulation*, *Phase*) on the *Time Production* revealed no significant effect of any of the factors. However, the effect of time modulators (either voluntary with the presence of *Modulation* or inherent to the task with the *Attack Phase*) might be short-lived after their end, therefore we repeated the ANOVA, but using *Task Accuracy(1)* instead of *Time Production* in the case of the *Post Attack Phase* (see Fig 3.10). In this case, the ANOVA reveals a significant main effect of *Phase* on *Production Accuracy* ($F = 11.141$, $p = 0.003$, $\eta_p^2 = 0.3$); post-hoc tests are suggesting that the effect of *Phase* was significant when *Modulation* is present for both *PU* trials ($F = 7.23$, $p = 0.012$, $\eta_p^2 = 0.217$) and *TU* trials ($F = 5.43$, $p = 0.028$, $\eta_p^2 = 0.173$).

Although limited due to the need to rely on *Task Accuracy(1)* instead of *Time Production* in the case of the *Post Attack*, these findings suggest that the presence of *Modulation* led to stronger under-productions (over-estimation) for both *TU* and *PU* immediately after the attack; which could indicate a successful time expansion during the seemingly stressful *Attack Phase*, which aligns to the objectives of this experiment.

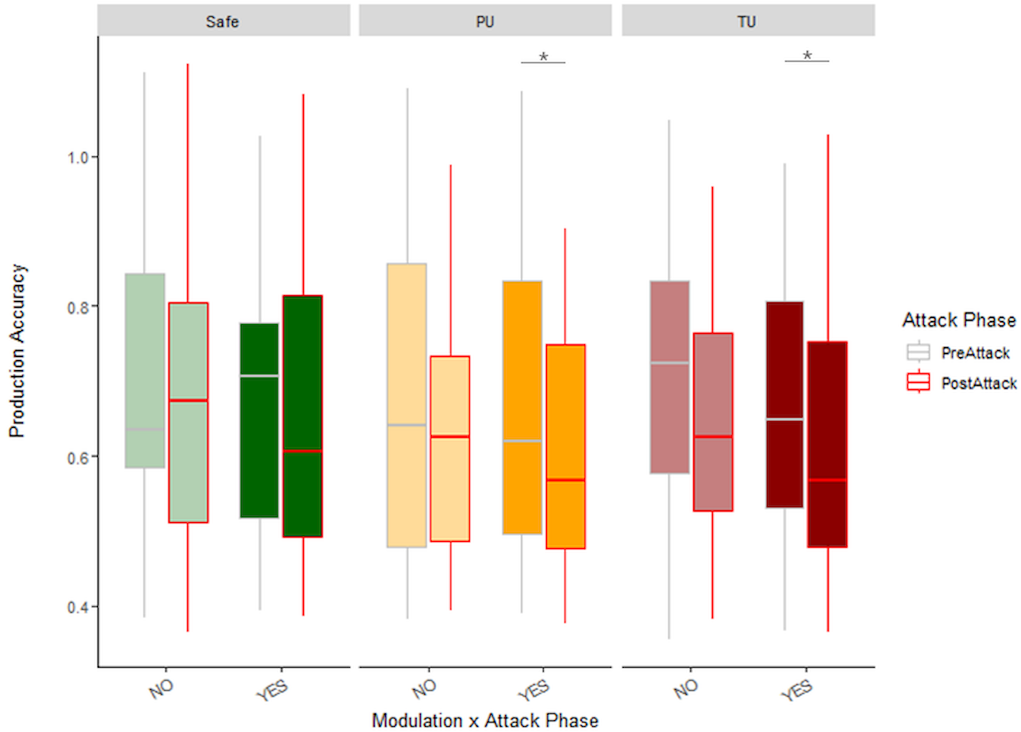


Figure 3.10 *Production Accuracy* (when *Phase = Pre Attack*) or *Task Accuracy(1)* (when *Phase = Post Attack*) per *Phase-Ambiguity-Modulation* conditions (illustration courtesy of E.Balta)

Outlook and perspectives

In this study, we investigated how the user experience can be affected through the introduction of scenario-specific time *Modulation* under different situations of *Ambiguity* (*Safe*, *PU*, and *TU*). We assumed that the introduced *Ambiguity* would lead to different situations of stress and perceived difficulty, which was supported by the analysis on *Questionnaire* data and physiological (*SCR Rate*) data. However, it remains unclear whether the effect of the *PU* situations differ from *TU* and *Safe* situations not because of its uncertainty in the apparition of an attack, but rather due to the frequency of attack apparitions across all *PU* trials.

For each situation of *Ambiguity*, trials were differentiated by the introduction (or absence) of stimuli (*Modulation*) during the *Attack Phase*. As a first attempt in the implementation of *ChronoPilot design principles*, a specific set of stimuli was designed for each situation of *Ambiguity* in order to counter-act, through time modulation, the expected decrease in *Relaxation* or increase in *Boredom* induced by the *Ambiguity*.

If we consider the *SCR Rate* data analysis, the addition of *Modulation* in either *PU* or *TU* leads to significant differences between the trial *Phases* of *Attack* and *Post Attack*, indicating either that the introduction of stimuli raised the *SCR Rate* or that the deactivation of stimuli between the *Phases* favored its drop. Meanwhile, the *Questionnaire* data yields a decrease in *Boredom*, but also a decrease in *Relaxation* with the introduction of *Modulation* for any of the *Ambiguity* situations.

These results, when put into perspective of the *ChronoPilot design principles*, contrast with the goals of this experiment as a demonstrator of these principles. The introduction of the different stimuli in each situation of *Ambiguity* was aimed to increase *Relaxation* in the case of *PU* or *TU* and decrease *Boredom* in *Safe* situations, but led to the same effect of decreased *Relaxation* and *Boredom* in all cases of *Ambiguity*. However, the overall decrease in *Boredom*, as well as the *Modulation*-specific drop in *SCR Rate* after the *Attack Phase*, may be indicative of an improved user experience due to the *Modulation*.

The use of *Modulation* also led to differences between the timing tasks of the participant between the *Phases* of *Pre Attack* and *Post Attack*, but only in the *PU* and *TU* trials. Although we can assume a time modulation effect from the stimuli introduced in *PU* and *TU* trials, the effect of *Modulation* in the *Questionnaire* data was also observed in the *Safe* trials. Therefore, while *Boredom* and *Relaxation* are affected by stimuli that are also able to modulate time perception in *PU* and *TU* situations, *Boredom* and *Relaxation* are also affected by a *Modulation* with no apparent effect on timing experience in the *Safe* situations.

These findings on *Boredom* and *Relaxation* could be explained by a novelty effect experienced by the participants due to the addition of any stimulus compared to the non-modulated situations, regardless of the intrinsic properties specific to the stimulus; some specific stimuli just happened to be time modulators (in *PU* and *TU*, but not in *Safe*). Because the study had specific sets of stimuli for each *Ambiguity* rather than a rotation of all possible stimuli, we cannot assess any stimulus-specific effect within *Ambiguity* situations nor causal effect between *time* modulation (as opposed to *any* modulation) and user experience.

Although this study serves as a demonstrator of the *ChronoPilot design principles* in improving user experience through the introduction of stimuli, by reducing the *Boredom* of the participants through *Modulation* in various situations of stress induced through trial *Ambiguity*, limitations of the study design prevent potential observations of the effect of each stimulus individually, as well as any causal effect between modulation of time perception and improved user experience.

Future studies could benefit from investigating the use of stimuli with opposite expected time modulation effects on *Boredom* and *Relaxation* within the same *Ambiguity* situations.

3.3.3 Bulb Experiments: Timing Studies Compared Between VR, AR and Real-Life

The following unpublished results explore the effect of oddball series and color properties as time modulators across VR (using a VIVE Pro Eye), AR (using a Microsoft Hololens 2) and *Real-Life* (RL) settings. Due to the unpublished nature of these results, the following section will be divided into distinct titled paragraphs.

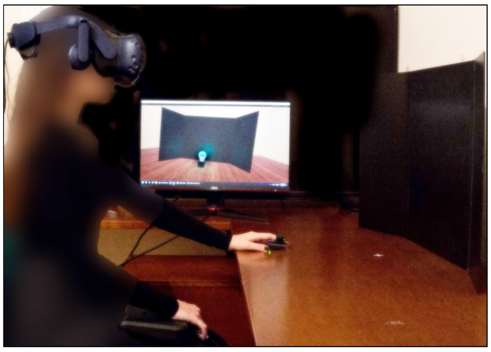
These results come from joint experiments conducted with *Panteion University*, while I designed the bulb settings and built the environments, the design and execution of the experiments were carried out by our colleagues in Greece. The results presented in this thesis focus only on the implications on time estimation across the different settings; more detailed use of the results and analysis of physiological data are planned for later articles.

Rationale and environment design

In Section 2.5.2 it was observed that the use of VR technology can induce bias in time perception, such as the same game leading to underestimation in VR compared to a flat screen [82] or a waiting room that leads to greater overestimation in VR [83]. Another observation is that when emotional content is involved, no significant differences emerge in VR time estimation when emotional content remains the same [86].

Stimuli such as the *oddball effect* (see the Paragraph on *Attention* in Section 2.3.1) or the *display of colors* (see the Paragraph on *Colors* in Section 2.3.3) are well known in the timing literature. However, since these stimuli are visually represented, it could be argued that the use of XR technologies such as VR or AR head-mounted displays may alter the perception of these stimuli. With the following experiments, our objective is to evaluate whether stimuli related to *oddball sequences* and *color properties* show similar effects on time perception in these different realities.

To this end, we designed an environment centered around a lightbulb; the lightbulb can be controlled to change its hue, contrast, value (or light intensity), and be switched on or off. This environment also has the advantage of being easy to implement on the three different settings we are studying, in AR and VR the lightbulb is a simple controllable 3D virtual object. For the RL setting, we are using a custom-built lightbulb controlled by an *Arduino* board, to which serial instructions can be sent from a computer. In all three settings, the participant interacts with the bulb using a three buttons device (*Elgato Stream Deck Pedal*) placed on their lap. The participants were also equipped with physiological sensors (*Emotibit*), however the physiological data are outside of the scope of the results presented in this thesis.



(a) VR

(b) Real-Life

Figure 3.11 The lightbulb environment in VR and Real-Life (pictures courtesy of E.Balta)

Two experiments were carried out in this environment: the *oddball* experiment, the *color effect* experiment.

Oddball experiment

In the case of the oddball experiment, the main goal was to compare the effect of oddball series on time perception in the three settings (AR, VR, RL). Oddball series imply that a series of the same stimulus (the standard) is shown to the participant, with one (the oddball) being different. An example study by Ongchoco, Wong, and Scholl suggests that in oddball series, the effect on timing may come from the uncertainty created by the sudden appearance of the oddball [113].

By nature, the use of oddball series is compatible with time-bisection tasks (see Section 2.4.1) as the standard stimuli can also be shown with a standard duration, while the oddball can be displayed with a varying duration. The oddball series depends on the *Contrast* of the bulb, with the standard being *Low* or *High* and the oddball being the opposite (see Fig. 3.12).

Considering that a confounding effect may arise from the position of the oddball in a series [114], the oddballs were positioned in the second half and before the last stimuli of the series. During a trial, the participant would then be shown a series of 10 standard stimuli of the same contrast with one oddball replacing the standard between the 5th and 8th stimuli; after the series, the participant must indicate by pressing a button if they think the oddball lasted *longer* or *shorter* than the standard.

Each of the 33 participants went through two consecutive sessions of 48 trials in each setting (8 repetitions per possible trial, 288 trials total per participant), through an entire session the standard stimulus would always be either *high* or *low* contrast. We thus have the following parameters of the experiment, with

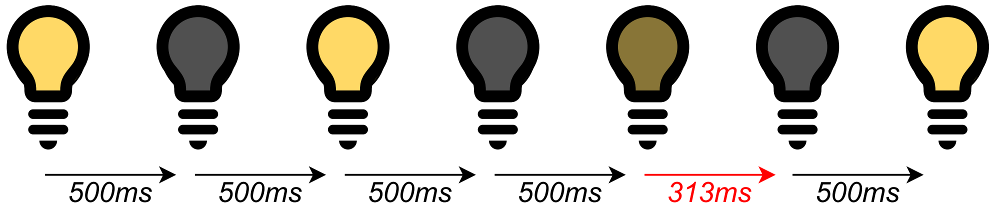


Figure 3.12 An oddball sequence using the lightbulb (contrast)

trials being combinations of all these parameters:

- **Setting:** VR, AR, Real-Life (RL)
- **Standard Contrast:** Low, High
- **Standard Duration (in milliseconds):** 500
- **Oddball Durations (in milliseconds):** 313, 412, 445, 555, 588, 687

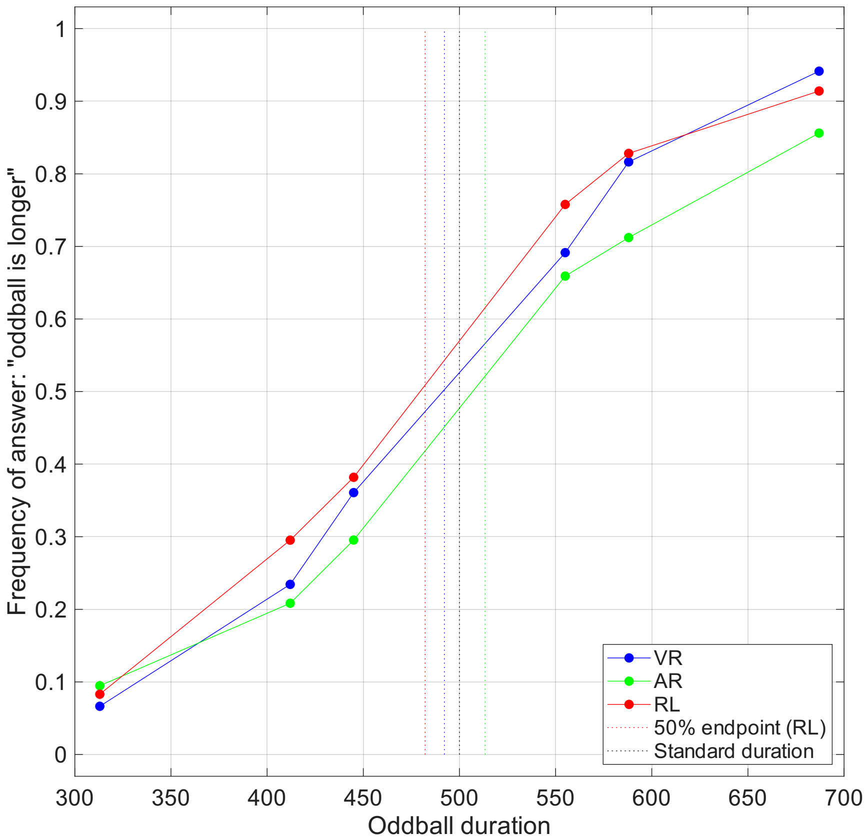
The only variable we get per trial is the *Answer* of the participant (*"oddball is shorter"* or *"oddball is longer"*). This design allows us to compute a *psychometric function*, and more specifically the *50% endpoint* which corresponds to the estimated oddball duration that would lead the participant to answer either *shorter* or *longer* 50% of the time. In the following analysis, we compute the *Spearman-Kärber Endpoint (SKE)* using the method and Matlab script provided in the *Timing and time perception: Procedures, measures, & applications* book [72].

	VR	**AR**	**RL**
Low Contrast	492.14	513.38	482.16
High Contrast	477.18	445.15	454.85

Table 3.2 Global SKEs per *Setting* and *Contrast*

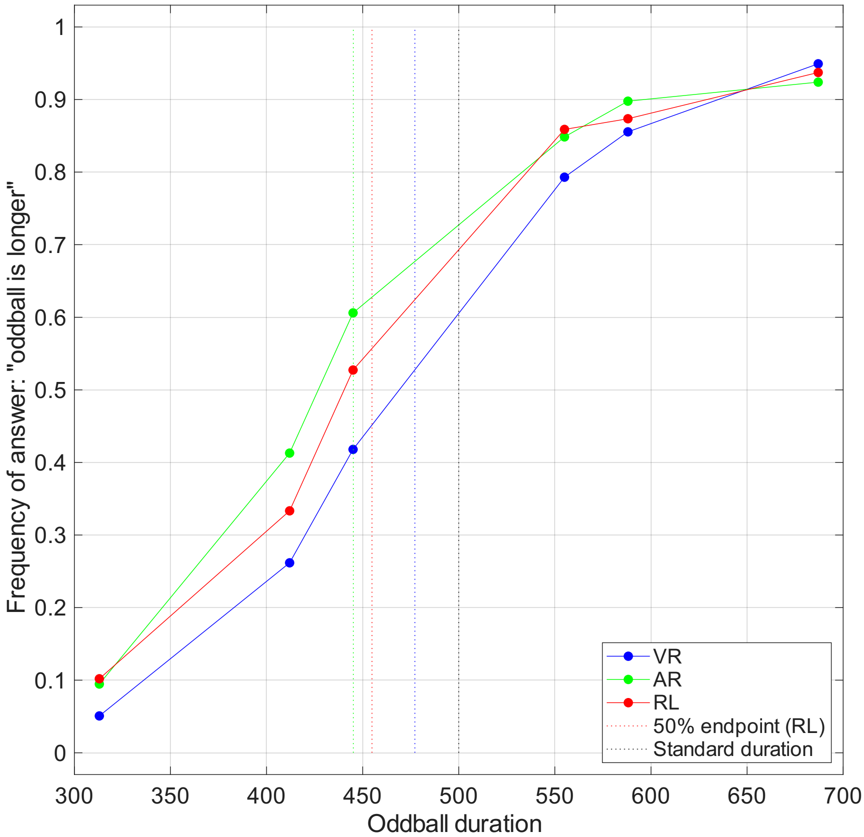
First, we compute the per-setting *global SKEs* for both *Contrasts (Low, High)*, the exact values can be found in Table 3.2. Except for the *AR-Low contrast* combination, all *global SKEs* are below the standard duration, indicating a tendency of the participants to overestimate the duration of the oddball which is the expected effect. However, since the SKEs are all between the *Oddball Durations* 445-555 which are the closest durations to the *Standard Duration*, the effect of the oddball on overestimation is relatively weak.

Fig. 3.13 represents the frequency of the answer *"oddball is longer"* by the participants for each possible *Oddball duration*; the sub-figures are separated according to the *Contrast*. In both sub-figures we can observe that all 3 settings follow a similar shape, however the *AR* frequencies change its vertical position relatively to *VR* and *RL* depending on the *Contrast* condition. These observations, coupled with the *global SKE* exception of the *AR-Low Contrast* condition being above the *Standard Duration*, could indicate that the perception of the oddball stimuli are inconsistent in *AR* relatively to *RL* and *VR*.



(a) Low Contrast

Figure 3.13 Answer frequency and global SKEs per *Setting*



(b) High Contrast

Figure 3.13 Answer frequency and global SKEs per *Setting*

To further investigate this inconsistency of *AR*, we can use a *per-session SKE*. As described earlier, the participants performed the tests within sessions of 48 trials of a *Setting-Contrast* combination. We can thus extract a *per-session SKE* for each participant (1 value per *Setting-Contrast-participant* combination, 6 values total per participant); the Fig. 3.14 represents the extracted values for each *Setting-Contrast*. Of the 33 participants, 3 did not finish the 6 sessions, and thus have been removed from the *per-session SKE* analysis.

The ANOVA between the *per-session SKEs* and the *Setting-Contrast* yield significant differences ($F = 7.94$, $p < 0.001$). Additional analysis using Tukey HSD (see Table 3.3) shows significant ($p < 0.05$) differences between all *AR* and *RL* conditions, at least a low significant ($p < 0.1$) difference between all *AR* and *VR* conditions, as well as a strong significant ($p < 0.001$) difference between *AR High* and *AR Low*.

These results reinforce the idea that the *AR* setting is actively altering the perception of the durations of the oddball stimuli. Although we can also observe

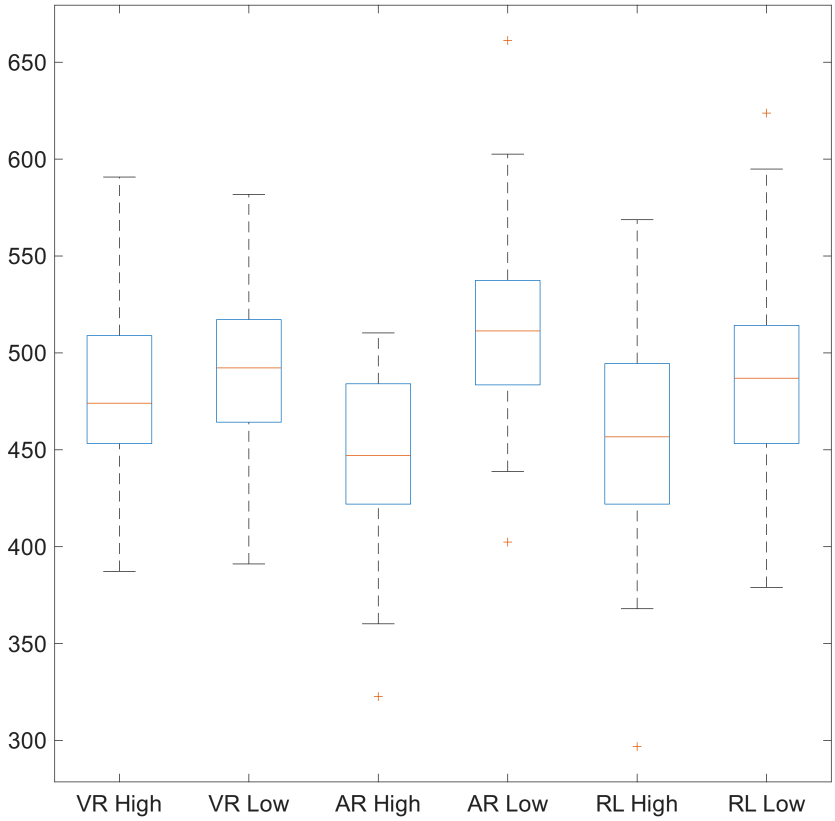


Figure 3.14 *Per-session SKEs* under each *Setting-Contrast*

a low significant ($p = 0.07$) difference between *VR Low* and *RL High*, this is not enough to suggest inconsistent timing perception between *VR* and *RL* as we cannot observe differences in symmetric *Contrast* conditions.

Color effect experiment

The color experiment had as a main goal the evaluation of the time perception bias induced by the display of different colors varying through *hue*, *saturation*, *value*, and how consistent it is between each setting.

In order to avoid the creation of an oddball series using a standard duration and stimulus, this experiment cannot use a time comparison task. As we are interested in the differences in perception from the color display, in this experiment a *time production* task (see Section 2.4.1) is used. During a trial, participants will face a turned off lightbulb and then hear a duration in seconds, they will then attempt to produce the target duration by pressing and holding a button. While pressing the button, the lightbulb will turn on with a specific color and then turn off after the button is released (see Fig. 3.15).

	diff	lwr	upr	p adj
VR High-VR Low	-10.77	-46.59	25.04	0.96
VR High-AR High	35.04	-0.77	70.85	0.06
VR High-AR Low	-34.88	-70.69	0.93	0.06
VR High-RL High	23.43	-12.38	59.24	0.42
VR High-RL Low	-5.71	-41.53	30.10	0.99
VR Low-AR High	45.81	10.00	81.62	0.004
VR Low-AR Low	-24.11	-59.92	11.70	0.39
VR Low-RL High	34.20	-1.61	70.01	0.07
VR Low-RL Low	5.06	-30.75	40.87	0.99
AR High-AR Low	-69.92	-105.73	-34.11	<0.001
AR High-RL High	-11.61	-47.42	24.20	0.94
AR High-RL Low	-40.75	-76.57	-4.94	0.02
AR Low-RL High	58.31	22.50	94.12	<0.001
AR Low-RL Low	29.17	-6.64	64.98	0.19
RL High-RL Low	-29.14	-64.95	6.67	0.19

Table 3.3 ANOVA's Tukey HSD between *Setting-Contrast* and SKE

A secondary task is also in place to prevent the participant from counting: during production, they will hear a repeated beeping sound, which has a random chance to be off-tone; in the case of an off-tone beep, the beeping will stop until they press a dedicated button on the controller.

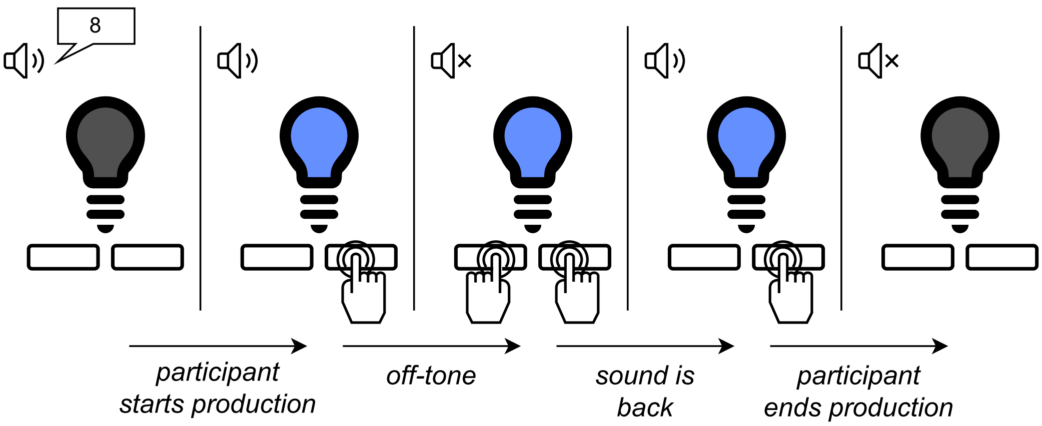


Figure 3.15 Example of a trial in the color experiment

Here colors are considered in the *HSV* (Hue, Saturation, Value) color space; specifically, a color is a combination of these 3 values. In this experiment, we consider 3 combinations of *Hue-Saturation* which are *Blue* (H=230,S=100%), *Red* (H=350,S=100%), *White* (H=60,S=6%), as well as 2 possible *Values* of either 100% (*Bright*) or 40% (*Dim*).

Each of the 20 participants went through one session of 90 trials per setting (15 repetitions per possible trial, 270 trials total per participant), here are the parameters of the experiment, with trials being combinations of all these parameters:

- **Setting:** VR, AR, Real-Life (RL)
- **Hue-Saturation:** Blue, Red, White
- **Value:** Bright, Dim
- **Production Target (in seconds):** 4, 6, 8

The direct variables we get per trial are:

- **Time Produced:** the time in milliseconds produced by the participant
- **Reaction Time:** the average time in milliseconds the participant took to report an off tone in the secondary task
- **Missed Delay:** if the participant had to validate an off tone when they finished the timing task, how much time they had to validate it

As we want to ensure that the participants were active in the secondary task to avoid counting strategies, we also recorded the *Reaction Times* on that task and the *Missed Delay*. With a mean *Reaction Time* of 852 milliseconds across all trials, we will filter out any trial with more than double that amount in *Missed Delay* as it could indicate that the participant was not actively performing the secondary task.

With the output data of the trials, we want to extract the participants' performance in the timing task; however, the interpretation of what is a second may vary per participant, requiring a normalization process.

Taking into account a participant p, all trials of p being $trials(p)$, the amount of trials of p being $count(trials(p))$, the target duration of a trial t being $target(t)$ and the duration produced in the same trial $production(t)$, we can compute the representation of a second $secondBias(p)$ for a participant as:

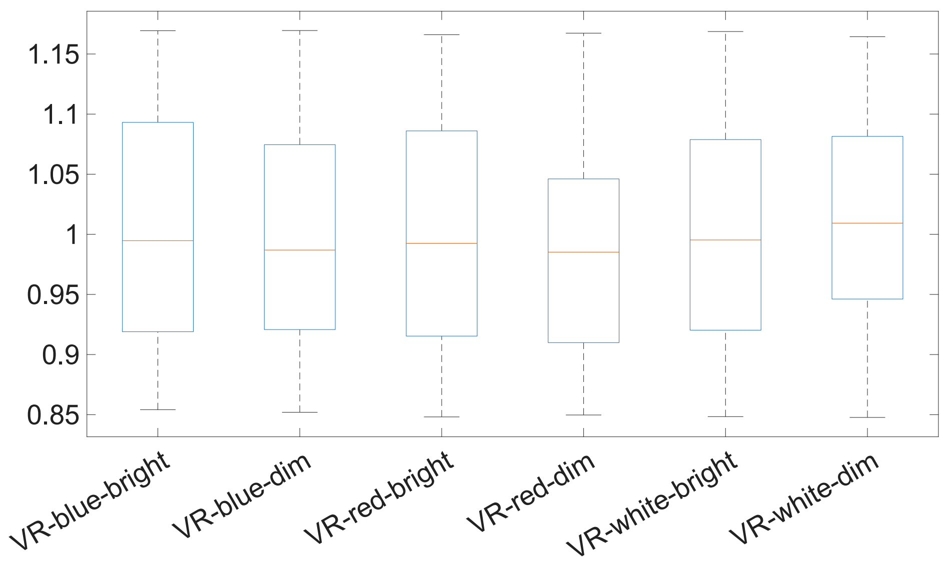
$$secondBias(p) = \frac{\sum_{t \in trials(p)} \frac{production(t)}{target(t)}}{count(trials(p))}$$

Considering the participant of a trial t being $participant(t)$, using $secondBias(p)$ we can compute the $normalizedTimeEstimation(t)$ for each trial as:

$$normalizedTimeEstimation(t) = \frac{production(t)}{target(t) * secondBias(participant(t))}$$

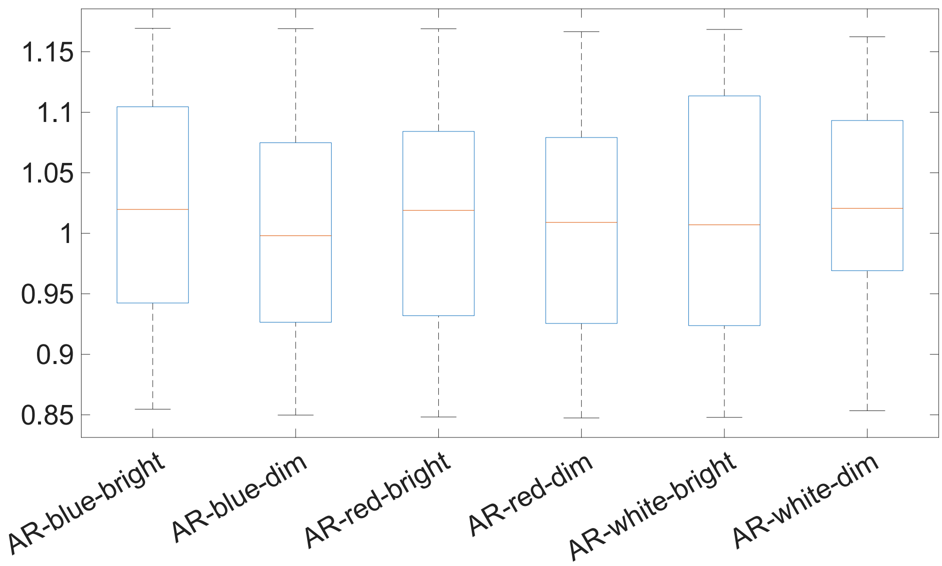
The $normalizedTimeEstimation$ is our main variable in the following analysis, outlier trials were filtered using the interquartile range (IQR); 2615 trials were remaining for the analysis after the IQR filter.

In this analysis, we investigate how the $normalizedTimeEstimation$ could be affected by the *Setting (VR, AR, RL)*, the *Hue-Saturation (Blue, Red, White)* and the color's *Value (Bright, Dim)*; a boxplot representation of the combined factors can be found in Fig. 3.16.

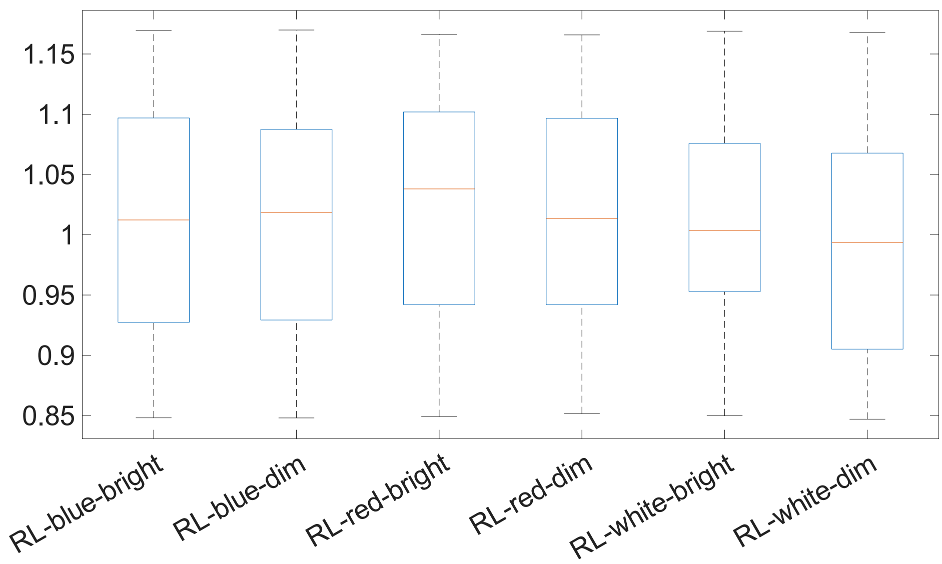


(a) VR

Figure 3.16 Combined factors on $normalizedTimeEstimation$



(b) AR



(c) RL

Figure 3.16 Combined factors on $normalizedTimeEstimation$

The N-way ANOVA between the different factors and *normalizedTimeEstimation* yields a significant p-value for the *Setting* ($F = 5.65$, $p = 0.004$) but not for the *Hue-Saturation* ($F = 0.08$, $p = 0.92$) or the *Value* ($F = 1.91$, $p = 0.17$). Additional analysis on the *Setting* using Tukey HSD (see Table 3.4) shows tendencies ($p < 0.06$) of differences whenever *VR* is involved.

These results seem to hint that color properties changes (*Hue-Saturation,*

	diff	lwr	upr	p adj
AR-RL	0.002	-0.008	0.012	0.90
AR-VR	0.01	0.003	0.024	0.006
RL-VR	0.01	0.002	0.021	0.017

Table 3.4 ANOVA's Tukey HSD between *Setting* and *normalizedTimeEstimation*

Value) of the bulb did not manage to induce time perception changes across all *settings*, however we also observe a general bias induced in the *VR Setting* on time perception.

Since the *Setting* affects the time experience, the ANOVAs of *Hue-Saturation* and *Value* with $normalizedTimeEstimation$ have been repeated after selecting the trials of each *Setting* separately. In the case of the *RL* setting, the ANOVA of *Hue* ($F = 2.92$, $p = 0.054$) reveals a near-significant ($p < 0.06$) effect, but not for the *Value* ($F = 1.35$, $p = 0.25$). Followup Tukey HSD on the *RL Hue* ANOVA reveals a significant difference only between *White-Red*($p = 0.04$). In the case of both the *AR* and *VR* settings, no significant results can be observed from the repeated ANOVAs. These results suggest that the XR settings failed to replicate the effect of the *Hue-Saturation* on the participants' time production task.

Outlook and perspectives

In both bulb experiments, while we do not observe a significant expected effect of the modulators on time estimation, we do observe an inconsistent timing experience from participants across settings. More specifically, the AR setting seems to yield a different time perception with (visual) *oddball sequences* and fails to transcribe the effect of *Hue-Saturation* in the color experiment. Meanwhile, the VR setting only differs within the color experiment.

It is important to remember that the hardware used for this experiment in the AR setting was a Microsoft Hololense 2, which is an AR headset with a *Laser Beam Scanning Display* (LBSD). These types of screen come with several limitations, such as a limited field of view and the inability to display black colors. These screens also require the wear of a tainted glass eye shield, which further affects color perception of the entire environment. Meanwhile, in the VR setting we used a Vive Pro Eye, which uses an *Organic Light-Emitting Diode* (OLED) display, which is known for its color accuracy.

While the design of the experiment attempts to minimize the caveats of LBSD displays (using a black background to counter the lack of possibility to display black colors), we can hypothesize that the perception of the virtual bulb's color might have heavily been affected by the display technology in the oddball experiment. Another hypothesis, while unlikely due to the absence of notable effect in the VR conditions, would be that this difference could come from AR as a paradigm.

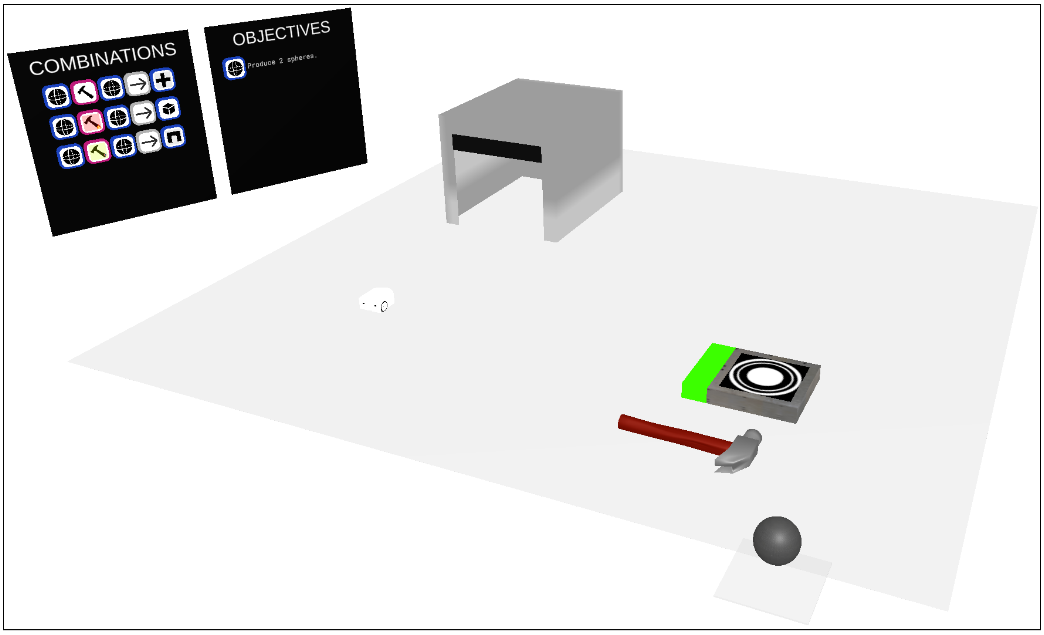
The differences observed in the color effect experiment between VR and the other settings seem to be a general underestimation of time within VR which was not unexpected; however, the absence of effect on time perception from the color in both XR settings is much more concerning. An assumption could be that the amount of luminosity changes in the RL setting, as the physical bulb is the only light source the participant directly looks at, while in both the AR and VR settings, the screen is always emitting light and turning on the virtual bulb would only change the illumination of a fraction of all pixels available on the screen.

Further research between display technologies and time perception might be needed as existing and upcoming XR technologies use a variety of displays (OLED, LBSD, Quantum Dot Display, Liquid-Crystal Display, etc.) and lenses (fresnel, pancake, aspherical, etc.) which may invalidate some time modulators within the same XR paradigm depending on the entire optical stack.

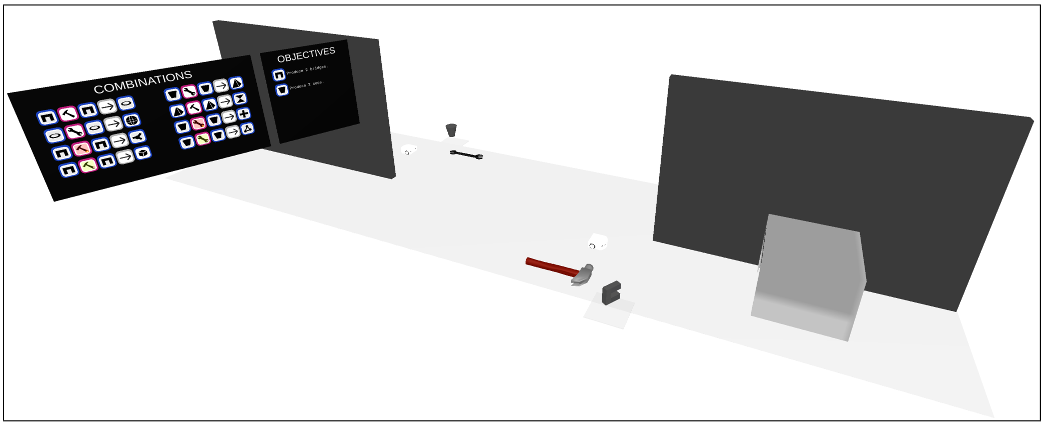
3.3.4 Active Modulation Controller Pilot Experiments

After the successful realization of our XR environment and framework and the cumulation of the ChronoPilot's collaborators' studies in the effect of various stimuli as time modulators, modulation strategies needed to be investigated to approach the ChronoPilot design goal of purposeful modulation. As such, two studies have been designed around opportunities to apply time modulators according to real-time strategies that involve participants' stress, performance, and current time experience. These studies are ongoing works currently managed by collaborators in the ChronoPilot project; therefore, this section is mostly a showcase of current use cases of the *XR MUSE* framework.

The main objective of the ChronoPilot project is to use time modulation techniques to induce flow and reduce stress (see Chapter 1), our collaborators designed a *time modulation controller*; an entity able to trigger stimuli (time modulators) according to both the participants and the state of the environment. More specifically, two modulation strategies have been designed, one for single-user scenarios and the other for multi-user scenarios. The testing of these modulation strategies led to two separate experiments: the *Single-user* experiment, the *Multi-user* experiment, which both required a custom-tailored test environment (see Fig. 3.17).



(a) Single-user



(b) Multi-user

Figure 3.17 Single-user and Multi-user environments

Although we could give details behind the modulation strategies, from a point of view of the implementation of the environment, the modulation strategy of the controller is a black box (see Fig. 3.18).

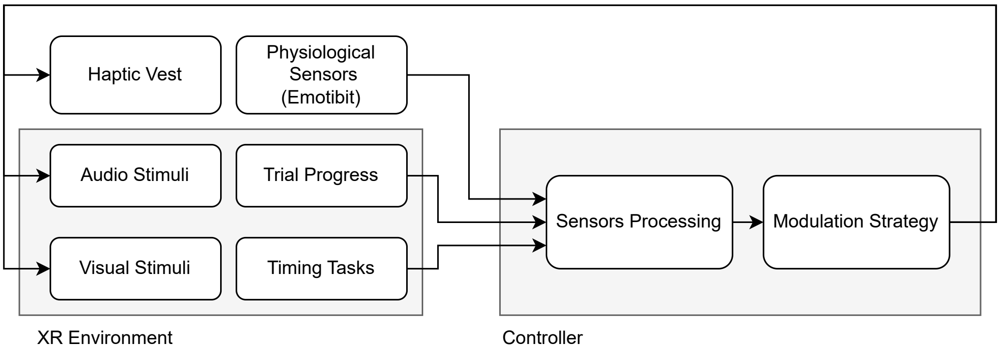
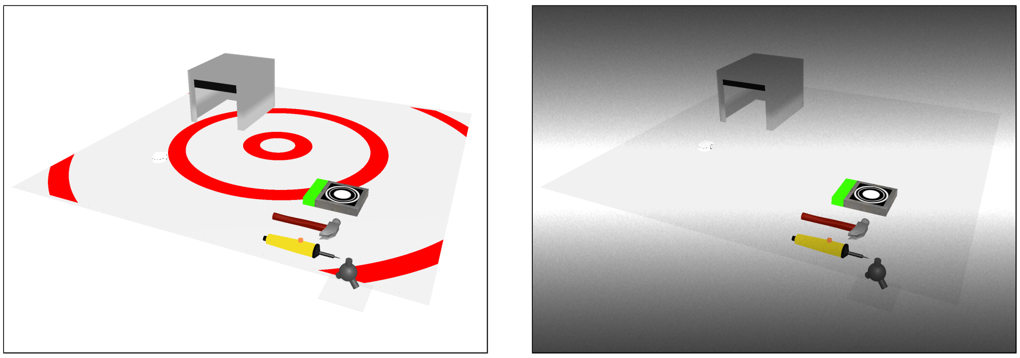


Figure 3.18 Relations between the Controller, the XR Environment, the Haptic Vest, and Physiological Sensors in the Single-User Experiment

These experiments are using the *XR MUSE* framework as a basis (see Section 3.2.2), with new requirements coming from both the controller's need and the piloting needs. For the controller's need, in-environment stimuli have been added in accordance to previous ChronoPilot works, in the form of emotional music, audio clicks, visual rhythmic pulses and screen vignette (see Fig. 3.19).

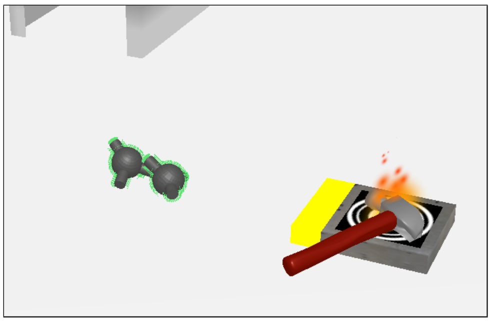


(a) Rhythmic Pulses **(b)** Vignette

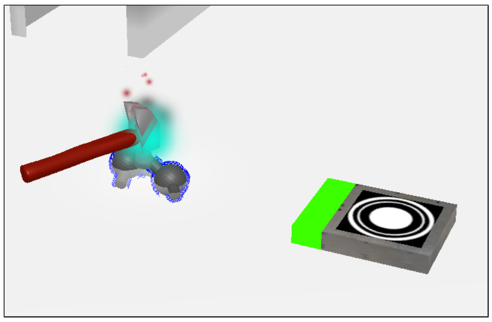
Figure 3.19 Visual stimuli available in the XR environment

As shown in Fig. 3.18, the controller also needed to be able to monitor and control the environment, which led to a custom bi-directional messaging format involving regular reports and accessible stimuli-triggering functions.

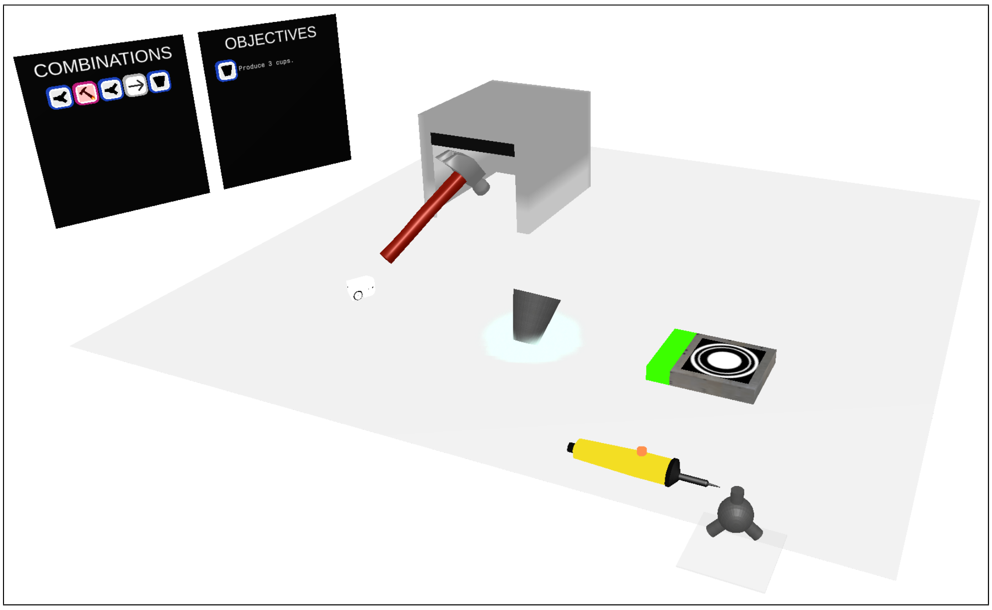
Although the controller uses physiological data to assess stress through a custom-built learning model, precise time experience metrics were needed. As such, new user interactions have been added that hide time experience measurements by embedding them as part of the in-environment tasks. The combination of material (from the *XR MUSE* example scene) can now have an additional *time* *production task* (by holding the tool for 5 seconds during the fusion), or a *time reproduction task* which involves reproducing the duration needed to heat up the tool on a fire plate (see Fig. 3.20).



(a) Heating the tool



(b) Holding the fusion



(c) Release

Figure 3.20 In-environment time reproduction task sequence

My involvement in these studies is mostly the technical implementation and design of the specific per-experiment requirements for the virtual environment.

Chapter 4

Discussion

Although the literature presented in the previous chapter proposes its own conclusions and perspectives, some new conclusions on these results can be drawn, both through linking of the results and through retrospective, toward the objectives (see Chapter 1) of this thesis on the use of ChronoPilot *design principles* in the design of interactive applications.

4.1 Multi-Sensory Time Modulation

Multiple stimuli studies have been conducted as part of the ChronoPilot project. With stimuli ranging from fundamental stimuli, such as our studies on rhythmic stimuli or studies from other members on vibro-tactile stimulation [115], to content-based approaches such as time perception in film viewing experiences [116], as well as task-based modulation with cognitive load [117].

These studies performed by the different members of ChronoPilot, coupled with the already existing approaches detailed in 2.3, give plenty of options to attempt purposeful time modulation techniques in non-timing tasks.

In this thesis, we conducted two experiments on the purposeful use of rhythmic stimuli. First in a browser version (3.1.1), then in a VR setting (3.1.2). The key observations include varying time perception effects of tempo on time perception, as well as influence of the presence or absence of either visual or audio stimuli. These effects on time perception were also, in the case of the longer VR study, depending on the context of task familiarity and fatigue, in addition to a positive effect of audio stimuli on task performance and a negative one of visual stimuli.

More specifically, in the browser version, significant *absolute time estimation errors* were observed when using a single-type stimulus, while *time passage* was significantly affected by the tempo when visual stimuli were present. Meanwhile, the effects observed in the VR version are more fine-grained when it comes to time estimation errors, but yield consistent results with a faster time passage depending on stimuli presence and tempo.

While these results may appear as positive toward the use of rhythmic stimuli for ChronoPilot, since we do have varying interesting effects, it also raises the issue that stimuli effects do not appear as cumulative. One of the main appeals of rhythm was its ability to orchestrate various stimuli (as discussed in the *Rhythm* paragraph of Section 2.3.3); for instance, the browser version shows significant differences in time estimation errors only on single-mode stimuli.

Instead, the combined stimuli gain new properties such as a higher effect on high fatigue situations in the VR study or direction of the time estimation error depending on tempo for the web one.

Moreover, effects were reported through the use of *absolute time estimation error*, while any transformation is compatible with the flow state, this value does not indicate if the transformations are a shortening or a lengthening of the time experience. However, due to the faster time passage reported, we can assume that the participant tends to experience a global shortening effect.

As such, in the context of the ChronoPilot project, we must consider this rhythmic audio-visual technique as its own individual stimulus, which offers various levers to modulate (tempo, audio presence, visual presence) rather than as a synchronization method. We could further hypothesize that any combination of stimuli becomes its own separate technique with potentially new modulating effects, at least if they are synchronized through rhythm.

Another important observation from these studies is how the visual stimulus was going against the user's performance, which supposedly is due to it impeding the user's task, which was object recognition.

With all of this in mind, attempting to apply multi-sensory stimuli with the intent to modulate time perception should be handled with care. The effects of stimuli may not be cumulative, and any new parameter introduced to a stimulus can drastically change its effect on the time experience. Moreover, stimuli should be designed to avoid overlaps with the non-timing aspects of the intended experience.

4.2 Hardware and Design Limitations in Time Perception Studies

The goal of the ChronoPilot project is to integrate time perception modulation as a proxy for flow induction and stress reduction. As this goal is tested within interactive applications, a question could be asked of how important is the choice of the technology delivering said applications. In the context of this thesis, the interactive applications uses both VR and AR.

We have established the possibility of VR being an amplifier of time distortions in the Paragraph *Virtual Reality Paradigm on Time Perception* of Section 2.5.2. Since in this thesis we are using mediated reality technologies and associated virtual environments as a playground for piloting active modulation, another question could arise on whether results we find in our virtual environment would transpose to a real one.

To this end, we conducted two experiments to specifically investigate different aspects of technological immersion. One about how the screen and HMD affect time experience, in the form of the *Bulb Experiments* (see Section 3.3.3; another about software design and interactive aspects of the technology, in the form of the *Some Times Fly* experiment (see Section 3.3.1. We also observe differences across the *Time Perception and Rhythmic Stimuli* experiments (see Section 3.1) that are performed in both a desktop and a VR setting.

Let us start with the implications of the *Bulb Experiments*, where *oddball* (see Paragraph *Attention* of Section 2.3.1) and *color properties* stimuli were being compared between *Real-Life* (RL), AR and VR setups. Observing the results, while VR was biased towards underestimation and lost the effect of *Hue* and *Saturation* on time perception in the *color properties* experiment; the time experience remained consistent in VR compared to RL in the *oddball series* experiment. Meanwhile, in both experiments, the AR setting was inconsistent with the RL setting with respect to time experience.

We made the assumption that these differences with the RL conditions from both VR and AR may come from the color accuracy of the screen, as well as the overall luminosity (or emitting light source) being a main difference with RL conditions. Although these experiments do not necessarily invalidate the use of color properties and oddball series in mediated realities, these results suggest that the representation of these stimuli needs to be mindful of both the limitations (color accuracy) and inherent properties (screen light emission) of the employed technology.

Within the *Some Times Fly* experiment, participants were subjected to different versions of the environment, with varying engagement (through letterboxing and in-environment interactions) and dynamism (through the presence or absence of both immersive sounds and background visual elements). In the results of that experiment, it can be observed that engagement levels will significantly affect time estimations and negatively impact memory task performances, and that the dynamism of the environment seems to amplify these effects from engagement.

As for the *Time Perception and Rhythmic Stimuli* experiments, the transition to a VR environment coupled with associated motion controls to sort objects and more trials per participant led to measurable fatigue in our participants, a fatigue that significantly affects the modulation of time perception.

Now, if we keep these insights from those three experiments in mind, a clear importance of design emerges, from technology choices to implementation of interactions. Pure hardware-related biases are observed from the *Bulb Experiments* results. Meanwhile, fatigue from the *Time Perception and Rhythmic Stimuli* as well as biases from *engagement* in *Some Times Fly* revolves solely around the implementation of interactions and representations of the virtual world.

These results may highlight how delicate the process of designing interactions is with the goal of modulating time perception towards flow and lower stress, as any change in implementation details will lead to its own task performance and time perception side effects. If the previous section focused on how the effectiveness of modulators on time perception and task performance depends on context (such as task repetition and fatigue), in this section we saw that implementation and design choices further affect this effectiveness.

In conclusion, investigating purposeful time perception modulation would require a great deal of attention to implementation details. If possible, with a series of experiments that uses comparable hardware and software implementation of test scenarios.

An approach we took to respond to this threat is the development of customizable environments with a *VR farming scenario* as well as with *XR MUSE*, which are built with the aim of allowing extended reproducibility and cross-experiments integrations; these environments were presented in the *Results* chapter (Sections 3.2.1 and 3.2.2) and the implementation is further discussed in the next section.

4.3 Software Design Considerations After Iterative Framework Development

In Section 3.2, two customizable environments have been presented, the first is a VR precision farming environment, while the second is for AR with an industrial setting. Then, multiple studies were conducted in these environments.

If environments differ at first glance in terms of user interactions and appearance, they share a similar basis that allows for their in-depth customization. The reality is that this basis has been evolving through iterative works after each environment's development, but also after their usage in experiments.

In this section, we will discuss in detail the design choices that lead to the evolution of that basis, as well as the potential for improving the *XR MUSE* framework after reflecting on its uses in experiments.

4.3.1 A Shared Software Architecture Between the Environments

Each of the papers that present our frameworks have a focus on customization possibilities. However, a weakness from these papers is that we did not properly define what we meant by customization, and only showed what the customizable elements are in these environments. One could argue that since these are Unity environments, anyone with access to the source code would be able to customize everything, and that it holds true for any Unity project. Since both customizable environments use the same basis and design philosophy, this section will act as both a reminder of the software architecture, and as an opportunity to cover missing details from the published papers.

By saying *customization*, we imply that the elements that make up the environment are not just atomic functionalities that are either implemented or not. We imply that they are complex systems that answer a problem, and that these systems are accessible either to create situations from them, or made to be expanded upon.

The two environments share the customization abilities in two aspects:

- Allowing to change the in-environment properties and task sequences without changing the code, through external files

- A code structure that facilitates the implementation of new functionalities, with the aim to be used with the external files customization

The former is called *Virtual Environment Definition* in the publication of the VR farming environment, and *Operation* for the AR industrial one, internally we call it *Scene Description.* Although the details of the system can be found in their respective papers (Appendices D, E), the implementation revolves around making referenceable instances of in-environment objects, stimuli, and events.

The latter is tied to the former, as the framework's systems are built so that they do not need to be changed if used within an experiment. By providing abstract classes and available methods (see Fig. 4.1) to allow a developer to build new referenceable instances compatible with the *Scene Description.*

The next Section 4.3.2 will go more in-depth for the specific implementation of the *Events* system, as it is a critical aspect of how our framework allows fast prototyping and high reproducibility across experiments. The evolution of the *Scene Description* from the *farming environment* to the *industrial environment* is also discussed in Section 4.3.3. Finally, a self-report on the use of the system within experiments is available through Section 4.3.4.

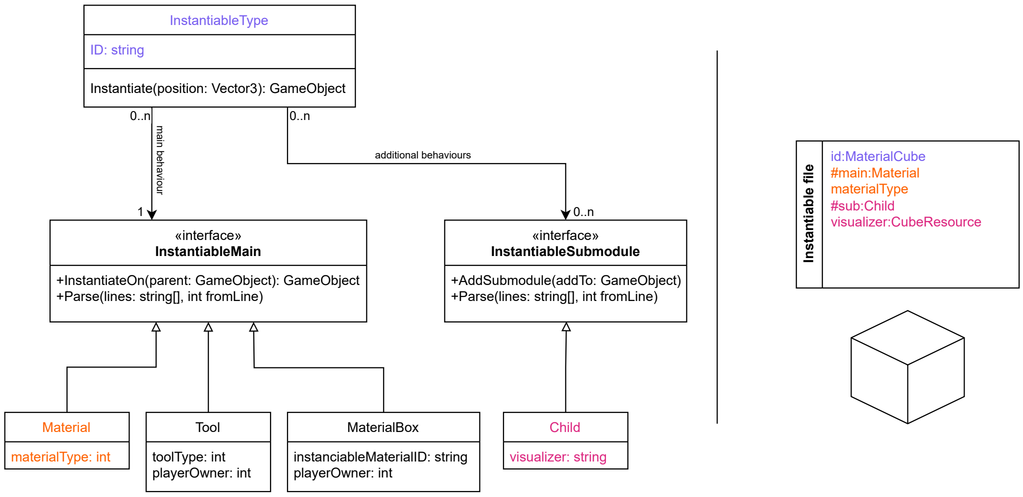


Figure 4.1 Simplified UML of *Instantiable* in *XR MUSE* and an example file to define a usable cube

4.3.2 In-Depth Explanation of the Events System

The two environments allow custom sequences using an *Events* system. Here we will see how it works in its latest iteration from *XR MUSE*, then what this system brings to software development. This *Event* system have been described as part of the *XR MUSE* paper (see Appendix E, *Section 3.2.2 Scripting*); however, if in the paper it was described from the point of view of the end user (how to use the system), here we will see how it works and discuss what it alleviates.

First, here is a reminder of the purpose of an *Event* for the end user, who is someone who will integrate experiments within a scenario. An *Event* allows to define a behavior that will happen during a trial of an experiment under specific conditions.

In the following example in Fig. 4.2, we see two events, one of which is a timer event that starts at the beginning of the trial, while the other one spawns an item after the first timer has ended. Each event has the following parameters: an *ID* which allows other events to use it within their parameters, a set of *start conditions* which must all be met for the event to start (more details later), and a *main behavior* with its own parameters that defines what the environment does to the environment.

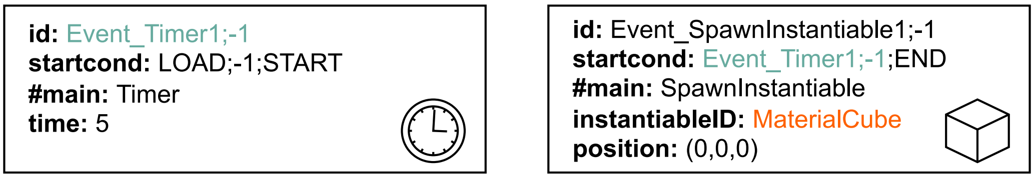
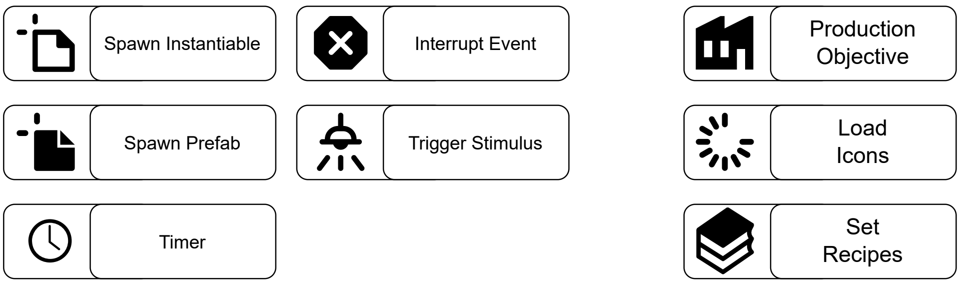


Figure 4.2 Two *XR MUSE* Events

These behaviors are meant to be written as an extension to an abstract class by the developer of the experiment (or any interactive application using the framework) if they wish to add new functionalities to their trials. However, a set of all-purpose behaviors are already available as part of the open-source *XR MUSE* framework (see Fig. 4.3), of which: timer events, *Unity Prefabs* and *Instantiables* spawning, triggering *Stimuli*, and *Event* interruption. In addition, some events specific to our *Example Scene*, such as the settings of the possible material combinations or the current objectives displayed on screen, are included. When an *Event* is defined by a developer, they must implement its behavior for when the event is started and when it is interrupted.



(a) Generic *Events* **(b)** Specific to *Example Scene*

Figure 4.3 Events available in the XR MUSE framework

Stimuli and *Instantiables* can also be expanded through the implementation of interfaces, as shown in Fig. 4.1 where *Material*, *Tool*, and *MaterialBox* are implementations specific to the *Example Scene* of *XR MUSE*.

Now let us look at the example events again in Fig. 4.2, both events have as a start condition the *ID* of an event, as well as either a *START* or *END*. This relies on a *Log* system; *Events* have as a *start condition* the need for a *Log* of a specific type to be produced by another specific *Event*. There are five logs that will be produced within an *Event*'s lifetime: *PENDING*, when the *Event* is loaded into the system but not yet started; *START*, which is produced when the event effectively start its behavior after all its start conditions are met; *SUCCESS*, which indicates that the event managed to finish its behavior properly; *FAIL*, which indicates that the event was ended abruptly (for instance, due to interruption from another event, a participant not finishing their task within a deadline, etc.); *END*, indicating that the event is finished, regardless of the end being due to a success or failure. In addition, there are automatic *Logs* related to the functioning of the framework, such as when a trial is started (see Fig. 4.4).

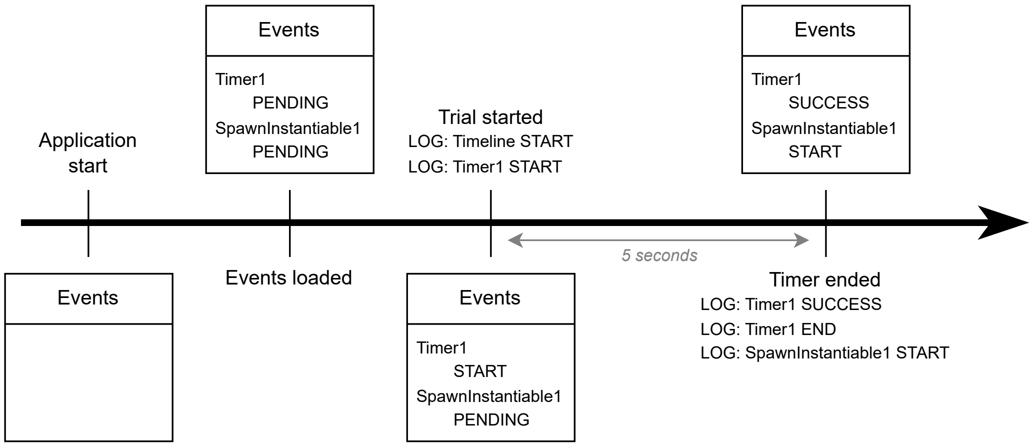


Figure 4.4 Timeline of the *Events* and *Logs* using the example in Fig. 4.2

In order to manage these *Events* and their *Logs*, the framework uses an *EventManager* which do not need any management from a developer using the framework. When *Events* are loaded within the application, regardless of whether it is before or after a trial, they are registered within the *EventManager* as *PENDING*. Then, any produced *Log* will be processed by the *EventManager* to validate any start condition from the awaiting events. Once an event is emptied from its start conditions, it will be started and produce its associated *Logs*. Therefore, if we take the example in Fig. 4.2 and Fig. 4.4, *EventManager* is the entity that starts the *Timer* when the *Timeline START Log* is produced and starts the *SpawnInstantiable* when *Timer1 END* is produced.

To its core, the *Events* system replaces the need to implement Observer [118] design patterns to handle trial sequences. More importantly, it allows for environment or framework behaviors to be linked without needing trial- or experiment-specific code changes for separate experiments. In the case of this thesis, the *Events* system allowed us to randomly and externally generate the sequences of the experiments trials seen in Sections 3.3.2, 3.3.4.

In its current implementation, this system is relatively limited to producing flowchart-like sequences of events, it even lacks the possibility to use incomplete or *either-or* start conditions. This is perfectly suited for restricted experiments and for quick prototyping, but may need to be expended in functionalities for more complex scenarios.

This kind of design, on top of simplifying the implementation work of an experiment, greatly improves the reproducibility of experiments, as a later version of the same experiment's program should stay compatible with previous sequences. Sequences within a trial no longer need to be a set element of the design, but can be seen as its own parameter.

4.3.3 Moving Away From Task-Specific Implementation

As implied in previous sections, the *XR MUSE* framework is a refinement of the functions developed for our *farming environment*. If the driving force behind this rework were the need for a multi-user environment, a main goal was to transform what was an environment and task-specific design into a generic framework that would be compatible across environments.

Within the *XR MUSE* paper, two main new features over the *VR farming environment* are presented: *Networking* and a custom *TypedCollision* system (used in conjunction with a *Dual-Transform Design*); however, the paper focuses mainly on the interactions built for networking, so we will focus a bit on the collision system here.

The *TypedCollision* system is an abstraction aiming to dissociate Unity physics from collision-based interactions, which facilitates the *Networking* implementation (see the *Dual-Transform Design* section of Appendix E) but also allows for class-agnostic implementations. A *TypedCollision* have a *collisionType* which is a simple integer hinting at the aim of the *Collider* of the object, custom events can then be registered to the *Typed Collision* as *UnityEvent Listeners* when the collider enters or exit another *TypedCollision*. This may sound familiar to the existing default *Collider* system of Unity; however, our extra layer, through its explicit *collisionType* and collision history management (access to all currently colliding objects, as well as the last entering and exiting collisions' types), is taking a step further in generalizing interactions. For example, the *Example Scene* uses *TypedCollision* for the *materials* and *tools* (both are interactive objects in the environment), where different *materials* can collide with each other to indicate whether a fusion can be attempted, and where the *tools* interact with the materials to trigger a fusion. Adding a new interaction, such as a visual effect when a tool touches the main table, would simply require to add a *TypedCollision* to the main table (see Fig. 4.5) with a registered *UnityEvent Listener* when a collision of the *collisionType* of a *tool* enters, and this does not require any knowledge on the implementation of the *Tool* class (see Listing 4.1).

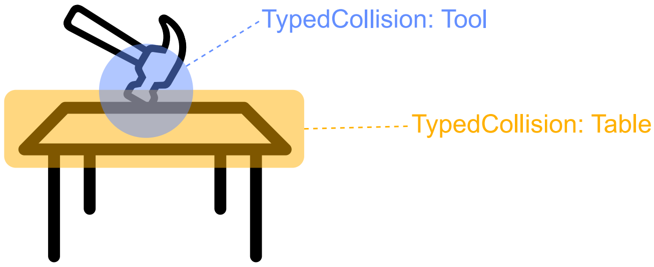


Figure 4.5 Use of *TypedCollision* between a table and a *Tool*

```
public class TableCollision : MonoBehaviour
{
    TypedCollision _tc;
    List<TypedCollision> _enteredTools = new List<TypedCollision>();
    CustomTable _table;

    void Start()
    {
        _tc = GetComponent<TypedCollision>();
        _table = GetComponent<CustomTable>();
        _tc.eventsEnter.AddListener(CheckOnEnter);
        _tc.eventsExit.AddListener(CheckOnExit);
    }

    void CheckOnEnter()
    {
        if(_tc._lastEnterType.collisionType ==
            TypedCollision.TYPE_TOOL)
        {
            _table.TriggerEnterEffect();
            _enteredTools.Add(_tc._lastEnterType);
        }
    }

    void CheckOnExit(){
        if(_enteredTools.Contains(_tc._lastEnterType)){
            _table.TriggerExitEffect();
            _enteredTools.Remove(_tc._lastEnterType);
        }
    }
}
```

Listing 4.1 C# code for a new interaction between the table and a *Tool*

Besides the new *TypedCollision* and *Networking* features, *XR MUSE* also moves away from the task-specific implementation of the *farming environment*. The process of adding new dynamic loading items (*Events, Instantiable, Stimuli*) now uses reflection for automatic registration instead of requiring a manual one for the parsing implementation. The *Instantiable* concept is a re-definition of the *Plants* and *Agents* definition from the *farming environment*. During both the design of the precision farming experiment (see Section 3.3.2) and building of the *XR MUSE* framework, it was clear that both *Plants* and *Agents* were derivatives of the same concept, as both allowed dynamic definitions of environmental objects to be loaded later within an experiment.

Through these iterations of what ultimately lead to the *XR MUSE* framework, it becomes quite clear how beneficial it can be to separate implementation logic from task design, as it allows for a more flexible and reusable implementation across experiments. However, finding the balance between what is task-specific and what can be generalized is quite challenging. For instance, both the *External Control* and *Task Types* of the *farming environment* (Appendix D) remained unused in later experiments due to how specific it revealed itself to be, and the *Plants* and *Agents* needed a new layer of generalization.

4.3.4 Accumulated Experience After Usage in Experiments

As discussed in Chapter 3, we effectively used both environments and their underlying systems to build experiments. This allows us to present a self-report on the experience of actually using the systems, and think of directions that could either improve them or warn future similar development projects on potential threats.

In *XR MUSE*, the networking uses a combination of automatic per-object synchronizations (through the *Dual Transform Design*) including object physics; however, the behaviors of the objects, if they need synchronization, require a deeper understanding of data synchronization. Within the *XR MUSE* paper, we describe an implemented system as *Data Synchronization* (Appendix E). Although we greatly simplify data registration through *XR MUSE*, new behaviors require the developer to manually manage which data needs to be synchronized, since each connected machine produces their own centralized *PlayerSyncedValues* (the per-machine synchronized data allowing for data registration and global aggregation, as well as remote procedure calls). Moreover, contrary to the synchronization of transforms, this manual synchronization does not provide any lag compensation techniques, which again require manual integration, especially if the system is used for animations. While providing a higher level of API on top of *Photon Unity Networking*, the *XR MUSE* framework could benefit from additional generic implementations of generic synchronization needs in multi-user applications, such as animation predictions and rollbacks, synchronized trigger systems, and optional *Remote-Procedure Calls* timeouts.

Another issue on the multi-user aspect of *XR MUSE* is its current use of *Photon* for its networking. During the development of the *XR MUSE* framework, *ExitGames* changed the terms of its licensing for *Photon Server* without a proper announcement, no longer allowing for free self-hosting in the case of non-commercial applications, this can be seen by the snapshots of the licensing page from *WebArchive* between March 2023[1] and September 2023[2]. Currently, the licensing page in non-english versions of the website[3] still mentions the free license. This new policy unfortunately forces the networking of *XR MUSE* to rely on *ExitGames*' dedicated server for our multi-user experiments, adding potential latency between the users as well as Internet access dependency. This issue, as well as Unity's recent failed tentative in runtime fee implementation[4], highlights the need for academic science to move to open-source alternatives. In our case, *XR MUSE* would greatly benefit from a re-implementation using *Mirror Networking*, and maybe a port to open *Unity* alternatives such as *Stride3D* or *GodotEngine*.

Right now, the *SceneDescription* (see Section 4.3.1) requires external production of text files to describe in-environment instances. Although this allows for quick prototyping, it can be quite confusing for someone not familiar with the framework to begin with. *Events sequences* also need extreme care to avoid dead ends or unexpected overlaps. In the case of the industrial scenario pilots, all trials used the same manually defined *Instantiables* (such as the *tools* and *materials* available to generate within the environment), and only the *Events* were randomly generated separately through a custom *Python* script. For simplification of this process, the system could heavily benefit from an external *Graphical User Interface* application allowing non-developer scientists to generate those files more naturally. Another potential improvement that could greatly help the experimenters, especially in the case of *XR MUSE*, which requires the installation of files on multiple devices, would be an automatic deployment and synchronization system for those files.

Currently, the *Stimuli* description has been under-used within our own experiments, as the *Stimuli* as defined in the framework are meant to be swapped or replaced between trials, which, in our use cases, was not ideal, as all stimuli were either a parameter of the full trial (thus, unused by the *Events*) or needed to be available at all times regardless of the trial due to external activation. However, this raises the question whether the external definition of *Stimuli* is desirable, as it could either be covered by the *Instantiables* or require too many specific per-stimulus parameters; the use of the *Stimuli* definition may, however, be desir-

[1] `https://bit.ly/4oHY0vL`
[2] `https://bit.ly/4fMlBYj`
[3] `https://bit.ly/47ydexh`
[4] `https://unity.com/products/pricing-updates`

able in the case of iterative development, if an environment is being reused with accumulated stimuli from previous versions.

In conclusion, it is clear that there are aspects where the current *XR MUSE* framework could see improvements, notably in its networking technology and accessibility to the use of *SceneDescription*. However, despite these issues, the developed technology was of great help in the development of interactive applications implementing intentional time perception modulation (Section 3.3.4), and can already be used for any interactive application prototyping. Considering that our framework has an open-source version, these issues could be resolved either through reuse of individual components of the project in other similar frameworks or through new forks of it.

4.4 Perspective on the Use of Active Time Modulation in Interactive Applications

After the results discussed in the previous sections and chapters, it is now time to raise the question about the driving force behind all this research: active and meaningful time perception modulation.

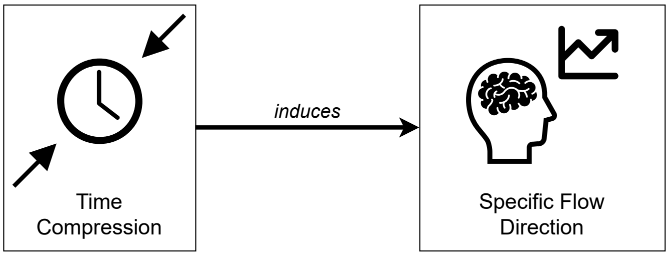
This thesis started with the definition of *ChronoPilot* and its *design principles* in the case of interactive applications. This assumes that by carefully implementing meaningful stimuli to strategically change time perception, this can lead to a flow state and lower stress. However, while on paper the *ChronoPilot design principles* are desirable, after consideration of the results, this idea is not flawless, because time modulation is flawed.

In our results, *time perception* and *task performance* are often affected by similar confounding effects. More specifically, with task familiarity, fatigue, hardware implementation, and software implementation. We also observed a case where time modulation stimuli (such as visual stimuli in *Time Perception and Rhythmic Stimuli* experiments) would interfere with the task. This is a huge issue, because flow states are by definition tied to the *challenge-skill ratio* (see Section 2.2.1), and thus with task performance. The expected effects of time modulators, as a result, become unpredictable between the activities, environments, and personal experience of the participants. A result that is unpredictable equally on the effectiveness of the time modulation, but also on the overall effectiveness of the stimulus towards flow.

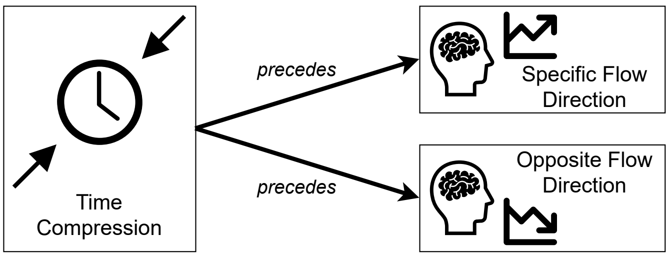
Another issue we observe with our results across stimuli on time estimations is that often stimuli would not have a specific effect in a specific direction (such as time compression or time expansion), but rather an effect on the *absolute* time estimation error's magnitude. It could be argued that, if we consider a time *clock* *model* (see Section 2.1.1), the introduction (or removal) of any stimulus will always to some degree affect both arousal levels and the attention resource, which could explain this unpredictability. For instance, if we take any introduction of a new visual element, regardless of whether it is in the background or not; we could assume that a participant noticing this element uses attention regardless of what it is to process it, but will also make their eyes react physically to the visual change.

A possible interpretation would be that in the case of aiming for ChronoPilot's goal, integrating a stimulus into an application does not need to answer *How will this stimulus affect time perception?* but rather *Why would, in that context, this stimulus affect time perception?*; for instance:

Are the audio stimuli in the *Time Perception and Rhythmic Stimuli* experiments affecting time perception due to their effect of rhythm on arousal, or is it because they were acting as a guide to keep track of performance as a participant? Is the letterboxing in *Some Times Fly* affecting time perception due to the effect on engagement, or was it due to the interpretation of the pure novelty coming from an unusual visual limitation?



(a) Assumption: in a given situation, time compression should induce a specific flow state



(b) *VR rhythm experiment*: time compression precedes both situations of increased and decreased performance

Figure 4.6 Assumption of ChronoPilot opposed to the results of the *VR rhythm experiment*

By integrating the *Why?* question to these results, we directly confront the assumptions of the *ChronoPilot design principles*, which are that a specific change in time experience would lead to another specific change in either flow or stress. However, the same changes in time perception can still be observed toward opposite flow directions through the task performance changes in the *VR rhythm sorter experiment* (see Fig. 4.6).

The lack of evidence, so far, for a causal effect between changes in time perception and changes in flow states or stress is the strongest threat to *ChronoPilot design principles*, especially when it is so difficult to dissociate the effect of a change in time perception with other changes, intrinsic to a specific stimulus, in user experience. As such, it would be hard to recommend trying to find a stimulus affecting time perception for a specific context or application to manage a participant's stress, especially compared to applying a stimulus directly affecting stress.

However, these threats do not necessarily invalidate all potential use cases of the *ChronoPilot design principles.* A first potential use case would be to use time experience as a proxy to assess changes in either stress or flow. But without assuming that the participant is indeed stressed or under flow based on the direction of a time perception change; due to how intertwined these states are, any change in time perception would indicate a potential (unknown) change on flow or stress. This would allow to use time perception measurement techniques as a gate on stimuli validation in cases where flow or stress are harder to measure directly. For instance, the *Active Modulation Controller Pilots Experiments* (see Section 3.3.4) and the *Precision Farming Scenario* (see Section 3.3.2) have hidden timing tasks that allow for accurate analysis of the time experience without needing any additional sensors or questionnaires. Proxy data (see Section 2.4.2) could also be used for this assessment, as we observed strong correlations through simple wearables (see Section 3.1.3), classification models could be a viable solution. This kind of design with time perception as a *validation gate* could be especially useful in the case of flow-induction intent, as time transformations (regardless of directions) are one of the aspects that define flow itself.

Another use case for the *ChronoPilot design principles* would be in situations where time perception is known in advance to be the source of stress or an obstacle to flow. For instance, a situation where a participant is waiting would lead to boredom as they would feel time passing, a situation where a participant fails to assess their remaining time in a time-limited task would likely be stressed. In this case, applying modulators specifically towards time perception would make sense, not necessarily only towards being compression or expansion, but also towards being accurate or inaccurate.

In conclusion, while time modulation does not seem, with current progress, as powerful as expected, it can still be a useful tool under the right conditions. Future potential progress in research, such as accurate assessment of time experience through proxy data or better classification of stimuli and their effect, could also lead to a broader potential of such *ChronoPilot design principles.*

Chapter 5

Conclusion and Perspectives

In this thesis, the focus is on the integration of time perception modulation as part of interactive applications and the challenges that come with integration. More specifically, motivated by the goals of the *ChronoPilot* research project, the objective was to explore to what extent and with which methods time modulation can be used to improve flow and reduce stress, in *eXtended Reality* (XR) applications. To this end, multiple works have been conducted: exploration of the possibilities of rhythmic stimuli, investigation of the bias introduced by XR on time perception, building an open-source framework for XR behavioral studies implementation, and a successful use of this framework as part of pilot applications which meaningfully modulated time perception.

Based on the results presented in the rhythm experiments, it was concluded that rhythmic stimuli work as time modulators, but also that rhythm should not be considered as a way to synchronize other stimuli, and that instead combining two different stimuli may lead to novel effects rather than a simple accumulation of effects. As such, designers of interactive applications who wish to affect or investigate time perception changes through rhythmic stimuli should consider that combination of a rhythmic pattern and its representations (such as audio only, visual only, or audio and visual combined) as one singular stimulus. However, it is important to note that the results presented only encompass metronome-like rhythmic patterns in combination to a specific audio and visual representation within an experiment-tailored task; it is not impossible that other stimuli combinations or rhythmic patterns may change the behavior of the effects of combined stimuli.

Regarding the use of XR, situations were observed where the time experience would differ in *Augmented Reality* (AR) and *Virtual Reality* (VR) compared to *Real-Life* (RL) likely due to the hardware in use; moreover, differences were also observed between VR situations depending on the implementation of the interactions. These findings imply that the implementation of a stimulus needs careful consideration on the reasons of its effect on time perception (e.g., amount of light emissions) and how sensitive that effect depends on the accuracy of representation (e.g., color accuracy), as both these aspects may change depending on the chosen hardware. However, the biases caused by the hardware were only investigated using a visual stimulus, but XR integrations can also include other types of inputs such as audio or haptics. Future research on other types of stimuli representations and with different hardware could be beneficial to timing research, as XR can be a powerful research tool if researchers are aware of its limitations.

Finally, the framework provided through this thesis allows for a simplified integration of single- and multi-user experiments in XR. However, while the framework has been successfully used in the development of complex experimental scenarios, it needs improvements on user-friendliness and ideally removing dependencies towards closed-source software. The usefulness of the framework is also difficult to evaluate due to its young age (as an open-source software) and lack of external projects and developers using it. However, the different experimental scenarios created with it for the ChronoPilot project have already been successfully used in experiments. These scenarios immerse users in a farming or industrial environment, integrating experimental requirements (e.g., diverse stress situations and time measurements) and modulation strategies (e.g., time modulators and controller integration) from ChronoPilot's multidisciplinary team and approaches.

Although the presented research provides useful information regarding the use of meaningful time modulation in interactive XR applications, it is difficult with current knowledge to advise using time perception as a proxy towards improved flow and lesser stress. This is due to the lack of observed causal relationships between time perception shifts and either stress or flow. However, it might still be possible to get immediate use of time perception measurements and time modulation stimuli in relation to stress and flow, by considering that a change in time perception within a given situation may confirm an unknown change in the target mental states. Modulation of time perception could also be considered in situations where a source of stress (or any other target mental state) is known to be directly related to time perception, such as deadline management.

The research presented in this thesis was part of a larger project. As such, a large portion of the work presented was designed towards testing specific use cases and scenarios of time modulation, rather than investigating the presence or absence of causal links between time perception changes and other mental states.

Future research on time modulation as any part of a user's experience needs further investigations on which mental states the time perception can have a causal effect to, both in the case of scenario-specific links, but also systemic links. As human time perception is an inherent aspect of any experience, with many approaches available to modulate it, investigating these potential causal effects from time perception shifts could allow novel design opportunities for interactive application designers.

List of Publications

Stéven Picard and Jean Botev. "Rhythmic Stimuli Effects on Subjective Time Perception in Immersive Virtual Environments". In *Proceedings of the 13th ACM Multimedia Systems Conference (MMSys)* (2022). doi:10.1145/3534086.3534330

Stéven Picard and Jean Botev. "Rhythmic Stimuli and Time Experience in Virtual Reality". In *Proceedings of the 20th EuroXR International Conference (EuroXR)* (2023), pp. 53-75. Springer, Cham. doi:10.1007/978-3-031-48495-7_4

Stéven Picard, Jean Botev, and Sahar Niknam. "A Dynamic and Scriptable Environment and Framework for Stimulus-Based Cognitive Research in Virtual Reality". In *Proceedings of the 22nd IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2023), pp. 387-392. IEEE. doi:10.1109/ISMAR-Adjunct60411.2023.00084

Stéven Picard, Ningyuan Sun, and Jean Botev. "XR MUSE: An Open-Source Unity Framework for Extended Reality-Based Networked Multi-User Studies". *Virtual Worlds, 3* (4) (2024). doi:10.3390/virtualworlds3040022

Stéven Picard and Jean Botev. "Psychophysiology of Rhythmic Stimuli and Time Experience in Virtual Reality". *Computers and Graphics, 124* (2024). doi:10.1016/j.cag.2024.104097

Sahar Niknam, Stéven Picard, Valentina Rondinelli, and Jean Botev. "Some Times Fly: The Effects of Engagement and Environmental Dynamics on Time Perception in Virtual Reality". In *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology (VRST)* (2024). ACM. doi:10.1145/3641825.3687726

Efthymia Lamprou, Stéven Picard, Eirini Balta, Jean Botev, Bora Celebi, Alvaro Garrido Perez, Pieter Simoens, Knut Drewing, Yara Khaluf, and Argiro Vatakis. "Shaping Time Perception Under Ambiguity: Multisensory Interventions in a VR Precision Farming Scenario". *In Submission*

Efthymia Lamprou, Stéven Picard, Eirini Balta, Jean Botev, and Argiro Vatakis. "Color Magnitude Effects with a Lightbulb in VR, AR, and Real-Life (*FINAL TITLE TBD*)". *In Preparation*

Eirini Balta, Stéven Picard, Efthymia Lamprou, Jean Botev, and Argiro Vatakis. "Oddball Effects with a Lightbulb in VR, AR, and Real-Life (*FINAL TITLE TBD*)". *In Preparation*

References

[1] Warrick Roseboom et al. “Activity in perceptual classification networks as a basis for human subjective time perception”. In: 10.1 (Jan. 2019). DOI: 10.1038/s41467-018-08194-7.

[2] Hudson Hoagland. “The Physiological Control of Judgments of Duration: Evidence for a Chemical Clock”. In: *The Journal of General Psychology* 9.2 (Oct. 1933), pp. 267–287. DOI: 10.1080/00221309.1933.9920937.

[3] Jean Botev et al. “ChronoPilot – Modulating Time Perception”. In: *Proceedings of the 4th IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR 2021)*. IEEE, Nov. 2021, pp. 215–218. ISBN: 978-1-6654-3225-2. DOI: 10.1109/AIVR52153.2021.00049.

[4] Amirhossein Ghaderi, Matthias Niemeier, and John Douglas Crawford. “Linear vector models of time perception account for saccade and stimulus novelty interactions”. 2021. DOI: 10.1101/2020.07.13.201087.

[5] Andrei Gorea. “Ticks per thought or thoughts per tick? A selective review of time perception with hints on future research”. In: *Journal of Physiology-Paris* 105.4 (2011), pp. 153–163. ISSN: 0928-4257. DOI: 10.1016/j.jphysparis.2011.09.008.

[6] Sylvie Droit-Volet et al. “Time, Emotion and the Embodiment of Timing”. In: *Timing & Time Perception* 1.1 (2013), pp. 99–126. DOI: 10.1163/22134468-00002004.

[7] Sylvie Droit-Volet et al. “Similar time distortions under the effect of emotion for durations of several minutes and a few seconds”. In: *Acta Psychologica* 210 (2020), p. 103170. ISSN: 0001-6918. DOI: 10.1016/j.actpsy.2020.103170.

[8] Mukesh Makwana and Narayanan Srinivasan. “Intended outcome expands in time”. In: *Scientific Reports* 7.1 (July 2017), p. 6305. DOI: 10.1038/s41598-017-05803-1.

[9] Dan Zakay, Devora Nitzan, and Joseph Glicksohn. "The influence of task difficulty and external tempo on subjective time estimation". In: 34.5 (Sept. 1983), pp. 451–456. DOI: 10.3758/bf03203060.

[10] Richard Fereday and Marc J. Buehner. "Temporal binding and internal clocks: No evidence for general pacemaker slowing." In: 43.5 (2017), pp. 971–985. DOI: 10.1037/xhp0000370.

[11] Tzu-Han Zoe Cheng and Sarah C. Creel. "The interplay of interval models and entrainment models in duration perception." In: 46.10 (Oct. 2020), pp. 1088–1104. DOI: 10.1037/xhp0000798.

[12] James Ash. "Technology, Technicity, and Emerging Practices of Temporal Sensitivity in Videogames". In: 44.1 (Jan. 2012), pp. 187–203. DOI: 10.1068/a44171.

[13] Alice Tomassini and Maria Concetta Morrone. "Perceived visual time depends on motor preparation and direction of hand movements". In: 6.1 (June 2016). DOI: 10.1038/srep27947.

[14] Jiaxin Shi and Xiting Huang. "The colour red affects time perception differently in different contexts". In: 52.1 (Aug. 2016), pp. 77–80. DOI: 10.1002/ijop.12384.

[15] Rose De Kock et al. "Slowing the body slows down time perception". In: 10 (Apr. 2021). DOI: 10.7554/elife.63607.

[16] Valérie Dormal et al. "Time perception is not for the faint-hearted? Physiological arousal does not influence duration categorisation". In: 19.3 (Dec. 2017), pp. 399–409. DOI: 10.1007/s10339-017-0852-3.

[17] Kyla Rankin, Kate Sweeny, and Sandra Xu. "Associations between subjective time perception and well-being during stressful waiting periods". In: 35.4 (Aug. 2019), pp. 549–559. DOI: 10.1002/smi.2888.

[18] John D. Watt. "Effect of Boredom Proneness on Time Perception". In: *Psychological Reports* 69.1 (1991). PMID: 1961817, pp. 323–327. DOI: 10.2466/pr0.1991.69.1.323.

[19] Hans Rutrecht et al. "Time Speeds Up During Flow States: A Study in Virtual Reality with the Video Game Thumper". In: *Timing & Time Perception* 9.4 (2021), pp. 353–376. DOI: 10.1163/22134468-bja10033.

[20] Shan Xu and Prabu David. "Distortions in time perceptions during task switching". In: *Computers in Human Behavior* 80 (2018), pp. 362–369. ISSN: 0747-5632. DOI: 10.1016/j.chb.2017.11.032.

[21] Susan A. Jackson and Robert C. Eklund. "Assessing Flow in Physical Activity: The Flow State Scale–2 and Dispositional Flow Scale–2". In: *Journal of Sport and Exercise Psychology* 24.2 (June 2002), pp. 133–150. DOI: 10.1123/jsep.24.2.133.

[22] Jean-Éric Pelet, Saïd Ettis, and Kelly Cowart. "Optimal experience of flow enhanced by telepresence: Evidence from social media use". In: 54.1 (Jan. 2017), pp. 115–128. DOI: 10.1016/j.im.2016.05.001.

[23] Fiona Nah and Brenda Eschenbrenner. "Flow Experience in Virtual Worlds: Individuals versus Dyads". In: *SIGHCI 2015 Proceedings* (Dec. 2015). URL: https://aisel.aisnet.org/sighci2015/19.

[24] Denise Reid. "A Model of Playfulness and Flow in Virtual Reality Interactions". In: *Presence: Teleoperators and Virtual Environments* 13.4 (Aug. 2004), pp. 451–462. DOI: 10.1162/1054746041944777.

[25] Stephen Kaplan. "The restorative benefits of nature: Toward an integrative framework". In: *Journal of Environmental Psychology* 15.3 (Sept. 1995), pp. 169–182. DOI: 10.1016/0272-4944(95)90001-2.

[26] Elizabeth Beverly et al. "A tranquil virtual reality experience to reduce subjective stress among COVID-19 frontline healthcare workers". In: *PLOS ONE* 17.2 (Feb. 2022). Ed. by Kelvin Ian Afrashtehfar, e0262703. DOI: 10.1371/journal.pone.0262703.

[27] Allison P. Anderson et al. "Relaxation with Immersive Natural Scenes Presented Using Virtual Reality". In: *Aerospace Medicine and Human Performance* 88.6 (June 2017), pp. 520–526. DOI: 10.3357/amhp.4747.2017.

[28] Gerhard Reese, Jasmin Stahlberg, and Claudia Menzel. "Digital shinrin-yoku: do nature experiences in virtual reality reduce stress and increase well-being as strongly as similar experiences in a physical forest?" In: *Virtual Reality* (Feb. 2022). DOI: 10.1007/s10055-022-00631-9.

[29] Dongrae Cho et al. "Detection of Stress Levels from Biosignals Measured in Virtual Reality Environments Using a Kernel-Based Extreme Learning Machine". In: 17.10 (Oct. 2017), p. 2435. DOI: 10.3390/s17102435.

[30] Christoph Rockstroh, Johannes Blum, and Anja S. Göritz. "Virtual reality in the application of heart rate variability biofeedback". In: 130 (Oct. 2019), pp. 209–220. DOI: 10.1016/j.ijhcs.2019.06.011.

[31] Hironori Hiraishi. "Heart rate analysis for pre-performance routine in darts game". In: *2020 IEEE 19th International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*. IEEE, Sept. 2020. DOI: 10.1109/iccicc50026.2020.9450215.

[32] B.H. Cho et al. "Attention Enhancement System using virtual reality and EEG biofeedback". In: *Proceedings IEEE Virtual Reality 2002.* IEEE Comput. Soc. DOI: 10.1109/vr.2002.996518.

[33] Peter Ulric Tse et al. "Attention and the subjective expansion of time". In: *Perception & Psychophysics* 66.7 (Oct. 2004), pp. 1171–1189. ISSN: 1532-5962. DOI: 10.3758/bf03196844.

[34] Ourania Tachmatzidou and Argiro Vatakis. "Attention and schema violations of real world scenes differentially modulate time perception". In: *Scientific Reports* 13.1 (June 2023). ISSN: 2045-2322. DOI: 10.1038/s41598-023-37030-2.

[35] Qingqing Li et al. "The effect of phasic alertness on temporal precision". In: 80.1 (Oct. 2017), pp. 262–274. DOI: 10.3758/s13414-017-1418-2.

[36] Clemens Wöllner and David Hammerschmidt. "Tapping to hip-hop: Effects of cognitive load, arousal, and musical meter on time experiences". In: *Attention, Perception, & Psychophysics* 83.4 (Jan. 2021), pp. 1552–1561. DOI: 10.3758/s13414-020-02227-4.

[37] Sven Thönes et al. "Boosting working memory with accelerated clocks". In: 226 (Feb. 2021), p. 117601. DOI: 10.1016/j.neuroimage.2020.117601.

[38] Fabian Christandl, Katja Mierke, and Corinna Peifer. "Time flows: Manipulations of subjective time progression affect recalled flow and performance in a subsequent task". In: 74 (Jan. 2018), pp. 246–256. DOI: 10.1016/j.jesp.2017.09.015.

[39] Ville Johannes Harjunen, Michiel Spapé, and Niklas Ravaja. "Anticipation of aversive visual stimuli lengthens perceived temporal duration". In: *Psychological Research* 86.4 (Aug. 2021), pp. 1230–1238. DOI: 10.1007/s00426-021-01559-6.

[40] Sylvie Droit-Volet and Michaël Dambrun. "Awareness of the passage of time and self-consciousness: What do meditators report?" In: 8.1 (Feb. 2019), pp. 51–65. DOI: 10.1002/pchj.270.

[41] J.H. Wearden and I.S. Penton-Voak. "Feeling the Heat: Body Temperature and the Rate of Subjective Time, Revisited". In: *The Quarterly Journal of Experimental Psychology Section B* 48.2 (1995), pp. 129–141. DOI: 10.1080/14640749508401443.

[42] Boris R.M. Kingma et al. "Time perception and timed decision task performance during passive heat stress". In: 8.1 (June 2020), pp. 53–63. DOI: 10.1080/23328940.2020.1776925.

[43] Kate Lambourne. "The Effects of Acute Exercise on Temporal Generalization". In: 65.3 (Mar. 2012), pp. 526–540. DOI: 10.1080/17470218.2011.605959.

[44] Emma Allingham, David Hammerschmidt, and Clemens Wöllner. "Time perception in human movement: Effects of speed and agency on duration estimation". In: 74.3 (Dec. 2020), pp. 559–572. DOI: 10.1177/1747021820979518.

[45] Nobuhiro Hagura et al. "Ready steady slow: action preparation slows the subjective passage of time". In: 279.1746 (Sept. 2012), pp. 4399–4406. DOI: 10.1098/rspb.2012.1339.

[46] Kielan Yarrow et al. "Illusory perceptions of space and time preserve cross-saccadic perceptual continuity". In: *Nature* 414.6861 (Nov. 2001), pp. 302–305. DOI: 10.1038/35104551.

[47] Jonas Knöll, M. Concetta Morrone, and Frank Bremmer. "Spatio-temporal topography of saccadic overestimation of time". In: *Vision Research* 83 (May 2013), pp. 56–65. DOI: 10.1016/j.visres.2013.02.013.

[48] Iona Hodinott-Hill et al. "Auditory Chronostasis". In: 12.20 (Oct. 2002), pp. 1779–1781. DOI: 10.1016/s0960-9822(02)01219-8.

[49] Kielan Yarrow and John C Rothwell. "Manual Chronostasis". In: 13.13 (July 2003), pp. 1134–1139. DOI: 10.1016/s0960-9822(03)00413-5.

[50] Keiko Ihaya, Takeharu Seno, and Yuki Yamada. "Più mosso: Fast self-motion makes cyclic action faster in virtual reality". In: 46.1 (2014), pp. 53–58. DOI: 10.1016/s0120-0534(14)70006-8.

[51] Stefan Weber, David Weibel, and Fred Mast. "How Self-Motion in Virtual Reality Affects the Subjective Perception of Time". In: *Timing & Time Perception* 8 (Jan. 2020), pp. 1–18. DOI: 10.1163/22134468-20191152.

[52] Ulrik Söderström, Martin Bååth, and Thomas Mejtoft. "The Users' Time Perception". In: ACM, Sept. 2018. DOI: 10.1145/3232078.3232092.

[53] Michele Fornaciai, Roberto Arrighi, and David C. Burr. "Adaptation-Induced Compression of Event Time Occurs Only for Translational Motion". In: 6.1 (Mar. 2016). DOI: 10.1038/srep23341.

[54] Gerald J. Gorn et al. "Waiting for the Web: How Screen Color Affects Time Perception". In: 41.2 (May 2004), pp. 215–225. DOI: 10.1509/jmkr.41.2.215.28668.

[55] S. Thönes et al. "Color and time perception: Evidence for temporal overestimation of blue stimuli". In: 8.1 (Jan. 2018). DOI: 10.1038/s41598-018-19892-z.

[56] Sebastian L. Kübel, Henrike Fiedler, and Marc Wittmann. "Red visual stimulation in the Ganzfeld leads to a relative overestimation of duration compared to green". In: 10.1 (Oct. 2020), pp. 5–19. DOI: 10.1002/pchj.395.

[57] Sylvie Droit-Volet et al. "Music, emotion, and time perception: the influence of subjective emotional valence and arousal?" In: *Frontiers in Psychology* 4 (2013), p. 417. DOI: 10.3389/fpsyg.2013.00417.

[58] Timothy Sanders and Paul Cairns. "Time Perception, Immersion and Music in Videogames". In: *Proceedings of the 24th BCS Interaction Specialist Group Conference*. BCS '10. Dundee, United Kingdom: BCS Learning and Development Ltd., 2010, pp. 160–167. ISBN: 9781780171302.

[59] Mariya Davydenko and Johanna Peetz. "Time grows on trees: The effect of nature settings on time perception". In: 54 (Dec. 2017), pp. 20–26. DOI: 10.1016/j.jenvp.2017.09.003.

[60] Li-Ann Leow, Taylor Parrott, and Jessica A. Grahn. "Individual Differences in Beat Perception Affect Gait Responses to Low- and High-Groove Music". In: *Frontiers in Human Neuroscience* 8 (Oct. 2014), p. 811. DOI: 10.3389/fnhum.2014.00811.

[61] Joanne E. Wittwer, Kate E. Webster, and Keith Hill. "Music and metronome cues produce different effects on gait spatiotemporal measures but not gait variability in healthy older adults". In: *Gait & Posture* 37.2 (Feb. 2013), pp. 219–222. DOI: 10.1016/j.gaitpost.2012.07.006.

[62] Li-Ann Leow, Kristina Waclawik, and Jessica A. Grahn. "The role of attention and intention in synchronization to music: effects on gait". In: *Experimental Brain Research* 236.1 (Oct. 2017), pp. 99–115. DOI: 10.1007/s00221-017-5110-5.

[63] Jan Stupacher. "The experience of flow during sensorimotor synchronization to musical rhythms". In: *Musicae Scientiae* 23.3 (July 2019), pp. 348–361. DOI: 10.1177/1029864919836720.

[64] Betty Yin et al. "Tour de Tune 2 - Auditory-Game-Motor Synchronisation with Music Tempo in an Immersive Virtual Reality Exergame". In: ACM, Feb. 2021. DOI: 10.1145/3437378.3437379.

[65] Carrie M. Peters and Cheryl M. Glazebrook. "Rhythmic auditory stimuli heard before and during a reaching movement elicit performance improvements in both temporal and spatial movement parameters". In: 207 (June 2020), p. 103086. DOI: 10.1016/j.actpsy.2020.103086.

[66] Michael J. Hove and Jane L. Risen. "It's All in the Timing: Interpersonal Synchrony Increases Affiliation". In: *Social Cognition* 27.6 (Dec. 2009), pp. 949–960. DOI: 10.1521/soco.2009.27.6.949.

[67] Jan Stupacher, Guilherme Wood, and Matthias Witte. "Synchrony and sympathy: Social entrainment with music compared to a metronome". In: *Psychomusicology: Music, Mind, and Brain* 27.3 (Sept. 2017), pp. 158–166. DOI: 10.1037/pmu0000181.

[68] David Hammerschmidt et al. "Disco Time: The Relationship Between Perceived Duration and Tempo in Music". In: *Music & Science* 4 (2021), p. 2059204320986384. DOI: 10.1177/2059204320986384.

[69] Sarah Lukas, Andrea M. Philipp, and Iring Koch. "Crossmodal attention switching: Auditory dominance in temporal discrimination tasks". In: *Acta Psychologica* 153 (2014), pp. 139–146. ISSN: 0001-6918. DOI: 10.1016/j.actpsy.2014.10.003.

[70] Xinyue Wang, Clemens Wöllner, and Zhuanghua Shi. "Perceiving Tempo in Incongruent Audiovisual Presentations of Human Motion: Evidence for a Visual Driving Effect". In: *Timing & Time Perception* 10.1 (2021), pp. 75–95. DOI: 10.1163/22134468-bja10036.

[71] Taeko Tanaka, Taiki Ogata, and Yoshihiro Miyake. "The Effect of Rhythmic Tactile Stimuli Under the Voluntary Movement on Audio-Tactile Temporal Order Judgement". In: 11 (Feb. 2021). DOI: 10.3389/fpsyg.2020.600263.

[72] Argiro Vatakis et al. *Timing and time perception: Procedures, measures, & applications*. BRILL, Apr. 2018.

[73] Charlotte Harrison et al. "Time-Order Errors in Duration Judgment Are Independent of Spatial Positioning". In: *Frontiers in Psychology* 8 (Mar. 2017), p. 340. DOI: 10.3389/fpsyg.2017.00340.

[74] Lara Orlandic, Adriana Arza Valdes, and David Atienza. "Wearable and Continuous Prediction of Passage of Time Perception for Monitoring Mental Health". In: *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, June 2021. DOI: 10.1109/cbms52027.2021.00050.

[75] Hao Chen et al. *Exploring Pupil Dilation in Emotional Virtual Reality Environments*. 2017. DOI: 10.2312/EGVE.20171355.

[76] Amir Hossein Ghaderi et al. "Time estimation and beta segregation: An EEG study and graph theoretical approach". In: *PLOS ONE* 13.4 (Apr. 2018). Ed. by Warren H Meck, e0195380. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0195380.

[77] Laetitia Grabot et al. "The Strength of Alpha–Beta Oscillatory Coupling Predicts Motor Timing Precision". In: *Journal of Neuroscience* 39.17 (2019), pp. 3277–3291. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2473-18.2018. URL: https://www.jneurosci.org/content/39/17/3277.

[78] Sahar Niknam et al. "Brain Signatures of Time Perception in Virtual Reality". In: *IEEE Transactions on Visualization and Computer Graphics* 31.5 (May 2025), pp. 2535–2545. ISSN: 2160-9306. DOI: 10.1109/tvcg.2025.3549570.

[79] Joschka Mütterlein. "The Three Pillars of Virtual Reality? Investigating the Roles of Immersion, Presence, and Interactivity". In: Jan. 2018. DOI: 10.24251/HICSS.2018.174.

[80] Fabian Unruh et al. "The Influence of Avatar Embodiment on Time Perception - Towards VR for Time-Based Therapy". In: *Frontiers in Virtual Reality* 2 (July 2021). DOI: 10.3389/frvir.2021.658509.

[81] Daehwan Kim and Yong Jae Ko. "The impact of virtual reality (VR) technology on sport spectators' flow experience and satisfaction". In: *Computers in Human Behavior* 93 (2019), pp. 346–356. ISSN: 0747-5632. URL: https://www.sciencedirect.com/science/article/pii/S0747563218306265.

[82] Grayson Mullen and Nicolas Davidenko. "Time Compression in Virtual Reality". In: *Timing & Time Perception* 9 (May 2021), pp. 1–16. DOI: 10.1163/22134468-bja10034.

[83] Federico Alvarez Igarzábal et al. "What Happens While Waiting in Virtual Reality? A Comparison Between a Virtual and a Real Waiting Situation Concerning Boredom, Self-Regulation, and the Experience of Time". In: *Technology, Mind, and Behavior* 2.2 (July 22, 2021). https://tmb.apaopen.org/pub/what-happens-while-waiting-in-virtual-reality. DOI: 10.1037/tmb0000038.

[84] Gerd Bruder and Frank Steinicke. "Time perception during walking in virtual environments". In: Mar. 2014, pp. 67–68. ISBN: 978-1-4799-2871-2. DOI: 10.1109/VR.2014.6802054.

[85] Steven C. Mallam, Jørgen Ernstsen, and Salman Nazir. "Accuracy of Time Duration Estimations in Virtual Reality". In: 64.1 (Dec. 2020), pp. 2079–2083. DOI: 10.1177/1071181320641503.

[86] Ineke J.M. van der Ham et al. "Elapsed time estimates in virtual reality and the physical world: The role of arousal and emotional valence". In: 94 (May 2019), pp. 77–81. DOI: 10.1016/j.chb.2019.01.005.

[87] Kazuma Shimokawa and Eriko Sugimori. “Using Virtual Reality to Study Subjective Time in Crowded Versus Uncrowded Environments”. In: *Perceptual and Motor Skills* 126.5 (2019). PMID: 31216235, pp. 737–752. DOI: 10.1177/0031512519857869.

[88] Christian Schatzschneider, Gerd Bruder, and Frank Steinicke. “Who turned the clock? Effects of Manipulated Zeitgebers, Cognitive Load and Immersion on Time Estimation”. In: 22.4 (Apr. 2016), pp. 1387–1395. DOI: 10.1109/tvcg.2016.2518137.

[89] Anastasia Ruvimova et al. “"Transport Me Away": Fostering Flow in Open Offices through Virtual Reality”. In: ACM, Apr. 2020. DOI: 10.1145/3313831.3376724.

[90] Luca Lo Verde et al. “Time dilation effect in an active observer and virtual environment requires apparent motion: No dilation for retinal- or world-motion alone”. In: 19.3 (Mar. 2019), p. 4. DOI: 10.1167/19.3.4.

[91] Ambika Bansal, Séamas Weech, and Michael Barnett-Cowan. “Movement-Contingent Time Flow in Virtual Reality Causes Temporal Recalibration”. In: 9.1 (Mar. 2019). DOI: 10.1038/s41598-019-40870-6.

[92] Victoria Wright and Genovefa Kefalidou. “Can you hear the Colour? Towards a Synaesthetic and Multimodal Design Approach in Virtual Worlds”. In: (Aug. 2021). URL: https://leicester.figshare.com/articles/conference_contribution/Can_you_hear_the_Colour_Towards_a_Synaesthetic_and_Multimodal_Design_Approach_in_Virtual_Worlds/15035064.

[93] Austin Erickson et al. “Examining Whether Secondary Effects of Temperature Associated Virtual Stimuli Influence Subjective Perception of Duration”. In: IEEE, Mar. 2020. DOI: 10.1109/vr46266.2020.00070.

[94] Katja Rogers et al. “The Potential Disconnect between Time Perception and Immersion: Effects of Music on VR Player Experience”. In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, Nov. 2020. DOI: 10.1145/3410404.3414246.

[95] Benjamin Outram et al. “Crystal Vibes feat. Ott: A psychedelic musical virtual reality experience utilising the full-body vibrotactile haptic synesthesia suit”. In: IEEE, Oct. 2017. DOI: 10.1109/vsmm.2017.8346269.

[96] John Reif and Wadee Alhalabi. *Advancing Attention Control Using VR-Induced Multimodal Artificial Synesthesia*. Aug. 2018. DOI: 10.20944/preprints201808.0134.v1.

[97] Immo Schuetz, Harun Karimpur, and Katja Fiehler. “vexptoolbox: A software toolbox for human behavior studies using the Vizard virtual reality platform”. In: *Behavior Research Methods* 55.2 (Feb. 2023), pp. 570–582. ISSN: 1554-3528. DOI: 10.3758/s13428-022-01831-6.

[98] Adam O. Bebko and Nikolaus F. Troje. “bmlTUX: Design and Control of Experiments in Virtual Reality and Beyond”. In: *i-Perception* 11.4 (2020). PMID: 32733664, p. 2041669520938400. DOI: 10.1177/2041669520938400.

[99] Jack Brookes et al. “Studying human behavior with virtual reality: The Unity Experiment Framework”. In: *Behavior Research Methods* 52.2 (Apr. 2020), pp. 455–463. ISSN: 1554-3528. DOI: 10.3758/s13428-019-01242-0.

[100] Marcus R Watson et al. “USE: An integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents”. In: *Journal of Neuroscience Methods* 326 (2019), p. 108374. ISSN: 0165-0270. DOI: 10.1016/j.jneumeth.2019.108374.

[101] Maximilian Landeck et al. “Metachron: A framework for time perception research in VR”. In: *26th ACM Symposium on Virtual Reality Software and Technology*. ACM, Nov. 2020. DOI: 10.1145/3385956.3422111.

[102] Divine Maloney, Guo Freeman, and Andrew Robb. “A Virtual Space for All: Exploring Children’s Experience in Social Virtual Reality”. In: *7th Annual Symposium on Computer-Human Interaction in Play (CHI PLAY)*. 2020, pp. 472–483.

[103] Jongyong Kim et al. “XR Framework for Collaborating Remote Heterogeneous Devices”. In: *1st IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2020, pp. 586–587.

[104] Vasco Pereira et al. “Extended Reality Framework for Remote Collaborative Interactions in Virtual Environments”. In: *2nd International Conference on Graphics and Interaction (ICGI)*. 2019, pp. 17–24.

[105] Georgi Kostov and Josef Wolfartsberger. “Designing a Framework for Collaborative Mixed Reality Training”. In: *Procedia Computer Science* 200 (2022), pp. 896–903.

[106] Erich K. Grommet et al. “Time estimation of fear cues in human observers”. In: *Behavioural Processes* 86.1 (Jan. 2011), pp. 88–93. ISSN: 0376-6357. DOI: 10.1016/j.beproc.2010.10.003.

[107] Sophie Fayolle, Sandrine Gil, and Sylvie Droit-Volet. “Fear and time: Fear speeds up the internal clock”. In: *Behavioural Processes* 120 (Nov. 2015), pp. 135–140. ISSN: 0376-6357. DOI: 10.1016/j.beproc.2015.09.014.

[108] Virginie van Wassenhove et al. “Psychological and Neural Mechanisms of Subjective Time Dilation”. In: *Frontiers in Neuroscience* 5 (2011). ISSN: 1662-4548. DOI: 10.3389/fnins.2011.00056.

[109] Jason Tipples, Michael Lupton, and David George. “The Effects of Time Pressure on Temporal Overestimation Due to Threat”. In: *Timing & Time Perception* 9.3 (2021), pp. 301–314. DOI: 10.1163/22134468-bja10027.

[110] Jessica I. Lake and Kevin S. LaBar. “Unpredictability and Uncertainty in Anxiety: A New Direction for Emotional Timing Research”. In: *Frontiers in Integrative Neuroscience* 5 (2011). ISSN: 1662-5145. DOI: 10.3389/fnint.2011.00055.

[111] Wanlu Yang et al. “Affective auditory stimulus database: An expanded version of the International Affective Digitized Sounds (IADS-E)”. In: *Behavior Research Methods* 50.4 (Mar. 2018), pp. 1415–1429. ISSN: 1554-3528. DOI: 10.3758/s13428-018-1027-6.

[112] Deane E. Aikins and Michelle G. Craske. “Cognitive Theories of Generalized Anxiety Disorder”. In: *Psychiatric Clinics of North America* 24.1 (Mar. 2001), pp. 57–74. ISSN: 0193-953X. DOI: 10.1016/s0193-953x(05)70206-9.

[113] Joan Danielle K. Ongchoco, Kimberly W. Wong, and Brian J. Scholl. “What’s next?: Time is subjectively dilated not only for ‘oddball’ events, but also for events immediately after oddballs”. In: *Attention, Perception, & Psychophysics* 86.1 (Oct. 2023), pp. 16–21. ISSN: 1943-393X. DOI: 10.3758/s13414-023-02800-7.

[114] Jordan J. Wehrman, John Wearden, and Paul Sowman. “The expected oddball: effects of implicit and explicit positional expectation on duration perception”. In: *Psychological Research* 84.3 (Sept. 2018), pp. 713–727. ISSN: 1430-2772. DOI: 10.1007/s00426-018-1093-5.

[115] Müge Cavdan, Bora Celebi, and Knut Drewing. “Simultaneous Emotional Stimuli Prolong the Timing of Vibrotactile Events”. In: *IEEE Transactions on Haptics* 16.4 (Oct. 2023), pp. 622–627. ISSN: 2334-0134. DOI: 10.1109/toh.2023.3275190.

[116] Lydia Liapi et al. “Time perception in film viewing: A modulation of scene’s duration estimates as a function of film editing”. In: *Acta Psychologica* 244 (Apr. 2024), p. 104206. ISSN: 0001-6918. DOI: 10.1016/j.actpsy.2024.104206.

[117] Eirini Balta, Andreas Psarrakis, and Argiro Vatakis. "The effects of increased mental workload of air traffic controllers on time perception: Behavioral and physiological evidence". In: *Applied Ergonomics* 115 (2024), p. 104162. ISSN: 0003-6870. URL: https://www.sciencedirect.com/science/article/pii/S0003687023002004.

[118] Erich Gamma et al. *Design patterns: elements of reusable object-oriented software.* USA: Addison-Wesley Longman Publishing Co., Inc., 1995, pp. 292–304. ISBN: 0201633612.

Appendix A

Paper: Rhythmic Stimuli Effects on Subjective Time Perception in Immersive Virtual Environments

Full paper of Section 3.1.1

Rhythmic Stimuli Effects on Subjective Time Perception in Immersive Virtual Environments

Stéven Picard
University of Luxembourg
Esch-sur-Alzette, Luxembourg
steven.picard@uni.lu

Jean Botev
University of Luxembourg
Esch-sur-Alzette, Luxembourg
jean.botev@uni.lu

ABSTRACT

Time perception is an essential component of a user's experience and interaction in immersive virtual environments. This paper explores the performance and subjective time perception when carrying out a cognitive task in a virtual environment while being exposed to unrelated rhythmic stimuli. To this end, we devised an experiment comprising a simple object sorting task with varying rhythmic stimuli, investigating time experience in the form of time estimation and time judgment. The results imply varying effects depending on the usage of single stimuli compared to synchronized audio-visual effects. Single stimuli can lead to more pronounced time perception variations regardless of tempo, but these variations are not specifically compression or dilation. Synchronized stimuli, in turn, can lead to time compression or dilation, depending on the tempo. The results further imply that time being judged as fast or slow correlates to stimuli' tempo, while the presence of visual stimuli can negatively impact task performance. An active, purposeful rhythmic stimuli modulation to tailor individual time experiences can open up exciting opportunities in virtual environment design.

CCS CONCEPTS

• **Human-centered computing → Interaction techniques**; **Interaction design**; **HCI design and evaluation methods**.

KEYWORDS

Rhythmic stimuli, time perception, virtual environments

ACM Reference Format:
Stéven Picard and Jean Botev. 2022. Rhythmic Stimuli Effects on Subjective Time Perception in Immersive Virtual Environments. In *International Workshop on Immersive Mixed and Virtual Environment Systems (MMVE '22), June 14, 2022, Athlone, Ireland*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3534086.3534330

1 INTRODUCTION

Time is an integral part of the human experience in real and virtual environments alike. Time perception is a subjective experience; therefore, it is interesting to investigate whether one can modulate that experience, and which factors affect time perception to what extent. The latter, in particular, has spawned many scientific studies this paper falls in line with. Specifically, we investigate relations between time perception and rhythmic stimuli in virtual environments. To isolate and study the effect of such stimuli on time experience, users in our experiment are asked to perform an unrelated cognitive task. Rhythm provides a simple way of synchronizing modalities and thus allows investigating the differences between single-modal and multi-modal, synchronized stimuli.

As part of the ChronoPilot project [1], which focuses on the active modulation of subjective time experience, a central objective of this study is to investigate whether rhythmic stimuli allow for influencing time experience in a cognitive task. Another objective is to investigate the interplay between synchronized stimuli modalities (auditory and visual). To this end, we investigate the following hypotheses:

- **H1** Performing a cognitive task, rhythmic visual (H1a) or auditory (H1b) stimuli affect time perception.
- **H2** The effect of rhythmic stimuli on time perception depends on the tempo of the rhythmic stimuli.
- **H3** Time distortions are more pronounced when using both stimuli types rather than individual types.
- **H4** Rhythmic stimuli affect task performance (H4a), varying depending on the tempo of the stimuli (H4b), which is amplified when using multiple stimuli (H4c).

Before focusing on the experiment, the data analysis, and the obtained results in Sections 3 and 4, the following section provides the necessary background by discussing related work. A discussion of the limitations, as well as a summary and outlook on challenges, opportunities, and research perspectives related to virtual environment design in Section 5 concludes the paper.

2 BACKGROUND AND RELATED WORK

Time perception encompasses different aspects such as time estimation, time judgment, and various representation models. In this paper, time perception is considered a personal interpretation of how time passed quantitatively (i.e., how much time one believes has passed) and qualitatively (i.e., how quickly that amount of time has passed).

The so-called clock model constitutes a classic and popular intrinsic model, which assumes the body or brain as a system dedicated to time perception [9]. The idea of an "internal clock" is often accompanied by a pacemaker-counter and oscillator model, suggesting that the brain has internal processes keeping track of time by producing "ticks" or "time units". The speed at which these time units are produced, the "clock speed", and whether some units are skipped can be influenced by external stimuli unrelated to time. In these models, attention or attentional resources can act as channels for the ticks

MMVE '22, June 14, 2022, Athlone, Ireland

ACM ISBN 978-1-4503-9382-9/22/06.
https://doi.org/10.1145/3534086.3534330

produced to be accounted for by the processes [5, 6]. Studies employing these models generally consider two sources of subjective temporal distortions: arousal levels and attention. Arousal levels would affect the tick rate or the clock speed, resulting in a stretched perception of time, while attention acts as a gate or switch between the produced ticks and the accumulator: the less one pays attention to time (paying attention to something else), the more compressed is time [2, 3, 14].

Studies related to time perception distortions often overlap with the notion of flow, which is a psychological state of full attention on a task defined by Csikszentmihalyi with nine dimensions: challenge-skill balance, action-awareness merging, clear goals, unambiguous feedback, concentration on task, sense of control, loss of self-consciousness, time transformation, and autotelic experience [10]. However, flow is much more prominent in research than pure time perception. Since we aim to regulate subjective time perception actively and not only try to understand the underlying process, examining what affects flow is necessary. In their study comparing flow states in a session of the rhythm game Thumper between a VR and non-VR setup, Rutrecht et al. observed that both scenarios lead to a flow state regardless of the VR setup's higher immersion level. Interestingly, they also note that the passage of time and time estimation reported were not correlated [15].

Studying subjective time can be complex as various time judgment tasks involve different processing mechanisms, especially when comparing time estimations (i.e., asking a participant to give an estimation of an event in time units, such as seconds) and the feeling of time passage (i.e., asking if time passes quickly or not). For instance, it was observed that depressed subjects underestimate time but feel that time passes slowly [3] or that the boredom-prone performing a boring task feel a longer subjective time passage but no difference to non-boredom-prone when it comes to time evaluation [19]. A more positive example is that flow states induced by the video game Thumper lead to the faster passage of time but without time estimation errors [15]. Generally, time passage appears to be more tied to attention, while time duration estimation would be tied to memory [3, 22].

Following the overview of time perception notions and their different dimensions, the subsequent discussion focuses on research related to rhythmic stimuli and how they affect user experience. Droit-Volet et al. [4] highlight the effect of a musical piece's emotional valence and orchestration on time perception, identifying tempo as one of the core elements of a musical piece that can affect time perception. Indeed, tempo tends to affect arousal; hence faster tempi lead to time dilation. A recent study by Hammerschmidt et al. [7] explored different timing evaluations (reproduction, estimation, and subjective rating) of instrumental excerpts of Disco songs at different tempi reveals interesting properties about tempo related to time perception. Faster tempi yield longer duration reproductions, but no effects are observed for duration estimations. Depending on the measure, a tempo difference of at least 20 BPM is required for stretching the duration.

In the above studies, tempo is represented in the form or at least as part of audio stimuli, which makes perfect sense since tempo can be associated with music and rhythm and, for instance, Lukas et al. [13] suggest that the auditory channel dominates the visual sense for time-processing tasks. However, Wang et al. [18] reconsider this in a later study by using visual stimuli containing more temporal information than the usual laboratory stimuli in the form of Point-Light-Display (PLD) dance motions. Presented with asynchronous PLD motions and simple audio tempos, participants had to indicate what they thought was the most fitting tempo globally. Visual stimuli have a more significant influence, suggesting that either under the right conditions, visuals dominate audio when it comes to time perception, or the prioritized channels only depend on the quantity of information conveyed by the stimuli.

Wöllner and Hammerschmidt [21] recently investigated the effect of cognitive load, arousal, and musical meter on time experiences. They varied the cognitive load (tasks) and arousal levels of music while keeping the tempo constant, finding that: (1) time passage is quicker under higher cognitive load (represented with a concurrent math task), (2) performing a concurrent motor task (tapping along with the music) results in a shorter perceived duration, and (3) for the tapping task, tapping to larger structures (half notes instead of eighth notes) of the same music resulted in both shorter duration estimates and faster time passage.

Rhythmic stimuli have also been shown to influence gait if patients are instructed to walk in synchrony with the stimuli, which can help with Parkinson's disease [12]. For older, healthy adults, Wittwer et al. observe that music cues lead to higher gait velocity and cadence [20]. However, in the context of gait, not all rhythmic cues are equal. In the above study, if music would affect both velocity and cadence, metronome cues would only affect cadence. Another study on healthy adults by Leow et al. [12] observes how effects vary depending on the groove of the music and the beat-perception abilities of the participants. More generally, synchronizing with a tapping task to medium and high rhythmic complexity (but not too low complexity) appears to be able to induce the "fluency of performance" dimension of flow but without the "absorption by activity" dimension (which includes the time passing distortions) [16]. Synchronizing or not with music also seems to affect social sympathy, as observed by Stupacher et al. [17] with synchronized tapping to music leading to more sympathy and unsynchronized tapping resulting in less sympathy. Like in [20], there is again a difference between music and metronome cues, as synchronous tapping does not lead to any difference.

The prospective paradigm appears more practical for experiments where as much subjective participant data as possible is gathered since, contrary to a retrospective approach, it allows for repeating similar tasks with changing parameters where the repetition of stimuli and their predictability affect time perception. Across studies, various effects can occur, such as time-order errors (TOEs) when judging stimuli depending on the presentation order. For instance, Harrison et al. [8] found that TOEs are greater when a stimulus involves a saccade while there is no effect of overlapping stimuli regarding time perception. Aside from TOEs, Makwana and Srinivasan [14] inquired what color participants wanted to see before showing the stimuli they needed for time estimation and observed that intention and predictability affect time perception. Moreover, Kruijne et al. [11] observe a repetition effect on the magnitude of sensory response due to neural repetition suppression, which leads to effects on time perception that are unrelated to changes in arousal or surprise. These notions highlight a potential challenge and source of bias regarding studies on time perception, as repeated scenarios and participants' preferences can affect time perception in a way that is out of our control.

3 EXPERIMENT

For our experiment, we recruited 50 participants through the online crowdsourcing platform Prolific with a balanced gender ratio, i.e., 50% male and 50% female, and no geographic restriction. The only technical requirement was that exclusively laptops and desktop computers were used, i.e., no mobile devices such as phones or tablets. The participants were provided with a link to a Unity WebGL build of the experimental app hosted on our servers. In a first step, we manually filtered out participants if their data was incomplete or if the data was incoherent and it was confirmed after inquiry that the participant incorrectly followed the instructions. Individual trials were then filtered from their output data if an asset did not load correctly or if the participant's computer was slowing down. Among the 39 retained participants, the female-male ratio was 51.3%-48.7%, with ages ranging from 18 to 65 (average 28.05 / median 24). Considering the retained trials, the female-male ratio is 51.7%-48.3% (age average 28.31 / median 24). The average number of trials retained per participant is 8.38 (median 9), corresponding to potential bootstrapping issues when loading the app data for the first trial. Table 1 indicates the total trials per stimulus mode.

	0	100	140	180	Total
None	38	-	-	-	38
AudioOnly	-	30	25	38	93
VisualsOnly	-	21	38	35	94
Both	-	31	35	36	102
Total	38	82	98	109	327

Table 1: Number of trials per stimulus mode

3.1 Trial Task and Design

The task during trials is to sort three-dimensional objects according to their shape (cube, capsule, or sphere). The objects are moved into two larger spheres acting as sinks that process only specific shapes indicated by a panel above; in the example shown in Figure 1, the left sphere accepts cubes or capsules, and the right sphere accepts capsules or spheres. New objects appear at the center of the virtual environment after the previous object is sorted away, i.e., disappearing with a small animation (cf. image sequence). The individual trials' duration is unknown to the participants; following a trial run, they were asked to complete a questionnaire asking for a time estimation in seconds, as well as an evaluation of how quick they felt the trial was and an evaluation of their fatigue, both on a Likert scale. Before the trials, we asked the participants to indicate their age and gender.

The conditions under which the participants are asked to perform the sorting task were defined by a combination of the following parameters:

- The **duration**, i.e., trial length (40, 50, or 60 seconds).
- The **tempo** of visual and auditory stimuli production (100, 140, or 180 beats per minute, or 0 for trials with no stimuli).
- Whether or not **visual stimuli** are produced, i.e., flashing pulses around the sortable object.
- Whether or not **audio stimuli** are produced. These resemble a metronome click sound.

All participants performed the task under 10 different random parameter combinations (excluding combinations of a 0 tempo value with any auditory or visual stimuli). The task order was also determined randomly for each participant.

3.2 Data

As indicated in Section 3.1, we collected demographic data such as age and gender per participant in a basic questionnaire prior to the experiment, i.e., outside of trials. In addition to the more detailed per-trial post-assessment questionnaire related to time estimation and fatigue, we recorded real-time data on (1) the evolution of correct and incorrect answers, as well as (2) the moments when visual or auditory stimuli were produced by the system, relative to both the audio file used as a basis for the effects and the start time of the trial. Recording the audio file time and the real-time trial time allows for correlating and filtering trials with either a bootstrap or slowdown issue; trials without audio would also use an audio file, but the sound was muted.

The next sections introduce and discuss the data extraction and processing in more detail.

3.2.1 Direct Variables and Functions. The following variables and functions can be directly extracted from the gathered data:

- $t \in T$: trial identifier t from all available trials T.
- $p \in P$: participant identifier p from all available participants P.
- $trials(p)$: all trial identifiers of a participant p.
- $participant(t)$: participant identifier of a trial t.
- $correct(t)$: number of correct sorts at the end of trial t.
- $trialLength(t)$: length of trial t in seconds.
- $reportedLength(t)$: reported length of trial t in seconds.
- $reportedFatigue(t)$: self-reported fatigue of a participant after performing trial t in an ordinal scale from 1 to 5.
- $reportedSpeedPerception(t)$: subjective participant rating of trial t's speed in an ordinal scale from 1 (slow) to 5 (fast).

3.2.2 Extracted Variables. To obtain a basis for comparison, the directly extracted participant data outlined is further normalized into the following variables:

- $correctPerSecondTrial(t)$: average number of correct answers per second during trial t.
$$\frac{correct(t)}{trialLength(t)}$$
- $correctPerSecondParticipant(p)$: average number of correct answers per second of a participant during trials.
$$\frac{\sum_{t' \in trials(participant(t))} correct(t')}{\sum_{t' \in trials(participant(t))} trialLength(t')}$$
- $correctNormalized(t)$: amount of correct answers per second of trial t normalized with 1, i.e., the average number of correct answers per second of a participant among all performed trials.
$$\frac{correctPerSecondTrial(t)}{correctPerSecondParticipant(p)}$$

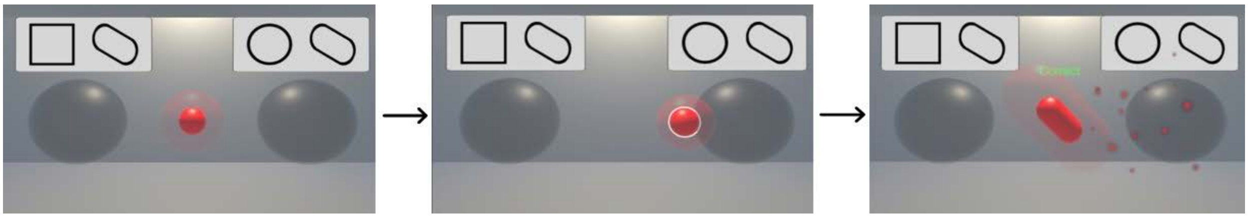


Figure 1: Example sequence of the experiment's sorting trial task, where users move geometric objects into matching sinks.

- $secondBias(p)$: ratio of the total of seconds of a participant p's trials and the total reported time, defining what the participant considers a second.

$$\frac{\sum_{t \in trials(p)} trialLength(t)}{\sum_{t \in trials(p)} reportedLength(t)}$$

- $deltaTimePerception(t)$: averaged delta per second between reported time (accounting participant bias) and trial length of a trial t.

$$\frac{secondBias(p) * reportedLength(t)}{trialLength(t)} - 1$$

- $abs(deltaTimePerception(t))$: absolute value of trial t's time perception delta.

3.2.3 Outcome Variables. The specific variables relevant to the analysis performed in this study are:

- $deltaTimePerception(t)$: if the difference between reported time and trial time shows too much individual bias, this variable represents a participant's variation in perception. A negative value indicates that the seconds of the trial t were reported as shorter than the other trial performed by this participant. A positive value means that the seconds were reported as longer.
- $abs(deltaTimePerception(t))$: instead of denoting how much longer or shorter a second is interpreted for a trial t compared to other trials performed by a participant, the absolute value represents the magnitude of the eventual time distortion.
- $correctNormalized(t)$: to simplify the analysis, we do not take into account negative answers to evaluate performance but only the amount of correct answers. A smaller number of correct answers would indirectly reflect the number of incorrect answers due to the time lost. Similar to a participant's variation in perception, this variable represents the variation in performance instead of the pure performance, with values < 1 indicating worse and > 1 better performances.
- $reportedSpeedPerception(t)$: the subjective interpretation of whether time drags or flies while doing a trial.
- $reportedFatigue(t)$: with relatively low trial numbers, reported fatigue should mostly depend on the participants, independent of trial parameters.

4 RESULTS AND DISCUSSION

As the primary goal of the experiment is to investigate the effect of rhythmic stimuli on time perception and task performance, two dimensions are considered for the stimuli: tempo (100bpm, 140bpm, 180bpm, or 0 for no stimulus) and type (visual pulse, audio kick, no stimulus, or both). Employing analysis of variance (ANOVA), an initial series of base ANOVAs is performed between the two dimensions and the outcome variables (Table 2). Additionally, stimuli ANOVAs between outcome variables and tempo for trials under each stimulus type further help investigate the effect of the tempo of that specific stimulus type (Table 3).

	Type		Tempo	
	F	*p*	*F*	*p*
deltaTimePerception(t)	1.33	0.26	1.35	0.26
abs(deltaTimePerception(t))	2.80	0.04	2.06	0.11
correctNormalized(t)	2.99	0.03	1.30	0.27
reportedSpeedPerception(t)	0.72	0.54	7.21	0.0001
reportedFatigue(t)	0.81	0.489	0.92	0.433

Table 2: ANOVAs on stimuli types and tempos.

	Audio		Visual		Both	
	F	*p*	*F*	*p*	*F*	*p*
deltaTimePerception(t)	0.27	0.76	0.67	0.51	2.86	0.06
abs(deltaTimePerception(t))	0.26	0.77	2.18	0.12	0.63	0.54
correctNormalized(t)	0.57	0.57	0.20	0.82	0.51	0.60
reportedSpeedPerception(t)	1.76	0.18	4.20	0.02	5.98	0.004
reportedFatigue(t)	3.48	0.04	0.89	0.41	0.25	0.78

Table 3: ANOVAs under specific stimuli types for tempos.

The base ANOVAs involving $deltaTimePerception(t)$ show no significant differences; however, the ANOVA between stimulus type and $abs(deltaTimePerception(t))$ reveals significant differences ($p = 0.04$), i.e., while no influence of the stimulus type was observed on the direction of time estimation's variations, the stimulus type in the experiment influenced the magnitude of the variations, indicating a possible validation of H1. Table 4 comprises a deeper analysis of this ANOVA using Tukey's range test (Tukey HSD), showing that compared to the absence of stimuli, each stimulus type leads to more substantial distortions (according to the diff value), thus validating H1a and H1b for the time estimation dimension of time perception. However, this is a near-significant tendency ($p < 0.1$), and only for single stimuli, the diff value is also in disfavor of the "Both" stimuli type compared to the single ones. This indicates that the combination of stimulus modalities (auditory and visual) potentially reduces the magnitude of time estimation variations under the influence of rhythmic stimuli.

	diff	lwr	upr	p adj
Both-AudioOnly	-0.04	-0.11	0.03	0.48
None-AudioOnly	-0.10	-0.20	0.001	0.053
VisualsOnly-AudioOnly	-0.003	-0.08	0.07	0.99
None-Both	-0.06	-0.15	0.04	0.43
VisualsOnly-Both	0.04	-0.04	0.11	0.54
VisualsOnly-None	0.10	-0.003	0.19	0.063

Table 4: Tukey HSD of ANOVA between stimulus type and abs(deltaTimePerception(t)).

As the stimulus type affects the magnitude of the variation of the time estimations, the absence of significant results in tempo might be due to these differences in stimulus type, which is why we also performed ANOVAs between time estimation variables and tempo (see Table 3). Here, differences appear between tempos on time estimations, revealing a near-significant tendency ($p = 0.06$) with the simultaneous use of audio and visuals with the variable $deltaTimePerception(t)$ but not with its absolute value. This suggests that when under the influence of both stimulus types, the tempo can modulate the direction of the subjective time estimation errors/variations but not the magnitude.

Further analysis employing Tukey HSD (see Table 5) shows that the difference is a tendency only between 180bpm and 140bpm (diff=-0.14, p=0.06); a tempo of 100bpm might not be enough for the rhythmic property to affect time estimation while the diff value suggests that the highest bpm value (180bpm) was rated as faster than a lower value of 140bpm for time estimation. These observations validate H2 for the time estimation dimension of time perception but only under the specific condition of the multi-modal stimuli while invalidating it for the single-modal ones.

	diff	lwr	upr	p adj
140-100	-0.11	-0.04	0.25	0.20
180-100	-0.03	-0.18	0.12	0.88
180-140	-0.14	-0.28	0.005	0.06

Table 5: Tukey HSD of ANOVA between tempo and correctNormalized(t).

The results regarding $abs(deltaTimePerception(t))$ indicate significant differences among stimulus types but not among tempos, while the opposite holds for the base ANOVAs involving $reportedSpeedPerception(t)$, thus invalidating H1 but validating H2 for the time passage aspect of time perception. From the analysis employing Tukey HSD (Table 6), the differences among tempos regarding $reportedSpeedPerception(t)$ are most significant between 180-100 (p=0.00005, diff=0.70), with tendencies for 180-0 (p=0.06, diff=0.50), 180-140 (p=0.102, diff=0.34) and 140-100 (p=0.102, diff=0.36). The different values among non-0 tempos (0 implying no stimuli and thus, no rhythm) suggest that higher tempos are rated faster than lower tempos. The stimuli ANOVAs suggest that this effect of tempo is significant for visual stimuli (p=0.02) and combined stimuli (p=0.004) but not necessarily for audio alone (p=0.18).

If the $correctNormalized(t)$ variable appears uncorrelated to tempo according to the base ANOVA, thus invalidating H4b, the significant p-value (p=0.03) between that variable and stimulus type indicates the effect of stimuli type on performance. Further analysis (Table 7) indicates tendencies of visual-only stimuli to decrease performances compared to audio-only (diff=-0.03, p=0.097) and no stimuli (diff=-0.45, p=0.056), which validates H4a in the case of visual stimuli but not H4c as for the "Both" stimuli no significant differences are observed to other stimuli types.

	diff	lwr	upr	p adj
100-0	-0.20	-0.73	0.33	0.77
140-0	0.16	-0.36	0.68	0.85
180-0	0.50	-0.02	1.01	0.06
140-100	0.36	-0.05	0.77	0.102
180-100	0.70	-0.30	1.10	0.00005
180-140	0.34	-0.04	0.71	0.102

Table 6: Tukey HSD of ANOVA between tempo and reportedSpeedPerception(t).

A t-test between the $correctNormalized(t)$ variable and two classes considering either the presence or the absence of visual stimuli shows a significant decrease in performance in the group with visual stimuli (p=0.006, mean in group with visual stimuli is 0.997, mean in group without is 1.025).

Since, like for the $deltaTimePerception(t)$, there is an effect of the stimulus type on the performance, we also performed ANOVAs between $correctNormalized(t)$ and tempo for each type (Table 3) which did not reveal significant results.

	diff	lwr	upr	p adj
Both-AudioOnly	-0.02	-0.05	0.02	0.55
None-AudioOnly	0.01	-0.03	0.06	0.87
VisualsOnly-AudioOnly	-0.03	-0.06	0.004	0.097
None-Both	0.03	-0.014	0.08	0.28
VisualsOnly-Both	-0.014	-0.05	0.02	0.73
VisualsOnly-None	-0.045	-0.09	0.001	0.056

Table 7: Tukey HSD of ANOVA between stimulus type and correctNormalized(t).

The absence of stimuli or tempo effects on the reported fatigue is expected since fatigue depends on the participant's well-being prior to the experiment.

Correlations between the various outcome variables are shown in Table 8. $deltaTimePerception(t)$ and $reportedFatigue(t)$ being uncorrelated, while a correlation (p=0.001, rho=0.175) between $abs(deltaTimePerception(t))$ and $reportedFatigue(t)$ exists, indicates that fatigue does not influence the direction of the subjective time distortions but influences their magnitude. A tendency yet non-significant correlation (p=0.06, rho=0.104) is observable between $reportedSpeedPerception(t)$ and $correctNormalized(t)$, suggesting that the speed judgment is influenced by a participant's performance, even though scores were not shown during trials.

4.1 Limitations

One potential limitation of this study stems from the reported fatigue. As Table 3 indicates, the fatigue significantly varies among tempos under the audio stimuli, which might imply that the random

	[1]	[2]	[3]	[4]	[5]
[1] deltaTimePerception(t)	-	-	p=0.80	p=0.74	p=0.40
[2] abs(deltaTimePerception(t))	-	-	p=0.51	p=0.81	p=0.001 rho=0.175
[3] correctNormalized(t)	p=0.80	p=0.51	-	p=0.06, rho=0.104	p=0.341
[4] reportedSpeedPerception(t)	p=0.74	p=0.81	p=0.06, rho=0.104	-	p=0.16
[5] reportedFatigue(t)	p=0.40	p=0.001 rho=0.175	p=0.341	p=0.16	-

Table 8: Pearson correlation coefficients; Spearman if reportedSpeedPerception(t) or reportedFatigue(t) are involved.

assignment of the most tired participants happened to be to a specific tempo with tasks under audio stimuli. The Tukey HSD of this ANOVA shows that the difference is significant between 140bpm and 100bpm; thus, the results might be biased under these conditions. Due to the online nature of the study, we can assume that the hardware and displays used by the participants varied, which in turn might have caused a bias in the perception of the stimuli (e.g., hardware delay for display and audio, or screen solution). Even in the worst-case ANOVA scenario, we still have a post-hoc statistical power >0.95 for the statistically significant results (effect size 0.4, total sample size 102, 3 groups); however, there are some disparities among trials (cf. Table 1) per stimulus mode due to random trial assignment and data discarded due to bootstrap issues.

5 CONCLUSION

This paper discussed an experiment exploring the effect of rhythmic stimuli on a cognitive task's experience in a virtual environment regarding time perception and task performance. The results suggest that the different types of stimuli (audio, visual, or combined) have varying effects on time estimations. Single-type stimuli lead to higher variations, but combined stimuli can influence the direction of these variations depending on the tempo of the stimuli. However, no effect of tempo on time estimations is observed across all types of stimuli. When it comes to reported time passage being perceived as slow or fast, no general effect of stimulus type is observed, but an effect of tempo, which appears only in conjunction with visual stimuli. However, the same visual stimuli appear to distract enough to lower the task performance among the participants. A general limitation is the constraints and hardware variations due to the online nature of the study. As some results indicated only tendencies, further lab-based studies are necessary to confirm these implications. To this end, a follow-up study reproducing this experiment in VR while monitoring physiological signals to confirm these results and explore if the observed changes in time perception have direct physiological implications is currently in preparation. Another possible extension to the experimental setup is incorporating more complex stimuli dimensions beyond modality and tempo, which might affect time perception differently depending on the employed audio-visual stimuli and rhythmic patterns. Regarding time perception research, the study presented in this paper already highlights the potential of employing rhythmic stimuli to influence specific aspects such as time estimation variation and time judgment. More generally, such active and purposeful rhythmic stimuli modulation to tailor individual time experiences can create novel and exciting opportunities for designing immersive virtual environments.

ACKNOWLEDGMENTS

This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964464.

REFERENCES

[1] Jean Botev, Knut Drewing, Heiko Hamann, Yara Khaluf, Pieter Simoens, and Vatakis Argiro. 2021. ChronoPilot – Modulating Time Perception. In *Proceedings of the 4th IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR 2021)*. IEEE, 215–218. https://doi.org/10.1109/AIVR52153.2021.00049

[2] Sylvie Droit-Volet, Ahmed El-Azhari, Samuel Haddar, Rémi Drago, and Sandrine Gil. 2020. Similar time distortions under the effect of emotion for durations of several minutes and a few seconds. *Acta Psychologica* 210 (2020), 103170. https://doi.org/10.1016/j.actpsy.2020.103170

[3] Sylvie Droit-Volet, Sophie Fayolle, Mathilde Lamotte, and Sandrine Gil. 2013. Time, Emotion and the Embodiment of Timing. *Timing & Time Perception* 1, 1 (2013), 99–126. https://doi.org/10.1163/22134468-00002004

[4] Sylvie Droit-Volet, Danilo Ramos, José L. O. Bueno, and Emmanuel Bigand. 2013. Music, emotion, and time perception: the influence of subjective emotional valence and arousal? *Frontiers in Psychology* 4 (2013), 417. https://doi.org/10.3389/fpsyg.2013.00417

[5] Amirhossein Ghaderi, Matthias Niemeier, and John Douglas Crawford. 2021. Linear vector models of time perception account for saccade and stimulus novelty interactions. (2021). https://doi.org/10.1101/2020.07.13.201087

[6] Andrei Gorea. 2011. Ticks per thought or thoughts per tick? A selective review of time perception with hints on future research. *Journal of Physiology-Paris* 105, 4 (2011), 153–163. https://doi.org/10.1016/j.jphysparis.2011.09.008

[7] David Hammerschmidt, Clemens Wöllner, Justin London, and Birgitta Burger. 2021. Disco Time: The Relationship Between Perceived Duration and Tempo in Music. *Music & Science* 4 (2021), 2059204320986384. https://doi.org/10.1177/2059204320986384

[8] Charlotte Harrison, Nicola Binetti, Isabelle Mareschal, and Alan Johnston. 2017. Time-Order Errors in Duration Judgment Are Independent of Spatial Positioning. *Frontiers in Psychology* 8 (March 2017), 340. https://doi.org/10.3389/fpsyg.2017.00340

[9] Hudson Hoagland. 1933. The Physiological Control of Judgments of Duration: Evidence for a Chemical Clock. *The Journal of General Psychology* 9, 2 (Oct. 1933), 267–287. https://doi.org/10.1080/00221309.1933.9920937

[10] Susan A. Jackson and Robert C. Eklund. 2002. Assessing Flow in Physical Activity: The Flow State Scale–2 and Dispositional Flow Scale–2. *Journal of Sport and Exercise Psychology* 24, 2 (June 2002), 133–150. https://doi.org/10.1123/jsep.24.2.133

[11] Wouter Kruijne, Christian N. L. Olivers, and Hedderik van Rijn. 2021. Neural Repetition Suppression Modulates Time Perception: Evidence From Electrophysiology and Pupillometry. *Journal of Cognitive Neuroscience* 33, 7 (06 2021), 1230–1252. https://doi.org/10.1162/jocn_a_01705

[12] Li-Ann Leow, Taylor Parrott, and Jessica A. Grahn. 2014. Individual Differences in Beat Perception Affect Gait Responses to Low- and High-Groove Music. *Frontiers in Human Neuroscience* 8 (Oct. 2014), 811. https://doi.org/10.3389/fnhum.2014.00811

[13] Sarah Lukas, Andrea M. Philipp, and Iring Koch. 2014. Crossmodal attention switching: Auditory dominance in temporal discrimination tasks. *Acta Psychologica* 153 (2014), 139–146. https://doi.org/10.1016/j.actpsy.2014.10.003

[14] Mukesh Makwana and Narayanan Srinivasan. 2017. Intended outcome expands in time. *Scientific Reports* 7, 1 (July 2017), 6305. https://doi.org/10.1038/s41598-017-05803-1

[15] Hans Rutrecht, Marc Wittmann, Shiva Khoshnoud, and Federico Alvarez Igarzábal. 2021. Time Speeds Up During Flow States: A Study in Virtual Reality with the Video Game Thumper. *Timing & Time Perception* 9, 4 (2021), 353–376. https://doi.org/10.1163/22134468-bja10033

[16] Jan Stupacher. 2019. The experience of flow during sensorimotor synchronization to musical rhythms. *Musicae Scientiae* 23, 3 (July 2019), 348–361. https://doi.org/10.1177/1029864919836720

[17] Jan Stupacher, Guilherme Wood, and Matthias Witte. 2017. Synchrony and sympathy: Social entrainment with music compared to a metronome. *Psychomusicology: Music, Mind, and Brain* 27, 3 (Sept. 2017), 158–166. https://doi.org/10.1037/pmu0000181

[18] Xinyue Wang, Clemens Wöllner, and Zhuanghua Shi. 2021. Perceiving Tempo in Incongruent Audiovisual Presentations of Human Motion: Evidence for a Visual Driving Effect. *Timing & Time Perception* 10, 1 (2021), 75–95. https://doi.org/10.1163/22134468-bja10036

[19] John D. Watt. 1991. Effect of Boredom Proneness on Time Perception. *Psychological Reports* 69, 1 (1991), 323–327. https://doi.org/10.2466/pr0.1991.69.1.323 PMID: 1961817.

[20] Joanne E. Wittwer, Kate E. Webster, and Keith Hill. 2013. Music and metronome cues produce different effects on gait spatiotemporal measures but not gait variability in healthy older adults. *Gait & Posture* 37, 2 (Feb. 2013), 219–222. https://doi.org/10.1016/j.gaitpost.2012.07.006

[21] Clemens Wöllner and David Hammerschmidt. 2021. Tapping to hip-hop: Effects of cognitive load, arousal, and musical meter on time experiences. *Attention, Perception, & Psychophysics* 83, 4 (Jan. 2021), 1552–1561. https://doi.org/10.3758/s13414-020-02227-4

[22] Shan Xu and Prabu David. 2018. Distortions in time perceptions during task switching. *Computers in Human Behavior* 80 (2018), 362–369. https://doi.org/10.1016/j.chb.2017.11.032

Appendix B

Paper: Rhythmic Stimuli and Time Experience in Virtual Reality

Full paper of Section 3.1.2

Rhythmic Stimuli and Time Experience in Virtual Reality

Stéven Picard( ) and Jean Botev

University of Luxembourg, 4364 Esch-sur-Alzette, Luxembourg
{steven.picard,jean.botev}@uni.lu
https://vrarlab.uni.lu

Abstract. Time experience is an essential part of one's perception of any environment, real or virtual. In this paper, from a virtual environment design perspective, we explore how rhythmic stimuli can influence an unrelated cognitive task regarding time experience and performance in virtual reality. The task involves sorting 3D objects by shape, with varying rhythmic stimuli in terms of their tempo and sensory channel (auditory and/or visual) in different trials, to collect subjective measures of time estimation and judgment. The results indicate different effects on time experience and performance depending on the context, such as user fatigue and trial repetition. Depending on the context, a positive impact of audio stimuli or a negative impact of visual stimuli on task performance can be observed, as well as time being underestimated concerning tempo in relation to task familiarity. However, some effects are consistent regardless of context, such as time being judged to pass faster with additional stimuli or consistent correlations between participants' performance and time experience, suggesting flow-related aspects. This could be of great interest for designing virtual environments, as purposeful stimuli can strongly influence task performance and time experience, both essential components of virtual environment user experience.

Keywords: Virtual Reality · UX · Time Experience · Rhythmic Stimuli

1 Introduction

While time itself is a concept, it is also something humans can perceive. However, the perception of time is subjective, and this experience is an integral part of the overall experience of any environment, with virtual environments being no exception. Therefore, acknowledging this in their conception and actively devising virtual environments to modulate the time experience of users would be an exciting design instrument. Time perception as an interdisciplinary topic is explored in numerous scientific studies in disciplines as diverse as psychology or

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964464 (ChronoPilot).

G. Zachmann et al. (Eds.): EuroXR 2023, LNCS 14410, pp. 53–75, 2023.
https://doi.org/10.1007/978-3-031-48495-7_4

neuroscience. In a previous study combining cognitive science and computer science, we already examined the relationships between time perception and rhythmic stimuli with a sorting game in a two-dimensional setting, which revealed varying time experience and performance effects depending on whether single or combined stimuli were used in relation to their tempo [19]. However, that initial experiment was limited to a crowd-sourced desktop setting, and we adapted the experiment for Virtual Reality (VR), which allows for extended control of the test environment plus extending the initial set of questions to include stimuli and time experience aspects in fully immersive environments. Therefore, while our main goal with this study is to investigate and interpret anything significant by having a correlation study process, we do come with assumptions coming from our initial study, which are different effects on time perception depending on the type of stimuli present (audio, visual, or both), tempo-related time estimation errors for combined stimuli, as well as decreased task performance by the presence of visual stimuli. Before detailing the experiment in Sect. 3, analyzing the data in Sect. 4, and discussing the results in Sect. 5, we will first provide the necessary background and review related work in the following Sect. 2. Section 6 then concludes the paper with an outlook and potential impact of our findings on general virtual environment design.

2 Background and Related Work

The most common time perception model in literature is the clock model, which assumes an "internal clock" as a body system dedicated to time perception [9]. This system is usually tied to a model producing "ticks" or "units", such as an oscillator or a pacemaker model, where the body keeps track of time by counting these ticks. In these models, it is assumed that time perception can be changed either through the speed at which these ticks are produced or by skipping some, with these effects potentially resulting from external stimuli unrelated to time. Counting the ticks can be delegated to attention, making attentional resources a key element of time perception [5,6]. Attention is believed to act as a gate or switch on accounting the time units, where paying less attention to time will result in compressed time experience as time units are likely to be skipped. Another source of subjective temporal distortions can be found in the use of arousal levels, which are believed to affect the clock speed, resulting in an altered time experience [2,3,13]. However, using an internal clock model is not necessary to predict time perception accurately [22]: on the context of watching videos in different scenarios, time perception was accurately predicted by a classification network using changes in perceptual content and visual spatial attention (more specifically, gaze position). Nevertheless, it is a natural way to interpret time experience based on the focus on attention and arousal that gives initial directions to time perception studies.

In the literature, we can observe different types of timing tasks, which involve different processing mechanics. The most common aspects are time estimation (i.e., asking for a duration estimate of an event with units, such as seconds) and the feeling of time passage (i.e., judging if time passes by quickly or slowly), which is also referred to as time judgment.

This difference can be observed in depressed subjects underestimating time but judging it as passing slowly [3], with boredom-prone people judging boring tasks as passing slowly but not necessarily overestimating those [27], or with players of the game "Thumper" reporting faster time passing without time estimation errors [23].

Having defined time perception, we can discuss the state of "flow", one of the applications of time perception alterations. Flow is a psychological state of full attention on a task defined by Csikzentmihalyi, represented through nine dimensions: challenge-skill balance, action-awareness merging, clear goals, unambiguous feedback, concentration on task, sense of control, loss of self-consciousness, time transformation, and autotelic experience. The psychological state of flow is a research subject in itself, centered around one's relation to a task as it primarily relies on the challenge-skill ratio aspect [11]; it is often considered a desirable state, and time perception alterations are one of its manifestations. In social media, the manifestation of flow seems to be influenced by the positive effect of telepresence on enjoyment, concentration, challenge, and curiosity; flow would then influence the presence of time distortions [17]. Delving further into social context, it was found that the concentration and time distortion components of flow, but not enjoyment, were affected by working in a group of two compared to working individually in virtual worlds (within the social game platform Second Life) [16]. More in line with our work, several studies on flow and VR have been conducted. The previously mentioned study on Thumper compared flow states between VR and non-VR setups, finding that despite VR's technical immersion, both scenarios could lead to a flow state [23]. Within VR activities, a model ostensibly associates flow and playfulness, defined by a combination of intrinsic motivation, control, and freedom to suspend reality, this association then influences competence in the activity [21].

VR studies on time perception, however, are not limited to flow. VR itself affects our senses due to what is being conveyed through sensory channels, but also due to the devices used and the eventual physical discomfort we can get from it. Simply comparing the time perception of the same game in a VR versus a desktop setup leads to an underestimation bias for VR [15]. It also seems that time perception changes when bored or waiting in VR compared to real life [10]. In another simple study about time perception comparing the time perception between VR and desktop while doing a task ranging from 30 s to 5 min, it was observed that if both situations yielded time overestimation, the VR scenario was overestimated more [14]. However, technological immersion alone might not be a sufficient explanation, as walking in VR does not seem to affect time perception significantly [1]. In another experiment about zeitgeber on time perception while doing a task conducted both in VR and on desktop, no significant difference was observed regarding time perception, but there was a difference in task performance (with the VR participant performing better) [24]. The Thumper study also observed the effect of performing better in a VR setup compared to a non-VR one [23]. When it comes to the effect of emotional content, VR itself appears not to yield any difference to real life in time distortions when the emotional content is the same [7]. Employing VR entails possibly unique content, such as having movements represented by an avatar. Differences were observed between avatar and no-avatar conditions where avatar presence leads to, in a retrospective paradigm, a significantly faster passage of time without an effect on time estimations [25].

When considering VR and time perception, we can thus regard both the technological immersion, i.e., the effect of being in VR through its dedicated hardware, as well as the virtual environment stimuli and transformations that can be induced through VR. A specific stimulus type we want to employ in VR is rhythmic stimuli, which we already identified as having notable effects in a desktop scenario [19]. Rhythmic stimuli and music generally have a high potential to modulate one's time experience. With music, it was observed that higher tempos induce longer subjective time, but emotional valence decreases (but does not suppress) the effect of tempo and affects time perception. These effects on time perception might be due to their effect on arousal; interestingly, using a different orchestration (piano only or full orchestra) does not affect time judgment and pleasantness while affecting arousal [4]. On timing evaluations of instrumental excerpts of Disco songs (including estimation and judgment of time passage) over different tempos, it was observed that faster tempos were correlated to longer reproduction duration; however, no effect on estimations was observed alongside the necessity of a tempo difference of at least 20 BPM required for timing measurement differences to appear [8]. By varying cognitive load through tasks and arousal levels through music choice while keeping music tempo constant, it was found that (1) time is judged as passing faster under higher cognitive load (presence of a math task), (2) presence of a concurrent motor task (tapping the music's tempo) yield shorter subjective durations, and (3) for the motor task, for the same music tapping to half notes instead of eighth notes ended up with smaller time estimations and time passage rated as faster [28]. Regarding timing and spatial movement, rhythmic auditory stimuli (RAS) have been observed to improve motor performance when vision is unavailable [18].

Rhythm, however, is not only tied to music and audio. Concerning temporal judgments, audio was believed to be dominant over visuals [12]. However, a later study found visual stimuli dominance using Point-Light-Display (PLD) dance motions compared to simple audio tempos [26]. Participants were presented with both the dance motion and the audio tempo and had to give a globally fitting tempo. The result of this study suggests that under the right conditions, visual stimuli can dominate audio in terms of timing, which may be due to the quantity of temporal information on a sensory channel rather than the channel itself.

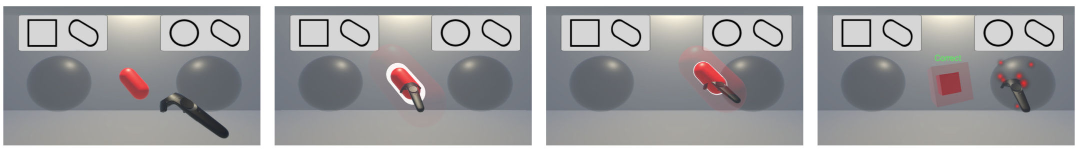
3 Experiment

We recruited 30 participants from a public, science-related event in Luxembourg City and from students and staff at the University of Luxembourg. The female-to-male ratio was 46.66%–53.33%, with ages ranging from 19 to 45 (mean age

25.63/median 24). Participants were received at the VR/AR Lab's test space at the University of Luxembourg and briefed about the experiment before the session, both verbally and via an informed consent form, in which we also collected demographic data. After the setup and familiarization phase, the participants performed the tasks, with each participant able to take breaks between trials if desired. Each session lasted a total of approximately one hour.

3.1 Trial Task and Design

Participants had to complete trials in which they had to sort three-dimensional objects according to their shape (spheres, capsules, or cubes). As shown in the screenshot sequence in Fig. 1, the objects must be grabbed with a VR controller and dragged into one of the two larger sinks, which only accept specific shapes displayed on a scoreboard above them (cf. Fig. 1a–c). Once sorted, an object disappears with a small animation, and a new object to be sorted appears in the center of the virtual environment (cf. Fig. 1d).



(a) (b) (c) (d)

Fig. 1. Sorting example from trial task, in sequential order from left to right.

The sorting attempt has a predefined duration unknown to participants. After this duration, the experiment ends with a questionnaire in which the participants are asked to estimate the time in seconds and rate on two Likert scales how fast the experiment felt and how tired they were (cf. Fig. 2).

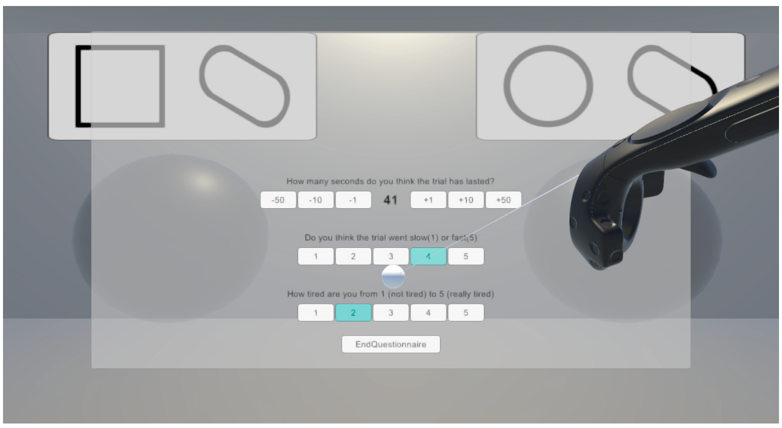


Fig. 2. Post-trial questionnaire on time estimation, time perception, and fatigue.

Trials were subjected to conditions that were a combination of these parameters:

- *trial length*: how long the trial lasted in seconds (either 40, 50, or 60)
- *tempo*: the rhythm of the audio/visual stimuli if present in beats per minute (either 100, 140, or 180; forced to 0 without stimuli)
- *visual stimuli*: whether or not there are flashing visual pulses around the object
- *audio stimuli*: whether or not metronome click sounds are produced

Each participant went through all 30 possible combinations in random order.

3.2 Technical Specifications

Participants used the VIVE Pro Eye head-mounted display (HMD) with one of its controllers, allowing them to enter VR and control virtual objects with six degrees of freedom (DoF). This particular HMD also allows for collecting precise eye-tracking data during the trials, specifically on gaze and pupil dilation. During the experiments, participants were also instructed to wear an Empatica E4 wristband to record further physiological data; our experiment included heart rate variability and skin temperature. Our custom application reads the data by receiving messages from Empatica's E4 Streaming Server software. However, the physiological data is out of the scope of this paper and will be discussed in a separate publication. The experimental application itself was developed in Unity and used SteamVR. In addition to the application, OVR advanced settings were employed to adjust the participant's height position.

4 Analysis and Results

4.1 General Methods and Data

As discussed previously, various data were collected from each trial; in this subsection, we will describe what data we effectively used for our analysis and how.

Disclaimed Trials. Some trials were disclaimed from the data set depending on our notes during trials. The reasons for removing trials were:

- Misunderstanding of task controls by the participant.
- Misunderstanding of trial questionnaire by the participant (verbally checked when giving incoherent values such as negative time estimation).
- Disturbance or interruption during the trial, either from the participants themselves (i.e., asking a question or talking during a trial) or external sources (i.e., technical issues, noise from a nearby room).

Variables/Functions. Variables and functions extracted from trial data are:

- $t \in T$: trial identifier t from all available trials T.
- $p \in P$: participant identifier p from all available participants P.
- *trials(p)*: all trial identifiers of a participant p.
- *participant(t)*: participant identifier of a trial t.
- *correct(t)*: number of correct sorts at the end of trial t.
- *trialLength(t)*: length of trial t in seconds.
- *reportedLength(t)*: reported length of trial t in seconds.
- *reportedFatigue(t)*: self-reported fatigue of a participant after performing trial t in an ordinal scale from 1 to 5.
- *reportedSpeedPerception(t)*: subjective participant rating of trial t's speed in an ordinal scale from 1 (slow) to 5 (fast).
- *trialIndex(t)*: the index of trial t, indicates how many trials were performed before t and, therefore, global repetition from the experiment session.

Since *reportedSpeedPerception(t)* and *reportedFatigue(t)* are purely subjective questions to the participant, we can use these values directly. However *correct(t)*, the task performance direct variable depends on the participants' individual performance and *reportedLength(t)*, while the direct time estimation variable depends on both the trial's duration as well as the participants' individual representation of a second. Therefore, these two variables need to undergo a normalization process. Normalizing the performance variable (*correct(t)*) for our analysis goes through a three-step process involving the following extracted variables:

- *correctPerSecondTrial(t)*: average number of correct answers per second during trial t.
$$\frac{correct(t)}{trialLength(t)}$$
- *correctPerSecondParticipant(p)*: average number of correct answers per second of a participant during trials.
$$\frac{\sum_{t' \in trials(participant(t))} correct(t')}{\sum_{t' \in trials(participant(t))} trialLength(t')}$$
- *correctNormalized(t)*: amount of correct answers per second of trial t normalized with 1, i.e., the average number of correct answers per second of a participant among all performed trials.
$$\frac{correctPerSecondTrial(t)}{correctPerSecondParticipant(p)}$$

For time estimation (*reportedLength(t)*), we employed the following process:

- *secondBias(p)*: ratio of the total of seconds of a participant p's trials and the total reported time, defining what the participant considers a second.
$$\frac{\sum_{t \in trials(p)} trialLength(t)}{\sum_{t \in trials(p)} reportedLength(t)}$$

- *deltaTimePerception(t)*: averaged delta per second between reported time (accounting participant bias) and trial length of a trial t.

$$\frac{secondBias(p) * reportedLength(t)}{trialLength(t)} - 1$$

However, in addition to *deltaTimePerception(t)* we also use its absolute *abs(deltaTimePerception(t))* as it represents the magnitude of time perception delta of a trial.

Outcome Variables. The specific variables relevant to the analysis performed in this study are:

- *deltaTimePerception(t)*: if the difference between reported time and trial time shows too much individual bias, this variable represents a participant's variation in perception. A negative value indicates that the seconds of the trial t were reported as shorter than the other trial performed by this participant. A positive value means that the seconds were reported as longer.
- *abs(deltaTimePerception(t))*: instead of denoting how much longer or shorter a second is interpreted for a trial t compared to other trials performed by a participant, the absolute value represents the magnitude of the eventual time distortion.
- *correctNormalized(t)*: to simplify the analysis, we do not take into account negative answers to evaluate performance but only the amount of correct answers. A smaller number of correct answers would indirectly reflect the number of incorrect answers due to the time lost. Similar to a participant's variation in perception, this variable represents the variation in performance instead of the pure performance, with values <1 indicating worse and >1 better performances.
- *reportedSpeedPerception(t)*: the subjective interpretation of whether time drags or flies after performing trial t in an ordinal scale from 1 to 5.
- *reportedFatigue(t)*: with relatively low trial numbers, reported fatigue should mostly depend on the participants, independent of trial parameters.
- *trialIndex(t)*: the index of trial t, indicates how many trials were performed before t and, thus, is an indicator of global repetition from the experiment session.

Parameters. The parameters used in this experiment are:

- *stimulusTrial(t)*: the type of stimulus used in trial t, possible values are: *None, VisualsOnly, AudioOnly, Both*.
- *hasAudioTrial(t)*: whether or not the trial t contains an audio stimulus.
- *hasVisualTrial(t)*: whether or not the trial t contains a visual stimulus.
- *hasStimulusTrial(t)*: whether or not the trial t contains any type of stimulus.
- *tempo(t)*: the tempo in beats per minute (BPM) of a trial t, possible values are: 0, 100, 140, 180. a tempo of 0 means that the trial had no stimuli.

Trial Filters. When performing analyses, we may want to include only subsets of the trials to investigate specific effects. The filters used are:

– *filterAUDIOONLY*: considers only trials that only have an audio stimulus, is equivalent to saying *t where stimulusTrial(t) == AudioOnly*.
– *filterVISUALSONLY*: considers only trials that have a visual stimulus, is equivalent to saying *t where stimulusTrial(t) == VisualsOnly*.
– *filterBOTH*: considers only trials that have both audio and visual stimuli, is equivalent to saying *t where stimulusTrial(t) == BOTH*.

Performed Tests. Using the above variables, parameters, and filters, we performed statistical tests on various data subsets to examine the effect of a stimuli presence on performance and time estimation with the following parameters:

– *stimulusTrial(t)*: ANOVAs on *deltaTimePerception(t)*, *abs(deltaTimePerception(t))* and *correctNormalized(t)* to see if any significant difference appears between possible stimuli situations.
– *hasAudioTrial(t), hasVisualTrial(t), hasStimulusTrial(t)*: t-tests on *deltaTimePerception(t), abs(deltaTimePerception(t))* and *correctNormalized(t)* to see if there is an effect on the presence or absence of a specific stimulus (since our t-tests are not pairwise, no t-test corrections have been performed).

To examine this effect on time judgment, we did the following as the time judgment variable is ordinal:

– *stimulusTrial(t)*: Kruskal-Wallis test on *reportedSpeedPerception(t)* to see if any significant difference appears between possible stimuli situations.
– *hasAudioTrial(t), hasVisualTrial(t), hasStimulusTrial(t)*: Wilcoxon test on *reportedSpeedPerception(t)* to see if there is an effect on the presence or absence of a specific stimulus.

In order to observe the effect of tempo on performance and time estimation, we performed ANOVAs between *tempo(t)* and the variables *deltaTimePerception(t), abs(deltaTimePerception(t))* and *correctNormalized(t)*. The ANOVAs were also repeated across the filters *filterAUDIOONLY, filterVISUALSONLY* and *filterBOTH* to see if differences in tempo appear only within stimuli conditions. As for the effect of tempo on time judgment (*reportedSpeedPerception(t)*), here we again replaced the ANOVAs with Kruskal-Wallis tests, including the repeated ones under filters. Correlations between time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) and performance (*correctNormalized(t)*) were investigated with Pearson tests. As the time judgment variable (*reportedSpeedPerception(t)*) is ordinal, correlation with time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) as well as performance (*correctNormalized(t)*) was made through Pearson tests. The confounding effect of fatigue was investigated by considering *reportedFatigue(t)* both as a nominal and as an ordinal variable; the former allows us to eventually observe differences between specific ratings and has been considered through ANOVAs with time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) and performance (*correctNormalized(t)*), the latter was considered through Spearman tests on time estimation variables, performance and time judgment (*reportedSpeedPerception(t)*). Regarding the confounding effect of trial index, *trialIndex(t)* is an ordinal data so it has been investigated with Spearman tests on time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*), performance (*correctNormalized(t)*) as well as time judgment (*reportedSpeedPerception(t)*). Each ANOVA with a p-value below 0.1 would lead to a subsequent Tukey HSD, Kruskal-Wallis tests would lead to subsequent paired Wilcoxon tests. Table 1 provides an overview of the different effects per subset with a significant p-value or tendency, while each is discussed in detail in the following sections. The complete data from our tests, including confidence intervals and average values, are available online [20].

Table 1. Subsets for which a significant p-value (•) or tendency (○) is observed for a combination of stimulus dimension and outcome variable group.

	Performance	Time Estimation	Time Judgment
Presence	—	—	• Full Set • Trials 11-20 ○ Fatigue Levels 1-2 • Fatigue Levels 3-4-5
Type	• Full Set • Trials 1-10 • Trials 21-30 • Fatigue Levels 3-4-5	—	• Full Set • Trials 11-20 • Fatigue Levels 1-2
Tempo	○ Full Set • Trials 11-20 • Fatigue Levels 1-2	• Full Set ○ Trials 11-20 ○ Fatigue Levels 3-4-5	• Full Set • Fatigue Levels 1-2

4.2 Across All Trials

Effects of Stimuli on Performance. One of the aims of this study is to investigate the effects of stimuli on task performance. Looking at performance across all trials, the ANOVA between task performance (*correctNormalized(t)*) and stimulus type (*stimulusTrial(t)*) revealed no significant difference. However, when performing a t-test between task performance (*correctNormalized(t)*) and the presence of visual stimuli (*hasVisualTrial(t)*), a significant difference ($p = 0.027$) can be observed alongside decreased performance when a visual stimulus is involved, as the mean with stimulus is lesser than without. No effect is observed when considering the t-test with the presence of an audio stimulus (*hasAudioTrial(t)*, $p = 0.167$) or any stimulus (*hasStimuluslTrial(t)*, $p = 0.431$).

Therefore, we can only observe a decrease in performance due to the presence of a visual stimulus but no effect on performance from the sole presence of any or of an audio stimulus. When stimuli have the dimension of type, they also have the dimension of tempo. The ANOVA between participant performance (*correctNormalized(t)*) and stimuli tempo (*tempo(t)*) generally finds no effect of tempo. However, tempo might have an effect under a specific stimulus type. Therefore, we performed the same ANOVA but only considering subsets of data where trials contained either audio stimuli only (*filterAUDIOONLY*), visual stimuli only (*filterVISUALSONLY*), or both simultaneously (*filterBOTH*). We can then observe a tendency when trials have audio stimuli only (*filterAUDIOONLY*, $p = 0.088$, $F = 2.461$). The Tukey HSD of this ANOVA reveals that the effect is significant between 180-100 ($p = 0.070$, $diff = 0.04$), with a diff value indicating the faster tempo leads to better trial performance with audio stimuli only. In the absence of interference from visual stimuli, the faster tempo for audio stimuli may have implicitly stimulated the participant to sort objects faster.

Effects of Stimuli on Time Estimation. Similarly to performance, we evaluated the effect of stimuli type and tempo on time estimation variables. Likewise, ANOVAs were effectuated regarding the type of stimuli (*stimulusTrial(t)*) and tempo (*tempo(t)*) on both the normalized time estimation error (*deltaTimePerception(t)*) and its magnitude (*abs(deltaTimePerception(t))*). The only significant result is a tendency between time estimation error (*deltaTimePerception(t)*) and tempo (*tempo(t)*) ($p = 0.073$, $F = 2.328$). Tukey's HSD of this ANOVA reveals that the effect is a tendency only between tempi of 180 and 140 ($p = 0.65$, $diff = 0.067$), with trials under a tempo of 180 being rated with a longer time per second than trials under a tempo of 140. We performed similar ANOVAs involving tempo considering subsets of data where the trials had either only audio stimuli (*filterAUDIOONLY*), only visual stimuli (*filterVISUALSONLY*), or both at the same time (*filterBOTH*). The only significant result comes from the ANOVA between time estimation error (*deltaTimePerception(t)*) and tempo (*tempo(t)*) across trials within the AudioOnly condition (*filterAUDIOONLY*) ($p = 0.039$, $F = 3.288$), where the Tukey HSD follow-up reveals a near-significant difference between tempi of 140 and 100 ($p = 0.051$, $diff = -0.106$) and a near tendency between 180 and 140 ($p = 0.107$, $diff = 0.091$). This means that in the case of trials with only an audio stimulus, trials with a BPM of 140 were evaluated as faster than others, which contradicts the analysis under all types of stimuli. This contradiction may indicate the confounded effect of tempo in time perception depending on stimuli types. Finally, t-tests between our time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) and the presence of audio stimuli (*hasAudioTrial(t)*), visual (*hasVisualTrial(t)*), or any (*hasStimuluslTrial(t)*), yielded no significant result, meaning no effect of any type of stimuli present can be observed on time estimation here.

Effects of Stimuli on Time Judgment. As the time judgment variable (*reportedSpeedPerception(t)*) is ordinal, we produced Kruskal-Wallis tests between it and the type of stimuli (*stimulusTrial(t)*) and tempo (*tempo(t)*). In the case of the test between time judgment (*reportedSpeedPerception(t)*) and the type of stimuli (*stimulusTrial(t)*), we can see a significant effect ($p = 0.016$, $chi2 = 10.267$), however a follow-up paired Wilcoxon test reveals statistical difference only between the stimulus "None" and each of the other stimuli types ($p = 0.048$ *for None-AudioOnly*, $p = 0.015$ *for None-Both*, $p = 0.015$ *for None-VisualsOnly*). As for the test on tempo (*tempo(t)*), we observe another correlation ($p = 0.001$, $chi2 = 15.432$) that, after a paired Wilcoxon, shows significant differences between 0–140 ($p = 0.006$), 0–180 ($p = 0.005$), 100–180 ($p = 0.031$) as well as a tendency between 100–140 (($p = 0.064$) and a near-tendency between 0–100 ($p = 0.12$). These two tests highlight a significant difference in time judgment depending on the presence of any stimuli (both by the differences from the "None" stimulus in the first test and the "0" BPM tempo in the second, which correspond to trials without stimuli). This is also verified by the Wilcoxon test between time judgment (*reportedSpeedPerception(t)*) and the presence of any stimulus (*hasStimuluslTrial(t)*) ($p = 0.003$, *mean(TRUE)>mean(FALSE)*), considering the mean values, we can say that the presence of a stimulus has a significant impact in making a trial judged as passing faster than one without any. The same test has been done on the presence of audio (*hasAudioTrial(t)*) ($p = 0.323$) and visuals (*hasVisualTrial(t)*) ($p = 0.023$, *mean(TRUE)>mean(FALSE)*), meaning no significant difference on the presence or not of an audio stimulus is observed but a fast-inducing effect is observed on the presence of a visual one is recorded. In the case of tempo, the results of the paired Wilcoxon discussed earlier also indicate a significant between 100BPM and other (non-0) tempi across all types of stimuli. However, running the same Kruskal-Wallis test under subsets on "AudioOnly" trials (*filterAUDIOONLY*), "VisualOnly" trials (*filterVISUALSONLY*) and trials with both (*filterBOTH(t)*) highlights a significant difference only across trials with both stimuli (*filterBOTH(t)*) ($p = 0.027$, $chi2 = 7.2343$) meaning that meanwhile tempo may have an effect across all stimuli, that effect might only be due to the combined stimuli scenario. Follow-up paired Wilcoxon tests indicate a significant difference between 100–180 ($p = 0.032$) and 140–180 ($p = 0.091$), which are the same conclusion as the tests without subsets.

Correlations Between Outcome Variables. In order to investigate the correlation between outcome variables, Spearman tests were used when time judgment (*reportedSpeedPerception(t)*) was involved as the data is ordinal; otherwise, Pearson tests were used. When comparing time estimation error (*deltaTimePerception(t)*) and performance (*correctNormalized(t)*), we see no correlation ($p = 0.218$, $cor = -0.043$), but we see a significant negative correlation with the magnitude of time estimation error (*abs(deltaTimePerception(t))*) ($p = 0.016$, $cor = -0.083$). This means the performance is correlated to the magnitude of time estimation errors but not to the direction; in other words, participants may generally be more error-prone in their estimations depending on their performance. When it comes to time judgment (*reportedSpeedPerception(t)*), it is negatively correlated to time estimation errors (*deltaTimePerception(t)*) (p = *1.724e-13*, $rho = -0.251$) and positively correlated to the magnitude of said error (*abs(deltaTimePerception(t))*) (p = *0.002*, rho = *0.107*). This means that the bigger the error, the faster the time is perceived and underestimated trials are rated at passing faster. As for time judgment (*reportedSpeedPerception(t)*) and performance (*correctNormalized(t)*), better performance is associated with faster passing trials (p = *1.265e-05*, rho = *1.50*).

Confounding Effect of Fatigue. While running the experiment, we noticed that participants were often exhausted at the end of the session. As exhaustion affects time perception and performance, we verified if it affected our outcome variables. For its effect on performance, a Spearman test between performance (*correctNormalized(t)*) and fatigue *(reportedFatigue(t))* reveals a significant correlation *(p = 2.924e-08, rho = 0.190)*. By considering the fatigue variable *(reportedFatigue(t))* nominal and performing an ANOVA with performance (*correctNormalized(t)*), we retrieve this correlation (p = *1.68e-11*, F = *14.56*), and subsequent Tukey HSD reveals that fatigue values of “3,4,5” are statistically different from values of “1,2” as the p-value is below 0.001 in all these situations, other situations (i.e., “3-4”, “1-2”, ...) have a p-value above 0.48. As for time estimations, signed error (*deltaTimePerception(t)*) is not correlated if we look through a Spearman test *(p = 0.191, rho = 0.045)*, but we retrieve statistical differences with the ANOVA *(p = 0.004, F = 3.86)*. Subsequent Tukey HSD indicates statistical differences between “3-2” *(p = 0.001)*, “5-2” *(p = 0.045)* and a tendency between “4-2” *(p = 0.06)*. No correlation is observed for the absolute error (*abs(deltaTimePerception(t))*) with both the Spearman test (p = *0.842*, rho = *0.107*) and the ANOVA (p = *0.893*, F = *0.277*), however it is observed for the Spearman test with time judgment (*reportedSpeedPerception(t)*) (p = *2.88e-11*, rho = *0.227*). From the results of the ANOVAs involving performance (*correctNormalized(t)*) and time estimation (*deltaTimePerception(t)*), we can identify two groups of reported fatigue values which are “1-2” and “3-4-5”. We thus decided to perform the same tests on subsets of our data according to these two groups on Sect. 4.4.

Confounding Effect of Trial Index. Similarly to fatigue, repeated trials can affect both performance and time perception due to learning effects and repetition. We thus evaluated correlations through Pearson tests between the number of a trial across the session *(trialIndex(t))* as a continuous variable and time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) as well as performance (*correctNormalized(t)*). When it comes to performance (*correctNormalized(t)*), the test reveals a correlation (p = *2.626e-14*, cor = *0.259*), which indicates a learning effect. Trial repetition also seems to affect time estimation as we retrieve a significant correlation with the signed time estimation error (*deltaTimePerception(t)*) (p = *1.407e-04*, cor = *0.131*) and a tendency with its absolute ($abs(deltaTimePerception(t))$) ($p$ = 0.089, cor = -0.059). We, therefore, decided to investigate different phases (beginning, middle, end) in the experiment defined by three subsets of the data based on the trial index, as shown in Fig. 3 and discussed in detail in the following Sect. 4.3.

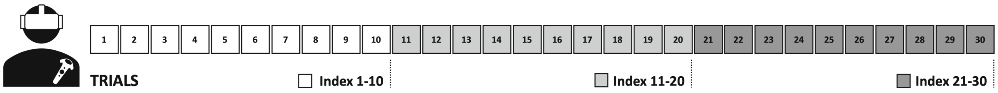


Fig. 3. Trial subset allocation for a participant by index.

4.3 Trial Index Subsets

Due to our results on the confounding effects on trial index as described in Sect. 4.2, we decided to investigate three subsets of the data based on the trial index with steps of ten (1–10, 11–20, 21–30). For each subset, we performed all the tests as on the full set of trials, which are available for download [20] and described in Sect. 4.1. However, the normalization process only considered the targeted subset when using the sum of data on trials. We go through each subset in the following subsections, focusing on the significant results.

Trials 1–10. This subset corresponds to each participant's first ten trials of the experiment, constituting a discovery phase. Regarding stimuli effects on performance, the results indicate a positive effect of audio stimulus presence (t-test performance∼audiopresence, p = 0.016; TukeyHSD performance∼stimuli, p = 0.099 on worse performance between visuals∼both). This can be linked to the results across the entire experiment as we have seen a negative impact of visual stimulus presence and a tendency for trials with just audio to have their performance led by the tempo (see Sect. 4.2). This difference might be due to a learning effect on the trials where the participants are not proficient enough to lose enough performance from visuals but may be eased by the presence of any leading audio rhythmic stimulus for this repetitive task. As for stimuli on time estimation, here we only observe a potential novelty effect on trials without stimuli as they are rarer than trials with any stimuli (t-test time estimation error stimuli presence, p = 0.093). The most notable difference with the analysis on all trials regarding this aspect is the absence of the effect of tempo on the time estimations. Surprisingly, no effect of stimuli concerning time judgment is observed from any of our tests. When it comes to correlation between performance and time estimations through Pearson correlation tests, contrary to the full set of trials, we observe a (negative) correlation with the signed time estimation error (p = 0.025, cor = -0.142) but not on the absolute error. Regarding time judgment concerning both time estimations and performance, we lost the correlation with the absolute time error; however, we retrieve the positive correlation from the Spearman tests with performance (p = 0.095, rho = 0.106) and the signed time estimation error ($p = -0.183$, $rho = -0.184$). Finally, regarding results on confounding effects of trial index and fatigue, we retrieve correlations of fatigue on performance and time experience, correlation of performance and trial index but none between trial index and time experience. This means that with this subset, we should have isolated an experiment phase based on trials for time perception but not for performance, which is expected as the participants were likely learning how to perform better during the first few trials.

Trials 11–20. This subset corresponds to each participant's ten trials in the middle of the experiment, representing a neutral phase as they no longer learn the task while not being in the experiment long enough to be bored. Regarding performance and stimuli, in this trial, we observe a performance increase from higher tempo within trials using combined stimuli (Tukey HSD 180-100, $p = 0.044$, $diff = 0.055$; 180-140, $p = 0.075$, $diff = 0.053$) As for the effect of stimuli on time estimation, we only observed a tendency between tempo 140–100 across trials solely using a visual stimulus (Tukey HSD 140-100, $p = 0.100$, $diff = -0.083$). When it comes to time judgment, strong evidence shows that under this subset, the presence of any stimulus heavily alters it (paired Wilcoxon on time judgment and type of stimuli, *p¡0.002 for all situations with "None"*; paired Wilcoxon on time judgment and tempo, *p¡0.02 for all cases with "0"*; Wilcoxon time judgment and stimuli presence, (*$p = 8.939e\text{-}05$, higher mean with stimulus*). We also observe an effect of visual stimulus specifically with the same Wilcoxon test on visual stimulus presence (*$p = 0.041$, higher mean with stimulus*) but not on audio presence. Therefore, the effect of stimuli on time judgment in this subset is consistent with the full set regarding the effect of the present stimuli type, but we lost the effect of the tempo. This time, no correlation has been observed between time estimation and performance. However, we retrieve the time judgment correlations from the full set with Spearman tests on the performance ($p = 0.004$, $rho = 0.168$), the signed time estimation error ($p = 1.988e\text{-}04$, $rho = 0.215$) and its absolute ($p = 4.20e\text{-}05$, $rho = -0.236$). Finally, on confounding effects, we find a correlation between fatigue and time judgment, which is expected, yet we also see a correlation tendency between trial index and signed time estimation error through a Pearson test ($p = 0.071$, $cor = 0.105$). Still, the absence of correlation with performance indicates a proper subset division on the trial index.

Trials 21–30. This subset corresponds to each participant's last ten trials, representing the end of the experiment and, thus, a phase where the participant is possibly tired or bored. Here, investigation of performance suggests that when there are stimuli, the presence of visual stimulus leads to worse task performance (Tukey HSD performance between stimuli modes "Both" and "AudioOnly", $p = 0.010$, $diff = -0.045$; "VisualsOnly" and "AudioOnly", $p = 0.026$, $diff = -0.041$; t-test on performance and visual presence, *$p = 1.449e\text{-}04$, lower mean when stimulus is present*). Under this subset, nothing significant has been observed in the relation between stimuli (type or tempo) and time perception (time estimation and judgment). Similarly to the previous set (Trials 11–20), no correlation between time estimation and performance is observed. The relation from Spearman tests between time judgment with both performance and time estimation is similar to the subset at the beginning of the experiment (Trials 1–10), where performance is positively correlated ($p = 0.002, rho = 0.176$) and time estimation error is negatively correlated ($p = 4.795e\text{-}04, rho = -0.202$) but the absolute error is not. On confounding effect, while finding effects of fatigue on time judgment as expected, unfortunately, we see tendencies on the effects of the trial index on both time estimation error ($p = 0.067, cor = 0.107$) and task performance ($p = 0.073, cor = -0.104$) from Pearson tests. This may indicate a transition between phases of boredom and tiredness relative to the time spent in the experiment.

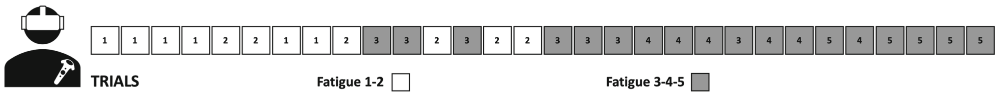


Fig. 4. Example trial subset allocation for a participant by fatigue.

4.4 Fatigue Subsets

Having obtained the results on the confounding effects of fatigue described in Sect. 4.2, we decided to investigate two subsets depending on the participants' answers on fatigue, one for fatigue at 1 or 2, and one for fatigue at 3, 4, or 5 (see Fig. 4). Similar to the previous subsets, for each, we performed all tests on the complete trial set, which can be found online [20], and the same modification on the variable normalization process by only considering the targeted subset when using the sum of data on trials. In the following sections, we will again focus exclusively on significant test results and will not re-elaborate the methodology.

Fatigue Levels 1-2. This subset corresponds to the participant experiencing "low" fatigue. First, concerning performance and stimuli tempo, we can observe the lesser performance of stimuli with a tempo of 100 (Tukey HSD on performance and tempo between 140 and 100 bpm, $p = 0.056, diff = 0.065$; 180 and 100 bpm, $p = 0.020, diff = 0.081$). This finding can be aligned to results from the full set (performance dependent on tempo for audio stimuli) and trials 11–20 (180 bpm leading to better performance under combined stimuli). No general effect of stimuli type on performance is observed, either from the specific situations possible or the presence of a modality. This subset yielded no significant insights regarding stimuli dimensions (type and tempo) and time estimation. Regarding time judgment and tempo, however, we observe significant differences between 180 and 100 bpm across all stimuli (Paired Wilcoxon on time judgment and tempo between 100 and 180 bpm, $p = 0.023$) as well as an effect of the presence of 180 bpm (Wilcoxon between 0 and 180, $p = 0.025$). When considering only trials with combined stimuli, the paired Wilcoxon shows a significant difference between 180–100 ($p = 0.017$) and a tendency between 100–140 ($p = 0.094$), which is consistent with the time judgment effects results on the complete set of trials. As for time judgment and stimuli type, we observe another consistent result from the full set as stimuli tend to be judged faster when there is any stimulus (Wilcoxon on the presence of any stimulus, $p = 0.073$) or if there is at least a visual one ($p = 0.020$). Looking for a correlation between performance, time estimation, and time judgment yielded similar results to the subset of Trials on index 1–10. With a tendency of a negative correlation ($p = 0.059$, $cor = -0.141$) from Pearson between time estimation error and performance, a significant negative correlation ($p = 0.022$, $rho = -0.171$) out of the Spearman between time estimation error and time judgment, a positive one ($p = 0.033$, $rho = 0.157$) between performance and time judgment, but no correlations from the absolute time estimation error. Confounding effects of trial index on performance (Pearson test, $p = 0.086$, $cor = 0.129$) are similar to trials 1–10, which is not surprising as early trials probably are low fatigue trials. A confounding effect of fatigue is not observed for both time estimation and judgment; however, we can observe it for performance (Spearman test, $p = 0.037$, $rho = 0.157$; Tukey HSD (more of a t-test considering we have two values in this subset), $p = 0.048$, $F = 3.977$, *diff2-1* $= 0.046$). We can assume that higher fatigue trials in this subset would be after the learning phase when the participant would be more proficient.

Fatigue Levels 3-4-5. This subset corresponds to the participant having a higher fatigue level. Concerning performance and stimuli type, like for the full set and trials 21–30, we observe a negative impact from the presence of visual stimuli (TukeyHSD on stimuli type and performance between VisualsOnly and Audio Only, $p = 0.080$, *diff* $= -0.024$; t-test between performance and presence of visuals, ($p = 0.020$)). Some observations converge towards contextual effect on tempo depending on the type of stimulus of the trial (TukeyHSD on absolute time estimation error between tempos 140 and 180 for audio trials, $p = 0.086$, *diff* $= 0.1$; TukeyHSD on signed time error between tempos 140 and 180 across all, $p = 0.080$, *diff* $= 0.087$). As for time judgment and stimuli, we only found evidence indicating an effect of general stimulus presence (Paired Wilcoxon on stimuli type and time judgment, $p < 0.07$ for pairs involving "None"; Wilcoxon on time judgment and stimulus presence, $p < 0.011$; paired Wilcoxon on tempos, $p = 0.032$) between 0–140 and $p = 0.040$ for 0–180). Looking for a correlation between performance, time estimation, and time judgment yielded similar results to the full set of trials. From Spearman tests with time judgment, we retrieve the negative correlation with the time estimation error ($p = 2.217e\text{-}11$, $rho = -0.256$), the positive correlation with the absolute error ($p = 4.366e\text{-}04$, $rho = 0.136$) and with the performance ($p = 3.915e\text{-}05$, $rho = 0.159$). We do not retrieve the significant p-value on the Pearson test between performance and absolute error, but a near-tendency ($p = 0.110$, $cor = -0.062$). While we do not observe a confounding effect from the ANOVAs between fatigue and time estimation variables, we see a tendency (p = *0.089, rho* = *0.066*) from the Spearman test on the absolute time estimation error. The effect of fatigue is more pronounced on performance (Pearson test, p = *0.029, rho* = *0.085*) and from the subsequent Tukey HSD of the ANOVA (p = *0.064*, F = *2.768*) the difference appears to be between 3–5 (p = *0.050, diff* = *0.023*). Fatigue also seems to have a significant effect on time judgment (p = *2.185e-08, rho* = *0.215*). In this subset, the fatigue level of 3 and 5 may be significantly different on both performance and time judgment; however, this is apparently due to the normalization on the subset and was not observable across all trials. Confounding effects of the trial index observed from Pearson tests are similar to those of the full set, which is not too surprising as the subset is rather large and was not made to minimize the effect of the index.

5 Discussion

We conducted a VR experiment where participants repeatedly performed a simple sorting task subjected to different stimuli conditions, allowing us to gather numerous data, including information related to task proficiency, subjective data from questionnaires, and physiological data. The overarching goal was to explore relationships between time experience, task performance, environment/stimuli conditions, and physiological cues. However, we found that for some of the results on the complete data set, it was necessary to investigate closer multiple subsets, which we will discuss together with their implications for VR application design. Looking at the entire data set, we can observe specific effects of stimuli type and tempo on different aspects of time perception and performance, as well as some interesting correlations between those variables, which appeared to be also heavily impacted by the trial index and fatigue through the experiment. Therefore, we defined subsets of data based on the trial index and difference values from ANOVAs for fatigue. As indicated in Table 1, we can observe effects of stimuli presence, type, and tempo on performance and time experience depending on the subset. Some of the data and correlations align between subsets while others do not, which may indicate contextual effects of stimuli on performance and time experience depending on task repetition and fatigue.

5.1 Observations on Task Performance and Stimuli

A central result from the analysis of the complete set of trials is how the presence of visual stimuli negatively impacts task performance. This is coherent with our previous study and is to be anticipated as the task requires visual attention, and those stimuli may be disturbing. However, within subsets, this result is observed only for later trials (index 21–30) and high fatigue (fatigue 3-4-5). Surprisingly, we see a positive effect on performance from the presence of audio but only in the early trials (index 1–10), and no effect of stimuli type presence in-between (index 11–20). This could be interpreted as the disturbance of visual stimulus not being impacting enough when one is learning the task or not physically tired. We can also interpret the presence of audio stimuli as beneficial for this task only when the participant is in a learning phase. Effects of tempo are observed on trials within trials with only the audio stimulus when considering all trials and within trials with combined stimuli within the subset of trials from index 11–20. In both cases, the faster tempo led to faster performance, which indicates an invitation to go faster in the task from the faster stimuli; however, that interpretation from the participant depends on the context.

5.2 Observations on Time Estimation and Stimuli

Time estimation variables are defined from the difference between (normalized) participants' estimation of time taken for a trial and the actual time of a trial; we thus talk about the time estimation error and its absolute, which represents the magnitude of error regardless of if the participant under- or overestimated the length of a trial. A global effect of tempo can only be observed with the complete set of trials between 180–140 (with 180 being overestimated). As for differentiation within stimuli situations, we see a time estimation error difference on the audio stimuli for the entire trial set and for high-fatigue trials, and absolute error difference in visuals for trials 11–20 as well as on combined stimuli for high fatigue. These results show a tendency of the 140 bpm tempo leading to fewer (absolute) estimation errors and being underestimated compared to 100–180. Another overestimating effect from stimulus presence is observed for trials 1–10. Overall, we also observe context-dependent effects of stimuli as the type of stimuli will affect one's time perception differently depending on the index or fatigue.

5.3 Observations on Time Judgment and Stimuli

Time judgment or time passage refers to the subjective evaluation of a participant on whether they think a trial is going by fast or slow. It differs from time estimation as the participant gives their subjective feeling about the time spent, whereas time estimation is an attempt of the participant to be objective about time. Time judgment has semi-constant results of the presence of any stimuli inducing faster perception; this is observed across all trials for both subsets on fatigue and the subset on stimuli 11–20. We can also observe a specific fastening effect of visual stimuli on all these sets affected by the presence of any stimuli except for the high fatigue one. The absence of these observations on subsets of trials either at the beginning or at the end of the experiment might indicate the participant needing to get used and, over time, getting too used to the presence of stimuli to be noticeable, regardless of fatigue levels. Another effect observed only on the complete set and for low fatigue is a difference between tempo in general and within trials with combined stimuli.

5.4 Time Judgment, Time Estimation, and Performance Balance

Two correlations were consistent across all sets: a negative correlation between time estimation error and time judgment and a positive correlation between performance and time judgment. The first means that under-estimation of time is reflected by a subjective faster trial and the second means that when the participant rated the trial as faster than usual, they would perform better. This could directly be tied to the notion of flow as two elements of flow states are the challenge-skill balance and time transformation. The similarity between time estimation error and time judgment is indicative that our time transformation was a general time experience shift and not a side effect of disorientation (i.e., a participant judging a trial as fast because they thought it was a higher amount of time that actually passed). Among low fatigue and early trials, we also retrieve a negative correlation between performance and performance, reinforcing the flow approach. As for the time estimation error magnitude and performance, under all trials and high fatigue, it is negatively correlated, which means that possibly in a specific context, higher time transformation generally was detrimental to performance. However, this is against the flow definition, and combined with previous observations, it may imply that we are approaching flow states only with time transformations that are an underestimation. We also observed positive correlations between this magnitude and time judgment with all trials, the 11–20, and high fatigue subsets, which could be interpreted as the presence of any time transformation potentially leading to faster time passage in general.

5.5 Limitations

It is important to remember that the effect of the rhythmic stimuli in our experiment is contextualized in the particular scenario of our sorting task. We can also see some limits from the confounding effects of task familiarity and fatigue, and even with the subsets, which unfortunately implies using fewer data and thus having lesser statistical power relevance (especially in the case of low fatigue), we can isolate the effect of at most one confounding effect but not both at the same time. Individual per-participant differences are also to be considered, as through casual talks with the participants, we know of varying degrees of VR experience between participants; however, this data was not recorded and is thus not included in our analysis.

6 Conclusion

In this paper, we used a simple sorting task to explore how rhythmic stimuli affect time experience and task performance in VR. We found that the context concerning the trial index (repetition of the action) and fatigue affected these aspects of the user experience. Depending on the familiarity with the task, the presence of a particular type of rhythmic stimulus under possible tempos will affect either performance or time experience. Both aspects can contribute significantly to a flow experience or even well-being in general, and the results of this study can thus inform the design of future interactive VR applications.

While the familiarity or repetition of a task or action can be easily assessed in any interactive application, using fatigue as a modulator could be a growing opportunity for VR developers as newer HMDs incorporate advanced sensors, e.g., for eye tracking. We observed effects of rhythmic stimuli under some fatigue and task familiarity, yet the more important finding is the presence of effect variation rather than the specific effect itself, highlighting the need for studies of time perception concerning context- and subject-dependent time modulations.

References

1. Bruder, G., Steinicke, F.: Time perception during walking in virtual environments. In: 2014 IEEE Virtual Reality (VR), pp. 67–68 (2014). https://doi.org/10.1109/VR.2014.6802054
2. Droit-Volet, S., El-Azhari, A., Haddar, S., Drago, R., Gil, S.: Similar time distortions under the effect of emotion for durations of several minutes and a few seconds. Acta Physiol. **210**, 103170 (2020). https://doi.org/10.1016/j.actpsy.2020.103170
3. Droit-Volet, S., Fayolle, S., Lamotte, M., Gil, S.: Time, emotion and the embodiment of timing. Timing Time Percept. **1**(1), 99–126 (2013). https://doi.org/10.1163/22134468-00002004
4. Droit-Volet, S., Ramos, D., Bueno, J.L.O., Bigand, E.: Music, emotion, and time perception: the influence of subjective emotional valence and arousal? Front. Psychol. **4**, 417 (2013). https://doi.org/10.3389/fpsyg.2013.00417
5. Ghaderi, A., Niemeier, M., Crawford, J.D.: Linear vector models of time perception account for saccade and stimulus novelty interactions. bioRxiv (2021). https://doi.org/10.1101/2020.07.13.201087
6. Gorea, A.: Ticks per thought or thoughts per tick? A selective review of time perception with hints on future research. J. Physiol.-Paris **105**(4), 153–163 (2011). https://doi.org/10.1016/j.jphysparis.2011.09.008
7. van der Ham, I.J., Klaassen, F., van Schie, K., Cuperus, A.: Elapsed time estimates in virtual reality and the physical world: the role of arousal and emotional valence. Comput. Hum. Behav. **94**, 77–81 (2019). https://doi.org/10.1016/j.chb.2019.01.005
8. Hammerschmidt, D., Wöllner, C., London, J., Burger, B.: Disco time: the relationship between perceived duration and tempo in music. Music Sci. **4**, 2059204320986384 (2021). https://doi.org/10.1177/2059204320986384
9. Hoagland, H.: The physiological control of judgments of duration: evidence for a chemical clock. J. Gen. Psychol. **9**(2), 267–287 (1933). https://doi.org/10.1080/00221309.1933.9920937
10. Igarzábal, F.A., Hruby, H., Witowska, J., Khoshnoud, S., Wittmann, M.: What happens while waiting in virtual reality? A comparison between a virtual and a real waiting situation concerning boredom, self-regulation, and the experience of time. Technol. Mind Behav. **2**(2) (2021). https://doi.org/10.1037/tmb0000038. https://tmb.apaopen.org/pub/what-happens-while-waiting-in-virtual-reality

11. Jackson, S.A., Eklund, R.C.: Assessing flow in physical activity: the flow state scale–2 and dispositional flow scale–2. J. Sport Exerc. Psychol. **24**(2), 133–150 (2002). https://doi.org/10.1123/jsep.24.2.133
12. Lukas, S., Philipp, A.M., Koch, I.: Crossmodal attention switching: auditory dominance in temporal discrimination tasks. Acta Physiol. **153**, 139–146 (2014). https://doi.org/10.1016/j.actpsy.2014.10.003
13. Makwana, M., Srinivasan, N.: Intended outcome expands in time. Sci. Rep. **7**(1), 6305 (2017). https://doi.org/10.1038/s41598-017-05803-1
14. Mallam, S.C., Ernstsen, J., Nazir, S.: Accuracy of time duration estimations in virtual reality. Proc. Hum. Factors Ergon. Soc. Ann. Meet. **64**(1), 2079–2083 (2020). https://doi.org/10.1177/1071181320641503
15. Mullen, G., Davidenko, N.: Time compression in virtual reality. Timing Time Percept. **9**, 1–16 (2021). https://doi.org/10.1163/22134468-bja10034
16. Nah, F., Eschenbrenner, B.: Flow experience in virtual worlds: individuals versus dyads. In: SIGHCI 2015 Proceedings (2015). https://aisel.aisnet.org/sighci2015/19
17. Pelet, J.É., Ettis, S., Cowart, K.: Optimal experience of flow enhanced by telepresence: evidence from social media use. Inf. Manag. **54**(1), 115–128 (2017). https://doi.org/10.1016/j.im.2016.05.001
18. Peters, C.M., Glazebrook, C.M.: Rhythmic auditory stimuli heard before and during a reaching movement elicit performance improvements in both temporal and spatial movement parameters. Acta Physiol. **207**, 103086 (2020). https://doi.org/10.1016/j.actpsy.2020.103086
19. Picard, S., Botev, J.: Rhythmic stimuli effects on subjective time perception in immersive virtual environments. In: Proceedings of the 14th International Workshop on Immersive Mixed and Virtual Environment Systems, MMVE 2022, pp. 5–11. Association for Computing Machinery, New York (2022). https://doi.org/10.1145/3534086.3534330
20. Picard, S., Botev, J.: Rhythmic stimuli and time experience in virtual reality - complete data. Analysis (2023). https://doi.org/10.5281/zenodo.8321053
21. Reid, D.: A model of playfulness and flow in virtual reality interactions. Presence Teleoperators Virtual Environ. **13**(4), 451–462 (2004). https://doi.org/10.1162/1054746041944777
22. Roseboom, W., Fountas, Z., Nikiforou, K., Bhowmik, D., Shanahan, M., Seth, A.K.: Activity in perceptual classification networks as a basis for human subjective time perception. Nat. Commun. **10**(1), 267 (2019). https://doi.org/10.1038/s41467-018-08194-7
23. Rutrecht, H., Wittmann, M., Khoshnoud, S., Igarzábal, F.A.: Time speeds up during flow states: a study in virtual reality with the video game thumper. Timing Time Percept. **9**(4), 353–376 (2021). https://doi.org/10.1163/22134468-bja10033
24. Schatzschneider, C., Bruder, G., Steinicke, F.: Who turned the clock? Effects of manipulated zeitgebers, cognitive load and immersion on time estimation. IEEE Trans. Visual Comput. Graphics **22**(4), 1387–1395 (2016). https://doi.org/10.1109/tvcg.2016.2518137
25. Unruh, F., Landeck, M., Oberdörfer, S., Lugrin, J.L., Latoschik, M.E.: The influence of avatar embodiment on time perception - towards VR for time-based therapy. Front. Virtual Reality **2**, 658509 (2021). https://doi.org/10.3389/frvir.2021.658509
26. Wang, X., Wöllner, C., Shi, Z.: Perceiving tempo in incongruent audiovisual presentations of human motion: evidence for a visual driving effect. Timing Time Percept. **10**(1), 75–95 (2021). https://doi.org/10.1163/22134468-bja10036

27. Watt, J.D.: Effect of boredom proneness on time perception. Psychol. Rep. **69**(1), 323–327 (1991). https://doi.org/10.2466/pr0.1991.69.1.323. pMID: 1961817
28. Wöllner, C., Hammerschmidt, D.: Tapping to hip-hop: effects of cognitive load, arousal, and musical meter on time experiences. Atten. Percept. Psychophys. **83**(4), 1552–1561 (2021). https://doi.org/10.3758/s13414-020-02227-4

Appendix C

Paper: Psychophysiology of Rhythmic Stimuli and Time Experience in Virtual Reality

Full paper of Section 3.1.3

Computers & Graphics 124 (2024) 104097

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on EuroXR 2023 Best Papers

Psychophysiology of rhythmic stimuli and time experience in virtual reality

Stéven Picard *, Jean Botev

VR/AR Lab, University of Luxembourg, L-4364 Esch-sur-Alzette, Luxembourg

A R T I C L E I N F O

Keywords:
Virtual reality
Human–computer interaction
User experience
Time experience
Rhythmic stimuli
Physiological data

A B S T R A C T

Time experience is an essential part of one's perception of any environment, real or virtual. In this article, from a virtual environment design perspective, we explore how rhythmic stimuli can influence an unrelated cognitive task regarding time experience and performance in virtual reality. This study explicitly includes physiological data to investigate how, overall, experience correlates with psychophysiological observations. The task involves sorting 3D objects by shape, with varying rhythmic stimuli in terms of their tempo and sensory channel (auditory and/or visual) in different trials, to collect subjective measures of time estimation and judgment. The results indicate different effects on time experience and performance depending on the context, such as user fatigue and trial repetition. Depending on the context, a positive impact of audio stimuli or a negative impact of visual stimuli on task performance can be observed, as well as time being underestimated concerning tempo in relation to task familiarity. However, some effects are consistent regardless of context, such as time being judged to pass faster with additional stimuli or consistent correlations between participants' performance and time experience, suggesting flow-related aspects. We also observe correlations between time experience with eye-tracking data and body temperature, yet some of these correlations may be due to a confounding effect of fatigue. If confirmed as separate from fatigue, these physiological data could be used as reference point for evaluating a user's time experience. This might be of great interest for designing virtual environments, as purposeful stimuli can strongly influence task performance and time experience, both essential components of virtual environment user experience.

1. Introduction

While time itself is a concept, it is also something humans can perceive. However, the perception of time is subjective, and this experience is an integral part of the overall experience of any environment, with virtual environments being no exception. Therefore, acknowledging this in their conception and actively devising virtual environments to modulate the time experience of users would be an exciting design instrument. Time perception as an interdisciplinary topic is explored in numerous scientific studies in disciplines as diverse as psychology or neuroscience. In a previous study combining cognitive science and computer science, we already examined the relationships between time perception and rhythmic stimuli with a sorting game in a two-dimensional setting, which revealed varying time experience and performance effects depending on whether single or combined stimuli were used in relation to their tempo [1]. However, that initial experiment was limited to a crowd-sourced desktop setting, and we adapted it for Virtual Reality (VR), which allows for extended control of the test environment plus extending the initial set of questions to include stimuli and time experience aspects in fully immersive environments. Therefore, while our main goal with this study is to investigate and interpret anything significant by having a correlation study process, we do come with assumptions coming from our initial study, which are different effects on time perception depending on the type of stimuli present (audio, visual, or both), tempo-related time estimation errors for combined stimuli, as well as decreased task performance by the presence of visual stimuli. A locally controlled environment furthermore allows us to let participants use hardware of our choice, including hardware with physiological sensors. Therefore, we also investigate the relationship between physiological data and subjective time experience. This article extends a paper originally published at the EuroXR 2023 conference [2], which did not comprise any physiological data analysis and correlations provided in this work to convey the full picture. Before detailing the experiment in Section 3, analyzing the data in Section 4, and discussing the results in Section 5, we will first provide the necessary background and review related work in the following Section 2. Section 6 then concludes the article with an outlook and potential impact of our findings on general virtual environment design.

2. Background and related work

The most common time perception model in literature is the clock model, which assumes an "internal clock" as a body system dedicated

* Corresponding author.
E-mail address: steven.picard@uni.lu (S. Picard).

https://doi.org/10.1016/j.cag.2024.104097
Received 13 March 2024; Received in revised form 8 August 2024; Accepted 24 September 2024
Available online 27 September 2024

to time perception [3]. This system is usually tied to a model producing "ticks" or "units", such as an oscillator or a pacemaker model, where the body keeps track of time by counting these ticks. In these models, it is assumed that time perception can be changed either through the speed at which these ticks are produced or by skipping some, with these effects potentially resulting from external stimuli unrelated to time. Counting the ticks can be delegated to attention, making attentional resources a key element of time perception [4,5]. Attention is believed to act as a gate or switch on accounting the time units, where paying less attention to time will result in compressed time experience as time units are likely to be skipped. Another source of subjective temporal distortions is the use of arousal levels, which are believed to affect the clock speed, resulting in an altered time experience [6–8]. However, using an internal clock model is not necessary to predict time perception accurately [9]: in the context of watching videos in different scenarios, time perception was accurately predicted by a classification network using changes in perceptual content and visual spatial attention (more specifically, gaze position). Nevertheless, it is a natural way to interpret time experience based on the focus on attention and arousal that gives initial directions to time perception studies.

In the literature, we can observe different types of timing tasks, which involve different processing mechanics. The most common aspects are time estimation (i.e., asking for a duration estimate of an event with units, such as seconds) and the feeling of time passage (i.e., judging if time passes by quickly or slowly), which is also referred to as time judgment.

This difference can be observed in depressed subjects underestimating time but judging it as passing slowly [6], with boredom-prone people judging boring tasks as passing slowly but not necessarily overestimating those [10], or with players of the game "Thumper" reporting faster time passing without time estimation errors [11].

Having defined time perception, we can discuss the state of "flow", one of the applications of time perception alterations. Flow is a psychological state of full attention on a task defined by Csikzentmihalyi, represented through nine dimensions: challenge-skill balance, action-awareness merging, clear goals, unambiguous feedback, concentration on task, sense of control, loss of self-consciousness, time transformation, and autotelic experience. The psychological state of flow is a research subject in itself, centered around one's relation to a task as it primarily relies on the challenge-skill ratio aspect [12]; it is often considered a desirable state, and time perception alterations are one of its manifestations. In social media, the manifestation of flow seems to be influenced by the positive effect of telepresence on enjoyment, concentration, challenge, and curiosity; flow would then influence the presence of time distortions [13]. Delving further into social context, it was found that the concentration and time distortion components of flow, but not enjoyment, were affected by working in a group of two compared to working individually in virtual worlds (within the social game platform Second Life) [14]. More in line with our work, several studies on flow and VR have been conducted. The previously mentioned study on Thumper compared flow states between VR and non-VR setups, finding that despite VR's technical immersion, both scenarios could lead to a flow state [11]. Within VR activities, a model ostensibly associates flow and playfulness, defined by a combination of intrinsic motivation, control, and freedom to suspend reality; this association then influences competence in the activity [15].

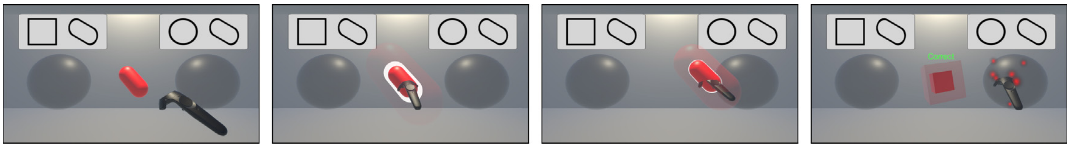
VR studies on time perception, however, are not limited to flow. VR itself affects our senses due to what is being conveyed through sensory channels, but also due to the devices used and the eventual physical discomfort we can get from it. Simply comparing the time perception of the same game in a VR versus a desktop setup leads to an underestimation bias for VR [16]. It also seems that time perception changes when bored or waiting in VR compared to real life [17]. In another study comparing time perception between VR and desktop while completing a task ranging from 30 s to 5 min, it was observed that if both situations yielded time overestimation, the VR scenario was overestimated more [18]. This seemingly contradicts the underestimation effect in [16] but may be due to VR affecting time perception in correlation to the task. However, technological immersion alone might not be a sufficient explanation, as walking in VR does not seem to affect time perception significantly [19]. In another experiment about zeitgeber on time perception while doing a task conducted both in VR and on desktop, no significant difference was observed regarding time perception, but there was a difference in task performance (with the VR participant performing better) [20]. The Thumper study also observed the effect of performing better in a VR setup compared to a non-VR one [11]. When it comes to the effect of emotional content, VR itself appears not to yield any difference to real life in time distortions when the emotional content is the same [21]. Employing VR entails possibly unique content, such as having movements represented by an avatar. Differences were observed between avatar and no-avatar conditions, where avatar presence in a retrospective paradigm leads to a significantly faster passage of time without affecting time estimation [22].

When considering VR and time perception, we can thus regard both the technological immersion, i.e., the effect of being in VR through its dedicated hardware, as well as the virtual environment stimuli and transformations that can be induced through VR. A specific stimulus type we want to employ in VR is rhythmic stimuli, which we already identified as having notable effects in a desktop scenario [1]. Rhythmic stimuli and music generally have a high potential to modulate one's time experience. With music, it was observed that higher tempos induce longer subjective time, but emotional valence decreases (but does not suppress) the effect of tempo and affects time perception. These effects on time perception might be due to their impact on arousal; interestingly, using a different orchestration (piano only or full orchestra) does not affect time judgment and pleasantness while affecting arousal [23]. On timing evaluations of instrumental excerpts of Disco songs (including estimation and judgment of time passage) over different tempos, it was observed that faster tempos were correlated to longer reproduction duration; however, no effect on estimations was observed alongside the necessity of a tempo difference of at least 20 BPM required for timing measurement differences to appear [24]. By varying cognitive load through tasks and arousal levels through music choice while keeping music tempo constant, it was found that (1) time is judged as passing faster under higher cognitive load (presence of a math task), (2) presence of a concurrent motor task (tapping the music's tempo) yield shorter subjective durations, and (3) for the motor task, for the same music tapping to half notes instead of eighth notes ended up with smaller time estimations and time passage rated as faster [25]. Regarding timing and spatial movement, rhythmic auditory stimuli (RAS) have been observed to improve motor performance when vision is unavailable [26].

Rhythm, however, is not only tied to music and audio. Concerning temporal judgments, audio was believed to be dominant over visuals [27]. However, a later study found visual stimuli dominance using Point-Light-Display (PLD) dance motions compared to simple audio tempos [28]. Participants were presented with both the dance motion and the audio tempo and had to give a globally fitting tempo. The result of this study suggests that under the right conditions, visual stimuli can dominate audio in terms of timing, which may be due to the quantity of temporal information on a sensory channel rather than the channel itself.

While our focus here is primarily on VR-related stimuli and associated in-environment monitoring, the physiological aspects of time perception are equally interesting.

Temperature is among the stimuli believed to be linked to arousal and affect time perception. Indeed, as shown by the literature review of Wearden and Penton-Voak [29], multiple previous works suggest "a parametric relation between the rate of subjective time and body temperature". Wearden and Penton-Voak further suggests that increases and decreases of the rate of subjective time that happens in temperature



(a) (b) (c) (d)

Fig. 1. Sorting example from trial task, in sequential order from left to right.

situations could be assumed to correspond to equivalent arousal level changes (with higher body temperature leading to arousal increase and thus time dilation, gradual cooling reducing arousal and therefore leading to time compression but cold stress leading to time dilation could be explained by an arousal increase).

Still on the topic of body temperature, with an experiment involving head-out water immersion conditions, Kingma et al. investigated the relation between time perception, reaction time, gastrointestinal temperature, heart rate and thermal sensation. From their observations, a change in the accuracy of an expanded judgment task was significantly associated with a change in gastrointestinal temperature, thermal sensation and heart rate. It was also observed that a change in response time was significantly associated with a change in gastrointestinal temperature. Finally, changes in time estimations were associated with a change in thermal discomfort or gastrointestinal temperature but not with any other measured physiological or subjective variables [30].

Dormal et al. investigated physiological arousal/heart rate and time processing in an offline arousal manipulation manner (separated from duration processing). According to the authors, the overestimation observed is not due to physiological arousal but to a distortion of memorized standard duration caused by time lag rather than a physiological arousal effect. Heart rate/arousal alone does not seem sufficient to explain time distortions at a supra-second scale, and other studies online (concurrently with duration processing) show that arousal manipulation's effect on time perception due to the allocation of attentional resources [31].

In the context of VR experiments, Cho et al. designed a stress-level classification algorithm using K-ELM based on physiological signals (Heart rate variability, skin conductance, measured by photoplethysmogram (PPG), electrodermal activity (EDA) and skin temperature (SKT)). In their presentation study, the algorithm had over 95% accuracy [32].

Emotional response seems correlated to pupil dilation, likely due to increased arousal. In this context, Chen et al. designed methods to use measure and normalize pupil dilation values depending on the brightness of the screen [33].

As for exercising and time perception and the relation of these two with arousal, Lambourne observed a leftward shift of temporal generalization gradients during exercise when compared to rest, indicating that the subject perceived intervals to elapse more slowly. They also observed no significant differences in the episodic timing or working memory tasks. These findings support the notion that exercise influences the internal clock similar to other arousal-inducing manipulations [34].

The stopped-clock illusion, or saccadic chronostasis, is a phenomenon that occurs when making a saccade (here referred to as an eye movement) to a silent clock. This causes the first–second feeling to last longer than the next ones [35,36].

Yarrow et al. [35] investigated through multiple experiments the effect of the duration of the saccade on chronostasis (using an incremental digital counter with the first value being displayed for a time between 400 ms and 1600 ms and the subsequent values being a constant 1s) if the illusion of chronostasis is due to eye movement or shift of the locus of visual attention. By comparing situations in which the participant had to either simply do a saccade to the target or shift their attention first towards the subject and then move their eyes, they observed what happens when moving the target of the saccade. The results of Yarrow et al.'s experiments suggest that the extra-time judgment of chronostasis depends on the time used to move the eye, that the shift of attention does not affect the length of chronostasis and regarding the target of the saccade moving, if the shift was noticed then no illusory effect was observed but if it was not noticed then the illusion had a magnitude between a control scenario and a full illusion scenario. It is then suggested that "chronostasis is an illusion occurring to fill in the perceptual gap during saccadic suppression" and "moving the target unpredictably during the saccade breaks the spatial continuity and thus, the illusion disappears" [35].

3. Experiment

We recruited 30 participants from a public, science-related event in Luxembourg City and from students and staff at the University of Luxembourg. The female-to-male ratio was 46.66%–53.33%, with ages ranging from 19 to 45 (mean age 25.63/median 24). Participants were received at the VR/AR Lab's test space at the University of Luxembourg and briefed about the experiment before the session, both verbally and via an informed consent form, in which we also collected demographic data. The study was approved by the local ethics committee in compliance with all ethical standards and guidelines. Among others, participants were allowed to abort immediately at any time, e.g., in case of cybersickness. After the setup and familiarization phase, the participants performed the tasks, with each participant able to take breaks between trials if desired. Each session lasted a total of approximately one hour.

3.1. Trial task and design

Participants had to complete trials in which they had to sort three-dimensional objects according to their shape (spheres, capsules, or cubes). As shown in the screenshot sequence in Fig. 1, the objects must be grabbed with a VR controller and dragged into one of the two larger sinks, which only accept the specific shapes displayed on a scoreboard above them (cf. Figs. 1(a)–1(c)). While cubes and spheres can only be placed in one of the sinks, capsules can go in either of the two. Once sorted, an object disappears with a small animation, indicating if the sorting is correct, and a new object to be sorted appears in the center of the virtual environment (cf. Fig. 1(d)).

The sorting attempt has a predefined duration unknown to participants. After this duration, the experiment ends with a questionnaire in which the participants are asked to estimate the time in seconds and rate on two Likert scales how fast the experiment felt and how tired they were (cf. Fig. 2).

Trials were subjected to conditions that were a combination of these parameters:

- *trial length*: how long the trial lasted in seconds (either 40, 50, or 60)
- *tempo*: the rhythm of the audio/visual stimuli if present in beats per minute (either 100, 140, or 180; forced to 0 without stimuli)
- *visual stimuli*: whether or not there are flashing visual pulses around the object
- *audio stimuli*: whether or not metronome click sounds are produced

Each participant went through all 30 possible combinations in random order.

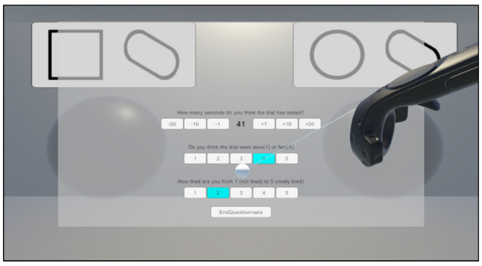


Fig. 2. Post-trial questionnaire (time estimation, time perception, fatigue).

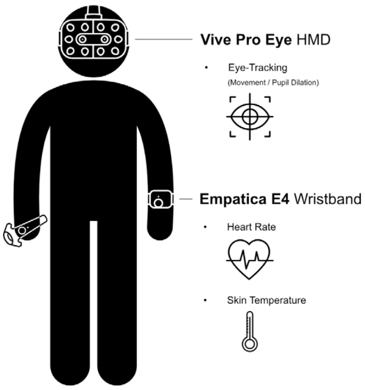


Fig. 3. Hardware used for its associated physiological data collection. Vive Pro Eye in the case of eye-tracking related data, Empatica E4 otherwise.

3.2. Technical specifications

Participants used the VIVE Pro Eye head-mounted display (HMD) with one of its controllers, allowing them to enter VR and control virtual objects with six degrees of freedom (DoF). This particular HMD also allows precise eye-tracking data to be collected during the trials, specifically on gaze and pupil dilation. During the experiments, participants were also instructed to wear an Empatica E4 wristband, as illustrated in Fig. 3, to record further physiological data; our experiment included heart rate variability and skin temperature. Our custom application reads the data by receiving messages from Empatica's E4 Streaming Server software. The experimental application itself was developed in Unity and used SteamVR. In addition to the application, OVR advanced settings were employed to adjust the participant's height position.

4. Analysis and results

4.1. General methods and data

As discussed previously, various data were collected from each trial; in this subsection, we will describe what data we effectively used for our analysis and how.

4.1.1. Disclaimed trials

Some trials were disclaimed from the data set depending on our notes during trials. The reasons for removing trials were:

- Misunderstanding of task controls by the participant
- Misunderstanding of trial questionnaire by the participant (verbally checked when giving incoherent values such as negative time estimation).
- Disturbance or interruption during the trial, either from the participants themselves (i.e., asking a question or talking during a trial) or external sources (i.e., technical issues, noise from a nearby room).

4.1.2. Variables/functions

Variables and functions extracted from trial data are:

- $t \in T$: trial identifier t from all available trials T.
- $p \in P$: participant identifier p from all available participants P.
- *trials(p)*: all trial identifiers of a participant p.
- *participant(t)*: participant identifier of a trial t.
- *correct(t)*: number of correct sorts at the end of trial t.
- *trialLength(t)*: length of trial t in seconds.
- *reportedLength(t)*: reported length of trial t in seconds.
- *reportedFatigue(t)*: self-reported fatigue of a participant after performing trial t in an ordinal scale from 1 to 5.
- *reportedSpeedPerception(t)*: subjective participant rating of trial t's speed in an ordinal scale from 1 (slow) to 5 (fast).
- *trialIndex(t)*: the index of trial t, indicates how many trials were performed before t and, therefore, global repetition from the experiment session.

Since *reportedSpeedPerception(t)* and *reportedFatigue(t)* are purely subjective questions to the participant, we can use these values directly. However *correct(t)*, the task performance direct variable depends on the participants' individual performance and *reportedLength(t)*, while the direct time estimation variable depends on both the trial's duration as well as the participants' individual representation of a second. Therefore, these two variables need to undergo a normalization process. Normalizing the performance variable (*correct(t)*) for our analysis goes through a three-step process involving the following extracted variables:

- *correctPerSecondTrial(t)*: average number of correct answers per second during trial t.
  $$\frac{correct(t)}{trialLength(t)}$$
- *correctPerSecondParticipant(p)*: average number of correct answers per second of a participant during trials.
  $$\frac{\sum_{t' \in trials(participant(t))} correct(t')}{\sum_{t' \in trials(participant(t))} trialLength(t')}$$
- *correctNormalized(t)*: amount of correct answers per second of trial t normalized with 1, i.e., the average number of correct answers per second of a participant among all performed trials.
  $$\frac{correctPerSecondTrial(t)}{correctPerSecondParticipant(p)}$$

For time estimation (*reportedLength(t)*), we employed the following process:

- *secondBias(p)*: ratio of the total of seconds of a participant p's trials and the total reported time, defining what the participant considers a second.
  $$\frac{\sum_{t \in trials(p)} trialLength(t)}{\sum_{t \in trials(p)} reportedLength(t)}$$
- *deltaTimePerception(t)*: averaged delta per second between reported time (accounting participant bias) and trial length of a trial t.
  $$\frac{secondBias(p) * reportedLength(t)}{trialLength(t)} - 1$$

However, in addition to *deltaTimePerception(t)* we also use its absolute *abs(deltaTimePerception(t))* as it represents the magnitude of time perception delta of a trial.

4.1.3. Outcome variables

The specific variables relevant to the analysis performed in this study are:

- *deltaTimePerception(t)*: this variable represents a participant's variation in perception, i.e, the difference between reported time and trial time. A negative value indicates that the seconds of the trial *t* were reported as shorter than the other trial performed by this participant. A positive value means that the seconds were reported as longer.
- *abs(deltaTimePerception(t))*: instead of denoting how much longer or shorter a second is interpreted for a trial *t* compared to other trials performed by a participant, the absolute value represents the magnitude of the eventual time distortion.
- *correctNormalized(t)*: to simplify the analysis, we do not take into account incorrect answers to evaluate performance but only the amount of correct answers. A smaller number of correct answers compared to other trials would indirectly reflect the number of incorrect answers due to the time lost. Similar to a participant's variation in perception, this variable represents the variation in performance instead of the pure performance, with values < 1 indicating worse and > 1 better performances.
- *reportedSpeedPerception(t)*: the subjective interpretation of whether time drags or flies after performing trial *t* in an ordinal scale from 1 to 5.
- *reportedFatigue(t)*:reported fatigue should mostly depend on the participants, independent of trial parameters.
- *trialIndex(t)*: the index of trial *t*, indicates how many trials were performed before *t* by the participant and, thus, is an indicator of global repetition from the experiment session.

4.1.4. Parameters

The parameters used in this experiment are:

- *stimulusTrial(t)*: the type of stimulus used in trial *t*, possible values are: *None, VisualsOnly, AudioOnly, Both*.
- *hasAudioTrial(t)*: whether or not the trial *t* contains an audio stimulus.
- *hasVisualTrial(t)*: whether or not the trial *t* contains a visual stimulus.
- *hasStimulusTrial(t)*: whether or not the trial *t* contains any type of stimulus.
- *tempo(t)*: the tempo in beats per minute (BPM) of a trial *t*, possible values are: 0, 100, 140, 180. A tempo of 0 means that the trial had no stimuli.

4.1.5. Trial filters

When performing analyses, we may want to include only subsets of the trials to investigate specific effects. The filters used are:

- *filterAUDIOONLY*: considers only trials that only have an audio stimulus, is equivalent to saying *t where stimulusTrial(t) == AudioOnly*.
- *filterVISUALSONLY*: considers only trials that have a visual stimulus, is equivalent to saying *t where stimulusTrial(t) == VisualsOnly*.
- *filterBOTH*: considers only trials that have both audio and visual stimuli, is equivalent to saying *t where stimulusTrial(t) == BOTH*.

4.1.6. Physiological variables and functions

The raw eye-tracking data recorded and used in this study are the pupil position in the sensor area and the pupil dilation. These data were being recorded at every frame the application was running, so aimed at about 90hz frequency, as this is the headset's refresh rate. The pupil dilation was used for each trial as follows: recording errors were filtered (detected with values below 0), and then values were weighted depending on the time difference between it and the previously recorded data point. We thus collect this Pupil dilation and use the mean *PupilDilation.mean* and variance *PupilDilation.var*. The pupil position was used to derive what we call the "eye speed", which refers to the distance per second the pupils are traveling. Similar to pupil dilation, we filter the incorrect recordings (using the interquartile method) and then get weighted values from subsequent values in time. We thus also get the mean (*EyeSpeed.mean(t)*) and variance (*EyeSpeed.var(t)*) of this eye speed. While the Empatica E4 also allows for the collection of electrodermal activity (EDA/GSR) data and inter-beat interval (IBI), we restricted our collection to heart rate and skin temperature. The reason is that the Empatica software already uses IBI to derive heart rate, and EDA/GSR would require minimal user movement as well as a 15-minute waiting phase, which was incompatible with the task. The temperature value uses the same filtering and weighting as the eye-tracking values and is used with both its mean and variance (*Temperature.mean(t)*, *Temperature.var(t)*); we also keep the temperature difference between the start and end of the trial as it shows the change across the trial (*Temperature.start-end(t)*. As heart rate is already derived from the Empatica's blood value pressure (BVP) over a longer period, we do not have enough data points. Thus, we take the last recorded value in a trial (*Heartrate(t)*) and filter trials where the value is not updated. Due to numerous technical issues with the Empatica hardware and software, 463 (44.88% of the data is retained) of the trials' heart rate data and 42 (95% of the data is retained) for the temperature have been removed.

4.1.7. Performed tests

Using the variables, parameters, and filters detailed in Section 4.1.2 to 4.1.6, we performed statistical tests on various data subsets to examine the effect of a stimuli presence on performance and time estimation with the following parameters:

- *stimulusTrial(t)*: ANOVAs on *deltaTimePerception(t)*, *abs(deltaTimePerception(t))* and *correctNormalized(t)* to see if any significant difference appears between possible stimuli situations.
- *hasAudioTrial(t), hasVisualTrial(t), hasStimulusTrial(t)*: t-tests on *deltaTimePerception(t), abs(deltaTimePerception(t))* and *correctNormalized(t)* to see if there is an effect on the presence or absence of a specific stimulus (since our t-tests are not pairwise, no t-test corrections have been performed).

To examine this effect on time judgment, we did the following as the time judgment variable is ordinal:

- *stimulusTrial(t)*: Kruskal–Wallis test on *reportedSpeedPerception(t)* to see if any significant difference appears between possible stimuli situations.
- *hasAudioTrial(t), hasVisualTrial(t), hasStimulusTrial(t)*: Wilcoxon test on *reportedSpeedPerception(t)* to see if there is an effect on the presence or absence of a specific stimulus.

In order to observe the effect of tempo on performance and time estimation, we performed ANOVAs between *tempo(t)* and the variables *deltaTimePerception(t), abs(deltaTimePerception(t))*, and *correctNormalized(t)*. The ANOVAs were also repeated across the filters *filterAUDIOONLY, filterVISUALSONLY*, and *filterBOTH* to see if differences in tempo appear only within stimuli conditions. For the effect of tempo on time judgment (*reportedSpeedPerception(t)*), we again replaced the ANOVAs with Kruskal–Wallis tests, including the repeated ones under filters. Correlations between time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) and performance (*correctNormalized(t)*) were investigated with Pearson tests. As the time judgment variable (*reportedSpeedPerception(t)*) is ordinal, Pearson tests were used to determine its correlation with time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) and performance (*correctNormalized(t)*). The confounding effect of fatigue was investigated by considering *reportedFatigue(t)* both as a nominal and as an ordinal variable; the former allows us to eventually observe differences

Table 1
Subsets for which a significant *p*-value (•) or tendency (∘) is observed for a combination of stimulus dimension and outcome variable group.

	Performance	Time Estimation	Time Judgment
Stimuli Presence	–	–	• Full Set • Trials 11-20 ∘ Fatigue Levels 1-2 • Fatigue Levels 3-4-5
Stimuli Type	• Full Set • Trials 1-10 • Trials 21-30 • Fatigue Levels 3-4-5	–	• Full Set • Trials 11-20 • Fatigue Levels 1-2
Stimuli Tempo	∘ Full Set • Trials 11-20 • Fatigue Levels 1-2	• Full Set ∘ Trials 11-20 ∘ Fatigue Levels 3-4-5	• Full Set • Fatigue Levels 1-2

Table 2
Physiological data for which a significant *p*-value (•) or tendency (∘) is observed for a target variable.

	Fatigue	Performance	Time Estimation	Time Judgment
Eye-Tracking	• EyeSpeed.std ∘ PupilDiameter.mean	• PupilDiameter.mean	∘ PupilDiameter.mean (abs)	• PupilDiameter.std
Temperature	• Temperature.mean • Temperature.std • Temperature.start-end	–	• Temperature.mean (abs)	• Temperature.std • Temperature.start-end
Heartrate	–	–	• Heartrate (signed)	–

between specific ratings and has been considered through ANOVAs with time estimation variables (*deltaTimePerception(t)*, *abs(deltaTimePerception(t))*), and performance (*correctNormalized(t)*), the latter was considered through Spearman tests on time estimation variables, performance, and time judgment (*reportedSpeedPerception(t)*). Regarding trial index's confounding effect, *trialIndex(t)* is ordinal data, so it has been investigated with Spearman tests on time estimation variables (*deltaTimePerception(t)*, *abs(deltaTimePerception(t))*), performance (*correctNormalized(t)*) as well as time judgment (*reportedSpeedPerception(t)*). Each ANOVA with a *p*-value below 0.1 would lead to a subsequent Tukey HSD, Kruskal–Wallis tests would lead to subsequent paired Wilcoxon tests. Table 1 provides an overview of the different effects per subset with a significant *p*-value or tendency, while each is discussed in detail in the following sections. For physiological data, all retained outcome variables (*PupilDilation.mean(t), PupilDilation.var(t), EyeSpeed.mean(t), EyeSpeed.var(t), Temperature.mean(t), Temperature.var(t), Temperature.start-end(t), Heartrate(t)*) have been investigated in correlation tests with the time estimation variables (Pearson tests with *deltaTimePerception(t)* and *abs(deltaTimePerception(t))*, the performance outcome variable (Pearson test with *correctNormalized(t)*), time judgment (Spearman test with *reportedSpeedPerception(t)*), and fatigue (*reportedFatigue(t)*).

Table 2 provides an overview of the observed correlations related to our physiological variables. The complete data from our tests, including confidence intervals and average values, are available online [37].

4.2. Across all trials

4.2.1. Effects of stimuli on performance

One of the aims of this study is to investigate the effects of stimuli on task performance. Looking at performance across all trials, the ANOVA between task performance (*correctNormalized(t)*) and stimulus type (*stimulusTrial(t)*) revealed no significant difference. However, when performing a t-test between task performance (*correctNormalized(t)*) and the presence of visual stimuli (*hasVisualTrial(t)*), a significant difference (*p=0.027*) can be observed alongside decreased performance when a visual stimulus is involved, as the mean with the stimulus is lesser than without. No effect is observed when considering the t-test with the presence of an audio stimulus (*hasAudioTrial(t),p = 0.167*) or any stimulus (*hasStimuluslTrial(t),p = 0.431*). Therefore, we can only observe a decrease in performance due to the presence of a visual stimulus but no effect on performance from the sole presence of any or an audio stimulus. When stimuli have the dimension of type, they also have the dimension of tempo. The ANOVA between participant performance (*correctNormalized(t)*) and stimuli tempo (*tempo(t)*) generally finds no effect of tempo. However, tempo might have an effect under a specific stimulus type. Therefore, we performed the same ANOVA but only considering subsets of data where trials contained either audio stimuli only (*filterAUDIOONLY*), visual stimuli only (*filterVISUALSONLY*), or both simultaneously (*filterBOTH*). We can then observe a tendency when trials have audio stimuli only (*filterAUDIOONLY,p = 0.088, F=2.461*). The Tukey HSD of this ANOVA reveals that the effect is significant between 180-100 (*p=0.070, diff=0.04*), with a diff value indicating that the faster tempo leads to better trial performance with audio stimuli only. In the absence of interference from visual stimuli, the faster tempo for audio stimuli may have implicitly stimulated the participant to sort objects faster.

4.2.2. Effects of stimuli on time estimation

Similar to performance, we evaluated the effect of stimuli type and tempo on time estimation variables. Likewise, ANOVAs were effectuated regarding the type of stimuli (*stimulusTrial(t)*) and tempo (*tempo(t)*) on both the normalized time estimation error (*deltaTimePerception(t)*) and its magnitude (*abs(deltaTimePerception(t))*). The only significant result is a tendency between time estimation error (*deltaTimePerception(t)*) and tempo (*tempo(t)*) (*p=0.073, F=2.328*). Tukey's HSD of this ANOVA reveals that the effect is a tendency only between tempi of 180 and 140 (*p=0.65, diff=0.067*), with trials under a tempo of 180 being rated with a longer time per second than trials under a tempo of 140. We performed similar ANOVAs involving tempo, considering subsets of data in which the trials had either only audio stimuli (*filterAUDIOONLY*), only visual stimuli (*filterVISUALSONLY*), or both at the same time (*filterBOTH*). The only significant result comes from the ANOVA between time estimation error (*deltaTimePerception(t)*) and tempo (*tempo(t)*) across trials within the AudioOnly condition (*filterAUDIOONLY*) (*p=0.039, F=3.288*), where the Tukey HSD follow-up reveals a near-significant difference between tempi of 140 and 100 (*p=0.051, diff=-0.106*) and a near tendency between 180 and 140 (*p=0.107,diff=0.091*). This means that in trials with only an audio stimulus, trials with a BPM of 140 were evaluated as faster than others, which contradicts the analysis under all types of stimuli. This contradiction may indicate the confounded effect of tempo in time perception depending on stimuli types. Finally, t-tests between our time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) and the presence of audio stimuli (*hasAudioTrial(t)*), visual (*hasVisualTrial(t)*), or any (*hasStimuluslTrial(t)*) yielded no significant result, meaning no effect of any type of stimuli present can be observed on time estimation here.

4.2.3. Effects of stimuli on time judgment

As the time judgment variable (*reportedSpeedPerception(t)*) is ordinal, we produced Kruskal–Wallis tests between the variable and the type of stimuli (*stimulusTrial(t)*) and tempo (*tempo(t)*). In the case of the test between time judgment (*reportedSpeedPerception(t)*) and the type of stimuli (*stimulusTrial(t)*), we can see a significant effect (*p=0.016, chi2=10.267*); however, a follow-up paired Wilcoxon test reveals statistical difference only between the stimulus "None" and each of the other stimuli types (*p=0.048 for None-AudioOnly,p = 0.015 for None-Both,p = 0.015 for None-VisualsOnly*). As for the test on tempo (*tempo(t)*), we observe another correlation (*p=0.001, chi2=15.432*) that, after a paired Wilcoxon, shows significant differences between 0–140 (*p=0.006*), 0–180 (*p=0.005*), 100-180 (*p=0.031*) as well as a tendency between 100-140 ((*p=0.064*) and a near-tendency between 0–100 (*p=0.12*). These two tests highlight a significant difference in time judgment depending on the presence of any stimuli (both by the differences from the "None" stimulus in the first test and the "0" BPM tempo in the second, which correspond to trials without stimuli). This is also verified by the Wilcoxon test between time judgment (*reportedSpeedPerception(t)*) and the presence of any stimulus (*hasStimuluslTrial(t)*) (*p=0.003, mean(TRUE)>mean(FALSE)*); considering the mean values, we can say that the presence of a stimulus has a significant impact in making a trial judged as passing faster than one without any. The same test has been done on the presence of audio (*hasAudioTrial(t)*) (*p=0.323*) and visuals (*hasVisualTrial(t)*) (*p=0.023, mean(TRUE)>mean(FALSE)*), meaning no significant difference in the presence or absence of an audio stimulus is observed but a fast-inducing effect is observed on the presence of a visual stimulus is recorded. In the case of tempo, the results of the paired Wilcoxon discussed earlier also indicate a significant between 100BPM and other (non-0) tempi across all types of stimuli. However, running the same Kruskal–Wallis test under subsets on "AudioOnly" trials (*filterAUDIOONLY*), "VisualOnly" trials (*filterVISUALSONLY*), and trials with both (*filterBOTH(t)*) highlights a significant difference only across trials with both stimuli (*filterBOTH(t)*) (*p=0.027, chi2=7.2343*) meaning that meanwhile tempo may have an effect across all stimuli, that effect might only be due to the combined stimuli scenario. Follow-up paired Wilcoxon tests indicate a significant difference between 100–180 (*p=0.032*) and 140-180 (*p=0.091*), the same conclusion as the tests without subsets.

4.2.4. Correlations between outcome variables

Spearman tests were used to investigate the correlation between outcome variables when time judgment (*reportedSpeedPerception(t)*) was involved as the data is ordinal; otherwise, Pearson tests were used. When comparing time estimation error (*deltaTimePerception(t)*) and performance (*correctNormalized(t)*), we see no correlation (*p=0.218,cor=-0.043*), but we see a significant negative correlation with the magnitude of time estimation error (*abs(deltaTimePerception(t))*) (*p=0.016, cor=-0.083*). This means the performance is correlated to the magnitude of time estimation errors but not to the direction; in other words, participants may generally be more error-prone in their estimations depending on their performance. When it comes to time judgment (*reportedSpeedPerception(t)*), it is negatively correlated to time estimation errors (*deltaTimePerception(t)*) (*p=1.724e−13, rho=-0.251*) and positively correlated to the magnitude of said error (*abs(deltaTimePerception(t))*) (*p=0.002, rho=0.107*). This means that the bigger the error, the faster the time is perceived, and underestimated trials are rated as passing faster. As for time judgment (*reportedSpeedPerception(t)*) and performance (*correctNormalized(t)*), better performance is associated with faster passing trials (*p=1.265e−05, rho=1.50*).

4.2.5. Confounding effect of fatigue

While running the experiment, we noticed that participants were often exhausted at the end of the session. As exhaustion affects time perception and performance, we verified if it affected our outcome variables. For its effect on performance, a Spearman test between performance (*correctNormalized(t)*) and fatigue *(reportedFatigue(t))* reveals a significant correlation *(p=2.924e−08, rho=0.190)*. By considering the fatigue variable *(reportedFatigue(t))* nominal and performing an ANOVA with performance (*correctNormalized(t)*), we retrieve this correlation (*p=1.68e−11, F=14.56*). Subsequent Tukey HSD reveals that fatigue values of "3,4,5" are statistically different from values of "1,2" as the *p*-value is below 0.001 in all these situations. Other situations (i.e., "3-4", "1-2", ...) have a *p*-value above 0.48. As for time estimations, signed error (*deltaTimePerception(t)*) is not correlated if we look through a Spearman test *(p=0.191, rho=0.045)*, but we retrieve statistical differences with the ANOVA *(p=0.004, F=3.86)*. Subsequent Tukey HSD indicates statistical differences between "3-2" *(p=0.001)*, "5-2" *(p=0.045)* and a tendency between "4-2" *(p=0.06)*. No correlation is observed for the absolute error (*abs(deltaTimePerception(t))*) with both the Spearman test (*p=0.842, rho=0.107*) and the ANOVA (*p=0.893, F=0.277*); however, it is observed for the Spearman test with time judgment (*reportedSpeedPerception(t)*) (*p=2.88e−11, rho=0.227*). From the results of the ANOVAs involving performance (*correctNormalized(t)*) and time estimation (*deltaTimePerception(t)*), we can identify two groups of reported fatigue values: "1-2" and "3-4-5". We thus decided to perform the same tests on subsets of our data according to these two groups on Section 4.4.

4.2.6. Confounding effect of trial index

Similarly to fatigue, repeated trials can affect both performance and time perception due to learning effects and repetition. We thus evaluated correlations through Pearson tests between the number of a trial across the session *(trialIndex(t))* as a continuous variable and time estimation variables (*deltaTimePerception(t), abs(deltaTimePerception(t))*) as well as performance (*correctNormalized(t)*). Regarding performance (*correctNormalized(t)*), the test reveals a correlation (*p=2.626e−14, cor=0.259*), which indicates a learning effect. Trial repetition also seems to affect time estimation as we retrieve a significant correlation with the signed time estimation error (*deltaTimePerception(t)*) (*p=1.407e−04, cor=0.131*) and a tendency with its absolute (*abs(deltaTimePerception(t))*) (*p=0.089, cor=-0.059*). Therefore, we decided to investigate different phases (beginning, middle, end) in the experiment defined by three subsets of the data based on the trial index, as shown in Fig. 4 and discussed in detail in the following Section 4.3.

4.3. Trial index subsets

Due to our results on the confounding effects on trial index as described in Section 4.2.6, we decided to investigate three subsets of the data based on the trial index with steps of ten (1–10, 11–20, 21–30). For each subset, we performed all the tests like on the full trial set, which are available for download [37] and detailed in Section 4.1. However, the normalization process only considered the targeted subset when using the sum of data on trials. We go through each subset in the following subsections, focusing on the significant results.

4.3.1. Trials 1–10

This subset corresponds to each participant's first ten trials of the experiment, constituting a discovery phase. Regarding stimuli effects on performance, the results indicate a positive effect of audio stimulus presence (t-test performance~audiopresence, *p=0.016*; TukeyHSD performance~stimuli, *p=0.099* on worse performance between visuals~both). This can be linked to the results across the entire experiment as we have seen a negative impact of visual stimulus presence and a tendency for trials with just audio to have their performance led by the tempo (see Section 4.2.1). This difference

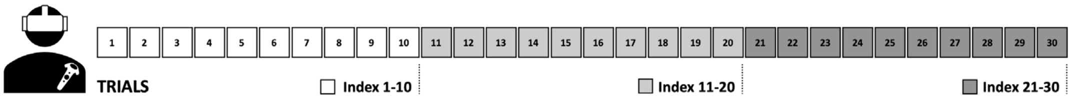


Fig. 4. Trial subset allocation for a participant by index.

might be due to a learning effect on the trials where the participants are not proficient enough to lose enough performance from visuals but may be eased by the presence of any leading audio rhythmic stimulus for this repetitive task. As for stimuli on time estimation, here we only observe a potential novelty effect on trials without stimuli as they are rarer than trials with any stimuli (t-test time estimation error stimuli presence, *p=0.093*). The most notable difference with the analysis on all trials regarding this aspect is the absence of the effect of tempo on the time estimations. Surprisingly, no effect of stimuli concerning time judgment is observed from any of our tests. When it comes to correlation between performance and time estimations through Pearson correlation tests, contrary to the full set of trials, we observe a (negative) correlation with the signed time estimation error (*p=0.025, cor=-0.142*) but not on the absolute error. Regarding time judgment concerning both time estimations and performance, we lost the correlation with the absolute time error; however, we retrieve the positive correlation from the Spearman tests with performance (*p=0.095, rho=0.106*) and the signed time estimation error (*p=-0.183, rho=-0.184*). Finally, regarding results on confounding effects of trial index and fatigue, we retrieve correlations of fatigue on performance and time experience, the correlation of performance and trial index, but none between trial index and time experience. This means that with this subset, we should have isolated an experiment phase based on trials for time perception but not for performance, which is expected as the participants were likely learning how to perform better during the first few trials.

4.3.2. Trials 11–20

This subset corresponds to each participant's ten trials in the middle of the experiment, representing a neutral phase as they no longer learn the task while not being in the experiment long enough to be bored. Regarding performance and stimuli, in this trial, we observe a performance increase from higher tempo within trials using combined stimuli (Tukey HSD 180-100, *p=0.044, diff=0.055*; 180-140, *p=0.075, diff=0.053*) As for the effect of stimuli on time estimation, we only observed a tendency between tempo 140-100 across trials solely using a visual stimulus (Tukey HSD 140-100, *p=0.100, diff=-0.083*). When it comes to time judgment, strong evidence shows that under this subset, the presence of any stimulus heavily alters it (paired Wilcoxon on time judgment and type of stimuli, *p<0.002 for all situations with "None"*; paired Wilcoxon on time judgment and tempo, *p<0.02 for all cases with "0"*; Wilcoxon time judgment and stimuli presence, (*p=8.939e–05, higher mean with stimulus*). We also observe an effect of visual stimulus specifically with the same Wilcoxon test on visual stimulus presence (*p=0.041, higher mean with stimulus*) but not on audio presence. Therefore, the effect of stimuli on time judgment in this subset is consistent with the full set regarding the effect of the present stimuli type, but we lost the effect of the tempo. This time, no correlation has been observed between time estimation and performance. However, we retrieve the time judgment correlations from the full set with Spearman tests on the performance (*p=0.004, rho=0.168*), the signed time estimation error (*p=1.988e–04, rho=0.215*), and its absolute (*p=4.20e–05, rho=-0.236*). Finally, on confounding effects, we find a correlation between fatigue and time judgment, which is expected, yet we also see a correlation tendency between trial index and signed time estimation error through a Pearson test (*p=0.071, cor = 0.105*). Still, the absence of correlation with performance indicates a proper subset division on the trial index.

4.3.3. Trials 21–30

This subset corresponds to each participant's last ten trials, representing the end of the experiment and, thus, a phase where the participant is possibly tired or bored. Here, investigation of performance suggests that when there are stimuli, the presence of visual stimulus leads to worse task performance (Tukey HSD performance between stimuli modes "Both" and "AudioOnly", *p=0.010, diff=-0.045*; "VisualsOnly" and "AudioOnly", *p=0.026, diff=-0.041*; t-test on performance and visual presence, *p=1.449e–04, lower mean when the stimulus is present*). Under this subset, nothing significant has been observed in the relation between stimuli (type or tempo) and time perception (time estimation and judgment). Similarly to the previous set (Trials 11–20), no correlation between time estimation and performance is observed. The relation from Spearman tests between time judgment with both performance and time estimation is similar to the subset at the beginning of the experiment (Trials 1–10), where performance is positively correlated (*p=0.002, rho=0.176*) and time estimation error is negatively correlated (*p=4.795e–04, rho=-0.202*) but the absolute error is not. On confounding effect, while finding effects of fatigue on time judgment as expected, unfortunately, we see tendencies on the effects of the trial index on both time estimation error (*p=0.067, cor=0.107*) and task performance (*p=0.073, cor=-0.104*) from Pearson tests. This may indicate a transition between phases of boredom and tiredness relative to the time spent in the experiment.

4.4. Fatigue subsets

Having obtained the results on the confounding effects of fatigue described in Section 4.2.5, we decided to investigate two subsets depending on the participants' answers on fatigue, one for fatigue at 1 or 2, and one for fatigue at 3, 4, or 5 (see Fig. 5). Similar to the previous subsets, for each, we performed all tests on the complete trial set, which can be found online [37], and the same modification on the variable normalization process by only considering the targeted subset when using the sum of data on trials. In the following sections, we will again focus exclusively on significant test results and will not re-elaborate the methodology.

4.4.1. Fatigue levels 1–2

This subset corresponds to the participant experiencing "low" fatigue. First, concerning performance and stimuli tempo, we can observe the lesser performance of stimuli with a tempo of 100 (Tukey HSD on performance and tempo between 140 and 100 bpm, *p=0.056, diff=0.065*; 180 and 100 bpm, *p=0.020, diff=0.081*). This finding can be aligned to results from the full set (performance dependent on tempo for audio stimuli) and trials 11–20 (180 bpm leading to better performance under combined stimuli). No general effect of stimuli type on performance is observed, either from the specific situations possible or the presence of a modality. This subset yielded no significant insights regarding stimuli dimensions (type and tempo) and time estimation. Regarding time judgment and tempo, however, we observe significant differences between 180 and 100 bpm across all stimuli (Paired Wilcoxon on time judgment and tempo between 100 and 180 bpm, *p=0.023*) as well as an effect of the presence of 180 bpm (Wilcoxon between 0 and 180, *p=0.025*). When considering only trials with combined stimuli, the paired Wilcoxon shows a significant

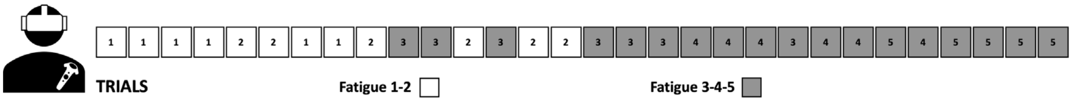


Fig. 5. Example trial subset allocation for a participant by fatigue.

difference between 180-100 ($p=0.017$) and a tendency between 100–140 ($p=0.094$), which is consistent with the time judgment effects results on the complete set of trials. As for time judgment and stimuli type, we observe another consistent result from the full set as stimuli tend to be judged faster when there is any stimulus (Wilcoxon on the presence of any stimulus, $p=0.073$) or if there is at least a visual one ($p=0.020$). Looking for a correlation between performance, time estimation, and time judgment yielded similar results to the subset of Trials on index 1–10. With a tendency of a negative correlation ($p=0.059$, $cor=-0.141$) from Pearson between time estimation error and performance, a significant negative correlation ($p=0.022$, $rho=-0.171$) out of the Spearman between time estimation error and time judgment, a positive one ($p=0.033$, $rho=0.157$) between performance and time judgment, but no correlations from the absolute time estimation error. Confounding effects of trial index on performance (Pearson test, $p=0.086$, $cor=0.129$) are similar to trials 1–10, which is not surprising as early trials probably are low fatigue trials. A confounding effect of fatigue is not observed for both time estimation and judgment; however, we can observe it for performance (Spearman test, $p=0.037$, $rho=0.157$; Tukey HSD (more of a t-test considering we have two values in this subset), $p=0.048$, $F=3.977$, $diff2\text{-}1=0.046$). We can assume that higher fatigue trials in this subset would occur after the learning phase when the participant is more proficient.

4.4.2. Fatigue levels 3–4–5

This subset corresponds to the participant having a higher fatigue level. Concerning performance and stimuli type, like for the full set and trials 21–30, we observe a negative impact from the presence of visual stimuli (TukeyHSD on stimuli type and performance between VisualsOnly and Audio Only, $p=0.080$, $diff=-0.024$; t-test between performance and presence of visuals, ($p=0.020$)). Some observations converge towards contextual effect on tempo depending on the type of stimulus of the trial (TukeyHSD on absolute time estimation error between tempos 140 and 180 for audio trials, $p=0.086$, $diff=0.1$; TukeyHSD on signed time error between tempos 140 and 180 across all, $p=0.080$, $diff=0.087$). As for time judgment and stimuli, we only found evidence indicating an effect of general stimulus presence (Paired Wilcoxon on stimuli type and time judgment, $p<0.07$ for pairs involving “None”; Wilcoxon on time judgment and stimulus presence, $p<0.011$; paired Wilcoxon on tempos, $p=0.032$) between 0–140 and $p=0.040$ for 0–180). Looking for a correlation between performance, time estimation, and time judgment yielded similar results to the full set of trials. From Spearman tests with time judgment, we retrieve the negative correlation with the time estimation error ($p=2.217e{-}11$, $rho=-0.256$), the positive correlation with the absolute error ($p=4.366e{-}04$, $rho=0.136$) and with the performance ($p=3.915e{-}05$, $rho=0.159$). We do not retrieve the significant p-value on the Pearson test between performance and absolute error, but a near-tendency ($p=0.110$, $cor=-0.062$). While we do not observe a confounding effect from the ANOVAs between fatigue and time estimation variables, we see a tendency ($p=0.089$, $rho=0.066$) from the Spearman test on the absolute time estimation error. The effect of fatigue is more pronounced on performance (Pearson test, $p=0.029$, $rho=0.085$) and from the subsequent Tukey HSD of the ANOVA ($p=0.064$, $F=2.768$) the difference appears to be between 3–5 ($p=0.050$, $diff=0.023$). Fatigue also seems to significantly affect time judgment ($p=2.185e{-}08$, $rho=0.215$). In this subset, the fatigue levels 3 and 5 may be significantly different on both performance and time judgment; however, this is apparently due to the normalization on the subset and was not observable across all trials. The confounding effects of the trial index observed from Pearson tests are similar to those of the full set, which is not too surprising as the subset is rather large and was not made to minimize the effect of the index.

4.5. Results on physiological data

As stated in earlier sections, one of our objectives is to investigate the interrelationship of physiological measures, fatigue, task performance, and time perception. To this end, we systematically performed a Pearson test between all extracted values discussed in Section 4.1.6 and *deltaPerception(t), abs(deltaPerception(t))*, and *performance(t)*, and a Spearman test with these values and both *subjFatigue(t)* and *subjTimeRating(t)*. Contrary to the analysis on trial parameters, we could not perform this analysis under the different subsets due to the amount of data lost on the Empatica data and because we are trying to detach physiological data from the activity itself. Looking at the correlation on subjective fatigue, we find correlations on the eye speed variance ($p=9.43e{-}5$, $rho=0.134$), all temperature values (mean $p=0.013$, $rho=0.088$; variance $p=1.38e{-}10$, $rho=-0.225$; start-end $p=0.005$, $rho=-0.099$) and a tendency for pupil diameter ($p=0.097$, $rho=-0.057$). Regarding temperature, we can observe correlations between the temperature mean and the magnitude of time estimation error ($p=0.029$, $cor=-0.077$), between temperature variance and subjective time rating ($p=2.11e{-}5$, $cor=-0.150$), as well as between the subjective time rating with temperature difference at the end and start of the trial ($p=0.001$, $cor=-0.114$). Pupil diameter also has a tendency on its mean with the magnitude of time estimation error ($p=0.088$, $cor=0.059$) and a significant correlation with participant performance ($p=0.023$, $cor=-0.079$). However, while the mean pupil diameter is tangled between multiple confounded variables, its variance only correlates to the subjective time rating ($p=1.71e{-}5$, $cor=0.148$). As for the heart rate, we only observe a tendency with the signed time error ($p=0.071$, $cor=-0.130$), which, while not being a significant p-value, is, with the variance of pupil diameter, one of the only values that are not also correlated with fatigue.

5. Discussion

We conducted a VR experiment in which participants repeatedly performed a simple sorting task subjected to different stimuli conditions. This allowed us to gather numerous data, including information related to task proficiency, subjective data from questionnaires, and physiological data. The overarching goal was to explore relationships between time experience, task performance, environment/stimuli conditions, and physiological cues. However, we found that for some of the results on the complete data set, it was necessary to investigate closer multiple subsets, which we will discuss together with their implications for VR application design. Looking at the entire data set, we can observe specific effects of stimuli type and tempo on different aspects of time perception and performance, as well as some interesting correlations between those variables, which appeared to be also heavily impacted by the trial index and fatigue through the experiment. Therefore, we defined subsets of data based on the trial index and difference values from ANOVAs for fatigue. As indicated in Table 1, we can observe effects of stimuli presence, type, and tempo on performance and time experience depending on the subset. Some of the data and correlations align between subsets while others do not, which may indicate contextual effects of stimuli on performance and time experience depending on task repetition and fatigue.

5.1. Observations on task performance and stimuli

A central result from the analysis of the complete set of trials is how the presence of visual stimuli negatively impacts task performance. This is coherent with our previous study and is to be anticipated as the task requires visual attention, and those stimuli may be disturbing. However, within subsets, this result is observed only for later trials (index 21–30) and high fatigue (fatigue 3-4-5). Surprisingly, we see a positive effect on performance from the presence of audio but only in the early trials (index 1–10) and no effect of stimuli type presence in between (index 11–20). This could be interpreted as the disturbance of visual stimulus not being impactful enough when one is learning the task or not physically tired. We can also interpret the presence of audio stimuli as beneficial for this task only when the participant is in a learning phase. Effects of tempo are observed on trials within trials with only the audio stimulus when considering all trials and within trials with combined stimuli within the subset of trials from index 11–20. In both cases, the faster tempo led to faster performance, which indicates an invitation to go faster in the task from the faster stimuli; however, the participant's interpretation depends on the context.

5.2. Observations on time estimation and stimuli

Time estimation variables are defined from the difference between (normalized) participants' estimation of time taken for a trial and the actual time of a trial; we thus talk about the time estimation error and its absolute, which represents the magnitude of error regardless of whether the participant under- or overestimated the length of a trial. A global effect of tempo can only be observed with the complete set of trials between 180–140 (with 180 being overestimated). As for differentiation within stimuli situations, we see a time estimation error difference on the audio stimuli for the entire trial set and high-fatigue trials, and an absolute error difference in visuals for trials 11–20 as well as on combined stimuli for high fatigue. These results show a tendency of the 140 bpm tempo leading to fewer (absolute) estimation errors and being underestimated compared to 100–180. Another overestimating effect from stimulus presence is observed for trials 1–10. Overall, we also observe context-dependent effects of stimuli as the type of stimuli will affect one's time perception differently depending on the index or fatigue.

5.3. Observations on time judgment and stimuli

Time judgment or time passage refers to the subjective evaluation of a participant on whether they think a trial is going by fast or slow. It differs from time estimation in that the participant gives their subjective feeling about the time spent, whereas time estimation is an attempt by the participant to be objective about time. Time judgment has semi-constant results of the presence of any stimuli inducing faster perception; this is observed across all trials for both subsets on fatigue and the subset on stimuli 11–20. We can also observe a specific fastening effect of visual stimuli on all these sets affected by the presence of any stimuli except for the high fatigue one. The absence of these observations on subsets of trials either at the beginning or at the end of the experiment might indicate that the participant needs to get used to and, over time, gets too used to the presence of stimuli to be noticeable, regardless of fatigue levels. Another effect observed only on the complete set and for low fatigue is a difference between tempo in general and within trials with combined stimuli.

5.4. Time judgment, time estimation, and performance balance

Two correlations were consistent across all sets: a negative correlation between time estimation error and time judgment and a positive correlation between performance and time judgment. The first means that a subjective faster trial reflects under-estimation of time. The second means that when the participant rated the trial as faster than usual, they would perform better. This could directly be tied to the notion of flow as two elements of flow states are the challenge-skill balance and time transformation. The similarity between time estimation error and time judgment is indicative that our time transformation was a general time experience shift and not a side effect of disorientation (i.e., a participant judging a trial as fast because they thought it was a higher amount of time that actually passed). Among low fatigue and early trials, we also retrieve a negative correlation between performance and performance, reinforcing the flow approach. The time estimation error magnitude and performance are negatively correlated under all trials and high fatigue, which means that possibly, in a specific context, higher time transformation was generally detrimental to performance. However, this is against the flow definition, and combined with previous observations, it may imply that we are approaching flow states only with time transformations that are an underestimation. We also observed positive correlations between this magnitude and time judgment with all trials, the 11–20, and high fatigue subsets, which could be interpreted as the presence of any time transformation potentially leading to faster time passage in general.

5.5. Observations on physiological values

The results on physiological values are strongly affected by fatigue, as most extracted values correlate with it. Overall, our results imply that the same physiological outputs may represent time estimation and feeling of time passage differently. However, we cannot overlook how fatigue confounds with all other target variables of time perception and performance. So, while we do observe correlations with time perception variables on the mean of pupil dilation and temperature-related variables, the fact that these physiological values are also correlated with fatigue weakens their significance. While we can consider these values promising for assessing one's time experience, especially in the case of body temperature, additional research designed around tasks that are not fatigue-inducing would benefit this. These findings may also allow for looking at the problem from the other side, as fatigue is easier to assess than time perception through noninvasive sensors such as eye-tracking and body temperature; due to its influence, fatigue may be used as a proxy for time perception. However, we have two variables that are not observed as correlated with fatigue, which are correlated with time perception measurements: variance of pupil diameter (significant with subjective time rating) and heart rate (tendency with time estimation error).

5.6. Limitations

It is important to remember that the effect of the rhythmic stimuli in our experiment is contextualized in the particular scenario of our sorting task. We can also see some limits from the confounding effects of task familiarity and fatigue, and even with the subsets, which unfortunately implies using less data and thus having lower statistical power relevance (especially in the case of low fatigue), we can isolate the effect of at most one confounding effect but not both at the same time. In the case of physiological data, as shown in Table 3, the inter-correlation between fatigue, time estimation, time passage, and performance limits the scope of these findings, as discussed in Section 5.5. Individual per-participant differences are also to be considered, as through casual talks with the participants, we know of varying degrees of VR experience between participants; however, this data was not recorded and is thus not included in our analysis. It should also be mentioned that due to the fact that the self-assessment of fatigue had to be answered quickly in this study, standard tests such as the NASA Task Load Index could not be employed. Consequently, we cannot distinguish between mental and physical fatigue, which could be considered in future studies.

6. Conclusion

In this article, we used a simple sorting task to explore how rhythmic stimuli affect time experience and task performance in VR. We

Table 3
Results of correlation tests on investigation of cofounding factors. With annotation of significant p-values (•).

	deltaPerception	Abs (deltaPerception)	performance	subjTimeRating	subjFatigue
deltaPerception	–	–	P = 0.2177 Cor = −0.04257087	• P = 1.724e−13 Rho = −0.2505479	P = 0.1918 Rho = 0.04507991
Abs (deltaPerception)	–	–	• P = 0.01602 Cor = −0.08308206	• P= 0.001995 Rho = 0.1065027	P = 0.8419 Rho = 0.1065027
performance	–	–	–	• P = 1.265e−05 Rho = 0.1499229	• P = 2.924e−08 Rho = 0.1897809
subjTimeRating	–	–	–	–	• P = 2.924e−08 Rho = 0.1897809

found that the context concerning the trial index (repetition of the action) and fatigue affected these aspects of the user experience. Depending on the familiarity with the task, the presence of a particular type of rhythmic stimulus under possible tempos will affect either performance or time experience. Both aspects can contribute significantly to a flow experience or even well-being in general, and the results of this study can thus inform the design of future interactive VR applications. We also found promising physiological variables to assess one's time experience, which would benefit from future research to untangle them from fatigue's effect or use fatigue as a proxy to assess time perception.

While the familiarity or repetition of a task or action can be easily assessed in any interactive application, using fatigue as a modulator could be a growing opportunity for VR developers as newer HMDs incorporate advanced sensors, e.g., for eye-tracking. We observed effects of rhythmic stimuli under some fatigue and task familiarity, yet the crucial finding is the presence of effect variation rather than the specific effect itself, highlighting the need for studies of time perception concerning context- and subject-dependent time modulations.

CRediT authorship contribution statement

Stéven Picard: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jean Botev:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964464 (ChronoPilot).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cag.2024.104097.

References

[1] Picard S, Botev J. Rhythmic stimuli effects on subjective time perception in immersive virtual environments. In: Proceedings of the 14th international workshop on immersive mixed and virtual environment systems. MMVE '22, New York, NY, USA: Association for Computing Machinery; 2022, p. 5–11. http://dx.doi.org/10.1145/3534086.3534330.

[2] Picard S, Botev J. Rhythmic stimuli and time experience in virtual reality. In: Proceedings of the 20th euroXR international conference. EuroXR '23, Springer, Cham; 2023, p. 53–75. http://dx.doi.org/10.1007/978-3-031-48495-7_4.

[3] Hoagland H. The physiological control of judgments of duration: Evidence for a chemical clock. J Gen Psychol 1933;9(2):267–87. http://dx.doi.org/10.1080/00221309.1933.9920937.

[4] Ghaderi A, Niemeier M, Crawford JD. Linear vector models of time perception account for saccade and stimulus novelty interactions. bioRxiv 2021. http://dx.doi.org/10.1101/2020.07.13.201087.

[5] Gorea A. Ticks per thought or thoughts per tick? A selective review of time perception with hints on future research. J Physiol Paris 2011;105(4):153–63. http://dx.doi.org/10.1016/j.jphysparis.2011.09.008.

[6] Droit-Volet S, Fayolle S, Lamotte M, Gil S. Time, emotion and the embodiment of timing. Timing Time Percept 2013;1(1):99–126. http://dx.doi.org/10.1163/22134468-00002004.

[7] Droit-Volet S, El-Azhari A, Haddar S, Drago R, Gil S. Similar time distortions under the effect of emotion for durations of several minutes and a few seconds. Acta Psychol 2020;210:103170. http://dx.doi.org/10.1016/j.actpsy.2020.103170.

[8] Makwana M, Srinivasan N. Intended outcome expands in time. Sci Rep 2017;7(1):6305. http://dx.doi.org/10.1038/s41598-017-05803-1.

[9] Roseboom W, Fountas Z, Nikiforou K, Bhowmik D, Shanahan M, Seth AK. Activity in perceptual classification networks as a basis for human subjective time perception. Nature Commun 2019;10(1). http://dx.doi.org/10.1038/s41467-018-08194-7.

[10] Watt JD. Effect of boredom proneness on time perception. Psychol Rep 1991;69(1):323–7. http://dx.doi.org/10.2466/pr0.1991.69.1.323, PMID: 1961817.

[11] Rutrecht H, Wittmann M, Khoshnoud S, Igarzábal FA. Time speeds up during flow states: A study in virtual reality with the video game thumper. Timing Time Percept 2021;9(4):353–76. http://dx.doi.org/10.1163/22134468-bja10033.

[12] Jackson SA, Eklund RC. Assessing flow in physical activity: The flow state scale–2 and dispositional flow scale–2. J Sport Exerc Psychol 2002;24(2):133–50. http://dx.doi.org/10.1123/jsep.24.2.133.

[13] Pelet J-É, Ettis S, Cowart K. Optimal experience of flow enhanced by telepresence: Evidence from social media use. Inf Manag 2017;54(1):115–28. http://dx.doi.org/10.1016/j.im.2016.05.001.

[14] Nah F, Eschenbrenner B. Flow experience in virtual worlds: Individuals versus Dyads. In: SIGHCI 2015 proceedings. 2015, URL: https://aisel.aisnet.org/sighci2015/19.

[15] Reid D. A model of playfulness and flow in virtual reality interactions. Presence Teleoperators Virtual Environ 2004;13(4):451–62. http://dx.doi.org/10.1162/1054746041944777.

[16] Mullen G, Davidenko N. Time compression in virtual reality. Timing Time Percept 2021;9:1–16. http://dx.doi.org/10.1163/22134468-bja10034.

[17] Igarzábal FA, Hruby H, Witowska J, Khoshnoud S, Wittmann M. What happens while waiting in virtual reality? A comparison between a virtual and a real waiting situation concerning boredom, self-regulation, and the experience of time. Technol Mind Behav 2021-07-22;2(2). http://dx.doi.org/10.1037/tmb0000038, URL: https://tmb.apaopen.org/pub/what-happens-while-waiting-in-virtual-reality.

[18] Mallam SC, Ernstsen J, Nazir S. Accuracy of time duration estimations in virtual reality. Proc Hum Factors Ergon Soc Annu Meet 2020;64(1):2079–83. http://dx.doi.org/10.1177/1071181320641503.

[19] Bruder G, Steinicke F. Time perception during walking in virtual environments. In: 2014 IEEE virtual reality. VR, 2014, p. 67–8. http://dx.doi.org/10.1109/VR.2014.6802054.

[20] Schatzschneider C, Bruder G, Steinicke F. Who turned the clock? Effects of manipulated zeitgebers, cognitive load and immersion on time estimation. IEEE Trans Vis Comput Graphics 2016;22(4):1387–95. http://dx.doi.org/10.1109/tvcg.2016.2518137.

[21] van der Ham IJ, Klaassen F, van Schie K, Cuperus A. Elapsed time estimates in virtual reality and the physical world: The role of arousal and emotional valence. Comput Hum Behav 2019;94:77–81. http://dx.doi.org/10.1016/j.chb.2019.01.005.

[22] Unruh F, Landeck M, Oberdörfer S, Lugrin JL, Latoschik ME. The influence of avatar embodiment on time perception - towards VR for time-based therapy. Front Virtual Real 2021;2. http://dx.doi.org/10.3389/frvir.2021.658509.

[23] Droit-Volet S, Ramos D, Bueno JLO, Bigand E. Music, emotion, and time perception: the influence of subjective emotional valence and arousal? Front Psychol 2013;4:417. http://dx.doi.org/10.3389/fpsyg.2013.00417.

[24] Hammerschmidt D, Wöllner C, London J, Burger B. Disco time: The relationship between perceived duration and tempo in music. Music Sci 2021;4:2059204320986384. http://dx.doi.org/10.1177/2059204320986384.

[25] Wöllner C, Hammerschmidt D. Tapping to hip-hop: Effects of cognitive load, arousal, and musical meter on time experiences. Atten Percept Psychophys 2021;83(4):1552–61. http://dx.doi.org/10.3758/s13414-020-02227-4.

[26] Peters CM, Glazebrook CM. Rhythmic auditory stimuli heard before and during a reaching movement elicit performance improvements in both temporal and spatial movement parameters. Acta Psychol 2020;207:103086. http://dx.doi.org/10.1016/j.actpsy.2020.103086.

[27] Lukas S, Philipp AM, Koch I. Crossmodal attention switching: Auditory dominance in temporal discrimination tasks. Acta Psychol 2014;153:139–46. http://dx.doi.org/10.1016/j.actpsy.2014.10.003.

[28] Wang X, Wöllner C, Shi Z. Perceiving tempo in incongruent audiovisual presentations of human motion: Evidence for a visual driving effect. Timing Time Percept 2021;10(1):75–95. http://dx.doi.org/10.1163/22134468-bja10036.

[29] Wearden J, Penton-Voak I. Feeling the heat: Body temperature and the rate of subjective time, revisited. Q J Exp Psychol Sect B 1995;48(2):129–41. http://dx.doi.org/10.1080/14640749508401443.

[30] Kingma BR, Roijendijk LM, Maanen LV, Rijn HV, Beurden MHV. Time perception and timed decision task performance during passive heat stress. Temp Austin 2020;8(1):53–63. http://dx.doi.org/10.1080/23328940.2020.1776925.

[31] Dormal V, Heeren A, Pesenti M, Maurage P. Time perception is not for the faint-hearted? Physiological arousal does not influence duration categorisation. Cogn Process 2017;19(3):399–409. http://dx.doi.org/10.1007/s10339-017-0852-3.

[32] Cho D, Ham J, Oh J, Park J, Kim S, Lee NK, Lee B. Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine. Sens Basel 2017;17(10):2435. http://dx.doi.org/10.3390/s17102435.

[33] Chen H, Dey A, Billinghurst M, Lindeman RW. Exploring pupil dilation in emotional virtual reality environments. In: ICAT-EGVE 2017 - International conference on artificial reality and telexistence and eurographics symposium on virtual environments. The Eurographics Association; 2017, http://dx.doi.org/10.2312/EGVE.20171355.

[34] Lambourne K. The effects of acute exercise on temporal generalization. Q J Exp Psychol Hove 2012;65(3):526–40. http://dx.doi.org/10.1080/17470218.2011.605959.

[35] Yarrow K, Haggard P, Heal R, Brown P, Rothwell JC. Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. Nature 2001;414(6861):302–5. http://dx.doi.org/10.1038/35104551.

[36] Knöll J, Morrone MC, Bremmer F. Spatio-temporal topography of saccadic overestimation of time. Vis Res 2013;83:56–65. http://dx.doi.org/10.1016/j.visres.2013.02.013.

[37] Picard S, Botev J. Rhythmic Stimuli and time experience in virtual reality - complete data analysis. Zenodo; 2023, http://dx.doi.org/10.5281/zenodo.10804031.

Appendix D

Paper: A Dynamic and Scriptable Environment and Framework for Stimulus-Based Cognitive Research in Virtual Reality

Full paper of Section 3.2.1

2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)

A Dynamic and Scriptable Environment and Framework for Stimulus-Based Cognitive Research in Virtual Reality

Stéven Picard* Jean Botev† Sahar Niknam‡

VR/AR Lab
University of Luxembourg

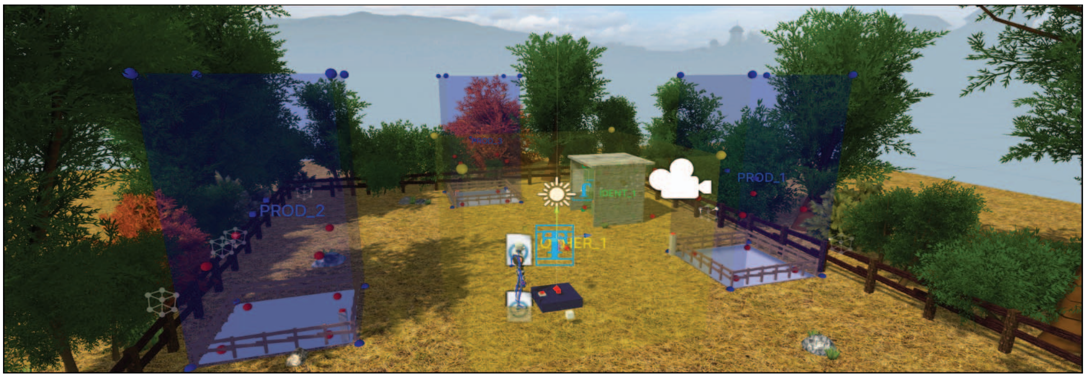


Figure 1: Example scene in the editor with various assets and experimenter-defined, scriptable task zones in different colors.

Abstract

Time perception is essential to immersive media experience, and particularly virtual reality. With the relevant technology becoming both readily available and affordable in recent years, there has been a corresponding growth in interdisciplinary research on time perception and virtual reality. This paper presents a fully customizable virtual environment and framework devised for such studies, which can furthermore be easily extended to accommodate any stimulus-based cognitive and behavioral research. The different elements of the environment can be defined in simple text-based configuration files to load and deploy new components and experimental setups quickly. Due to the generic architecture, experimenters can also control all elements externally via hardware-agnostic network messages.

Index Terms: Human-centered computing—Virtual reality Human-centered computing—User studies Software and its engineering—Software development techniques—Reusability

1 Introduction

In recent years virtual reality (VR) has become increasingly accessible and revealed itself as an excellent tool for behavioral studies, providing experimenters with an easy way to reproduce fully controlled environments and tasks across participants. However, creating these requires much technical effort, so making this work reusable across different experiments is a valuable resource.

Within the ChronoPilot project [4], which aims at active subjective time modulation, a substantial part of our activities is focused on stimulus-based behavioral studies of the user experience and time perception. Therefore, we developed a dynamic and scriptable environment and framework that, while tailored to the specific needs of our group in terms of time perception studies, can support any type of behavioral research scenario employing VR.

Before focusing on the framework itself, the following section will cover relevant concepts and notions related to time perception and its possible applications, VR behavioral studies, and existing frameworks. We will then discuss the experimental design of our environment, specifically its controls, behaviors, and settings, and explore the technical aspects that enable scriptability and dynamics.

2 Background and Related Work

Time perception can be considered from a design point of view as it is an integral part of the user experience with an application or an activity. Studies have been trying to identify ways to use the non-accurate perception of time to the user's advantage; for instance, a faster-loading animation (such as a rotating circle animation) yields a more compressed time perception than slower ones [24]. A similar example is related to downloading data which, according to Gorn et al.'s study testing fake download web pages that differ in color, that parameter appears to influence relaxation, which in turn influences the perceived download speed [10].

As such, different conditions or stimuli affect time perception aspects. For instance, when it comes to the setting, environments tend to have temporal cues or "zeitgebers" such as clocks or the sun's movement. Schatzschneider et al. investigated the latter's effect on time perception with and without cognitive loads as a zeitgeber in a virtual environment (VE). They found that the absence or presence of sun movement, as well as the presence of a task, affect time perception; however, the zeitgeber's speed did not affect time perception, and if its presence affected time perception, it did not affect task performance [21]. Davydenko and Peetz's work furthermore suggested that walks in nature felt longer than in an urban environment, showing that the environment affects time perception and nature's properties influence mood and behavior [6].Other conditions that can affect time perception can be tied to specific properties of stimuli. In a series of experiments, Droit-Volet et al. observed

*e-mail: steven.picard@uni.lu
†e-mail: jean.botev@uni.lu
‡e-mail: sahar.niknam@uni.lu

DOI 10.1109/ISMAR-Adjunct60411.2023.00084

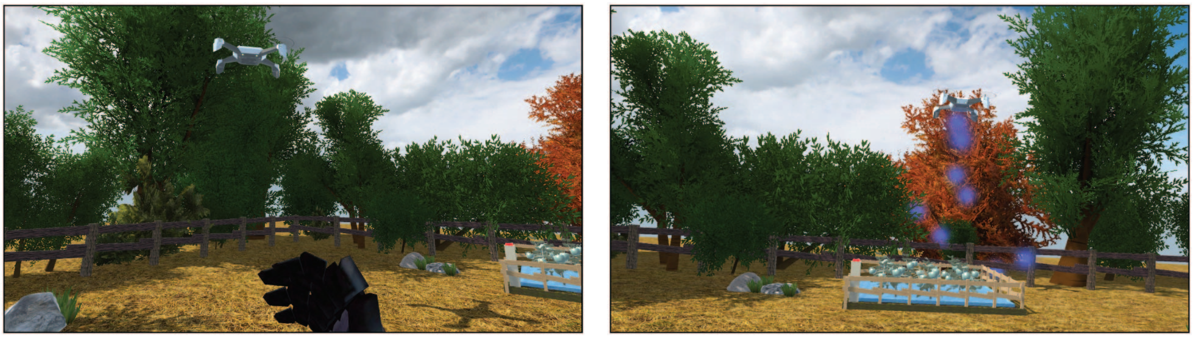


Figure 2: Drone moving towards a zone (left) and watering the plants in the target zone (right).

how higher tempos induce longer subjective time, and emotional valence decreases (but does not suppress) the effect of tempo while affecting time perception [7].A recent study by Hammerschmidt et al. [11] explored different timing evaluations (reproduction, estimation, and subjective rating) of instrumental excerpts of Disco songs at different tempi, revealing interesting tempo properties regarding time perception. Non-rhythmic stimuli properties also affect time perception such as motion. For instance, a study by Fornaciai et al. suggests the adaptation (i.e., the effect of repetition) of fast translation motion compresses time, but this effect was not observed for radial and circular motions [8].

VR, particularly when experienced through fully enclosing head-mounted displays (HMDs), is an ideal medium for psychological studies and testing new stimuli. A notable example of the potential is how Outram et al. employed artificial synesthesia in VR to demonstrate the possibilities of a full-body vibrotactile haptic suit using vibrations and synesthetically linked visuals, audio, and vibrations to increase immersion and satisfaction. The authors suggest that "as being highly enjoyable, the use of such environments may have implications for the exploration of altered states including flow, meditation and ecstasies, and the next stages of related research will be to measure these psychological aspects" [3]. Another example of artificial VR synesthesia can be found in Reif and Alhalabi's study [20], where it was used to control attention and increase immersion in the context of pain therapy. This possibility of immersing the users in fully controllable environments makes VR a perfect tool to investigate stimuli-based perceptual time transformations.

Because creating VEs to study these stimuli is software intensive, various frameworks to facilitate the development of user behavioral studies have emerged in recent years. Open-source frameworks like the Unity Experiment Framework (UXF) [5], the Biomotion-Lab Toolkit for Unity Experiments (bmlTUX) [2], or the vexptoolbox [22], aim to alleviate the coding work for experimenters wanting to use VR-compatible 3D engines such as Unity or Vizard for their experiments. These frameworks primarily provide tools to simplify or automate generic workloads typical to most behavioral studies, such as trial generation and sequencing, independent variables conditions, or data collection. Some frameworks, e.g., the Unified Suite for Experiments (USE) [25], have more advanced features like replaying trials, millisecond precision, and hardware integration. For time perception studies specifically, Landeck et al. created an Unreal Engine 4 framework mainly focused on zeitgebers, i.e., environmental time cues, by giving the option to modify available events on three dimensions (velocity, synchronicity, and density) [12]. These above frameworks aim to either test specific scenarios and stimuli [12] or simplify experimental designs [2,5,22,25]; however, they do not necessarily provide the tools to build a timeline of in-trial events and stimuli, as does our environment framework.

3 Example Design

We will illustrate the framework using an example VE prototype covering an initial precision farming application scenario developed in the context of the ChronoPilot project in order to test various mechanisms and stimuli for subjective time modulation in VR.

3.1 Objectives

The main focus of the VE is on enabling the execution of single-user evaluation tasks. The precision farming scenario naturally integrates time-modulating stimuli and time estimation metrics, emphasizing complex overall problem-solving. Users must perform a series of tasks in the VE characterized by specific deadlines, such as watering, harvesting, and sowing crops. They can assign (virtual) robotic agents, such as drones, to perform these tasks. The various robot types differ in their execution time for tasks, which arrive online, and the user must, for instance, allocate resources to the various tasks to maximize the number of completed tasks. Consequently, cognitive load constitutes one of the main variables in the current scenario, tuned by different parameters such as the number of tasks ready to schedule, the arrival rate of tasks, task criticality, or the number and composition of available resources and solutions. The fundamental idea behind precision farming is to support traditional farming practices with data-driven and AI-enabled technologies (cf. [9], [17], [19], or [27]). With this in mind, precision farming is a promising scenario to realistically test the potential of novel time modulation techniques to improve overall performance and reduce the anxiety of working with advanced technologies altering time perception. The main objectives in developing the VE were its generalizability and scalability while providing an immersive user experience at a practical yet sufficient level.

3.2 Tasks and Task Types

In the experiments and user studies conducted with the current VE, users perform control and learning tasks. The two major VE task types, production and identification, reflect these categories. Experimenters can generally define arbitrary task types depending on what is needed in the specific experimental environment. Tasks can be cumulative, i.e., multiple tasks may be active simultaneously, which allows for adjusting the cognitive load during the experiment and inducing stress or boredom.

The main goal of the production task is to meet a plant's needs and harvest it when it is fully grown. In this task type, the needs of the plants are known a priori, making it essentially a resource-allocation task, with the users deploying drones remotely via a specialized, arm-based interface (see Section 3.2) and assigning the desired parameter values, e.g., to water plants (see Figure 2). Currently, the production task only allows the deployment of drones to irrigate plants, but it can be extended to more dimensions and to include further steps (e.g., powering up the drones).

The command to harvest a crop is given by pressing a nearby button (see Figure 3), with the crop being replaced as soon as the users lift their hand, which forces them to move and navigate within the VE instead of sending drones during the growth phase.

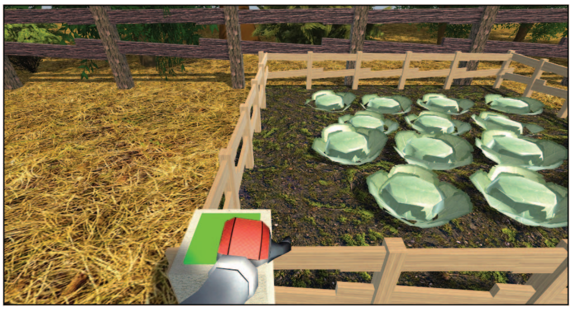


Figure 3: Manual activation of the harvesting sequence.

The user's primary goal regarding the identification task type is to explore and find information about an unknown plant. The retrievable information can be the necessary water levels for a plant to grow or the plant's health status. For this purpose, the user has access to manual controls affecting the properties of the plant; in the present example, they allow to increase or decrease the water level. The plant state can be ”reset” by harvesting and planting a new seed. Once the user feels confident knowing the plant, they can answer a questionnaire to complete the task (see Figure 4).

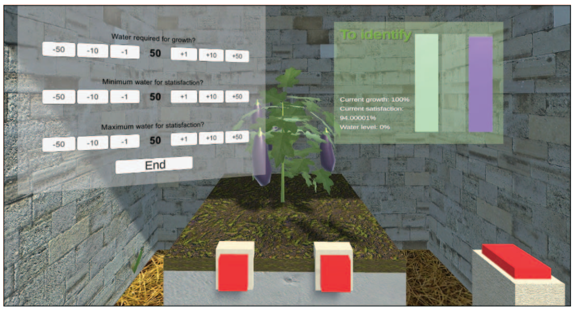


Figure 4: Prefab with plant status panel and questionnaire.

3.3 User Interface and Experience

To provide a compelling, immersive, and seamless user experience (UX) in VR, several aspects must be considered in the user interface (UI) design. Due to the three-dimensional nature of the environment, it is essential to provide spatially consistent feedback that matches the user's position and orientation in the VE. In general, the VE and its embedded UI should be intuitive and easy to navigate, with clear affordances, i.e., cues about where the user should look and where they can interact, employing VR-appropriate interaction methods such as gaze-based selection, hand gestures, and physical controllers. During navigation, relevant contextual information and feedback related to the user's location and actions in the VE must be presented with sufficiently large and clearly visible text and graphics to ensure readability [13, 23]. A specific goal in designing the UI for our test environment is to fully realize the desired level of interaction while making the UI cognitively and physically comfortable to use.

Two main categories of interaction are implemented in the VE: integrated objects in the environment that function as actuators, such as harvest buttons next to the production zones, and a control panel to operate or assist agents and obtain information. Recent human-computer interaction (HCI) studies show that implementing UI elements as integrated objects in the environment reduces cognitive complexity and is consistent with natural modes of behaviour [1]. Consequently, part of the interaction with the environment happens through integrated objects. However, to meet the perspectives of the precision farming scenario, participants need immediate access to the environment's status and remote control of the assisting agents. For this reason, we have devised a control panel that can be activated with a hand gesture performed with an outstretched arm and the open palm facing upward (see Figure 5).While dropping the arm deactivates the control panel, wrist flexion serves as a reset.

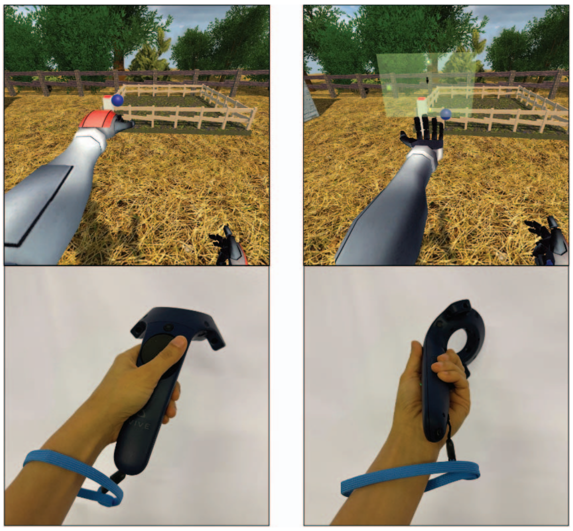


Figure 5: Control panel activation via stretch-and-turn hand gesture.

When designing the panel, instead of having interfaces pinned to the screen that are uncomfortable for the eyes, especially in the corners or during longer interactions, we defined the interface panel as a semi-transparent object that attaches to the (virtual) wrist at a comfortable distance from the eyes [1]. With hand gestures, we wanted to make the overall UX more natural and follow research-supported UI trends and best practices [14–16, 26]. The control panel rolls out progressively to keep complexity manageable, starting with a main map marking the production zones and the user's location. The user can select a zone with a virtual tap. In this case, selected data, such as the plant specifications and soil in the zone, as well as live plant statistics, are first displayed in an attached, smaller panel. Selecting the zone also activates a set of sliders representing the provided resources, such as water, fertilizer, and pesticides. Considering the plant specifications and the live statistics of the zone, the user can then allocate available resources to the selected zone. Selecting a value on at least one of the sliders activates the drone button, which can be tapped to send out one of the drones for resource distribution.

Different levels of UI complexity can be implemented with the current design in line with the experimental requirements. For instance, increasing the volume of presented task-related information is possible, which constitutes the basis of decision-making for the subject, while direct visual feedback can be given, for instance, by animating assets, e.g., to display the plant growth and representing plant states (see Figure 6 with a corn plant as an example).

Figure 6: Animated corn plant asset states; from healthy seed to fully grown (left) and sick (right) plant.

4 Virtual Environment Definition

To allow customization of the VE by different experimenters, we provide a system for flexible environment definition, which we call scene description. Experimenters can create scenes in their Unity project based on our VE and change the environment's layout by moving or replacing visual elements and defining custom zones.

Figure 1 shows an example scene in the editor with different experimenter-defined zones (production tasks in blue, an identification task in green, and another task type in yellow). Zones define physical spaces within the environment and have types and identifiers (cf. Figure 7). The type tells the VE what can happen in the respective zone, while the identifier is mainly used by events (that we will discuss later in this section) to select a specific zone. Once a scene is defined, part of the contents of the scene can be specified at application runtime by user-defined description files that can be divided into the two main subtypes of scene elements and events.

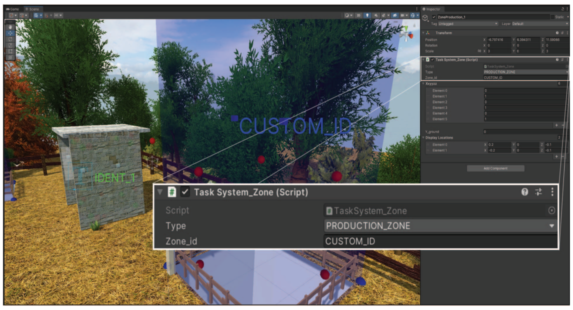


Figure 7: Example zone configuration in editor.

There are three scene element types in the example VE: plants, agents, and stimuli. The description files for plants define a plant template (e.g., how it looks and how much water it needs to grow) that can be used for a task. The agent description files define a user-controlled agent (e.g., a drone for watering plants) created in the scene. Finally, the stimuli description files define a stimulus template (behavior, parameters like frequency or intensity, and so on). All scene element files (see examples in Figure 8) follow a similar structure in which the experimenter describes the scene element identifier (e.g., "id:Plant_1"), the main behavioral specification used with its parameters (e.g., "main:plantwater"), and any sub-behaviors (e.g., "sub:plantbasevisualizer, visualizer:cabbage_production").

Events constitute a more abstract concept. The main idea is that after the scene is loaded, all events are waiting to start; an event can be anything an experimenter wants to happen in the scene. Events produce logs during their lifecycle, which are a combination of a log type and the id of the event that produced it. These logs are a core component of the event system since events have, as a start condition, a list of logs that must be produced before their start. The currently available log types are *START*, *END*, *SUCCESS* (also produces *END*), and *FAIL* (also produces *END*). One log is not produced by an

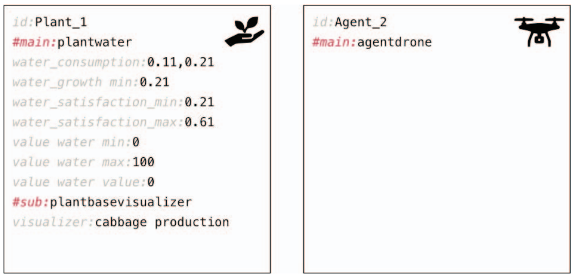


Figure 8: Example descriptions for a plant and a drone.

explicit user-defined event, which is the *TIMELINE_START* once the scene has finished loading. Like scene elements files, experimenters specify an identifier in an event description file and select the main behaviour of the event. In addition, the log values are listed, which represent the start conditions. Standard event behaviours include a timer, production and identification tasks, and starting or stopping a stimulus. While events can refer to other events via the logs that are part of their start conditions, the behaviour of both scene elements and events can refer to other described elements as well as the zones defined in the Unity scene. For example, a "production task" event can start when a timer produces its "end" and employ a user-defined "plant" in a particular "zone" (see Figure 10). Currently, the number of possible behaviors for each element is limited. What is provided is primarily a template or basic framework that can be expanded according on the experimenter's needs.

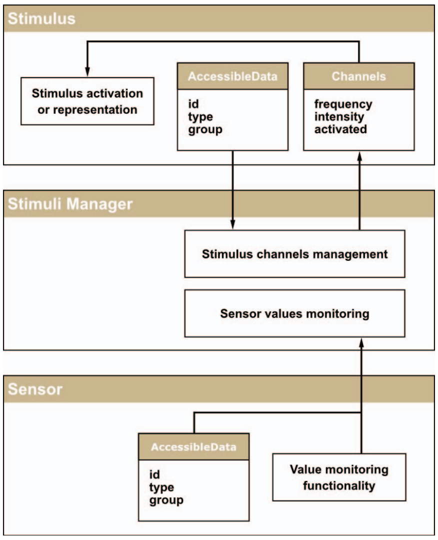


Figure 9: Framework component structure.

5 Component Architecture

To enable the implementation of reusable modulation approaches beyond the current VE and for the entire project, we have developed an underlying component architecture. In order to achieve the desired level of extensibility and flexibility, we opted for a tripartite structure, as shown in Figure 9, consisting of (1) a stimulus component that represents a time-modulating stimulus (e.g., (2) a sensor component that generates values that can be used to modify

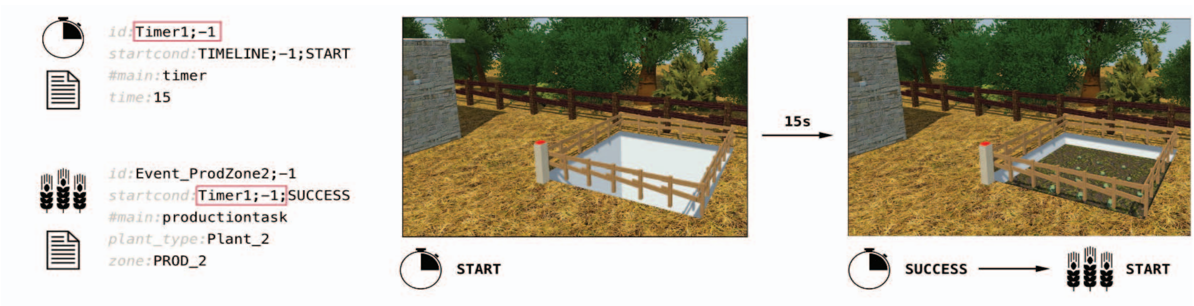


Figure 10: Event-based start of production task from a timer; after 15 seconds, a new production task is initiated.

the stimuli channels (from physiological sensors to more abstract modes such as task progress or time remaining), and (3) a stimuli manager component that processes the values from the sensors to affect the stimuli channels. In other words, the stimuli manager only defines the processing of the sensor values and how this processing will affect the stimuli.

Figure 11 shows an example of the applied component architecture with three stimulus patterns, each using a different modality (audio, visual, or haptic). The patterns have (at least) two common channels: Intensity and Frequency. The visual pattern could, for instance, be a golden flashing screen, as illustrated in Figure 12, which is a modulator type we recently employed in a study related to audio-visual rhythmic stimuli effects on subjective time perception [18]. The frequency would then correspond to the time interval between flashes, while the intensity would indicate the color transparency of the flash. Furthermore, there are three sensors, two of which record physiological values while another monitors task progress, stressing that a sensor component is not necessarily tied to a physical sensor but anything that can produce data output. The two stimuli managers must process and apply these sensor values, each affecting different channels of overlapping stimuli based on the various values.

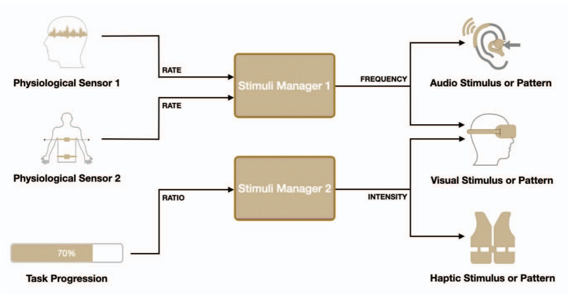


Figure 11: Application example for the generic component architecture; the rates of two different physiological sensors are processed by Stimuli Manager 1 to change the frequency of an audio-visual output stimulus (top), while a task progression ratio feeds into Stimuli Manager 2 resulting in an intensity change of a visio-haptic output stimulus (bottom).

5.1 External Control

Components of the architecture can be accessed externally to control their parameters and influence the VE in real-time. In the current VE, they can be accessed via a UDP interface. Experimenters can connect their own custom software to the VE to retrieve data from Sensors or send their own data "remotely" into the environment. The VE can then interpret and use this new data to control the environmental stimuli.

For example, the screenshot in Figure 12 shows a console application implemented in Python that is connected to the VE, sending sensor data constituting parameters of a color-flashing stimulus, as discussed before. The sensor data received from the console application is then processed by a stimuli manager within the VE, altering the stimulus (intensity/duration) according to the received user input.

Figure 12: External VE control via example Python console application controlling a visual stimulus (golden flashing).

6 Data Collection and Logging Capabilities

As the VE constitutes a testbed, it must integrate ways to collect and record various data in the background related to experimenters' needs. Within a scene of our VE, there are several points where data can be stored: the events system (including explicit data from questionnaires), virtual sensors, and custom behaviors. As described in Section 4, event systems store logs that associate a particular type with an event, as well as an optional message. After a log is processed, it is still available in memory and can then be written out cumulatively at the end of a session and stored in text form. Another possibility is to use a virtual sensor. The experimenter defines these sensors, which are then placed in the scene, and the system automatically retrieves their data. Technically, a virtual sensor is just an object that implements an interface. Virtual sensors can listen to data from real sensors, e.g., external physiological sensors and the eye-tracking data of the VR HMD, or from mechanisms in the environment, e.g., the user's position in the VE or the status of a plant in a zone. Finally, since the environment definition (cf. Section 4) is tied to implementing custom behaviors, these can also be the subject of user-defined data storage points. For instance, experimenters may add a questionnaire at the end of a task. Another example is when experimenters implement a custom stimulus, where it is possible to also add data monitoring in the stimulus behavior.

7 Conclusion

This paper presented a fully customizable VR environment and framework that, while initially developed for time perception studies, can be easily adapted to accommodate any stimulus-based cognitive research and behavioral studies. Time perception and other cognitive aspects are integral to user experience in VR, and such behavioral studies are crucial to understanding the intricacies and cross-effects, ultimately helping in the design of more immersive experiences and purposeful systems. The environment definition can be reused to produce any standard experimental setup and extended with custom tasks, while the external control system allows environment- and hardware-agnostic control of the VE.

We have been successfully employing the framework within the ChronoPilot project to investigate different factors related to time perception, and the framework substantially reduces the coding effort usually associated with creating and deploying new VR-based behavioral studies. Since the framework may be of great interest to other research groups, we are open to collaboration and open-sourcing it in the near future.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964464.

REFERENCES

[1] S. Alves, A. Callado, and P. Jucá. Evaluation of Graphical User Interfaces Guidelines for Virtual Reality Games. In *2020 19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, pp. 71–79, 2020. doi: 10.1109/SBGames51465.2020.00020

[2] A. O. Bebko and N. F. Troje. bmlTUX: Design and Control of Experiments in Virtual Reality and Beyond. *i-Perception*, 11(4):2041669520938400, 2020. doi: 10.1177/2041669520938400

[3] O. Benjamin, K. Yukari, S. Aria, S. Reiko, M. Kouta, S. Ayahiko, and M. Tetsuya. Crystal Vibes feat. Ott: A psychedelic musical virtual reality experience utilising the full-body vibrotactile haptic synesthesia suit. IEEE, 2017. doi: 10.1109/vsmm.2017.8346269

[4] J. Botev, K. Drewing, H. Hamann, Y. Khaluf, P. Simoens, and V. Argiro. ChronoPilot – Modulating Time Perception. In *Proceedings of the 4th IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR 2021)*, pp. 215–218. IEEE, 2021. doi: 10.1109/AIVR52153.2021.00049

[5] J. Brookes, M. Warburton, M. Alghadier, M. Mon-Williams, and F. Mushtaq. Studying human behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods*, 52(2):455–463, 2020. doi: 10.3758/s13428-019-01242-0

[6] M. Davydenko and J. Peetz. Time grows on trees: The effect of nature settings on time perception. 54:20–26, 2017. doi: 10.1016/j.jenvp.2017.09.003

[7] S. Droit-Volet, D. Ramos, J. L. O. Bueno, and E. Bigand. Music, emotion, and time perception: the influence of subjective emotional valence and arousal? *Frontiers in Psychology*, 4:417, 2013. doi: 10.3389/fpsyg.2013.00417

[8] M. Fornaciai, R. Arrighi, and D. C. Burr. Adaptation-Induced Compression of Event Time Occurs Only for Translational Motion. 6(1), 2016. doi: 10.1038/srep23341

[9] R. Gebbers and V. I. Adamchuk. Precision Agriculture and Food Security. *Science*, 327(5967):828–831, 2010. doi: 10.1126/science.1183899

[10] G. J. Gorn, A. Chattopadhyay, J. Sengupta, and S. Tripathi. Waiting for the Web: How Screen Color Affects Time Perception. 41(2):215–225, 2004. doi: 10.1509/jmkr.41.2.215.28668

[11] D. Hammerschmidt, C. Wöllner, J. London, and B. Burger. Disco Time: The Relationship Between Perceived Duration and Tempo in Music. *Music & Science*, 4:2059204320986384, 2021. doi: 10.1177/2059204320986384

[12] M. Landeck, F. Unruh, J.-L. Lugrin, and M. E. Latoschik. Metachron: A framework for time perception research in VR. In *26th ACM Symposium on Virtual Reality Software and Technology*. ACM, 2020. doi: 10.1145/3385956.3422111

[13] J. J. LaViola, D. A. Bowman, E. Kruijff, I. Poupyrev, and R. P. McMahan. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 2 ed., 2017.

[14] U. Lee and J. Tanaka. Finger Identification and Hand Gesture Recognition Techniques for Natural User Interface. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*, p. 274–279. Association for Computing Machinery, 2013. doi: 10.1145/2525194.2525296

[15] X. Lou, X. A. Li, P. Hansen, and P. Du. Hand-adaptive user interface: improved gestural interaction in virtual reality. *Virtual Reality*, 25(2):367–382, 2021. doi: 10.1007/s10055-020-00461-7

[16] K. Lupinetti, A. Ranieri, F. Giannini, and M. Monti. 3D Dynamic Hand Gestures Recognition Using the Leap Motion Sensor and Convolutional Neural Networks. In *Augmented Reality, Virtual Reality, and Computer Graphics*, pp. 420–439. Springer International Publishing, 2020. doi: doi.org/10.48550/arXiv.2003.01450

[17] A. McBratney, B. Whelan, T. Ancev, and J. Bouma. Future Directions of Precision Agriculture. *Precision Agriculture*, 6(1):7–23, 2005. doi: 10.1007/s11119-005-0681-8

[18] S. Picard and J. Botev. Rhythmic Stimuli Effects on Subjective Time Perception in Immersive Virtual Environment. In *Proceedings of the 14th International Workshop on Immersive Mixed and Virtual Environment Systems*, p. 5–11. Association for Computing Machinery, 2022. doi: 10.1145/3534086.3534330

[19] F. J. Pierce and P. Nowak. Aspects of Precision Agriculture. vol. 67, pp. 1–85. Academic Press, 1999. doi: 10.1016/S0065-2113(08)60513-1

[20] J. Reif and W. Alhalabi. Advancing Attention Control Using VR-Induced Multimodal Artificial Synesthesia, 2018. doi: 10.20944/preprints201808.0134.v1

[21] C. Schatzschneider, G. Bruder, and F. Steinicke. Who turned the clock? Effects of Manipulated Zeitgebers, Cognitive Load and Immersion on Time Estimation. 22(4):1387–1395, 2016. doi: 10.1109/tvcg.2016.2518137

[22] I. Schuetz, H. Karimpur, and K. Fiehler. vexptoolbox: A software toolbox for human behavior studies using the Vizard virtual reality platform. *Behavior Research Methods*, 55(2):570–582, 2023. doi: 10.3758/s13428-022-01831-6

[23] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Pearson, 6 ed., 2016. doi: 10.5555/3033040

[24] U. Söderström, M. Bååth, and T. Mejtoft. The Users' Time Perception: The Effect of Various Animation Speeds on Loading Screens. In *Proceedings of the 36th European Conference on Cognitive Ergonomics*. Association for Computing Machinery, 2018. doi: 10.1145/3232078.3232092

[25] M. R. Watson, B. Voloh, C. Thomas, A. Hasan, and T. Womelsdorf. USE: An integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents, journal = Journal of Neuroscience Methods. 326:108374, 2019. doi: doi.org/10.1016/j.jneumeth.2019.108374

[26] B. Yoo, J.-J. Han, C. Choi, K. Yi, S. Suh, D. Park, and C. Kim. 3D User Interface Combining Gaze and Hand Gestures for Large-Scale Display. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, p. 3709–3714. Association for Computing Machinery, 2010. doi: 10.1145/1753846.1754043

[27] Q. Zhang. *Precision Agriculture Technology for Crop Farming*. CRC Press, 2015. doi: 10.1201/b19336

Appendix E

Paper: XR MUSE: An Open-Source Unity Framework for Extended Reality-Based Networked Multi-User Studies

Full paper of Section 3.2.2

virtual worlds

Article

XR MUSE: An Open-Source Unity Framework for Extended Reality-Based Networked Multi-User Studies

Stéven Picard *, Ningyuan Sun and Jean Botev

VR/AR Lab, Department of Computer Science, University of Luxembourg, L-4364 Esch-sur-Alzette, Luxembourg; ningyuan.sun@uni.lu (N.S.); jean.botev@uni.lu (J.B.)
* Correspondence: steven.picard@uni.lu

Abstract: In recent years, extended reality (XR) technologies have been increasingly used as a research tool in behavioral studies. They allow experimenters to conduct user studies in simulated environments that are both controllable and reproducible across participants. However, creating XR experiences for such studies remains challenging, particularly in networked, multi-user setups that investigate collaborative or competitive scenarios. Numerous aspects need to be implemented and coherently integrated, e.g., in terms of user interaction, environment configuration, and data synchronization. To reduce this complexity and facilitate development, we present the open-source Unity framework XR MUSE for devising user studies in shared virtual environments. The framework provides various ready-to-use components and sample scenes that researchers can easily customize and adapt to their specific needs.

Keywords: extended reality; distributed systems; collaboration; interaction; software frameworks

check for updates

Citation: Picard, S.; Sun, N.; Botev, J. XR MUSE: An Open-Source Unity Framework for Extended Reality-Based Networked Multi-User Studies. *Virtual Worlds* **2024**, *3*, 404–417. https://doi.org/10.3390/virtualworlds3040022

Academic Editors: Thiago Malheiros Porcino and Jorge C. S. Cardoso

Received: 9 August 2024
Revised: 16 September 2024
Accepted: 27 September 2024
Published: 2 October 2024

1. Introduction

With the growing proliferation and capabilities of immersive media technologies, extended reality (XR) systems have been increasingly employed in user and user-agent collaboration studies [1–4]. On the one hand, such systems allow for utilizing the full simulation potential of virtual reality (VR) to decouple users from reality and previous experience. On the other hand, the advantages of including computer-generated objects and behaviors can also be combined with familiar mechanisms from the real environment, as in Mixed Reality (MR) settings. The latter is particularly important for continuing to use a wide array of established communication methods between users in their usual form instead of laboriously recreating them. The use of XR in research generally offers many advantages, for example, in terms of controllability and consistency [5]. In addition to allowing entirely new forms and explorations of multi-user collaboration, scenarios that would otherwise pose a serious risk, such as emergencies [6] and disasters [7], can also be easily replicated.

However, creating such XR environments can be challenging as it requires numerous technical skills, such as 3D development, user interface (UI) design, network programming, and other XR-specific knowledge, to ensure high presence and immersion levels. Even for experienced developers, it would be time-consuming and labor-intensive, as the various aspects must be individually implemented and coherently integrated. To streamline the development process for researchers, we introduce XR MUSE (https://github.com/vrar-lab/XRMUSE, accessed on 16 September 2024), an open-source Unity framework for XR-based multi-user collaboration studies. The framework comprises several dedicated modules for different functionalities, including data synchronization, spatial calibration, or environment configuration, and also provides an example scene researchers can use and adapt to other contexts.

Before introducing the XR MUSE framework and specific concepts and features in Section 3, we will provide further background on XR-based collaboration and related user studies, as well as existing toolkits for XR development, in Section 2. We conclude this article with a summary of the framework's key features and an outlook on the next development steps in Section 4.

2. Background

Over the last few decades, there has been a growing body of research on XR-based collaboration. The literature shows several unique benefits of XR-based collaboration compared to non-XR approaches. For users located in different places, collaborating in XR allows them to share a mediated reality in real time as if they were in the same space. This immersivity facilitates collaboration by providing a heightened sense of spatial presence and co-presence [8–10] as opposed to, for example, teleconferencing, where members see each other through screens. Furthermore, with XR, computer-generated objects can be included to create new ways of collaboration that otherwise would be difficult or infeasible in reality. To give an example, the design review meeting of an engineered system can be conducted in XR, so that the digital model of the system can be directly viewed, modified, and commented on in a shared space [11].

XR-based collaboration also benefits researchers as an experiment platform to investigate the dynamics underlying user–user or user–agent collaboration. Compared to physical reality, XR environments are more controllable since part of the environments are computer-generated. This allows researchers to repeat an experiment setup consistently across participants [5]. This consistency is advantageous for different types of studies, such as those involving confederates to create a specific collaborative scenario. Human actors' behaviors may fluctuate across sessions and, consequently, become an extraneous factor confounding the results [12,13]. Whereas in XR, collaborators can be replaced with virtual agents that can precisely replicate their behaviors following pre-defined scripts [14].

2.1. Technical Challenges in Collaborative XR Environments

Despite the above-mentioned advantages, creating a collaborative XR environment remains technically challenging since a complex of modules need to be individually implemented and coherently integrated. The following three modules, user interactions, spatial calibration, and networking, are common to most XR-based collaborative activities.

User Interactions. To maximize immersion and presence, XR-based applications usually allow users to grab, throw, and poke virtual objects by simulating or tracking users' hands. Compared to traditional WIMP-based (i.e., windows, icons, menus, and pointer) user interfaces, hand-based interactions feel more natural and intuitive [15] but involve more technical endeavors as they combines animations, physics simulation, computer vision, etc. The complexity further grows in multi-user scenarios, where a virtual object may be passed from one user's hand to another, during which the concept of ownership needs to be applied and synchronized. Furthermore, designing appropriate hand-based interactions can be demanding and knowledge-extensive. Numerous factors need to be optimized to maximize their positive effects on collaboration, ranging from static aspects such as the visual appearance of hands (e.g., shapes, skin tones, textures) to more dynamic ones like hand-tracking fidelities.

Spatial Calibration. When multiple MR devices are involved in an XR-based collaboration, they usually share the same physical space. However, each MR device has its own coordinate system to map the surrounding environment, and, consequently, the same coordinate may refer to different points in reality for different devices. To solve this issue, each individual mapping needs a unique transformation matrix to calibrate itself under a common coordinate system. While this calibration process can be boiled down to a mathematical problem, its implementation is not as straightforward. There are different approaches to spatial calibration, of which the most common one is by using fiducial markers such as QR or ArUco codes. The marker serves as a reference point for different MR devices to calculate the transformation between their spatial mappings. While being convenient, marker-based calibration is not persistent and requires re-calibration if the application restarts. Cloud-based anchors (e.g., Microsoft Azure Spatial Anchors (https://azure.microsoft.com/services/spatial-anchors, accessed on 16 September 2024)) provide more persistent calibration by extracting feature points from the surrounding environment and storing them in the cloud. When the application is running, it automatically detects and compares feature points and calibrates its coordinate system accordingly.

Networking Networking is a fundamental and the most challenging part of creating collaborative XR environments. Most interactions in XR happen in real time and need to be constantly synchronized among different clients to sustain immersion. There are numerous libraries (e.g., Photon (https://www.photonengine.com, accessed on 16 September 2024), Mirror (https://www.github.com/mirrornetworking/mirror, accessed on 16 September 2024), Netcode (https://docs-multiplayer.unity3d.com, accessed on 16 September 2024)) that provide commonly used networking functionalities such as object synchronization (e.g., ownership, spawn, pooling), networked physics, and remote procedure calls (RPCs, an inter-process communication protocol to invoke subroutines in a remote client). Based on these libraries, developers can focus on building the application without having to handle low-level protocols and networking frameworks. However, despite the streamlined process, these libraries can still be challenging to use depending on developers' expertise and experiences. Specifically in scenarios involving XR-based interactions, they often need to be combined with other XR-related packages (cf. Section 2.2), which are natively built for single-user scenarios.

2.2. XR Development Toolkits

Various approaches exist to create an XR experience, through game engines (e.g., Unity, Unreal Engine, Godot Engine (https://godotengine.org, accessed on 16 September 2024), modeling software (e.g., Blender, Maya), or platforms (e.g., Nvidia Omniverse (https://www.nvidia.com/omniverse, accessed on 16 September 2024). Unity is a popular choice as a game engine due to its ease of use, numerous third-party plug-ins, and large user base. Developers can find a variety of useful toolkits for XR development in Unity, often built and supported by leading IT companies, such as the XR Interaction Toolkit (https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.0, accessed on 16 September 2024) (Unity), AR Foundation (https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@6.0 accessed on 16 September 2024) (Unity), MRTK (https://github.com/mixedrealitytoolkit/mixedrealitytoolkit-unity, accessed on 16 September 2024) (Microsoft), Meta XR SDKs (https://developers.meta.com/horizon/documentation/unity/unity-package-manager, accessed on 16 September 2024), and others.

The widely used XR Interaction Toolkit provides a variety of versatile cross-platform components for managing XR inputs and interactions. These components cover a broad spectrum of functionalities such as locomotion, object manipulation (e.g., grab, throw), haptic feedback, and more. Developers can directly use this high-level interface without tailoring their code base to the APIs of individual target platforms. Similarly, the AR Foundation package is designed specifically for augmented reality development, wrapping commonly used functionalities (e.g., plane detection, face recognition) from vendors' APIs (e.g., ARKit (https://developer.apple.com/augmented-reality/arkit, accessed on 16 September 2024), and ARCore (https://developers.google.com/ar, accessed on 16 September 2024) into an overarching higher-level interface.

Some toolkits are provided by hardware vendors mainly for their own products. For example, Meta creates and maintains the Meta XR SDKs, primarily targeting its headsets. It contains basic XR support, such as locomotion and input management, and features specific to Meta's vision—e.g., lip sync, body tracking, and avatar assets—to create a social environment. There are also community toolkits developed by individuals or small commercials like VRIF (https://assetstore.unity.com/packages/templates/systems/vr-interaction-framework-161066, accessed on 16 September 2024) or HurricaneVR (https://assetstore.unity.com/packages/tools/physics/hurricane-vr-physics-interaction-toolkit-177300, accessed on 16 September 2024), which are usually built upon the ones mentioned above to provide out-of-the-box but opinionated implementations for XR game development. The VRIF package, for example, allows users to climb ladders, grab and fire a gun, and more.

Overall, these different toolkits provide common functionalities such as locomotion, object interaction, and more, significantly facilitating XR development. However, creating a shared XR experience based on these toolkits remains challenging. Most only provide fundamental features but do not wrap them into useful out-of-the-box components or provide a ready-to-use framework for behavioral studies. For instance, these toolkits do not include user connectivity and require developers to integrate third-party networking plugins into their applications.

2.3. Existing Frameworks for XR-Based Collaboration

The literature has documented several frameworks (e.g., [16–20]) for creating a shared collaborative space, to which users can connect via the device of their preference, including but not limited to VR headsets, MR glasses, hand-held AR devices, tablets, and desktops. These frameworks share similar technology stacks—with Unity and a subset of the above-mentioned toolkits—but target different scenarios with their individual implementations.

For example, most of the frameworks employ a server–client structure to connect different XR devices. Some of them are built upon existing solutions such as Unity Netcode or Photon [18–20], while others implemented their own server–client communication using, for instance, the UDP protocol [17]. The latter approach is more challenging but provides full control of the server logic. This liberty in server design can be advantageous for experiments on XR-based collaboration, since researchers can monitor all the clients and collect their data simply on the server side. Other than the server–client structure, one study opted for a peer-to-peer structure to allow for collaboration without a reliable network connection [16].

The aforementioned frameworks were all demonstrated to effectively synchronize different XR devices. However, none of the frameworks is open-source, leaving other researchers mostly theoretical insights without a usable code base. XR MUSE stands out from these existing frameworks for being publicly accessible on GitHub, providing not only several ready-to-use packages but also an example scene to work with. There also exist other open-sourced frameworks, such as the Unity Experiment Framework [21] and the BiomotionLab Toolkit for Unity Experiments [22]. However, they mainly target behavioral studies in general rather than focusing on XR-based collaboration.

Outside academia, various commercial products allow users to interact and collaborate in a shared XR space. Many are video games (e.g., Cook-Out (https://www.resolutiongames.com/cookout, accessed on 16 September 2024)), where players must act toward common goals in dedicated, pre-defined environments. Content-centric platforms, where users create objects and experiences, such as VRChat (https://hello.vrchat.com/, accessed on 16 September 2024) and Rec Room (https://recroom.com/, accessed on 16 September 2024), offer a higher level of customization, allowing users to create multi-user XR-based environments via APIs. While such commercial products and platforms can also be used for observational or field studies on multi-user collaboration in XR [23], running controlled experiments with them is challenging. Commercial products are usually optimized for socializing or entertainment and contain various elements (e.g., eye-catching animations) that are not controllable or are irrelevant to the experiment goal. These elements may distract participants or confound experiments, potentially leading to validity issues. In comparison, as an open-source solution, XR MUSE offers a minimal, neutral environment that researchers can fully control.

3. XR MUSE

The XR MUSE framework consists of the following four central packages that are based on existing and widely used technologies, as illustrated in Figure 1:

- *XRMUSE.Networking* (cf. Section 3.2) contains a collection of scripts for networking-related functionalities, including server connection, data synchronization, object pooling, networked physics, and more. At this point, these scripts are built upon the Photon Unity Network (PUN2 (https://assetstore.unity.com/packages/tools/network/pun-2-free-119922, accessed on 16 September 2024)), which uses the free remote server to connect XR devices.

- *XRMUSE.SceneDescription* (cf. Section 3.3) provides a system that allows researchers to describe the environment with text files following a specific syntax. When the application starts, these files are loaded and parsed by the application to instantiate the environment. To change the environment's parameters, only the configuration files need to be modified instead of the full scene within Unity. This dynamic loading system not only provides a convenient and flexible way to contextualize and adapt the environment but also hides the technical details behind scene creation in Unity.
- *XRMUSE.Utilities* encapsulates other generic functionalities that do not fit the packages mentioned above, like custom collisions (cf. Section 3.1.1) or spatial calibration.
- *XRMUSE.ExampleScene* (cf. Section 3.1.2) provides an example scene where two users work collaboratively to accomplish an industrial assembly task. The package includes various assets used for building the scene, such as meshes, textures, and materials, and a batch of scripts for realizing the logic behind the scene. The key elements in the scene are modularized into individual Unity Prefabs, which are reusable assets that can store a so-called GameObject along with its components and properties. The scene is deliberately kept generic to offer a baseline environment that can be easily customized for other collaborative tasks.

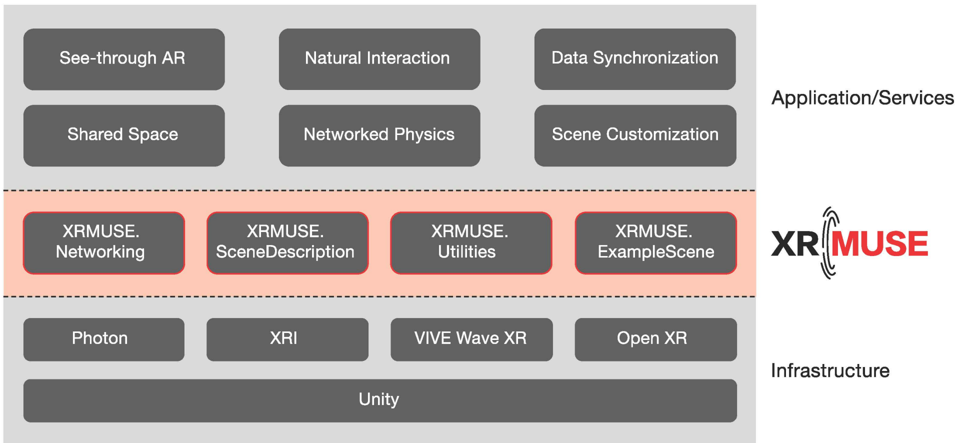


Figure 1. XR MUSE framework structure. The packages act as an intermediate between the infrastructure (i.e., common toolkits for XR development) and application (XR-based collaboration) layers.

3.1. Environment

XR MUSE features a workbench setting as an example scenario, where users share and interact with virtual objects on a real table. Figure 2 illustrates this setting in a dual-user case, where the physical space is conceptually divided into shared and individual zones. The area on and around the workbench is designed to be the main collaboration space. In contrast, the areas to the left and right are accessible only to the respective user.

This central workbench setting enables organic interaction between users within the designated scenarios. It is a particularly practical arrangement since any table can be used, and the physical separation between users avoids risks associated with co-location. As illustrated in Figure 3, users can access different resources either shared between them to collaborate, or only accessible to an individual for competitive advantage or informational asymmetry. This could be both tangible (e.g., the cubic objects on the table surface or the different tools of the users) or intangible (e.g., the sink seemingly altering the table's geometry or the information panel, which here is visible only to one of the users).

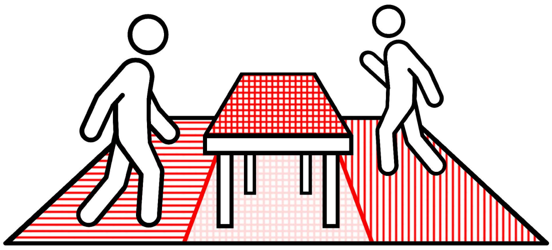


Figure 2. Conceptual division of the example environment. The table surface (checkered) is shared, whereas the left (horizontally ruled) and right (vertically ruled) areas belong to individual users.

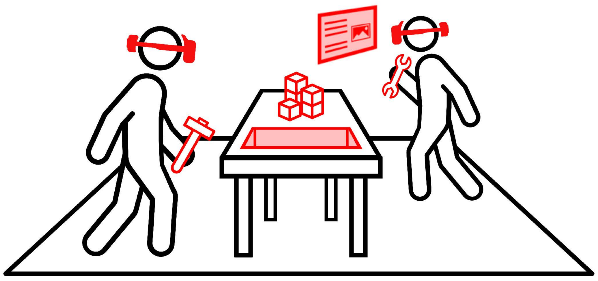


Figure 3. Digital content in XR includes objects and tools, as well as intangible items like virtual cavities or information panels that might be shared between or accessible only to individual users.

This setup assumes the users' individual space is in the same coordinate system that aligns with physical reality. For calibration, XR MUSE offers a headset-specific implementation. Once calibration is completed, the virtual objects will rest accurately on the table. Users can grab, release, and throw them using controllers. To make the user experience seamless, all three actions are bound exclusively to the so-called grip or trigger buttons, leaving the other buttons unused. This facilitates operation and significantly flattens the learning curve for employing the controllers, allowing users to quickly master the interaction regardless of their technical prowess and prior knowledge. This is essential for experiments and user studies, as changing experiences with the controllers can distort the results, and participants might be bored or frustrated by a lengthy familiarization process. All interactions in the example environment are collision-based and utilize virtual objects.

3.1.1. Collision System

To allow in-environment objects to interact with each other, XR MUSE provides a typed collider system on top of Unity's collision system. Each typed collider has an identifier of a custom type, which hints at the nature of the collision (in the example scenario, e.g., a material with another material or with a tool). Developers can add custom behaviors to these colliders that react to the type of colliding or existing objects. For instance, in the example scenario, two colliders typed as material colliding with each other will trigger a glow animation; when these colliders exit each other, the animation is toggled off (cf. Figure 4). In conjunction with the dual-transform design (cf. Section 3.2.1), this system allows one to dissociate the Unity physics from custom collision-based interactions.

(**a**) (**b**)

Figure 4. Example usage of custom collisions triggering the material outline for objects in proximity. (**a**) Screenshot, custom colliders set to invisible. (**b**) Screenshot, custom colliders set to visible.

3.1.2. Example Scene and Setup

Similar to the illustration in Figure 3, two users are tasked to produce a certain amount of products using the provided materials in the example scene. Each user has a unique tool to combine two materials into an intermediate or final product. Several combinations are possible, with each user only capable of realizing a subset, so they have to work collaboratively to produce the final product. All the materials and products are represented abstractly as basic primitives like cubes or spheres, whereas the tools have more specific visuals than wrenches or hammers. The virtual cavity on the table allows users to submit their products by throwing them into the cavity.

To combine two materials, users need to put them together and hit one of the materials with their tools, as illustrated in Figure 5. Internally, this is achieved through the typed collision system, where the collision between materials and tools triggers several networked events. These events handle and synchronize the task logic and animation of the combination process across users.

(**a**) (**b**)

Figure 5. Example material combination sequence, from two source objects to the final product. (**a**) Screenshot, hammering together source objects. (**b**) Screenshot, resulting target sphere object.

Apart from the interactive elements discussed above, there are also static information panels (cf. Figure 6) in the environment showing users the overall objective (i.e., which and how many products to produce), possible combinations, and the current status (i.e., how many products have been produced). As discussed in Section 3.1, these panels can be shared or only visible to an individual user. Informational asymmetry, as in the latter case, can be crucial to specific task types. For example, one user may see "two cubes with a hammer produce a sphere", while the other sees "a cube and a sphere with a wrench produce a cylinder". The objective in this case can be set to produce five cylinders, such that the two users must communicate and work collaboratively to reach the goal.

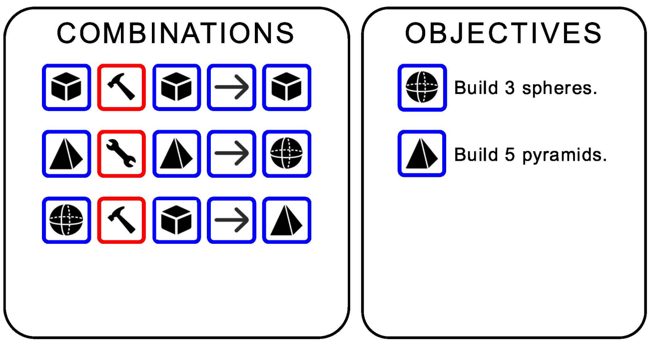


Figure 6. Outline view of the panels in the example scene with available combinations and objectives.

In the example setup, we employ two VIVE XR Elite (video-see-through mixed-reality headsets) to maximize the perception of presence and immersion. The choice of devices, however, is flexible and can also include virtual reality headsets such as the Meta Quest 3. The two headsets are connected through a remote third-party server via the internet. The communication among headsets follows a customizable protocol (cf. Section 3.2) built to facilitate experiment design. This introduces a small but acceptable latency, which can be further reduced by setting up a local server. This example was deployed using Unity

2022.3.16f1 with the support of numerous packages, including both commonly used ones such as XR Interaction Toolkit and dedicated packages for networking (e.g., Photon) and XR development (e.g., VIVE Wave XR Plugin). Along with the example Unity project, a batch of configuration files are also provided (cf. Section 3.3), allowing researchers to easily customize the example without the need to modify the Unity project directly.

The users' virtual spaces are calibrated to the real-life table using the XR Elite's ArUco marker detection feature. Experimenters can simply place a marker at the center of the table, which then serves as a reference point.

3.2. Networking

Networking is an essential but technically challenging aspect of multi-user XR environments. The XR MUSE framework reduces the complexity by providing a ready-to-use, configurable networking module developed based on the Photon Unity Network (PUN2) for ease of use and similarity to UNet. It should be noted that the network solution requires a platform-specific implementation involving concrete code and design, which will be briefly discussed before proceeding to XR MUSE's other systems. In developing the framework, we considered the PUN2 requirements that (1) a Unity GameObject is authored (or owned) by a user and ownership may be transferred between users, and (2) each networked GameObject may synchronize its data through de/serialization. Building on these two concepts, the module provides specific implementations for ownership transfer and transform synchronization, as well as a pooling system for GameObject instantiation and recycling. These specific implementations, which are required for using PUN2, are not the focus of this article and are therefore not discussed in detail. However, they are necessary foundations for the following subsections on the dual-transform design (Section 3.2.1) and data synchronization (Section 3.2.2).

XR MUSE was developed with co-located shared environments in mind, i.e., mixed-reality settings where users share a physical and digitally augmented space. The possibility of also interacting in a non-mediated fashion is an essential aspect of collaborative studies. However, the framework's networking components already support completely remote scenarios where users can be anywhere. In such cases, only an audiovisual representation of the users with an avatar, including voice, needs to be added.

3.2.1. Dual-Transform Design

In the virtual environment, users manipulate physical objects by moving them, placing them on the workbench or elsewhere, and interacting with other objects. In a networked environment, however, other users can also influence these objects. For example, physical interactions (e.g., gravity) can overlap with another user's actions. In Unity, physical objects are usually divided by a hierarchy of elements, and collisions can occur on an element at a higher point in the hierarchy. Each element of this hierarchy has a transformation that determines the positions of the subordinate elements.

In XR MUSE's dual-transform design (cf. Figure 7), two objects with a parent–child relationship in the same hierarchy are considered to be on the same level, with the authoring determined by one of the two transformations. The parent transform typically is the networked GameObject; its transform is synced, and when the user does not own the object, that parent transform is in charge of the dual-transform's authoring. The child transform is not synced but takes the dual-transform's authoring when the user owns the synced GameObject, typically having the physics, local controls, and gestures over the object (cf. Figure 7).

This allows the dissociation of the synchronizing behavior from the local interactions and assigning unity physics computation only to the owning user. Since both transforms can have different colliders, the parent transform also uses the custom collision system (cf. Section 3.1.1), which enables the synchronization of data in the case of collision-based interactions on networked objects.

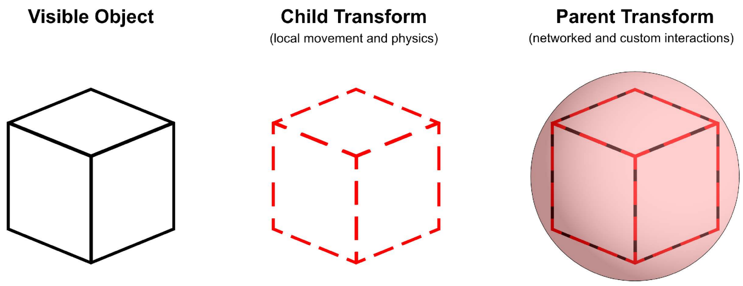


Figure 7. Example dual-transform design with the typed collision system on the parent.

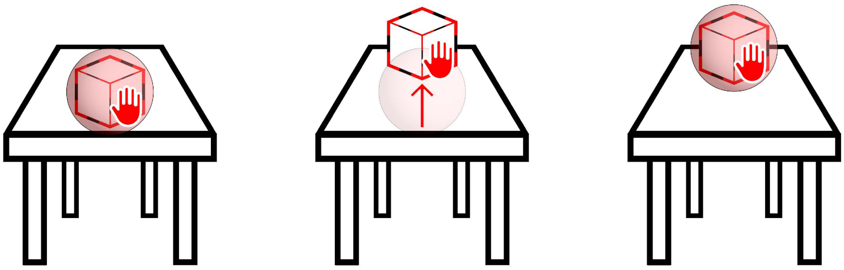


Figure 8. User moving object through child transform, parent transform snapping to new position.

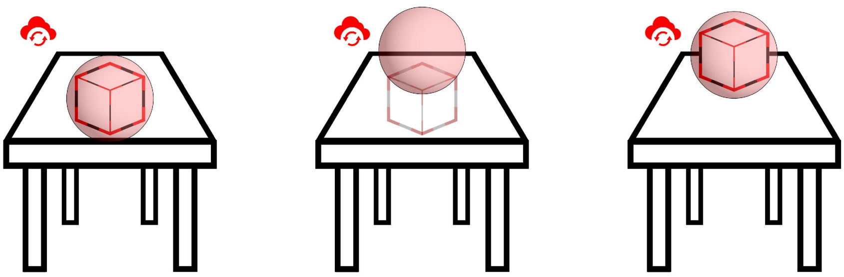


Figure 9. Network update of parent transform's position, child transform snapping to new position.

In the examples shown in Figures 8 and 9, the parent object uses XR MUSE's custom collision management system that does not trigger physical interactions. The geometrically simpler sphere of the parent transform is deliberately larger in diameter and completely encloses the actual object, enabling the following two aspects: first, separating the interactions between objects from the physics, which allows it to abstract tasks from the Unity engine's physics system, and second, activating physical authoring when a user needs the network object (e.g., when grabbing the object). Since only the position and rotation of the real parent are synchronized over the network, the networking does not affect the physics.

3.2.2. Data Synchronization

Not all synchronization needs fall into the category of GameObject synchronization. In an experiment, for instance, it may be necessary to synchronize environment states and record per-user actions. Therefore, XR MUSE includes a class created per user that shares its user-specific data with each connected user. The aggregation of these per-user data permits the deduction of environment states. For instance, if the objective is to produce together a certain amount of materials, this class will record the amount of materials each user produced, the aggregation representing the progress of this task (cf. Figure 10).

When using PUN or similar services, the network design assumes synchronization and serialization per virtual object, and different RPCs can be utilized to trigger events. This can be problematic when new data are added, where a new network object is created on both computers (or an existing object is edited). The calls must be triggered according to the chosen network solution, which requires the associated ownership and authoring permissions. This system leverages this need for RPCs as local scripts can trigger events in response to updated values on the remote users' dedicated objects. Figure 10 shows an animation in reaction to one user's synchronized value of the CubeAnimation variable.

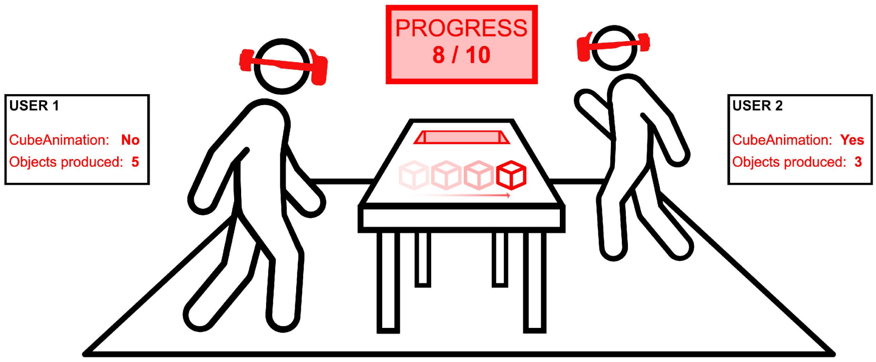


Figure 10. Aggregated per-user values affecting the scene for both users (progress display, animation).

More concretely, these user-specific data are created through a dedicated GameObject, to which we provide interfaces to append any custom type of data, which may also define accessible custom functions. That dedicated GameObject has a main component, *UserSyncedValues*, which handles the order of registered data, triggering serialization, and providing access to the various data and associated functions. This is realized using the C# programming language's reflection feature, which means that when loading, the object builds internal dictionaries of references to the corresponding exposed resources. Then, any number of user-defined *DataRegister* components that add data and functions for synchronization with the UserSyncedValues component can be added to it. The DataRegister components can mark each function as accessible via UserSyncedValues with a specific identifier (cf. diagram in Figure 11).

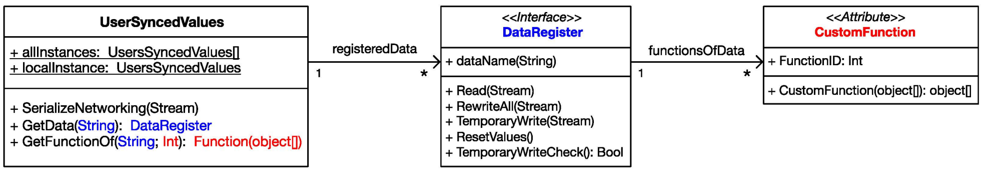


Figure 11. Simplified UML diagram for UserSyncedValues and DataRegister components reflection.

3.3. Operation

User studies [1–4] often create similar situations where only some elements of the environment or events are changed, e.g., to test a stimulus or collect data across several similar situations. While this can be achieved by creating multiple Unity scenes and special scripts for an experiment, the more scenes are added, the more complex maintenance becomes. Therefore, we propose a versatile scene loading system that has evolved from earlier work [24] towards a more generalized setting. This system is designed to adapt to a wide range of user studies, offering a solution that can be tailored to specific needs. The core elements of a scene can be defined in text files loaded at runtime from three types: instantiables, events, and stimuli.

3.3.1. Object Loader

An instantiable generates a template or Unity Prefab that other instantiables or events can use. Instantiables have an obligatory main type, which will set the main object generated but may also obtain optional sub-modules. An example of a set of instantiable files can be found in Figure 12, with declarations of the materials used in the scene and boxes using these materials. These instantiables may be declared networked and compatible for automatic integration through XR MUSE's custom pooling system. A stimulus constitutes a similar template concept; however, it is explicitly separated from other instantiables as a stimulus may not be re-instantiated but activated or de-activated on demand. Implementation-wise, all declared instantiables are loaded by the module manager, which in the Unity scene can be observed with the creation of inactive Unity Prefabs, typically at scene load. Those inactive prefabs are the template than can be cloned by events described in the following subsection (Section 3.3.2).

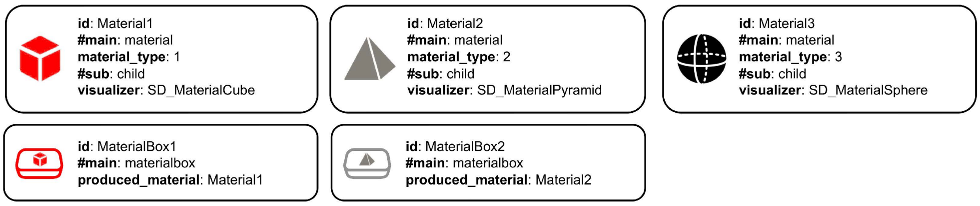


Figure 12. Example set of scene elements loaded from text files.

3.3.2. Scripting

Text files that are loaded at run-time also allow a form of scripting through a so-called events sequence. To handle the events sequence, XR MUSE uses a custom timeline, which also includes the production of log data. Event-specific logs may be produced during a trial except for the TIMELINE_LOAD and TIMELINE_START events. Events are either loaded from files or produced by other events They have an ID, a behavioral type (possibly custom), and a set of start conditions. The start conditions set is a collection of event-produced logs that all must be recognized. A log is a combination of an event's ID (or the default IDs for timeline load and start) and log type. The currently available log types are START, END, SUCCESS (which also produces END), and FAIL (which also produces END). Several generic events are available as examples and are sufficient for basic experiments, such as timers, instantiable spawning, or stimuli activation. Events can be declared as text files, just like instantiables and stimuli, with behavior-specific parameters that may refer to IDs of other declared files (such as an instantiable ID in the case of spawning instantiables) and the start condition logs. The example material boxes depicted in Figure 12 can be loaded by spawner events at loading, as shown in Figure 13.

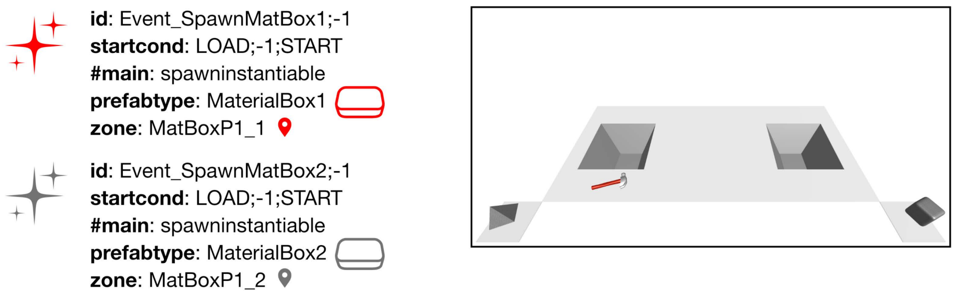


Figure 13. Scene after the TIMELINE_LOAD event.

As an example of more advanced scripting, a timer event may start with the automatic event TIMELINE_LOAD while another production task event has the success of the timer event as a starting condition (cf. Figure 14). When the timer generates its end log, the material production task can be started.

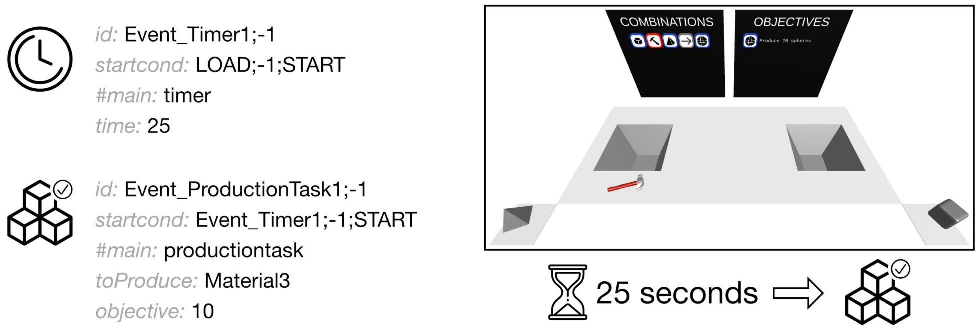


Figure 14. File-based definition of task start following a timer.

4. Summary and Outlook

This article introduces XR MUSE, an open-source framework that facilitates the creation of XR-based networked multi-user applications. Although it is designed for user studies in shared virtual environments, it can be used for any type of collaborative XR application. The framework significantly reduces some of the technical efforts by providing open solutions for dealing with XR-based interactions (Section 3.1.1), networking (Section 3.2), and scripting (Section 3.3), as well as a sample workbench environment integrating these elements (Section 3.1). Dedicated concepts, such as a custom collision system and dual transform design (Section 3.2.1), are provided to support real-time interaction. Implementing these together with the object loader (Section 3.3.1) unlocks the full potential of the corresponding XR MUSE packages.

Currently, the framework is based on Photon's PUN2, a free third-party Unity package that has some limitations due to its commercial nature. For example, XR MUSE uses Photon's free remote server for connecting different devices. Consequently, participants' personal data will go through their servers and risk exploitation. Although Photon offers a local server solution, it is not free. As a next step, we are thus considering porting to alternative open-source solutions such as Mirror.

At this stage, XR MUSE only supports two specific video see-through headsets (VIVE XR Elite and Meta Quest 3), which will be expanded in future versions of the framework to accommodate further or yet-to-be-released devices. Users also need to set up the devices by following relatively lengthy instructions; this setup process can be partially automated using scripts to improve XR MUSE's usability further.

It should be noted that XR MUSE was developed with functionality over performance in mind. Some of our advanced functionalities, like the scripting system (Section 3.3) and data registration (Section 3.2.2), which allow for extensive customization, may induce more processing than a purely custom solution. However, since most behavioral studies use minimal, neutral scenes for variable controls, the performance issues should be negligible. Still, in certain situations, such as when researchers try to load high-resolution assets, performance may suffer and affect the user experience. To improve XR MUSE, further stress tests with different devices are needed to identify potential performance issues.

XR MUSE's main application area is XR-based multi-user studies, in which users collaborate or compete in the same scene. For instance, we are currently employing the framework to conduct experiments on time perception in such scenarios. However, experimenters do not need to use the provided example scene. They may use only certain XR MUSE components, such as object synchronization, to handle distributed state. This is equally relevant for XR-based games or training environments. Application makers can use the framework as a prototyping tool, as the scene-loading features allow for the quick creation and testing of different scenarios with the same items. Different sets of configuration files can, for instance, represent various difficulties, including objects and their properties (e.g., physics and speed) to load as different stages and balance level the design quickly.

While XR MUSE is not a universal solution, it can be a valuable resource for diverse multi-user XR applications, either in its entirety or only in parts. The framework is freely available to anyone from researchers to XR application vendors. We look forward to potential collaborations and seeing XR MUSE as the basis for a new generation of customized multi-user XR software.

Author Contributions: Conceptualization, S.P. and J.B.; software, S.P., N.S. and J.B.; resources, J.B.; writing—original draft preparation, S.P., N.S. and J.B.; writing—review and editing, J.B.; visualization, J.B., S.P. and N.S.; supervision, J.B.; project administration, J.B.; funding acquisition, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964464 (ChronoPilot).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mourtzis, D.; Angelopoulos, J. Development of an Extended Reality-Based Collaborative Platform for Engineering Education: Operator 5.0. *Electronics* **2023**, *12*, 3663. [CrossRef]
2. Lucas Bravo, P.P.; Fasciani, S. A Human-Agents Music Performance System in an Extended Reality Environment. In Proceedings of the 23rd International Conference on New Interfaces for Musical Expression (NIME), Mexico City, Mexico, 31 May–2 June 2023.
3. Pidel, C.; Ackermann, P. Collaboration in Virtual and Augmented Reality: A Systematic Overview. In Proceedings of the 7th International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Lecce, Italy, 7 September 2020; pp. 141–156.
4. Santhosh, S.; De Crescenzio, F. A Mixed Reality Application for Collaborative and Interactive Design Review and Usability Studies. In Proceedings of the 11th International Joint Conference on Mechanics, Design Engineering & Advanced Manufacturing, Ischia, Italy, 1–3 June 2022; pp. 1505–1515.
5. Pan, X.; Hamilton, A.F.C. Why and How to Use Virtual Reality to Study Human Social Interaction: The Challenges of Exploring a New Research Landscape. *Br. J. Psychol.* **2018**, *109*, 395–417. [CrossRef] [PubMed]
6. Lorusso, P.; De Iuliis, M.; Marasco, S.; Domaneschi, M.; Cimellaro, G.P.; Villa, V. Fire Emergency Evacuation from a School Building Using an Evolutionary Virtual Reality Platform. *Buildings* **2022**, *12*, 223. [CrossRef]
7. Feng, Z.; González, V.A.; Amor, R.; Spearpoint, M.; Thomas, J.; Sacks, R.; Lovreglio, R.; Cabrera-Guerrero, G. An Immersive Virtual Reality Serious Game yo Enhance Earthquake Behavioral Responses and Post-Earthquake Evacuation Preparedness in Buildings. *Adv. Eng. Inform.* **2020**, *45*, 101118. [CrossRef]
8. Nguyen, H.; Bednarz, T. User Experience in Collaborative Extended Reality: Overview Study. In Proceedings of the 17th EuroVR International Conference, Valencia, Spain, 25–27 November 2020; pp. 41–70.
9. Reinhard, R.; Telatar, E.; Humayoun, S.R. Comparison of Object Detection in Head-Mounted and Desktop Displays for Congruent and Incongruent Environments. *Big Data Cogn. Comput.* **2022**, *6*, 28. [CrossRef]
10. Ochs, M.; Mestre, D.; De Montcheuil, G.; Pergandi, J.M.; Saubesty, J.; Lombardo, E.; Francon, D.; Blache, P. Training Doctors' Social Skills to Break Bad News: Evaluation of the Impact of Virtual Environment Displays on the Sense of Presence. *J. Multimodal User Interfaces* **2019**, *13*, 41–51. [CrossRef]
11. Sateei, S.; Roupé, M.; Johansson, M. Collaborative Design Review Sessions in Virtual Reality: Multi-Scale And Multi-User. In Proceedings of the 27th International Conference of the Association for Computer Aided Architectural Design Research in Asia (CAADRIA), Sydney, Australia, 9–15 April 2022; pp. 9–15.
12. Doyen, S.; Klein, O.; Pichon, C.L.; Cleeremans, A. Behavioral Priming: It's All in the Mind, But Whose Mind? *PLoS ONE* **2012**, *7*, e29081. [CrossRef] [PubMed]
13. Kuhlen, A.K.; Brennan, S.E. Language in Dialogue: When Confederates Might Be Hazardous to Your Data. *Psychon. Bull. Rev.* **2013**, *20*, 54–72. [CrossRef] [PubMed]
14. Koda, T.; Ruttkay, Z. Eloquence of Eyes and Mouth of Virtual Agents: Cultural Study of Facial Expression Perception. *AI Soc.* **2017**, *32*, 17–24. [CrossRef]
15. Kang, H.J.; Shin, J.H.; Ponto, K. A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality. In Proceedings of the 27th IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR), Atlanta, GA, USA, 22–26 March 2020; pp. 275–284.

16. Huh, S.; Muralidharan, S.; Ko, H.; Yoo, B. XR Collaboration Architecture Based on Decentralized Web. In Proceedings of the 24th International Conference on 3D Web Technology, Los Angeles, CA, USA, 26–28 July 2019; pp. 1–9.
17. Tümler, J.; Toprak, A.; Yan, B. Multi-user Multi-platform XR collaboration: System and Evaluation. In Proceedings of the 24th International Conference on Human-Computer Interaction, Virtual Event, 26 June–1 July 2022; pp. 74–93.
18. Kim, J.; Song, J.; Seo, W.; Ihm, I.; Yoon, S.H.; Park, S. XR Framework for Collaborating Remote Heterogeneous Devices. In Proceedings of the 1st IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Atlanta, GA, USA, 22–26 March 2020; pp. 586–587.
19. Pereira, V.; Matos, T.; Rodrigues, R.; Nóbrega, R.; Jacob, J. Extended Reality Framework for Remote Collaborative Interactions in Virtual Environments. In Proceedings of the 2nd International Conference on Graphics and Interaction (ICGI), Faro, Portugal, 21–22 November 2019; pp. 17–24.
20. Kostov, G.; Wolfartsberger, J. Designing a Framework for Collaborative Mixed Reality Training. *Procedia Comput. Sci.* **2022**, *200*, 896–903. [CrossRef]
21. Brookes, J.; Warburton, M.; Alghadier, M.; Mon-Williams, M.; Mushtaq, F. Studying Human Behavior with Virtual Reality: The Unity Experiment Framework. *Behav. Res. Nethods* **2020**, *52*, 455–463. [CrossRef] [PubMed]
22. Bebko, A.O.; Troje, N.F. bmlTUX: Design and Control of Experiments in Virtual Reality and Beyond. *i-Perception* **2020**, *11*. [CrossRef] [PubMed]
23. Maloney, D.; Freeman, G.; Robb, A. A Virtual Space for All: Exploring Children's Experience in Social Virtual Reality. In Proceedings of the 7th Annual Symposium on Computer-Human Interaction in Play (CHI PLAY), Virtual Event, 2–4 November 2020; pp. 472–483.
24. Picard, S.; Botev, J.; Niknam, S. A Dynamic and Scriptable Environment and Framework for Stimulus-Based Cognitive Research in Virtual Reality. In Proceedings of the 2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Sydney, Australia, 16–20 October 2023; pp. 387–392.

Appendix F

Paper: Some Times Fly: The Effects of Engagement and Environmental Dynamics on Time Perception in Virtual Reality

Full paper of Section 3.3.1

Some Times Fly: The Effects of Engagement and Environmental Dynamics on Time Perception in Virtual Reality

Sahar Niknam
University of Luxembourg
Esch-sur-Alzette, Luxembourg
sahar.niknam@uni.lu

Stéven Picard
University of Luxembourg
Esch-sur-Alzette, Luxembourg
steven.picard@uni.lu

Valentina Rondinelli
University of Luxembourg
Esch-sur-Alzette, Luxembourg
valentina.rondinelli.001@student.uni.lu

Jean Botev
University of Luxembourg
Esch-sur-Alzette, Luxembourg
jean.botev@uni.lu

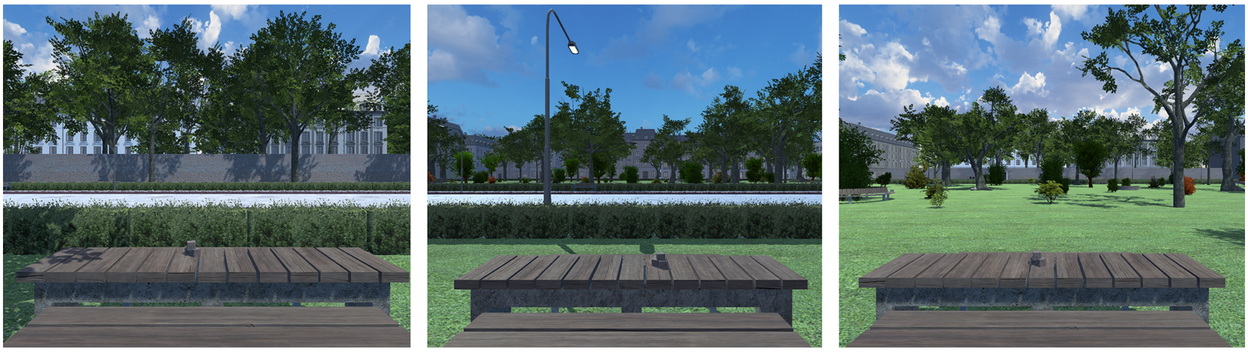


Figure 1: Randomly selected locations in the virtual environment for testing the experiment conditions with varying durations.

ABSTRACT

An hour spent with friends seems shorter than an hour waiting for a medical appointment. Many physiological and psychological factors, such as body temperature and emotions, have been shown to correlate with our subjective perception of time. Experiencing virtual reality (VR) has been observed to make users significantly underestimate the duration. This paper explores the effect of virtual environment characteristics on time perception, focusing on two key parameters: user engagement and environmental dynamics. We found that increased presence and interaction with the environment significantly decreased the users' estimation of the VR experience duration. Furthermore, while a dynamic environment lacks significance in shifting perception toward one specific direction, that is, underestimation or overestimation of the durations, it significantly distorts perceived temporal length. Exploiting these two factors' influence smartly constitutes a powerful tool in designing intelligent and adaptive virtual environments that can reduce stress, alleviate boredom, and improve well-being by adjusting the pace at which we experience the passage of time.

VRST '24, October 09–11, 2024, Trier, Germany

ACM ISBN 979-8-4007-0535-9/24/10
https://doi.org/10.1145/3641825.3687726

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Empirical studies in interaction design**.

KEYWORDS

Virtual Reality, Time Perception, User Engagement, Environmental Dynamics

ACM Reference Format:
Sahar Niknam, Stéven Picard, Valentina Rondinelli, and Jean Botev. 2024. Some Times Fly: The Effects of Engagement and Environmental Dynamics on Time Perception in Virtual Reality. In *30th ACM Symposium on Virtual Reality Software and Technology (VRST '24), October 09–11, 2024, Trier, Germany.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3641825.3687726

1 INTRODUCTION

Humans perceive the passage of time subjectively. Two healthy individuals may not share a similar temporal concept of *one minute*.

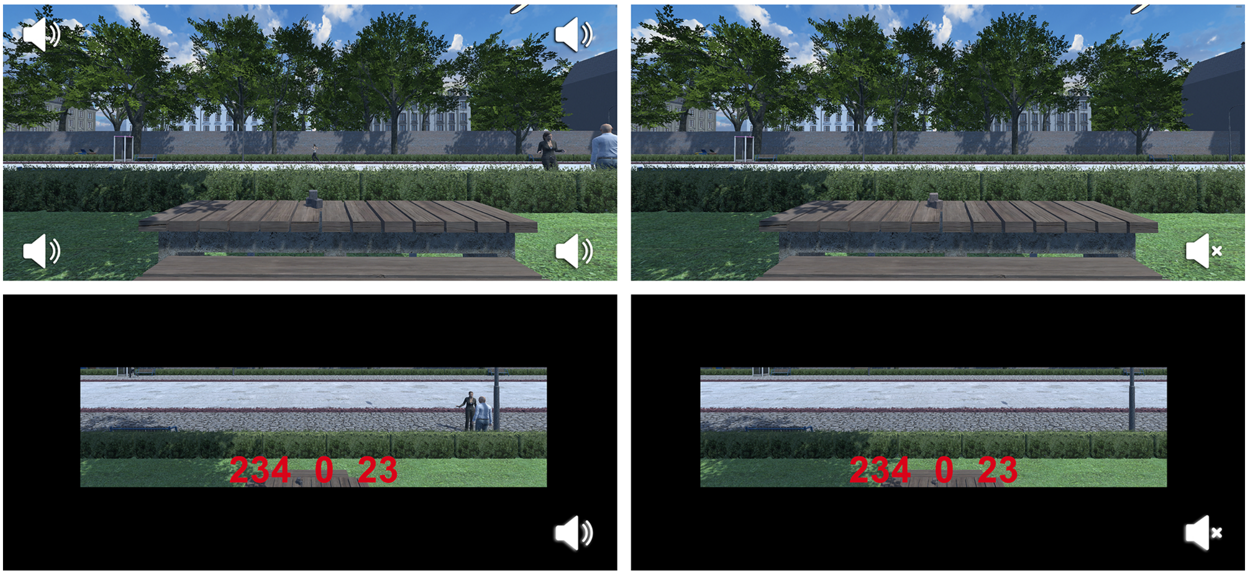


Figure 2: The study's four conditions; top left: dynamic environment w/ active user (DA), top right: static environment w/ active user (SA), bottom left: dynamic environment w/ passive user (DP), bottom right: static environment w/ passive user (SP).

Even more critically, a person under different physical or psychological conditions may perceive the same duration longer or shorter. A slew of parameters have been identified as being correlated with the pace at which we perceive the passage of time, including but not limited to age [18, 46, 47], body temperature [36, 43, 44], emotions [1, 8, 35], cognitive load [3, 4, 24], attention [28, 31, 34], and the amount of processed information [16, 39, 49].

Playing video games has also been shown to significantly alter users' perception of time, to the point where *time loss* during gaming can have detrimental effects on individuals' daily lives and well-being, often associated with mood changes and addictive behaviors. Researchers have attempted to explain this phenomenon by highlighting gamers' strong attentional focus and intense emotional arousal during gameplay, driven by highly engaging content and narratives [27, 38, 48]. However, beyond the game content, design features like immersive environments also significantly contribute to this temporal distortion. For example, Virtual Reality (VR), in general, has been reported to have a compressing effect on time perception [25, 26, 33]; time spent in VR is perceived shorter than the objective duration.

1.1 Related Work

Despite its significance, research on time perception modulation in VR is in its nascent stages and mostly limited to observation reports. These have been primarily justified using purely psychological literature and concepts, lacking a primarily VR perspective.

For example, Schatzschneider et al. [32] investigated the efficacy of real-world temporal clues simulated in VR. Weber et al. [45] used VR as a platform to validate the transferability of *magnitude* modulation of time perception [9, 17, 20] in terms of speed affordance, from observed objects to self. In another work, Malpica et al. [22] studied the magnitude modulator in affordances, including luminance, field of view, and visual complexity in VR, by displaying real images and videos. While these works paved the research path and made valuable contributions to understanding time perception in VR, the absence of VR perspective is noticeable.

However, in the last few years, a handful of studies have begun to explore the modulation of time perception using fundamental VR concepts and features. Lugrin et al. [21] researched the effect of embodiment and environment visualization on perceived waiting time. They recreated a plain waiting room and a seated avatar, once using 360-degree pictures and again with 3D models, and compared it with waiting in an identical but real space. Their results showed no significant difference in participant estimation of their 7.5-minute waiting time among the three conditions. Nevertheless, the same team of researchers continued their work by adding a complementary condition of a 3D model of the same waiting room, but without the embodiment and also focusing separately on the duration estimation versus the perceived pace of the passage of time [40]. This time, they observed a significant effect of the no-avatar condition on slowing down the perceived passage of time, but not on estimating duration. The authors refined their results by adding user interaction and reproduced the same results [41].

Read et al. [30, 31] has taken another significant step in understanding the intricacies of time perception in VR. They investigated time perception by asking participants to estimate the duration of their 3.5-minute experience in three virtual environments. These environments, sourced from two commercial games, elicited different levels of user engagement and varied spatial characteristics. Using these settings, the authors studied three conditions including a passive user in a dynamic environment, an active user in a static environment, and an active user in a dynamic environment.

The work's findings indicate a significant influence of a dynamic environment over a static one, while no difference was observed between the active and passive user conditions.

This paper follows suit by hypothesizing that user engagement and environmental dynamics have the potential to modulate the perception of time in virtual environments. However, despite intriguing results, the work of Read et al. [30, 31] suffers from specific methodological shortcomings, which we aim to address. Firstly, we developed a custom-designed environment, with complete control over parameters, as opposed to the commercial games used by Read et al. [30, 31]. Secondly, we set up all conditions within a single environment, which makes them more comparable. Additionally, Read et al. [30, 31] described and ranked the virtual environments used in their study as dynamic or engaging in a general and non-systematic manner. However, we offer a clearly delineated description for the settings of these two parameters, supported by VR literature. Finally, we also tested the missing condition in the work of Read et al. [30, 31], that is, a passive user in a static environment.

As much as accurate timing plays an essential role in different aspects of everyday and professional life, modulating the perception of time can have a critical impact on improving general well-being. For example, feeling that time is passing quickly while working under a tight timeframe can induce stress and negatively affect decision-making. Boredom, on the other hand, makes us perceive time as dragging and consequently can decrease focus and attention. By modulating time perception, we can improve decision-making, performance, and harmony in collaborative settings [5]. Understanding the precise ways through which VR influences users' perception of time, provides us with a powerful tool for developing adaptive virtual environments and workplaces for enhancing user experience.

2 MATERIALS AND METHODS

The experiment was designed to test the hypothesis that *user engagement* and *environmental dynamics* can modulate the perception of time. To test this hypothesis, we developed a dedicated virtual environment simulating an outdoor scene in a public park as the base platform. This environment was configured with two levels of user engagement and environmental dynamics in a 2×2 study design, with three repeated measurements.

User engagement. This parameter takes two values: *active user* and *passive user*. Active user in our work is defined by three features: partial embodiment (arms and hands), interaction with the environment, and a supported sense of presence by situating participants at a natural height and making ambient sound spatial (Figure 2, first row). In contrast, in the passive user case, hand-tracking is disabled, and therefore the user does not interact with the environment. In addition, the camera simulating the user's perspective was positioned at an elevated height, creating an observer viewpoint. The user's field of view (FOV) in the passive case was also limited and enclosed by a black frame to add to the user's sense of detachment from the environment. Environmental sounds in this case were 2D and muffled by applying a low-pass filter (Figure 2, second row).

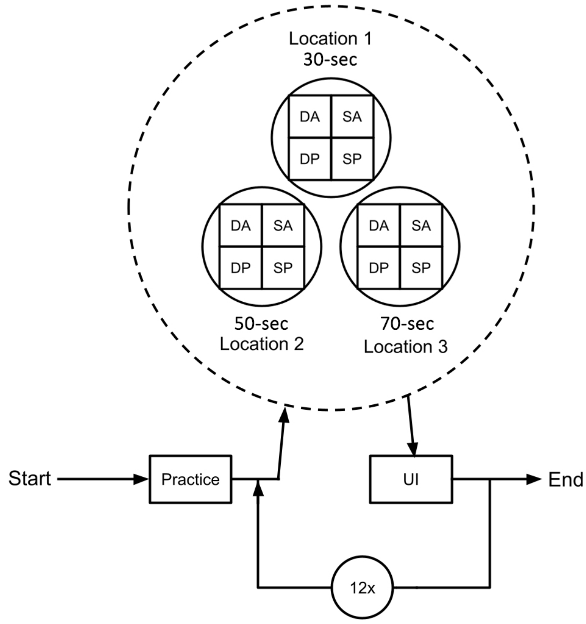


Figure 3: Flowchart of the experimental procedure.

Environmental dynamics. This parameter also takes two values: *dynamic environment* and *static environment*. We defined the dynamic environment as an environment with variation in both visual and auditory modalities. In the dynamic environment, we implemented animated human avatars, such as a jogger passing by and two office workers talking on the sidewalk. Furthermore, ambient sounds like birds chirping and a distant siren were added to the environment (Figure 2, first column). On the contrary, the static environment was a silent scene without any animated object (Figure 2, second column).

Consequently, we had four conditions, representing permutations of these two parameters: a static environment with a passive user (*SP*), a static environment with an active user (*SA*), a dynamic environment with a passive user (*DP*), and a dynamic environment with an active user (*DA*). All four conditions were tested three times, each round with a different duration (30, 50, and 70 seconds) and at a different location in the same environment (Figure 1). We repeated conditions in three different locations to prevent the repetition from diminishing the dynamism effect; however, the locations were randomly selected from the same environment to keep repeated measures comparable. That resulted to a total of 12 trials, which were presented to all participants in random order, ensuring that each participant experienced a different sequence of trials (Figure 3).

During the experiment, participants were seated on a swivel chair and informed that they were free to rotate and look around 360 degrees. For the VR experience, we used the HTC Vive Pro Eye[1] headset with a mounted Leap Motion Controller[2] for high-precision hand-tracking. The virtual environment in this experiment was

[1] https://www.vive.com/sea/product/vive-pro-eye/overview/
[2] https://www.ultraleap.com/product/

built with the Unity Game Engine[3], using ready-made assets, such as the Microsoft Rocketbox Avatars library[4] [13].

2.1 Tasks

During each trial, participants had two tasks to complete: estimating the duration of the trials with *verbal estimation* method [2, 7, 42], and memorizing three random numbers: a single-, a double-, and a triple-digit. We implemented the memory task, firstly, as a parallel task to the timing task to prevent participants from using explicit timing techniques, like counting the seconds. Secondly, participants' performance on the memory task provides us with further insight into possible side effects of time perception modulation, such as impaired cognitive function and unproductive distraction. In the active scenario, participants were placed seated behind a wooden table. In front of them, three cubes were placed on the table, each with a number engraved on one side. The participants were required to grab and turn the cubes to see the numbers. In the passive scenario, numbers were presented to participants printed on a screen attached to the camera.

After each trial, a user interface (UI) appeared, where participants entered their estimations of the trial duration and the three numbers. This gesture-based UI consisted of a 1-second resolution slider from 10 to 120 seconds for the timing task, and a numeric pad for the memory task (Figure 4). Using a slider for a verbal estimation task has the advantage of associating temporal length with physical length, reducing the likelihood of wild guessing that can occur with a numeric pad that limits input to numbers. The slider's range was selected primarily to avoid centrality bias by making the range asymmetrical around the mean trial duration (50 seconds). Additionally, the lower range (0-10 seconds) was removed as an improbable range, to provide a wider slider within the limited field of view, enhancing user control.

Figure 4: The study's gesture-based input user interface.

Before the study, participants had a short practice session with two trials to get familiar with the tasks and the UI (Figure 3). The first trial simulated an active user condition for 90 seconds, and the second one a passive user for 10 seconds; both trials were implemented in an empty environment.

[3] https://unity.com/
[4] https://github.com/microsoft/Microsoft-Rocketbox

2.2 Sample Size

We conducted an *a priori* sample size calculation using G*Power software [10, 11] for a repeated measures analysis of variance (RM-ANOVA) within subjects. The parameters were set as follows: effect size (Cohen's f) of 0.25, significance level (α) of 0.05, and statistical power ($1-\beta$) of 0.8, while the dependent variable, *perceived duration*, is measured three times for each of the four conditions (SP, SA, DP, and DA). Based on these parameters, the required sample size was 28 participants.

2.3 Recruitment

Participants were recruited via flyers posted on campus and nearby social venues, as well as through university email lists. The inclusion criteria were normal or corrected-to-normal vision and proficiency in English, while the exclusion criteria included a history of neurological disorders and susceptibility to motion or cybersickness. A total of 30 participants (8 identifying as women, 22 identifying as men), aged 18 to 46 years (Mean=28.3, SD=6.7), were recruited. Participants were informed both verbally and in writing of their right to ask questions, pause or exit the study at any time, and withdraw their consent without any consequences. Written informed consent was obtained from all participants before the experiment, and a debriefing session explaining the study's objectives was provided afterward. The full procedure took 30 minutes on average to complete, and upon completion, the participants were compensated with a gift card worth 15 euros. The study received ethical approval from the Ethics Review Panel of the University of Luxembourg, under approval number ERP 23-059.

3 RESULTS

Participant interval timings were, first, normalized for variation in trial duration by calculating the *normal estimation error* (NEE) using the following formula:

$$\text{NEE} = \frac{\text{Estimation - Objective duration}}{\text{Objective duration}}$$

such that a positive deviation means overestimation, negative, underestimation, and accurate estimation returns 0 (Figure 5). The results of a one-sample t-test on the NEE values show that, in general, the participants significantly underestimated durations ($t(359)$ = -2.92, p = 0.0037).

In addition, estimation errors needed to be adjusted for individual differences in emotional state, working memory capacity, or the level of familiarity with VR, as these factors could indirectly affect the performance on the timing task in the experiment. Therefore, to remove the individuals' baseline error, we calculated *zero-mean NEEs* (ZM-NEE) for each participant:

$$\text{ZM-NEE} = \text{NEE} - \text{mean}(\mathbf{NEE}_x) \quad \text{for participant } x,$$

where $\mathbf{NEE}_x$ is the set of 12 NEE values recorded for the participant x (Figure 6 and Figure 7).

Afterward, the *Shapiro-Wilk* test was conducted to assess normality, yielding a p-value of 0.5422, indicating the normal distribution of the residuals ($p > 0.05$). The RM-ANOVA on ZM-NEE values revealed a significant effect of experimental conditions ($F(3, 87)$ = 11.22, $p < .0001$, $\eta p^2 = 0.28$). The following *Bonferroni*-corrected post-hoc t-tests indicate significant differences between every possible

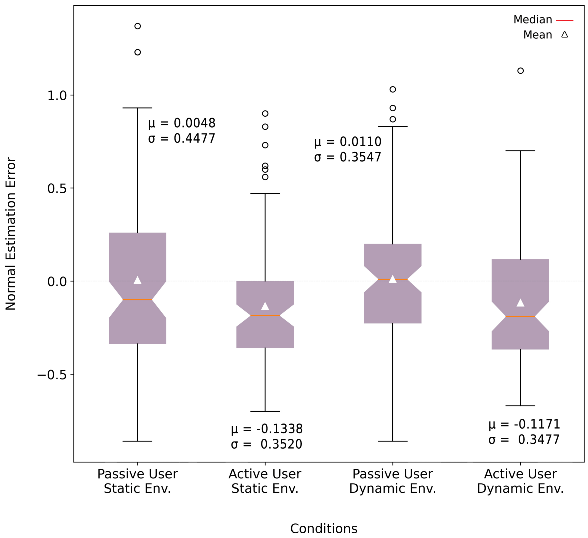


Figure 5: Normal duration estimation errors, grouped by conditions.

pair of active versus passive users. The contrast is most pronounced first in the DP-SA contrast, then in the dynamic environment, and least pronounced in the static environment. However, the environmental dynamics pairs under the same level of user engagement did not show a significant effect on ZM-NEE (Table 1 and Figure 7).

Table 1: Comparison of the conditions with ZM-NEE values

Contrast	$t(29)$	p-corr (bonf)
DA - DP	−4.5409	**0.0005**
DA - SA	0.6150	1.0000
DA - SP	−4.0017	**0.0024**
DP - SA	5.5416	**<.0001**
DP - SP	0.1522	1.0000
SA - SP	−3.4831	**0.0010**

Furthermore, we performed an RM-ANOVA on the absolute values of ZM-NEE to evaluate the intensity of time perception distortion, regardless of the direction of distortion toward underestimation or overestimation of durations. With absolute values, the RM-ANOVA exhibited a significant influence of conditions on the dependant value ($F(3, 87) = 4.65$, $p = 0.0046$, $\eta p^2 = 0.06$). However, the following post-hoc tests indicate the DA-SP contrast as the only significant difference among the conditions ($t(29) = -3.21$, $p = 0.0193$).

We also evaluated performance on the memory task by subtracting 1 point for each forgotten digit, resulting in scores ranging from -6 to 0 for one trial, and -72 to 0 in total. Participants scored -0.56 points on average for the memory task. Similar to the NEE values, each participant scores on the memory task were centered to have a mean of 0, too. In addition, to account for different memorizing time span that participants had in different trials, we calculated the zero-mean scores for each of the three trial durations (30, 50, and 70 seconds). Afterward, an RM-ANOVA was run on the memory task's scores recorded under the four experimental conditions. The results show a significant impact of the conditions on the memory performance ($F(3, 87) = 5.14$, $p = 0.0025$, $\eta p^2 = 0.15$). The following post-hoc tests indicate two significant contrasts in DA-SP ($t(29) = -3.16$, $p = 0.0222$) and SA-SP ($t(29) = -3.28$, $p = 0.0164$) pairs (Figure 8). But the evidence is insufficient to reject or accept a correlation between memory task performance and ZM-NEE (*Pearson*'s $r = -0.0514$, $p = 0.3311$).

However, the *Pearson* analysis shows a significant negative correlation between ZM-NEE and duration of the trials (*Pearson*'s $r = -0.4023$, $p < .0001$). See Figure 9. To further investigate the nature of the relationship between the duration of the trials and ZM-NEE, we ran RM-ANOVA on dataset subsets grouped by duration. The results show user engagement as the only parameter with a significant effect on estimated duration in all three duration groups, consistent with the observations on the whole dataset (Table 2 and Figure 6).

4 DISCUSSION

This study found that user engagement with and within the virtual environment significantly affects duration estimation. Active users with partial embodiment who interacted with the environment and had a supported sense of presence tend to underestimate the duration of their VR experience (Table 1). Conversely, despite the observation that the increased dynamism in the environment made the trial duration perceived longer, the effect is not significant (Figure 6 and Figure 7, second row). Nevertheless, the influence of environmental dynamics on time perception cannot be ruled out; on the contrary, it may be significant but in another dimension. We observed the effect of user engagement level to be significant in both dynamic and static environments. However, it is more pronounced in the dynamic environment. Furthermore, the most significant deviation of time perception happens in the DP-SA contrast, that is, when both parameters have opposite values, but most importantly when the dynamic environment is coupled with the passive user (Table 1). These results can be interpreted by considering that time perception modulation operates along two axes: direction and intensity (Figure 10).

By increasing user engagement, we can move along the *direction* axis and shift time perception from the feeling that time is passing fast to slow or vice versa. On the other hand, by changing the level of dynamism in a virtual environment, we can move along the *intensity* axis and increase or decrease time perception deviation.

In general, performance on the memory task shows a weak negative correlation with user engagement (Figure 11).

Similarly, Figure 8 shows slightly lower performance scores in both active user conditions compared to the passive user's conditions, especially the passive user in a static environment. This confirms the well-known fact that having successful interaction within a virtual environment consumes attentional resources and adds to the cognitive load. Nevertheless, the average score of participants on the memory task (-0.56 on a scale of -6 to 0) indicates that the cognitive load imposed by the task was not significant enough

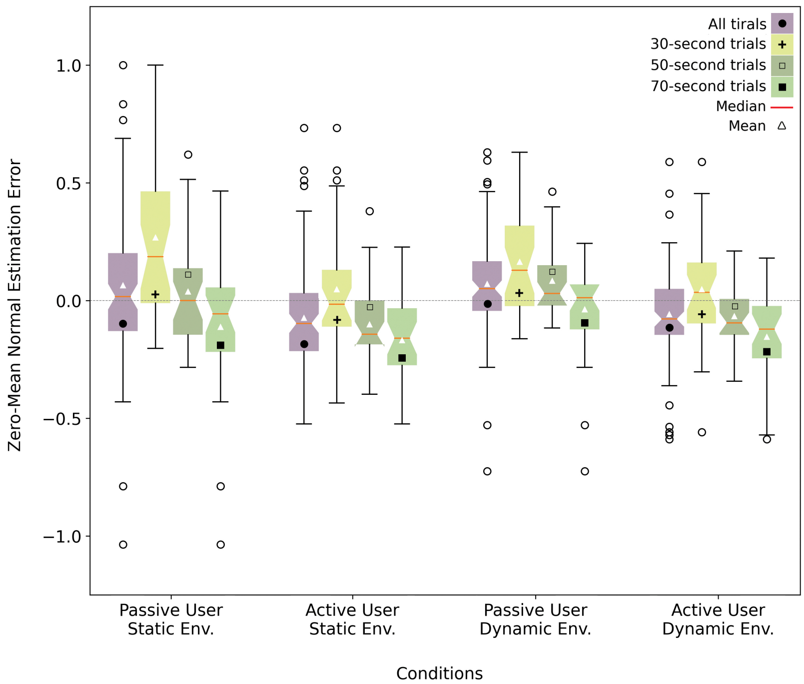


Figure 6: Zero-mean of normal duration estimation errors, grouped by conditions.

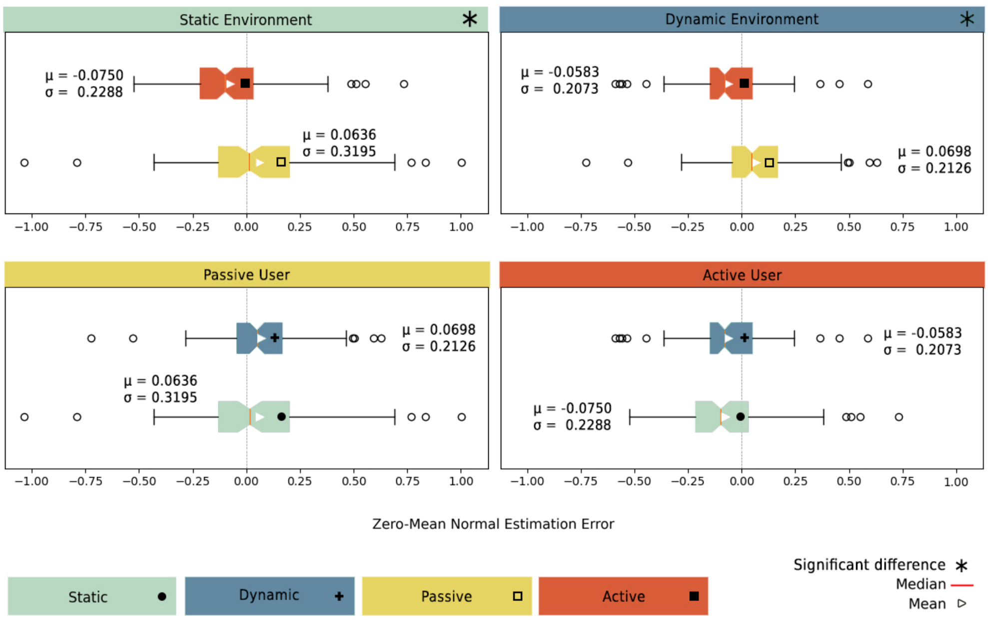


Figure 7: Pairwise comparisons of parameters in different experimental conditions.

Table 2: Results of RM-ANOVA on ZM-NEE values for experimental parameters grouped by trial duration

	All trials		30-sec trials		50-sec trials		70-sec trials	
Parameter	*F*-value	*p*-value	*F*-value	*p*-value	*F*-value	*p*-value	*F*-value	*p*-value
dynamism	0.3198	0.5761	1.0254	0.3196	2.1827	0.1504	1.7603	0.1949
engagement	44.9322	**<.0001**	17.9251	**0.0002**	22.0180	**0.0001**	4.3237	**0.0465**
dynamism:engagement	0.0344	0.8542	1.2309	0.2764	0.0251	0.8753	0.5787	0.4530

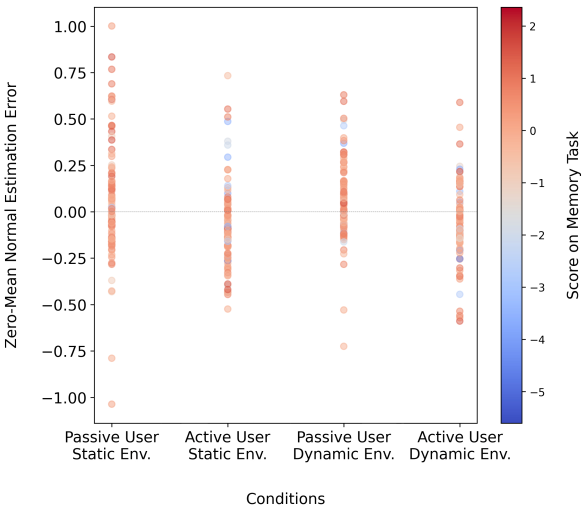


Figure 8: Performance on the memory task, grouped by conditions.

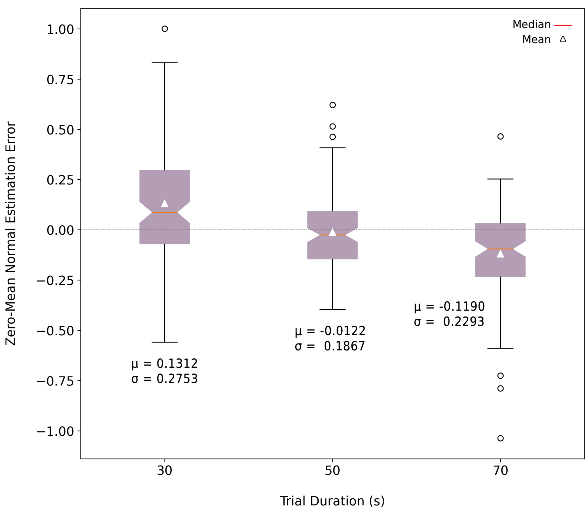


Figure 9: Zero-mean of normal duration estimation errors, grouped by trial duration.

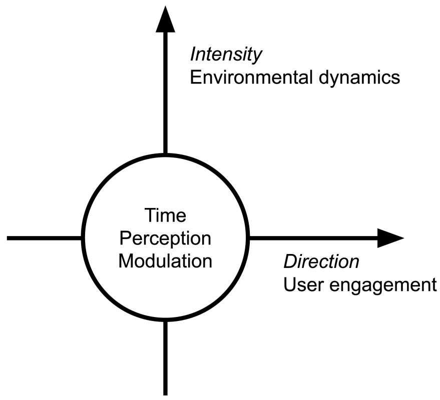


Figure 10: Time perception in VR can be modulated along two axes; *direction*: time passes fast or slowly, and *intensity*: time passes slightly fast/slowly or very fast/slowly.

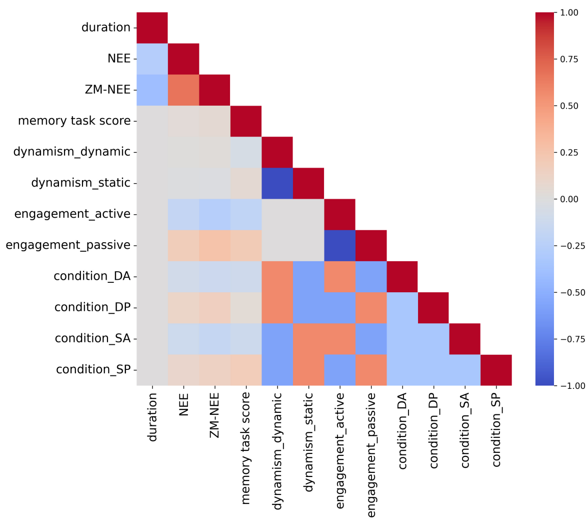


Figure 11: Study parameters and results correlation matrix.

to skew the study's results. This is also confirmed by Figure 8 which shows an even distribution of lower scores across both the negative (underestimation) and positive (overestimation) ranges of ZM-NEE axis, and Figure 11 which does not indicate a correlation between performance on the memory task and ZM-NEE. However, we observed two significant contrasts between the conditions regarding the score on the memory task: DA-SP and SA-SP. Both contrasts are also among the significant cases of shifting time perception (Table 1). This demonstrates the importance of having a well-defined monitoring system along with time perception modulation to ensure no undesired side effects on user performance or well-being.

Further analyses of the results grouped by trial duration indicate the findings are independent of both the time spent in the virtual environments and the specific location within the environment (Table 2 and Figure 6). However, even after normalizing to account for varying duration of trials, we notice a significant negative correlation between the duration and estimation errors (Figure 9). This observation is consistent with the literature on time perception, which states shorter intervals are typically overestimated, while longer intervals tend to be underestimated, which is in line with the literature [6, 14, 29]. Furthermore, we also observe a general tendency to underestimate the trial duration, which is usually attributed to VR's general effect on time perception (Figure 6).

Overall, the findings support our hypothesis that user engagement and environmental dynamics can modulate users' perception of time. The correlation between user engagement and perceived duration is negative, significant, and independent, while the effect of environmental dynamics is limited and vague in direction, serving as a complementary factor that can intensify the impact of user engagement. However, these findings sharply contrast with those of Read et al. [30, 31]. The results of their work indicate that user interaction with the environment does not affect the judgment of duration, while a more complex and dynamic environment tends to distort perceived duration. Nonetheless, general literature on time perception backs the idea that any processes that consume attentional resources contract perceived duration [12, 15, 19, 29]. Thus, in our case, significant underestimation of duration due to goal-driven engagement with the virtual environment seems natural and anticipated.

On the other hand, in our experiment, environmental dynamics played its most significant role when coupling with the passive user, in contrast with an active user in a static environment. This observation is also consistent with the existing literature on time perception. It is often reported that receiving visual information can expand the experienced time, and the vividness of visual input is positively correlated with the perceived duration [23, 37, 50]. Matthews and Meck [23] frame this phenomenon as the *the ease of extracting information from the stimulus*, covering not only syntactical features like magnitude and salient color of stimuli, but also semantic aspects such as simplicity and authenticity. In our dynamic environment, ambient sound, animated objects, and the natural user position and FoV contribute to the realism of the environment and consequently, to its *perceptual vividness* [23], thereby expanding the perceived duration compared to the static environment.

4.1 Future Work and Limitations

This study is a starting point, offering a framework and perspective for future, more in-depth research. Several aspects of this work could benefit from further investigation and improvement. For example, future work could investigate the influence of specific features of each condition, such as spatial ambient sound in the dynamic environment or the limited field of view in the static environment, separately. Another approach involves surveying participants on their experience of presence or detachment to validate the effective implementation of the intended effects within the virtual environments. Finally, less than 27% of the sample population for this study identified themselves as women, which biases the dataset towards males. A balanced distribution of gender could produce more reliable findings in terms of generalizability.

5 CONCLUSION

Based on our findings, a user actively engaged in the virtual environment experiences a compression of time. On the other hand, an observer within a dynamic virtual environment perceives an expansion of time. These findings, combined with the unprecedented opportunities offered by multisensory VR technology helps us develop more effective and intelligent user interfaces. We can explore the trade-off between the *interactor* and *observer* roles in user status. For example, when making time-sensitive decisions, reducing interaction with the environment will shift the user's role to that of an observer, thereby alleviating time pressure that could eventually improve performance. Similarly, reducing interaction could help gamers become more aware of elapsed time and mitigate or possibily prevent game addiction. On the other hand, a more engaging virtual environment can shorten perceived duration, which can be leveraged in designing unpleasant yet unavoidable experiences, such as chemotherapy (see Schneider et al. [33]).

However, the growing trend of perception research in VR needs to naturalize a VR perspective and exercise caution in interpreting the results through VR concepts and terminology. VR is not a simple tool in the social science lab; it is an entire laboratory.

ACKNOWLEDGMENTS

This study has been supported by the European Union's Horizon 2020 research and innovation programme under grant numbers 964464 (ChronoPilot).

REFERENCES

[1] Alessandro Angrilli, Paolo Cherubini, Antonella Pavese, and Sara Manfredini. 1997. The influence of affective factors on time perception. *Perception & Psychophysics* 59 (1997), 972–982. https://doi.org/10.3758/bf03205512
[2] Hamit Basgol, Inci Ayhan, and Emre Ugur. 2021. Time Perception: A Review on Psychological, Computational, and Robotic Models. *IEEE Transactions on Cognitive and Developmental Systems* 14, 2 (2021), 301–315. https://doi.org/10.1109/TCDS.2021.3059045
[3] Florian Block and Hans Gellersen. 2010. The impact of cognitive load on the perception of time. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. ACM, New York, NY, USA, 607–610. https://doi.org/10.1145/1868914.1868985
[4] Richard A. Block, Peter A. Hancock, and Dan Zakay. 2010. How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica* 134, 3 (2010), 330–343. https://doi.org/10.1016/j.actpsy.2010.03.006
[5] Jean Botev, Knut Drewing, Heiko Hamann, Yara Khaluf, Pieter Simoens, and Argiro Vatakis. 2021. ChronoPilot – Modulating Time Perception. In *2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, Taichung, Taiwan, 215–218. https://doi.org/10.1109/AIVR52153.2021.00049
[6] Atser Damsma, Nadine Schlichting, and Hedderik van Rijn. 2021. Temporal Context Actively Shapes EEG Signatures of Time Perception. *Journal of Neuroscience* 41, 20 (2021), 4514–4523. https://doi.org/10.1523/JNEUROSCI.0628-20.2021
[7] Atser Damsma, Nadine Schlichting, Hedderik van Rijn, and Warrick Roseboom. 2021. Estimating Time: Comparing the Accuracy of Estimation Methods for Interval Timing. *Collabra: Psychology* 7, 1 (2021), 21422. https://doi.org/10.1525/collabra.21422

[8] Sylvie Droit-Volet and Warren H. Meck. 2007. How emotions colour our perception of time. *Trends in Cognitive Sciences* 11, 12 (2007), 504–513. https://doi.org/10.1016/j.tics.2007.09.008

[9] Marco Fabbri, Jennifer Cancellieri, and Vincenzo Natale. 2012. The A Theory Of Magnitude (ATOM) model in temporal perception and reproduction tasks. *Acta Psychologica* 139, 1 (2012), 111–123. https://doi.org/10.1016/j.actpsy.2011.09.006

[10] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

[11] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191. https://doi.org/10.3758/bf03193146

[12] Aristides I. Ferreira. 2021. When Time Flies: How Timelessness Determines Creative Process Engagement. *Creativity Research Journal* 33, 2 (2021), 180–190. https://doi.org/10.1080/10400419.2020.1849486

[13] Mar Gonzalez-Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts-Escolano, et al. 2020. The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in virtual reality* 1 (2020), 20. https://doi.org/10.3389/frvir.2020.561558

[14] Simon Grondin and Marilyn Plourde. 2007. Judging Multi-Minute Intervals Retrospectively. *Quarterly Journal of Experimental Psychology* 60, 9 (2007), 1303–1312. https://doi.org/10.1080/17470210600988976

[15] Rashmi Gupta. 2022. Motivational salience, not valence, modulates time perception. *Emotion* 22, 2 (2022), 283.

[16] Robert E. Hicks, George W. Miller, and Marcel Kinsbourne. 1976. Prospective and retrospective judgments of time as a function of amount of information processed. *The American Journal of Psychology* 89, 4 (1976), 719–730. https://doi.org/10.2307/1421469

[17] Sae Kaneko and Ikuya Murakami. 2009. Perceived duration of visual motion increases with speed. *Journal of vision* 9, 7 (2009), 14–14. https://doi.org/10.1167/9.7.14

[18] Mathilde Lamotte and Sylvie Droit-Volet. 2017. Aging and Time Perception for Short and Long Durations: A Question of Attention? *Timing & Time Perception* 5, 2 (2017), 149–167. https://doi.org/10.1163/22134468-00002086

[19] Elizabeth Larson and Alexander von Eye. 2006. Predicting the Perceived Flow of Time From Qualities of Activity and Depth of Engagement. *Ecological psychology* 18, 2 (2006), 113–130. https://doi.org/10.1207/s15326969eco1802_3

[20] Aitao Lu, Lei Mo, and Bert H. Hodges. 2011. The Weight of Time: Affordances for an Integrated Magnitude System. *Journal of Experimental Psychology: Human Perception and Performance* 37, 6 (2011), 1855. https://doi.org/10.1037/a0024673

[21] Jean-Luc Lugrin, Fabian Unruh, Maximilian Landeck, Yoan Lamour, Marc Erich Latoschik, Kai Vogeley, and Marc Wittmann. 2019. Experiencing Waiting Time in Virtual Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. ACM, New York, NY, USA, 1–2. https://doi.org/10.1145/3359996.3364807

[22] Sandra Malpica, Belen Masia, Laura Herman, Gordon Wetzstein, David M. Eagleman, Diego Gutierrez, Zoya Bylinskii, and Qi Sun. 2022. Larger visual changes compress time: The inverted effect of asemantic visual features on interval time perception. *PloS one* 17, 3 (2022), e0265591. https://doi.org/10.1371/journal.pone.0265591

[23] William J. Matthews and Warren H. Meck. 2016. Temporal Cognition: Connecting Subjective Time to Perception, Attention, and Memory. *Psychological bulletin* 142, 8 (2016), 865. https://doi.org/10.1037/bul0000045

[24] Michele Migliore, Linda Messineo, and Maurizio Cardaci. 2000. A model of the effects of cognitive load on the subjective estimation and production of time intervals. *Biosystems* 58, 1-3 (2000), 187–193. https://doi.org/10.1016/S0303-2647(00)00122-2

[25] Ross Miller. 2016. Oculus founder thinks VR may affect your ability to perceive time passing. https://www.theverge.com/2016/3/17/11258718/palmer-luckey-oculus-time-vr-virtual-reality-gdc-2016

[26] Grayson Mullen and Nicolas Davidenko. 2021. Time Compression in Virtual Reality. *Timing & Time Perception* 9, 4 (2021), 377–392. https://doi.org/10.1163/22134468-bja10034

[27] Filip M. Nuyens, Daria J. Kuss, Olatz Lopez-Fernandez, and Mark D. Griffiths. 2020. The Potential Interaction Between Time Perception and Gaming: A Narrative Review. *International Journal of Mental Health and Addiction* 18 (2020), 1226–1246. https://doi.org/10.1007/s11469-019-00121-1

[28] Ian Phillips. 2012. Attention to the Passage of Time. *Philosophical Perspectives* 26 (2012), 277–308. https://doi.org/10.1111/phpe.12007

[29] Ignacio Polti, Benoît Martin, and Virginie van Wassenhove. 2018. The effect of attention and working memory on the estimation of elapsed time. *Scientific reports* 8, 1 (2018), 6690. https://doi.org/10.1038/s41598-018-25119-y

[30] Tyler Read, Christopher A. Sanchez, and Raffaele De Amicis. 2021. Engagement and time perception in virtual reality. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 65. SAGE Publications Sage CA, Los Angeles, CA, USA, 913–918. https://doi.org/10.1177/1071181321651337

[31] Tyler Read, Christopher A. Sanchez, and Raffaele De Amicis. 2023. The influence of attentional engagement and spatial characteristics on time perception in virtual reality. *Virtual Reality* 27, 2 (2023), 1265–1272. https://doi.org/10.1007/s10055-022-00723-6

[32] Christian Schatzschneider, Gerd Bruder, and Frank Steinicke. 2016. Who turned the clock? Effects of Manipulated Zeitgebers, Cognitive Load and Immersionon Time Estimation. *IEEE transactions on visualization and computer graphics* 22, 4 (2016), 1387–1395. https://doi.org/10.1109/TVCG.2016.2518137

[33] Susan M. Schneider, Cassandra K. Kisby, and Elizabeth P. Flint. 2011. Effect of virtual reality on time perception in patients receiving chemotherapy. *Supportive Care in Cancer* 19 (2011), 555–564. https://doi.org/10.1007/s00520-010-0852-7

[34] Katharina A. Schwarz and Lisa Weller. 2023. Distracted to a fault: Attention, actions, and time perception. *Attention, Perception, & Psychophysics* 85, 2 (2023), 301–314. https://doi.org/10.3758/s13414-022-02632-x

[35] Stephen D. Smith, Theresa A. McIver, Michelle S. J. Di Nella, and Michelle L. Crease. 2011. The Effects of Valence and Arousal on the Emotional Modulation of Time Perception: Evidence for Multiple Stages of Processing. *Emotion* 11, 6 (2011), 1305. https://doi.org/10.1037/a0026145

[36] Maria Tamm, Ainika Jakobson, Merle Havik, Saima Timpmann, Andres Burk, Vahur Ööpik, Jüri Allik, and Kairi Kreegipuu. 2015. Effects of heat acclimation on time perception. *International Journal of Psychophysiology* 95, 3 (2015), 261–269. https://doi.org/10.1016/j.ijpsycho.2014.11.004

[37] Ewart A. C. Thomas and Wanda B. Weaver. 1975. Cognitive processing and time perception. *Perception & psychophysics* 17, 4 (1975), 363–367. https://doi.org/10.3758/bf03199347

[38] Stefano Triberti, Luca Milani, Daniela Villani, Serena Grumi, Sara Peracchia, Giuseppe Curcio, and Giuseppe Riva. 2018. What matters is when you play: Investigating the relationship between online video games addiction and time spent playing over specific day phases. *Addictive Behaviors Reports* 8 (2018), 185–188. https://doi.org/10.1016/j.abrep.2018.06.003

[39] Peter U. Tse. 2010. *Attention and time.* Oxford University Press Oxford, England, Oxford, England, Chapter Attention underlies subjective temporal expansion, 137–150. https://doi.org/10.1093/acprof:oso/9780199563456.003.0010

[40] Fabian Unruh, Maximilian Landeck, Sebastian Oberdörfer, Jean-Luc Lugrin, and Marc Erich Latoschik. 2021. The Influence of Avatar Embodiment on Time Perception - Towards VR for Time-Based Therapy. *Frontiers in Virtual Reality* 2 (2021), 658509. https://doi.org/10.3389/frvir.2021.658509

[41] Fabian Unruh, David Vogel, Maximilian Landeck, Jean-Luc Lugrin, and Marc Erich Latoschik. 2023. Body and Time: Virtual Embodiment and its Effect on Time Perception. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (2023), 2626–2636. https://doi.org/10.1109/TVCG.2023.3247040

[42] John Wearden. 2014. Modelling behaviour on classical timing tasks: Verbal estimation, production, and reproduction. *Procedia-Social and Behavioral Sciences* 126 (2014), 34–35. https://doi.org/10.1016/j.sbspro.2014.02.307

[43] John H. Wearden. 2019. Body Temperature and the Sense of Time: A Translation of François (1927) with a commentary. *Timing & Time Perception* 7, 2 (2019), 87–107. https://doi.org/10.1163/22134468-20191142

[44] John H. Wearden and Ian S. Penton-Voak. 1995. Feeling the Heat: Body Temperature and the Rate of Subjective Time, Revisited. *The Quarterly Journal of Experimental Psychology Section B* 48, 2b (1995), 129–141. https://doi.org/10.1080/14640749508401443

[45] Stefan Weber, David Weibel, and Fred W. Mast. 2020. How Self-Motion in Virtual Reality Affects the Subjective Perception of Time. *Timing & Time Perception* 8, 2 (2020), 119–136. https://doi.org/10.1163/22134468-20191152

[46] Isabell Winkler, Katja Fischer, Kristin Kliesow, Tina Rudolph, Carolin Thiel, and Peter Sedlmeier. 2017. Has it really Been that Long? Why Time Seems to Speed up with Age. *Timing & Time Perception* 5, 2 (2017), 168–189. https://doi.org/10.1163/22134468-00002088

[47] Marc Wittmann and Sandra Lehnhoff. 2005. Age Effects in Perception of Time. *Psychological reports* 97, 3 (2005), 921–935. https://doi.org/10.2466/pr0.97.3.921-935

[48] Richard T. A. Wood and Mark D. Griffiths. 2007. Time Loss Whilst Playing Video Games: Is there a Relationship to Addictive Behaviours? *International Journal of Mental Health and Addiction* 5 (2007), 141–149. https://doi.org/10.1007/s11469-006-9048-2

[49] Andreas Wutz, Anuj Shukla, Raju S. Bapi, and David Melcher. 2015. Expansion and Compression of Time Correlate with Information Processing in an Enumeration Task. *Plos one* 10, 8 (2015), e0135794. https://doi.org/10.1371/journal.pone.0135794

[50] Meichao Zhang, Aitao Lu, and Bert H. Hodges. 2019. Lifting, tasting, and carrying: The interaction of magnitude and valence effects in time perception. *Acta psychologica* 193 (2019), 1–10. https://doi.org/10.1016/j.actpsy.2018.11.010