Journal of Building Engineering 70 (2023) 106280

Contents lists available at ScienceDirect

Journal of Building Engineering

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

Development and application of an experimental framework for the use of virtual reality to assess building users' productivity, comfort, and adaptive-behaviour

Arianna Latini, Elisa Di Giuseppe *, Marco D'Orazio

Department of Construction, Civil Engineering and Architecture (DICEA), Università Politecnica delle Marche, Ancona, Italy

ARTICLE INFO

Keywords:
Virtual reality
Validity
Experimental framework
Human comfort and behaviour
Productivity

ABSTRACT

The use of Virtual Reality (VR) to assess comfort, productivity, and behaviour in built environments is currently emerging. However, an effective and standardised methodology to support researchers in performing VR tests to collect reliable data is lacking. Thus, the study aims to develop a novel and comprehensive experimental framework for organising the application of VR in this research domain, based on five validity types: content, internal, face, ecological, and criterion validity. To illustrate its suitability and applicability, the framework was applied to a single-domain case study of a virtual office room. 52 volunteers were recruited to perform cognitive tests (Stroop test, OSPAN test, Magnitude-Parity test) and answer questionnaires. Tests and surveys have been developed to support the content validity of the research. Each test was performed under different thermal stimuli: 24 °C and 16 °C, randomly assigned, to pursue internal validity. The first goal was to demonstrate that the framework allowed the creation of highly immersive scenarios. Findings confirmed the ecological validity of the model with an excellent sense of presence, graphical satisfaction, involvement and realism and non-significant levels of cybersickness. The second aim concerned the evaluation of the ability of the VR environment to capture the influence of the temperature set point on the dependent variables. As expected, a statistically significant influence was detected only on thermal comfort votes and adaptive behaviour, thus supporting the criterion validity. Results highlighted the values and potentialities of applying the present framework in the context of the emerging multi-domain research in support of user-centred design.

Nomenclature

GS	Graphical Satisfaction
INV	Involvement
IVE	Immersive Virtual Environment
REAL	Experienced Realism
SP	Spatial Presence

* Corresponding author.
E-mail address: e.digiuseppe@staff.univpm.it (E. Di Giuseppe).

https://doi.org/10.1016/j.jobe.2023.106280
Received 22 September 2022; Received in revised form 6 March 2023; Accepted 10 March 2023
Available online 21 March 2023

VE	Virtual Environment
VR	Virtual Reality
TCV	Thermal Comfort Vote
TSV	Thermal Sensation Vote
TPV	Thermal Preference Vote
VCV	Visual Comfort Vote
VSV	Visual Sensation Vote
VPV	Visual Preference Vote

1. Introduction

Thanks to the technological advances in audio, graphics and rendering, Immersive Virtual Environments (IVEs) have been developed and widely used from simple means of entertainment into effective and reliable allies in research. Designers and researchers are facilitated in collecting complex data in highly realistic one-to-one environments [1], properly manipulating several variables and easily testing research hypotheses, thus greatly shortening the design procedure

Over the past years, Virtual Reality (VR) has been increasingly used in the research area of human well-being, comfort, productivity, and behaviour in buildings because it is a low-cost technology enhancing the speed of execution and repeatability of tests. It makes it possible to quickly simulate a larger number of environments and provides a reliable method for presenting the physical space to users through the immersive display.

Thus, VR technology allows researchers to overcome the many limitations of in situ research and laboratory environmental control studies [2]. Indeed, in situ studies often use real-life settings for a prolonged time and do not easily allow the creation of specific correlations for each design case. Laboratory studies are time and cost-consuming because strongly depend on the construction of setups with low flexibility in terms of layout and conditions to be tested and reproduced.

Recent advancements in VR technology offer a unique opportunity to support and expand the research in the built environment as properly designed indoor spaces are vital components for living and working. The overall aim is to discover, through the VR tool, the relationships among the changes and characteristics of the indoor spaces, in terms of layout (windows size, walls and light colour) and stimuli (thermal, visual, acoustical), with users' comfort and perceptions, productivity and adaptive behaviour.

Even if VR has been widely presented as a proper tool, a correct experimental procedure to effectively and efficiently structure the VR application is still missing [3]. Concerning the need and relevance of a strict measure protocol, this research aims to develop a comprehensive experimental framework for organising the application of virtual reality to enhance comfort, productivity, and adaptive behaviour research.

The rest of the paper is structured as follows: Section 2 presents a literature review on studies employing IVEs to evaluate human comfort, adaptive behaviour, and productivity. Section 3 describes the step-by-step experimental framework, which is later illustrated and applicated to a case study (Section 4).

2. Literature review on studies employing IVE to evaluate comfort, adaptive behaviour and productivity

To identify the gaps in the current state-of-the-art and guide the development of the framework on the topic, this paper focused on the available literature reviews. In particular, the application of IVEs for occupants' comfort and adaptive behaviour research is addressed in the recent comprehensive review by Alamirah et al. [1] that focused on the technologies, and the related concepts (i.e., immersion, presence, users experience, ecological validity), validation experiments, single and multi domains studies. Other systematic reviews focused on VR applications in the architecture, engineering and construction sectors [4,5] or on more specific subjects, such as biophilic design [6], visual perception [7], and energy-related behaviour [8].

The analysis revealed that visual comfort and subjective perception are the most investigated factors in the VR field. In particular, a wide range of studies addressed single-domain research considering the influence of different artificial and natural lighting statuses [9–16], correlated colour temperature [17,18], viewing location [19,20] or façade characteristics [21–26] and indoor settings [27–30]. Moreover, the main effect of different temperature set-point [31–33] on thermal comfort is also investigated in virtual environments, while acoustic comfort is addressed only in a few studies concerning different noise conditions [34,35]. It is well established that single-domain research is more understood than the multi-domain and crossed-effect approaches which remain still in part understudied [1]. Indeed, few multi-domain studies were carried out by combining thermal-visual stimuli [36,37], as for the crossed-effect research concerning visual stimuli on thermal/visual comfort [25,38] and combined audio-visual ones [39].

It is essential to discuss how data are collected during these experiments. The most widespread methods are subjective approaches. Individuals' comfort is typically assessed through onsite post-occupancy surveys that focus on self-reported perception. The surveys are typically developed according to well-established standards, such as ISO 10551:2019 [40], ANSI/ASHRAE 55:2004 [41] for sensation, preference and comfort ratings or ISO 12913-2: 2018 [42] and ITU-R BS 1116-3: 2015 [43] for the acoustical field-study.

In particular, the literature analysis revealed inhomogeneous approaches to the administration of the survey. Most of the studies administrated questionnaires at the end of the experiment, after removing the head-mounted display [9,10,16,18,20,23,28–31,36,44,45], others during the experiment session verbally [12,17,19,21,24,33,34] or, in a few cases, interactively using the controllers [27,38,39,46]. Concerning the former approach, participants' perceptions might not be realistic because their ratings, which refer to the virtual environment, are given in the real world [1]. The latter could potentially harm the immersion level throughout the tests, thus, influencing the test's reliability [1].

In addition, some studies [19–21,27–30,32–34,46] integrated the subjective approach of surveys with objective methods. Indeed, users' physiological responses throughout the experimental session were also recorded, such as the heart rate, skin conductance and temperature, and electroencephalogram, to be correlated with comfort and overall perceptions.

Concerning experiments studying users' productivity, this is usually objectively assessed via reading passage tasks, which are common obligations for office work and easy to be reproduced in IVEs [9–11,14,15,31,36,45]. Only few studies focused on objective measures across different tasks in cognitive performance using the Stroop test [19,23,30,44], visual [9,29,30,35,47] or auditory memory task [34,39], general visual tests [20] and arithmetic sums [48,29]. Sometimes these tests are combined with subjective measures asking users' to estimate their productivity or the influence of the environmental condition on it, by using, for example, the NASA-Task Load Index survey [15,23].

Productivity tasks are presented with different methods, e.g., on the virtual computer monitor [9,14,23,31,35,44,45], on virtual papers [10,11,15,36], slides [20] or via headphones [34,39].

Finally, to better understand the influence of preliminary design decisions on occupants' adaptive behaviour, some experiments in IVE addressed the complexity of the human-building interaction domain. The overall aim was to collect occupancy-related data to understand the link between behaviour and building design alternatives in both pre and post-occupancy assessments. The results will allow for better integration of occupancy information in the building energy design process, thus, reducing the performance gap between the simulated and real energy consumption. The majority of the reviewed papers focused on blinds and lighting adjustment [10–12,14,15,48]. Meanwhile, only a few papers relied on physical monitoring activity of indoor environmental conditions (air temperature, air velocity, relative humidity) while detecting participants' control decisions on climatic systems (heater, fans, air conditioning) to achieve better indoor comfort [31,45]. In general, participants experienced a dynamic visual change after expressing the intended behaviour and could also perceive the thermal outcome.

Within the above-mentioned studies, only some experiments compared IVEs and real environments for validation purposes [9,16,20,22,30–33,35,36,44,46,49]. The adequacy of virtual environments in replicating physical settings was assessed by looking for the reliability of the collected data.

On the other side, most of the reviewed articles applied IVE in isolation, without its real counterpart, and evaluated its suitability via questionnaires asking participants to self-report their sense of presence and immersivity. Well-established surveys schemes in this field are the *Slater-Usoh-Steed* (used in Refs. [22,31,45,50,51]), the *Igroup Presence Questionnaires* (*IPQ*) (used in Refs. [18,20,22,23,33]) and *ITC-Sense of Presence Inventory* (used in Ref. [12]). Another crucial aspect to evaluate the VE is "cybersickness", which occurs when there is any inconsistency between the visual sensations and the vestibular ones [52,53]. It is measured through the *Virtual Reality Sickness Questionnaire, VRSQ,* which investigates typical symptoms such as discomfort, eye fatigue, difficulty in focusing, headache, the feeling of heaviness in the head, blurred vision, dizziness and vertigo with closed eyes (used in Refs. [19,33,36,44,45]).

To sum up, the following key aspects, resulting from the studies reviewed in this section, were considered as a basis for the definition of the experimental framework.

- IVE experiments have been effectively carried out to study mainly visual comfort and perception in single-domain research: the main varied parameters were artificial and natural lighting statuses (e.g. illuminance level, correlated colour temperature).
- Surveys are administrated with inhomogeneous approaches: in the real environment, verbally or interactively during the tests.
- Most of the experiments analysing participants' productivity adopted reading passage tasks: none of them focused on the overall evaluation of human cognitive functions.
- Like the comfort studies, adaptive behaviour research primarily focused on blinds and lighting adjustment.
- Most studies adopted VEs in isolation, without the comparison with a real counterpart, and analysed users' sense of presence and immersivity and cybersickness disorders to support the suitability of the tool.

3. Design of the experimental framework

The proposed experimental framework aims to overcome the many limitations mentioned in the literature review, thus enhancing the application of VEs to investigate individuals' comfort, productivity and adaptive behaviour. It establishes a pre- and post-trials procedure, based on well-known validation types of the research study (internal and external validity) and measurement (content, face, and criterion validity) (Fig. 1). The former allows to get a strict experimental design to produce trustworthy data, the latter to organise the measurements, correlate the results with existing studies, and make general conclusions looking for a cause-effect relationship between variables.

In the early design stage, after the definition of the research purpose, the researcher should build the tool to measure the construct of interest. For instance, there is no objective, observable entity called "productivity" or "comfort/well-being" that can be directly measured, thus, specific tests and surveys should be properly produced.

To get valid results, the content of these instruments must cover all the relevant parts. Indeed, according to existing literature, standards, and theories.

- to deal whit the time constraint of VR exposure, individuals' productivity should be assessed with objective measures: asking workers to self-estimate their productivity should be a more suitable practice in long-term studies typical of RE experiments. Well-established literature [54–56] suggests measuring the three domain-general cognitive functions that are important across a

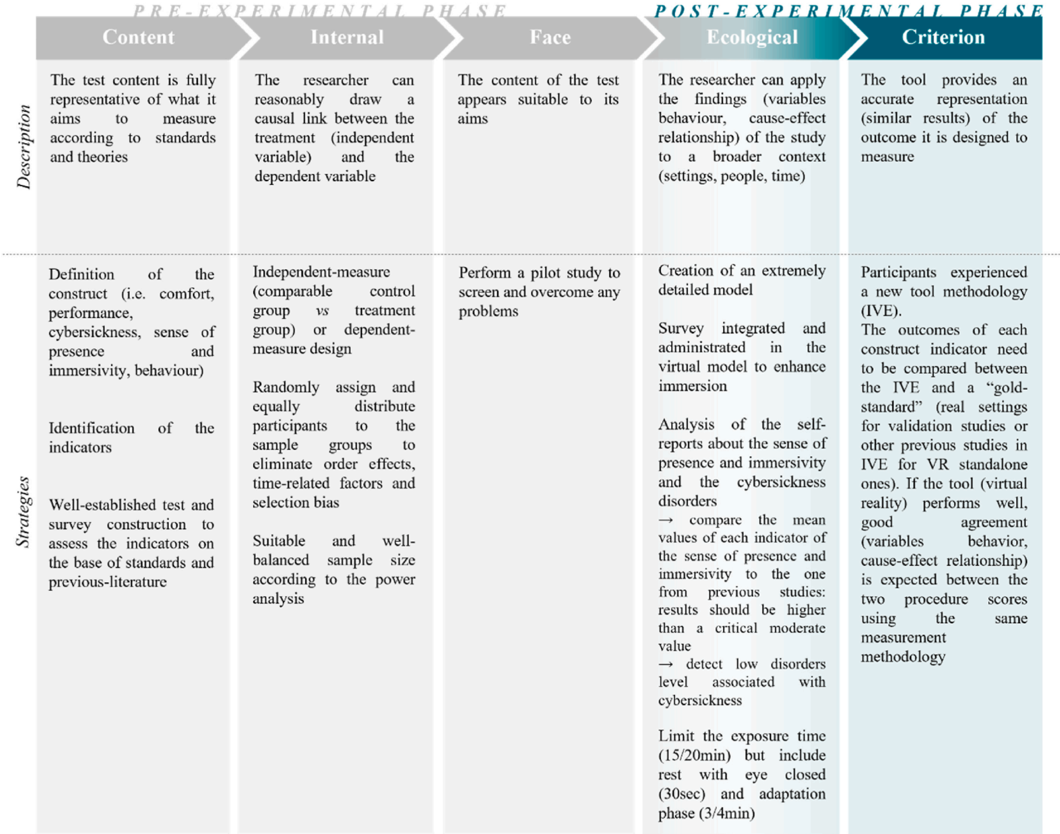

Fig. 1. The experimental framework: description and strategies.

variety of short-term tasks: *working memory*, which refers to the ability to keep information in mind and manipulate it, *inhibition*, which refers to the ability to control attention and override habits and impulses, and *task switching,* which refers to the ability to flexibly switch from one activity to another.

- Self-reports from participants should be based on the indicators of each construct: sensation, evaluation, and preference for comfort; spatial presence, graphical satisfaction, realism, and involvement for the sense of presence and immersivity factors; disorders for the cybersickness measure. To face the limit in VR time exposure, the researcher might need to revise the number of questions, meanwhile paying attention to developing effective and complete surveys coherent with the research purpose.
- Questionnaires should collect users' personal information (gender, age, educational level) which may explain the findings or some cause-effect relationships as potential confounding variables.

Thus, the outcomes of the developed instruments (tests and subjective votes) become the indicators of the constructs that the researcher wants to study and the more they represent the constructs, the greater the **content validity** [57].

Then, a strict study protocol should be developed to ensure the suitability of the experimental design supporting the **internal validity**. Researchers can conduct *repeated-measures* or *independent-measure* design research having the same group of individuals participate in all of the different treatments or have a different group in each treatment condition, respectively. Each experimental procedure has advantages and disadvantages that must be weighed in the early planning phase. Indeed, it is necessary to counterbalance time-related factors, order effect, individual differences and other biases that can be a threat to internal validity by managing the order of the tests [58]. In addition, the participants' group should be well-balanced in terms of gender and demographics, for example, to allow the generalisation of the results to a wider population. The sample size should be also adequate, due to its relevant effect on the results. Indeed, researchers should carry out a power analysis to ensure its adequacy in detecting significant effects based on the statistical power and the effect size.

Before the experimental session begins, content and internal validity should be assessed through a pilot study able to screen and overcome any problems with the created tool and protocol (**face validity**).

Once the trial design is concluded, the experiment sessions take place. Between the early definition of the experimental design from a theoretical point of view (content, internal, and face validity) and the empirical analysis of the data after the trials, researchers can find the fourth step: **external-ecological validity.** It refers to the ability of virtual environments to adequately represent real settings. This validity is an intermediate point because there are two inner steps: the need to provide a higher level of immersion during the test through the experimental procedure and virtual model development, and, afterwards, the analysis of the self-reports about the sense of presence, immersivity and the cybersickness disorders. For example, it is suggested to.

- Create an extremely detailed model in all its relevant features, such as objects, materials and lights, to be experienced through a suitable and effective head-mounted display in terms of resolution image per eye.
- Integrate and administrate the questionnaires inside the virtual model to ensure a fully immersive experience for the user and then reliable collected data. Since subjects provide feedback inside the virtual environment, authors will be allowed to avoid the *break-in-presence* [1] which can threaten the validity of the experiment.
- Test all the relevant constructs for the evaluation of the VE, thus, supporting also the content validity: cybersickness, sense of presence and immersivity according to the well-established surveys: *Virtual Reality Sickness Questionnaire*, *Igroup Presence Questionnaires* and *Slater-Usoh-Steed.* To verify the ecological validity is important that the majority of the subject do not experience high disorder levels and to detect a good level of sense of presence and immersivity. Concerning the latter, the authors suggest comparing the average value of each indicator to other literature studies that used the same tool, such as [22,23,44,59] and afterwards to ensure that they are higher than the value equivalent to the moderate-high level (i.e., 4) on a five-point Likert scale ranging from 1 to 5.
- Limit the exposure time to the virtual model below a recommended time interval of 25/30 min, to avoid the occurrence of any disturbances that could invalidate the test due to discomforting participants [8,47,60]. VR-induced symptoms are a recurrent issue in VR experiments and might depend on poor virtual experiments [33], the quality of hardware (HMD characteristics) and personal susceptibility to motion sickness [60]. Longer exposure time can be adopted but researchers should be aware of the risk of carrying out long-term studies that might generate higher disorder levels. Long-term experiments could be divided into shorter test sessions providing a break between them, but the researcher should avoid the removal of the head-mounted display not to affect participants' immersivity.

Despite that, the experimental schedule should always include, at the beginning of the experiment, 30 s of rest with eyes closed to reduce any psychological fluctuations related to the VE exposure, and 3/4 min of virtual scene adaptation to facilitate immersion (visual and acoustical) [27].

After the trials, once the ecological validity is established, the researchers should establish the **criterion-concurrent validity**. Within IVE validation studies to examine the adequacy in replicating real experimental situations and the reliability of the research, the outcomes of the constructs (i.e. comfort, productivity and behaviour) need to be compared between the two procedures (RE – IVE). Otherwise, VR standalone studies can be compared to the ones from previous research activities and accepted as being valid using a similar measurement methodology [61]. An example is a thermal comfort and adaptive behaviour correlation, the hue-heat hypothesis, or the disrupting effect of noise on productivity. If the tool (virtual reality) performs well a high correlation should be discovered in both cases, in terms of variables' behaviour or cause-effect relationships.

4. Application of the framework to a case study

4.1. Methods

To illustrate the suitability and applicability of the proposed framework, this section presents a single-domain case study to investigate the effect of indoor air temperature on users' comfort, productivity and adaptive behaviour in an immersive virtual office environment. Hence, two typical winter discomfort scenarios were set up: high (24 °C) and low (16 °C) indoor air temperatures.

A group of volunteers were recruited to perform three productivity tests and answer questionnaires about thermal comfort, visual comfort, sense of presence and immersivity cybersickness disorders. Occupant-building system intention of interactions was also recorded. Each individual performed the test twice under both different thermal stimuli.

In particular, using G*Power software [62], a minimum of 34 participants was estimated to detect a medium effect size (Cohen's d = 0.5) with a probability of 80% and confidence of 95%. To address potential data loss, 52 participants were recruited, thus, achieving 94% statistical power.

The sample was randomly divided into two balanced groups in terms of age and gender, with one group being tested at the lower temperature followed by the higher ones, and vice versa for the other group (24 °C and then 16 °C). Thus, there was no systematic bias in how participants are assigned to the groups, and also the groups the participants are in during experimentation are hidden. Moreover, there was one week between trials for each subject to support internal validity. As a result, any external effects on the two temperature conditions were evenly distributed.

Before the execution of the official trials, researchers carried out a pilot study with few participants to screen and overcome any problems (e.g., models, surveys, procedures), then achieved the face validity of the present study. Indeed, a newer script was added to the virtual model to go easily and sequentially back and forth between the scenes, minimising the interactions with the operators, some of the survey questions were modified, and in others, their number was reduced.

After the arrival of the volunteers, a pre-experimental phase (15 min) was carried out, to allow the adaptation to the environmental climatic conditions [63]. Meanwhile, the participants received information about the test procedure, read and signed a consent form and completed the pre-experimental questionnaire. Then, the operative phase began. After putting on the VR equipment and properly setting up the interpupillary distance, the subjects were asked to rest with their eyes closed for 30s and to adapt to the virtual scene for 3 min before the execution of the three cognitive tests (3min). Finally, subjects answered a sequence of questions (2min). The answers were given verbally and recorded by the researchers. To balance the fatigue level and IVE exposure of the subjects, each test session lasted about 25 min and was carried out as shown in Fig. 2.

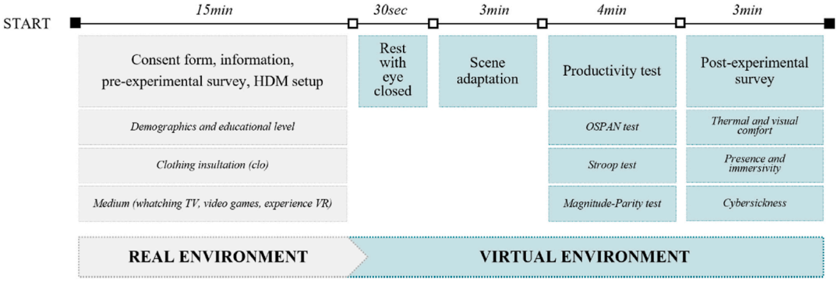

Fig. 2. Experimental schedule.

4.1.1. Test room

Each test was performed under different thermal stimuli (hot and cold air temperatures). Hence, the related experimental sessions were carried out by setting up two rooms with different indoor air temperatures, respectively, one inside the Building Science & Technology Laboratory (BS&T Lab) and the other inside the Department of Engineering, Civil, Construction and Architecture (DICEA) of the Università Politecnica delle Marche (Ancona, Italy). Both rooms were equipped with a computer station managed by the researcher to carry out the tests, and the equipment for the IVE visualisation.

The indoor air temperature, relative humidity and air velocity were recorded by several probes (temperature range:-40° - +125 °C, accuracy ±0.3 °C; relative humidity range: 0–100 %RH, accuracy ± 2%RH, air velocity range: 0.1–5 m/s, accuracy ± 0.2/0.3 m/s). The sensors were arranged on a tripod immediately behind the seated participants while performing the test (Fig. 3a) and were used to collect data at different heights corresponding to the common position of feet (0.10 m), waist (0.60 m) and head (1.10 m). The average values during the cooler test session were 15.84 °C (SD = 0.90), 50.36% RH and 24.45 °C (SD = 0.37), 32.43 %RH, in the warmer one. On average the air velocity has a constant value of 0.03 m/s in both temperature conditions.

4.1.2. IVE scenario

The IVE of this study is an extremely detailed 3D computer-generated graphics modelled based on a double-occupancy office space of about 20 square meters with two workstations (Fig. 3c). The initial 3D model was developed using *Rhinoceros,* which is one of the most used software for geometrical modelling, and *Unity* software [64], which is usually employed for the definition of the virtual model to apply materials, lights and cameras. To create the most realistic virtual model, it was necessary to address the correct representation of surfaces' colour and materials. Thus, the luminance parameter (L*) and chromatic components (a*, b*) of the CIELab model were performed through a spectrophotometer (*CM-2500d Konica Minolta*). For each office surface (walls, desk, chair, floor), 5 measurements were carried out (measurement area diameter 8 mm). Then, the resulting L*a*b* parameters were converted into RGB coordinates and entered into the material properties of the Unity model.

In particular, two basic scenarios were created: one corresponding to the adaptation phase located far from the virtual desk, to have a complete view of the virtual office room, and the other for the operative phase, where the participants were virtually seated at their desks to perform the performance tasks and the questionnaires. In neither of the two basic scenarios was it possible for subjects to view virtual hands, as no direct interaction with the model was allowed and the overall test session was verbally solved. Indeed, the productivity tests and questions of the surveys were shown through the virtual computer monitor, respectively, as attached timed videos and sequences of images, to provide the highest level of realism, to support external-ecological validity (Fig. 3d).

The IVE model was experienced through the *SteamVR* software plugin [65] combined with one of the most common technologies regarding head-mounted displays: the *HTC Corporation VIVE PRO Eye*. It has a 1440 × 1600 resolution image per eye, a pixel density of 615 PPI, a screen Dual OLED 3.5" diagonal, a maximal illumination of 125.73 cd/m^2, a refresh rate of 90 Hz, a field of view of 110° per eye and an adjustable interpupillary distance from 60.7 to 73.5 mm). To visualise the scenes, the height of the camera was set at 1.10 m height from the ground based on the average eye height of a seated person.

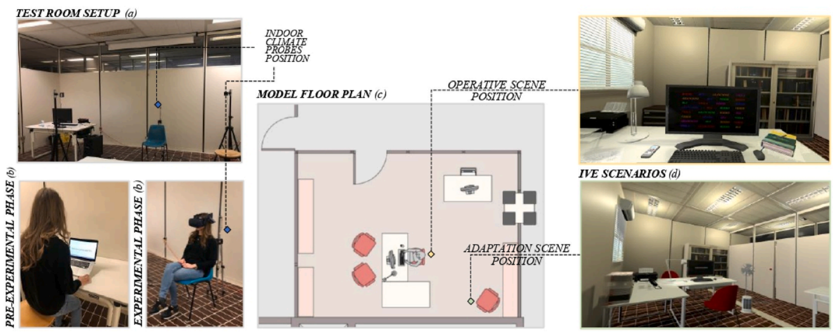

Fig. 3. (a) Test room setup; (b) A subject during the experiment; (c) the office room plan with the positions for the adaptation phase and the operative phase (d) the virtual model.

4.1.3. Performance test

To objectively assess the productivity supporting the content validity of the developed model, three specific tests were carried out to investigate the related executive functions: the *Magnitude-parity test* [66] for *task switching,* the *OSPAN test* [67] for the evaluation of the working memory, and the *Stroop test* [68] for inhibition.

The *Magnitude-Parity test* was performed through a timed video composed of a sequence of slides (200 ms). The digits from "1" to "9" except "5" were used as stimuli. All numbers were presented in black colour on a white background, in the middle of the screen preceded by coloured (red and blue) disks, presented in the centre of the screen too for 200 ms. Participants were asked to tell whether the displayed number is odd or even after the red dot is displayed (parity stimulus), and whether the displayed number is smaller or larger than 5 after the blue disk (magnitude stimulus). The dot-number strings were presented eight times for each stimulus, for a total of 16 digits to be ranked.

Also, the *OSPAN test* consists of a sequence of images of short duration: in the first slide (3sec) the participants saw one simple math operation, centred on the computer monitor, to be solved by the mind; the second slide (3sec) shows a number proposed as a possible solution to the previous equation. The subject was asked to tell aloud whether the solution is true or false. Immediately after, in the last slide (800 ms) they saw a letter to be memorised. The operation-true/false-letter strings were presented in a set of five items. At the end of each complete set, the participant was instructed to recall all the letters in the order presented.

Finally, the *Stroop test* was presented as a picture composed of a total of eight rows with four words on each row. Four colours: Red, Green, and Blue, representing the three components of the RGB model, pink and orange were displayed on a black background. Subjects were asked to name the colours of the words as fast as they can, while their execution time was collected. For example, if the word "red" is printed in blue ink, the correct answer is "blue".

Responses to all three cognitive tests were given by voice (Fig. 4).

To support internal validity, the execution of the three tests was randomised (e.g. *OSPAN, M-P, Stroop* vs *M-P, OSPAN, Stroop*) to counterbalance the order effect and time-related factors. Moreover, since repeated-measure design research was carried out and the subject performed the test under two thermal stimuli, a different version of the three tests was developed, with the same difficulty level.

4.1.4. Questionnaire

The pre-experiment survey was composed of 11 questions divided into two sections to retrieve information about demographics (gender, age, height, eyesight problems, educational level), and garments worn during the test to estimate the thermal insulation of the clothing (*clo* value) according to standard UNI EN ISO 9920:2007 [69]**.** To understand an individual's experience within the IVE, the authors included questions (from "never" to "every day" scores), that focused on daily sedentary activity and habits with some medium, such as watching TV, playing video games, and experience with VR [9]. The pre-experiment questionnaire was completed through an online platform to minimise interactions with the researcher avoiding any influence on the subject's answers.

On the other side, the post-experiment questionnaires included five sections: the intention of interactions, thermal comfort, visual comfort, cybersickness, sense of presence and immersivity. The first one focused on the adaptive strategies that subjects would have carried out to modify their well-being within the thermal environment. Indeed, according to the Theory of Planned Behaviour, the intention is assumed to be the immediate antecedent of the behaviour itself because people are expected to carry out their intentions when the opportunity arises [70]. A heater, fan, window, and air conditioning system were highlighted with red icons in the virtual office. However, participants' choices were not displayed in the virtual office or implemented in the physical environment to carry out a real change as a result of the interaction (switching devices on/off, opening/closing, etc.). Then, questions about thermal comfort were included to evaluate both temperature conditions through the parameters *Thermal Sensation Vote* (*TSV*), *Thermal Comfort Vote* (*TCV*), and *Thermal Preference Vote* (*TPV*). While visual comfort was assessed in the third section to understand the participants' perception of the simulated artificial lights. Subjects were asked about their *Visual Sensation Vote* (*VSV*), *Visual Comfort Vote* (*VCV*), and *Visual Preference Vote* (*VPV*). The questions were constructed based on standard UNI EN ISO 10551:2019 [40].

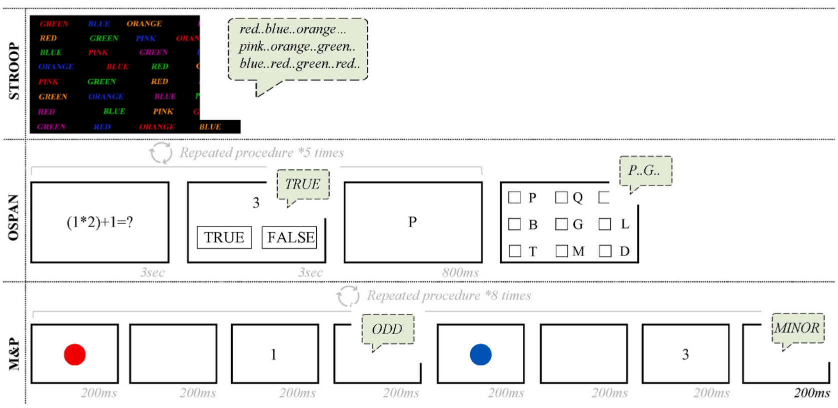

Fig. 4. Instructions and sequences of the slides for the three cognitive tests.

Two sections were also included to test the ecological validity of the developed model. The *Virtual Reality Sickness Questionnaire* (*VRSQ*) was used to assess motion sickness [71] on a five-point scale. The authors decide to reduce the investigated symptoms from nine to six: general discomfort, fatigue, eye strain, difficulty in focusing, headache, and vertigo. The last section was composed of a combination of the most popular questionnaires for the sense of presence and immersivity in IVEs: the *Slater-Usoh-Steed* and the *Igroup Presence Questionnaires* (*IPQ*). The construct was assessed according to four indicators: *Graphical Satisfaction (GS*), *Spatial Presence (SP*), *Involvement (INV*), and *Experienced Realism (REAL*), on a seven-point Likert scale.

The 16 questions and rating scales related to the post-experimental questionnaire are reported in Appendix A.

4.2. Results and discussion

The following paragraphs present the sample's data analyses and the dependent variables' results according to two research objectives: the ecological validity of the created virtual model, and the criterion validity in terms of the reliability of the effect of the independent variable (temperature set-point) on interactions, thermal and visual comfort, and productivity. All datasets were analysed with parametric and non-parametric statistical tests through *RStudio* software [72]. The significance level was set equal to 0.05 (5%).

4.2.1. Participants

The sample consisted of 52 participants (26 male, and 26 females) and it was mostly composed of young people as follows: 29% between 20 and 25 years old ($\mu = 22.6$; SD = 4.55), 44% between 26 and 30 ($\mu = 27.92$; SD = 7.10), 21% between 31 and 39 ($\mu = 33.1$; SD = 2.12) and only the 6% over 50 years old ($\mu = 54.33$; SD = 2.08). Most of the subjects were already graduated from university (58%), 23% were selected among university students and 17% had a higher educational level (PhD, graduate school). In particular, the majority of them (73%) reported doing sedentary activities every working day.

52% of the participants had eyesight problems, such as myopia and astigmatism, and all of them wore corrective lenses during the tests, thus not affecting the test performance or the model visualisation.

Moreover, to understand whether participants' experience within the IVE might be affected by other variables besides the sense of presence and immersivity and cybersickness, the authors focused on the mediums. In particular, the analysis revealed that: only 58% and 19% of the subjects watch television and play video games from “often” to “every day”; 62% had at least one previous experience with virtual reality. Data were collected for two weeks in December 2021. Thus, participants wore typical winter garments with about 0,82 clo and 0.81 clo values, in the warmer and colder temperature condition, respectively. The distributions of participants’ clothing insulation values are presented in Fig. 5: the box and the mean values highlight the correspondence between the two conditions in terms of garments worn.

4.2.2. Ecological validity

According to the first objective of this study, the self-reports from participants concerning the sense of presence and immersivity and cybersickness disorders were analysed.

Generally, the higher the values of the indicators *Graphical Satisfaction (GS*), *Spatial Presence (SP*), *Involvement (INV*), and *Experienced Realism (REAL*), the higher the sense of presence and immersivity of the participants within the IVE. The scores from the four indicators were compared with the ones from the previous studies [22,23,44,59]. The number of response options for each question in the questionnaires may vary depending on the experiment. As a consequence, a re-scaling of the average scores was initially carried out to convert the results of the answers expressed in seven-point scales into five-point scales and then make a direct comparison. According to Table 1 the scores in this study were generally higher than a moderate level (i.e. 4) on a five-point scale ranging from 1 to 5. The mean scores are reported in relevance order, as follows: graphic satisfaction (*GS* = 4.63), experienced realism (*REAL* = 4.55), spatial presence (*SP* = 4.23) and involvement (*INV* = 4.22). Moreover, the values are higher than those in the four previous studies. In particular, a very good sense of presence was obtained as the mean value for *SP* is 4.23 which is higher than [44] (3.74) and [59]

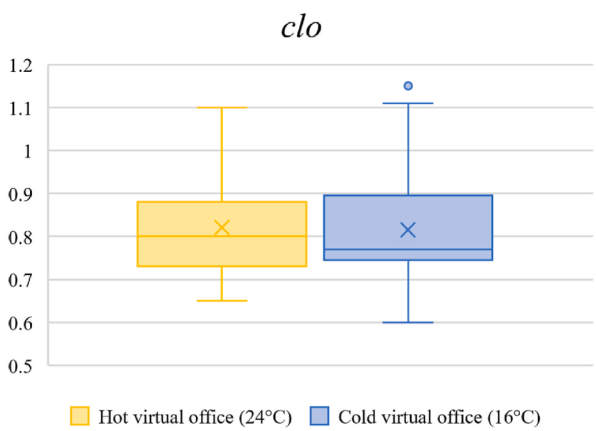

Fig. 5. *clo* values within the two experimental conditions.

Table 1
Comparison of scores on a 5-point scale of the Slater-Usoh-Steed and Igroup Presence Questionnaires for the four factors investigated: graphic appreciation (SP), spatial presence (SP), involvement (INV), realism (REAL).

		Year	*GS*	*SP*	*INV*	*REAL*
This study		2022	4.63	4.23*	4.22	4.55
Previous studies	[22]	2019	3.65	3.39	3.23	2.73
	[44]	2019	–	3.74	–	3.21
	[59]	2019	–	3.68	–	3.75
	[23]	2020	–	4.24*	4.11	3.54

(3.68) and almost similar to Ref. [23] (4.24). As the difference is negligible, this study developed a VE that offered an excellent sense of presence and immersivity (see Table 2).

Then, the authors analyse the relationship between the *mediums* and the ratings of the most significant indicators for the sense of presence and immersivity: spatial presence (*SP)* and realism (*REAL*). Somers' delta (d) was computed to analyse the association between the dependent variable (*IPQ* grades) and the independent variable (*medium*). The analysis revealed a low agreement concerning *SP* ($d_{\text{video games}} = 0.02$, $d_{VR} = 0.12$), and a low discordance concerning *REAL* ($d_{\text{video games}} = -0.001$, $d_{VR} = -0.14$). This result is in contrast with the baseline study [9] where it was found that subject gamers perceived the virtual environment as more realistic and similar to a real environment than non-gamers. However, based on the findings of the present study, the absence of a relationship between *SP, REAL* and the mediums seems to highlight that a high sense of presence (*SP*) and realism (*REAL*) could be achieved regardless of the level of familiarity with video games and virtual environments.

Finally, concerning *VRSQ* results (Fig. 6), the qualitative analysis of the cybersickness symptoms revealed that no subject suffered from vertigo since the test was carried out in static conditions, thus, without any difference between perceived and experienced movement. Other symptoms, such as general discomfort, fatigue and headaches, were negligible since the majority of the subjects (between 96% and 100%) assigned a score of «not at all» and «slightly». Moreover, only 7% of them reported at least « moderate » eye fatigue and 23% «difficulty in focusing » due to slightly blurred images presented by the head-mounted display.

4.2.3. *Temperature effect on thermal and visual comfort*

The second objective of this study (criterion validity) focused on the effect of the two temperature set-points on thermal (*Thermal Sensation Vote, Thermal Comfort Vote, Thermal Preference Vote*) and visual comfort (*Visual Sensation Vote, Visual Comfort Vote, Visual Preference Vote*). Thus, this study evaluated both a same-modality effect (thermal domain – thermal response) and a cross-modal effect (thermal domain and visual response).

Fig. 7 shows the subjective ratings on the left-hand side for the thermal votes and on the right-hand side for the visual votes. Within each temperature level, the plots represent participants' percentage of votes.

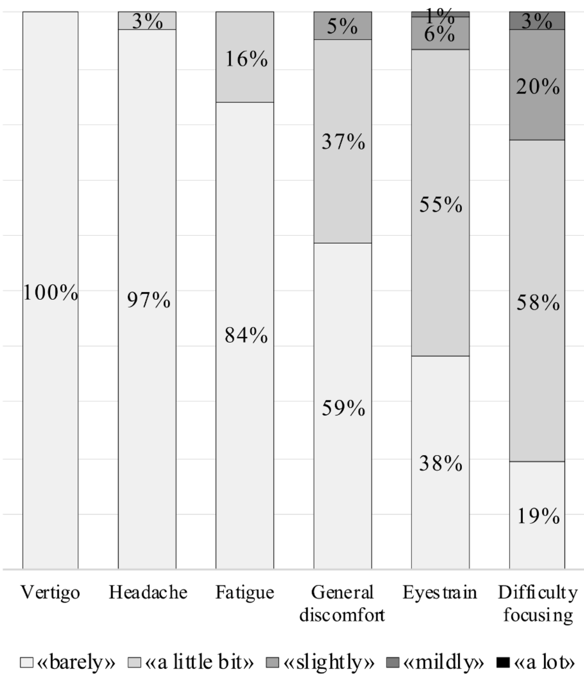

Fig. 6. Results of the cybersickness disorders.

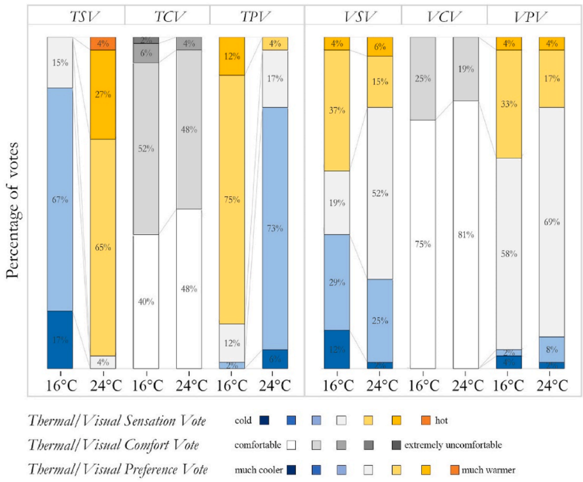

Fig. 7. Thermal and visual comfort percentage of votes vote within the two experimental conditions (16–24 °C).

As expected, the temperature has a significant influence on TSV: at 16 °C, 84% of the subjects felt at least « slightly cold» and «cold», while at 24 °C, 90% felt « slightly warm» and «warm». As a consequence, both thermal conditions were evaluated as not fully comfortable (from «slightly uncomfortable » to « uncomfortable»), by 52%–60% of the subjects. Thus, according to the TPV, at 16 °C the majority (87%) of the subjects would have wanted to feel at least « slightly warmer» in the cooler conditions, while 77% would have preferred a «slightly cooler» within the warmer environmental condition. To support the previous assumptions, a statistical analysis was carried out on the ordinal data using the non-parametric Wilcoxon signed-rank test. The null hypothesis stated that there was no difference in terms of thermal comfort votes ($\mu_\Delta = 0$). Because the sample data produced T values lower than the critical values in two out (TSV, TPV) of the three cases, the authors concluded that there was a significant difference (p-value < 0.00001). On the other side, there was no significant difference in TCV (p-value > 0.05). These results are not surprising. Indeed, the qualitative analysis suggested that the tested thermal conditions were discovered to be almost equally uncomfortable because both temperatures were ± 4 °C away from the winter thermal comfort temperature (20 °C).

Then, the same methodological approach was used to analyse visual comfort. The qualitative analysis reveals a congruence in VSV for both temperature set points. Indeed, between 60 and 79% of the subjects evaluated the correlated colour temperature of the lighting between « neutral» and «slightly cool». Interestingly, the researchers noted that, while filling out the surveys in the IVE, the participants focused on the virtual ambient light by a quick check looking away from the computer monitor. They focused more on the light colour of the work lamp (4000 K, cool white light) than on the colour of the LED ceiling lighting (warm white light). This supports the effectiveness and suitability of administrating questionnaires integrated into the virtual model. The same agreement was discovered also for the VCV parameter: most of the subjects (75%–81%) felt comfortable.

As a result, between 58% and 69% of them would have not changed their visual condition. Therefore, the qualitative analysis does not reveal a significant influence of temperature on visual comfort, thus, supporting previous literature studies in both IVEs [36,37] and real settings [73]. This hypothesis is supported by the statistical analysis (Wilcoxon signed-rank test, see Table 2) for all three visual parameters (p-value > 0.05).

4.2.4. *Temperature effect on adaptive behaviour*

To investigate adaptive strategies for improving well-being within the indoor environment, the authors analysed participants' intention of interaction with typical thermal environment control systems (heater, fan, window, air conditioning). The analysis evaluated the number, order and type of interaction intentionality within both temperature set points. Generally, a maximum of one intention per participant was recorded in both thermal conditions: at least 77% of participants would have modified their thermal condition by interacting with one of the highlighted components. This result agrees with the TPV scores. The Wilcoxon test also confirmed that there was no significant difference in the number of interactions for both temperature set-points (p-value$_{16\text{-}24}$ = 0.68, T-

Table 2
The results of the Wilcoxon signed-ranks test on the thermal and visual comfort.

	Parameter	T-value	T-critic	p-value	$\Delta\mu_{16\text{-}24}$
Thermal comfort	TSV	0	434	< 0.00001	−2.29
	TCV	130	116	0.97	0.23
	TPV	43	361	< 0.00001	0.05
Visual comfort	VSV	271.5	170	0.87	−0.05
	VCV	78.5	10	0.42	0.05
	VPV	110	46	0.70	0.17

value = 76, T-critic (18) = 40). Since only one interaction intention per subject was detected, it was not possible to study their order.

Moreover, the authors compared the type of interactions. The qualitative analysis in Fig. 8 showed a difference between the two thermal conditions: opening the window (65%) and turning on the heater (58%) were found to be the best strategies that the subjects would use to improve their thermal sensation in the higher temperature environment, to decrease the indoor temperature and enhance air change, and in the lower temperature environment, respectively. The significant differences across the type of interaction were also confirmed by the Wilcoxon signed-rank test (see Table 3), which generated a p-value < 0.00001 (T-value = 181, T-critic (49) = 416).

4.2.5. Temperature effect on productivity

The data from the three cognitive tests are assumed to be normal because they belong to a ratio scale distribution and the sample size is greater than 30 (n = 52), thus, the qualitative hypotheses were tested with parametric statistical analysis on a dependent-measure research design.

The Magnitude-Parity test results were analysed by calculating the number of times the subject wrongly classified the digits even/odd and greater/lower than "5". The analysis of the mean and standard deviations revealed no difference between the errors between the two thermal conditions: $mean_{16°C} = 0.58$ (dev.st = 1.02), $mean_{24°C} = 0.40$ (dev.st = 0.60). The outcomes of the *t*-test allowed not to reject the null hypothesis (no difference in terms of productivity, $\mu_\Delta = 0$): the t-value falls within the critical region (±2.011 for df = 51, α = 0.05), therefore the task-switching of the subjects was not influenced by the thermal conditions.

According to the automated OSPAN test [67] the working memory was evaluated as follows: the number of errors in the true/false combination ($mean_{24°C} = 0.29$, dev.st = 0.64; $mean_{16°C} = 0.35$, $dev.st_{16°C} = 0.84$), the correct order of the letters recalled ($mean_{24°C} = 4.48$, dev.st = 0.98; $mean_{16°C} = 4.60$, $dev.st_{16°C} = 0.91$), and the OSPAN score, i.e. the sum of the number of the right true/false and the letters correctly reported ($mean_{24°C} = 9.33$, dev.st = 1.08; $mean_{16°C} = 9.33$, $dev.st_{16°C} = 1.40$). The mean and standard deviation values, following the statistical analysis, revealed the absence of statically significant difference (p-value >0.05) also for the OSPAN test.

For the Stroop test, the number of errors in the colour recall, and the time in seconds participants took to pronounce all 32 colours (speed of execution) were measured. The qualitative analysis revealed approximately double the error in the warmer thermal condition ($mean_{24°C} = 0.31$, $dev.st_{24°C} = 0.94$), compared to the cooler one ($mean_{16°C} = 0.15$, $dev.st_{16°C} = 0.46$). Despite this, the *t*-test did not reveal any statistically significant differences in the inhibition function between them. The same result was found in the speed of execution ($mean_{24°C} = 28.85$, dev.st = 5.49; $mean_{16°C} = 29.71$, $dev.st_{16°C} = 6.98$), as confirmed by the *t*-test.

In conclusion, based on the outcome of the qualitative and statistical analyses, no influence of temperate on the subjects' executive functions was detected.

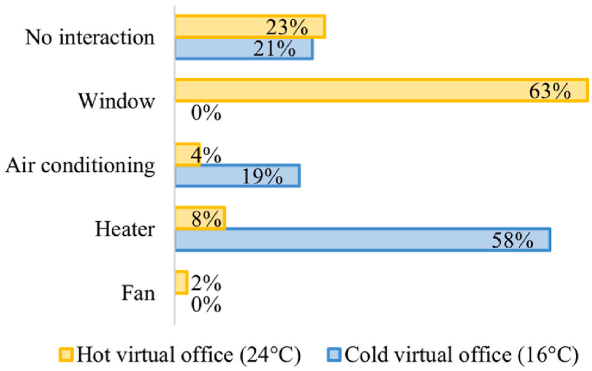

Fig. 8. Comparison of the intention of interaction within the two temperature conditions.

Table 3
The results of the paired-sample *t*-test on productivity within the three cognitive tests.

Cognitive function test	Parameter	t-value	p-value	$\Delta\mu_{16\text{-}24}$
Magnitude-Parity	number of errors in the classification of the digits even/odd and greater/lower than "5"	+1.46	0.15	0.17
OSPAN	the number of errors in the true/false string	0.28	0.78	0.02
	the correct order of the letters recalled	0.09	0.93	0.02
	OSPAN score (the sum of the number of the right true/false and the letters correctly reported)	0	1	0
Stroop	number of errors in the colour recall	−1.01	0.31	−0.15
	speed of execution	0.99	0.33	0.86

5. Conclusions

This study develops a comprehensive conceptual framework to efficiently conduct experiments using the Virtual Reality tool in the comfort, productivity, and adaptive behaviour research domain.

The framework was developed based on the five main types of validation of the experimental design and measures. The authors present a step-by-step description of the pre and post trials procedure which allows the researcher to establish a strict experimental design, produce trustworthy data (content – internal – face validity), correlate the results with existing studies looking for a cause-effect relationship between variables (criterion-validity) and make general conclusions (external-ecological validity).

To illustrate the suitability and applicability of the proposed framework, this was applied to a single-domain case study of a virtual office room. The aim was to investigate the effect of indoor air temperature on users' comfort, productivity and adaptive behaviour. A set of 52 participants performed three cognitive short-term tests (OSPAN test, Stroop test, Magnitude-Parity test) and answered questionnaires (thermal comfort, visual comfort, sense of presence and immersivity, cybersickness disorders) developed to support the content validity. The intention of interactions with several building systems (heater, fan, window, air conditioning) was also recorded. Each test was performed under different thermal stimuli: hot (24 °C) and cold (16 °C) indoor air temperatures to simulate typical winter discomfort scenarios (randomly assigned, according to internal validity). The tests and questions were administered automatically and integrated into the virtual computer monitor to limit interactions with operators and ensure a constant level of immersivity limiting the break-in presence.

The data were analysed according to two objectives: the external-ecological validity of the model created and the evaluation of the effect of the independent variable (temperature set-point) on the dependent variables (interactions, thermal and visual comfort, productivity). The main results of the case study are.

- The analysis and the comparison of the four indicators (graphical satisfaction, experienced realism, involvement, and spatial presence), with similar past studies, revealed that the virtual environment created an excellent level of presence and immersivity. The cybersickness analysis showed that most subjects did not report high disorder levels, except 23% who presented at least moderate « difficulty focusing » due to slightly blurred images. Moreover, the absence of a correlation with the mediums showed that, regardless of the level of familiarity with video games and other virtual environments, users can experience a high level of sense of presence and realism. As a result, these outcomes allowed the authors to support ecological validity.
- The independent variable (temperature set-point) did not influence the subjects' executive functions (task-switching, inhibition, working memory) and visual comfort. As expected it had a significant impact on thermal comfort (TSV) and the interactions with the building system. In particular, participants rated both the thermal conditions to be equally « uncomfortable» (TCV), because the set points were about ±4 °C far from the winter thermal comfort temperature (20 °C). Indeed, at least 77% of them would have to interact with opening the window (65%) and turning on the heater (60%), in warmer and cooler environments, respectively, thus, identified as the best feasible strategies. The reliability of these results confirms the criterion validity.

In this study, an accurate and reliable experimental method was developed. However, a validation process was not carried out to verify the correspondence between the virtual and real environments in terms of productivity, comfort and adaptive behaviour. Hence, the authors foresee using the proposed framework to carry out similar tests in real environments and then look for a correspondence with the results above presented. Future studies should also compare the data collected in these "extreme" thermal conditions (16 °C and 24 °C) with more traditional ones (around 20 °C). Moreover, the demographics of the sample size should be extended to involve a wider range of adults, thus ensuring the homogeneity of the participants' group. Even if the ecological validity revealed an excellent sense of presence and immersivity, another issue needing further attention is the modelling of the light characteristics and parameters via proper lighting software to reproduce more photorealistic renders that will enhance users' virtual experience. Future studies should assess users' productivity on mid-term tests, such as integrating the three cognitive tests with comprehension tasks, while paying attention to VR exposure without discomforting participants. In addition, to face the time constraints of VR exposure, a pilot study with different exposure levels should be carried out to improve the face validity of VR studies and detect unforeseen problems that could compromise the quality of the whole research study in terms of cybersickness disorders.

In conclusion, according to the outcomes, the proposed framework allows the development of an effective and efficient testing procedure to apply VR tools in further analysis to deliver valid and generalisable measures and results. The framework will allow to strengthen and facilitate research about human comfort through multi-domain and cross-effect analyses carried out in virtual environments. Indeed, thanks to its many advantages, different stimuli (thermal, visual, acoustic, olfactory) and environmental layouts (colours of walls, lighting, size of windows, indoor design, indoor geometry) could be easily tested and combined to improve a user-centred design approach and develop models about occupants' behaviour to be integrated into building performance simulation software.

CRediT authorship contribution statement

Arianna Latini: Methodology, Software, Investigation, Formal analysis, Conceptualization, Writing – original draft, Writing – review & editing. **Elisa Di Giuseppe:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Marco D'Orazio:** Methodology, Supervision, Funding acquisition.

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.

Appendix A. Post-experimental questions and rating scales

Factor	Question	Rating scale
Intention	Would you interact with the highlighted building systems to improve your wellbeing? If yes, please state your willing interactions	yes - no
Thermal comfort	*TSV* How do you judge this environment? *TCV* Do you find this..? *TPV* Please state how would you prefer to be now.	very cold/very warm comfortable/extremely uncomfortable much colder/much warmer
Visual comfort	*VSV* How do you judge artificial lighting? *VCV* Do you find this..? *VPV* Please state how would prefer the artificial lighting to be now.	very cold/very warm comfortable/extremely uncomfortable much colder/much warmer
Graphical satisfaction (GP)	I appreciate the graphics and images of the virtual model	totally disagree/totally agree
Spatial presence (SP)	I perceived the office space as a place I visited rather than a photo I saw During the experience, I felt present in the office space I perceived the virtual model as immersive	totally disagree/totally agree
Involvement (INV)	During the experience, I was not aware of the real world around me	totally disagree/totally agree
Experienced realism (REAL)	I perceived the objects inside the virtual office as proportionally correct (i.e., they had about the right size and distance from me and other objects) I had the feeling of being able to interact with the office space (e.g. grab objects) How realistic did you find the virtual model of the office space?	totally disagree/totally agree
Cybersickness	Did you experience..? GENERAL DISCOMFORT – FATIGUE - EYE STRAIN - DIFFICULTY IN FOCUSING – HEADACHE - VERTIGO	not at all/a lot

References

[1] H. Alamirah, M. Schweiker, E. Azar, Immersive virtual environments for occupant comfort and adaptive behavior research – a comprehensive review of tools and applications, Build. Environ. 207 (2021) 108396, https://doi.org/10.1016/j.buildenv.2021.108396.

[2] J. Li, W. Wu, Y. Jin, R. Zhao, W. Bian, Research on environmental comfort and cognitive performance based on EEG + VR + LEC evaluation method in underground space, Build. Environ. 198 (2021) 107886, https://doi.org/10.1016/j.buildenv.2021.107886.

[3] F. Vittori, C. Chiatti, I. Pigliautile, A.L. Pisello, The NEXT.ROOM: design principles and systems trials of a novel test room aimed at deepening our knowledge on human comfort, Build. Environ. 211 (2022) 108744, https://doi.org/10.1016/j.buildenv.2021.108744.

[4] Y. Zhang, H. Liu, S.C. Kang, M. Al-Hussein, Virtual reality applications for the built environment: research trends and opportunities, Autom. ConStruct. 118 (2020) 103311, https://doi.org/10.1016/j.autcon.2020.103311.

[5] J.M. Davila Delgado, L. Oyedele, P. Demian, T. Beach, A research agenda for augmented and virtual reality in architecture, engineering and construction, Adv. Eng. Inf. 45 (2020) 101122, https://doi.org/10.1016/j.aei.2020.101122.

[6] M. Mollazadeh, Y. Zhu, Application of virtual environments for biophilic design: a critical review, Buildings 11 (2021), https://doi.org/10.3390/buildings11040148.

[7] A. Bellazzi, L. Bellia, G. Chinazzo, F. Corbisiero, P. D'Agostino, A. Devitofrancesco, F. Fragliasso, M. Ghellere, V. Megale, F. Salamone, Virtual reality for assessing visual quality and lighting perception: a systematic review, Build. Environ. 209 (2022) 108674, https://doi.org/10.1016/j.buildenv.2021.108674.

[8] Y. Zhu, S. Saeidi, T. Rizzuto, A. Roetzel, R. Kooima, Potential and challenges of immersive virtual environments for occupant energy behavior modeling and validation: a literature review, J. Build. Eng. 19 (2018) 302–319, https://doi.org/10.1016/j.jobe.2018.05.017.

[9] A. Heydarian, J.P. Carneiro, D. Gerber, B. Becerik-Gerber, T. Hayes, W. Wood, Immersive virtual environments versus physical built environments: a benchmarking study for building design and user-built environment explorations, Autom. ConStruct. 54 (2015) 116–126, https://doi.org/10.1016/j.autcon.2015.03.020.

[10] A. Heydarian, E. Pantazis, J.P. Carneiro, D. Gerber, B. Becerik-Gerber, Lights, building, action: impact of default lighting settings on occupant behaviour, J. Environ. Psychol. 48 (2016) 212–223, https://doi.org/10.1016/j.jenvp.2016.11.001.

[11] S. Saeidi, T. Rizzuto, Y. Zhu, R. Kooima, Measuring the effectiveness of an immersive virtual environment for the modeling and prediction of occupant behavior, in: Sustain. Human-Building Ecosyst. - Sel. Pap. From 1st Int. Symp. Sustain, . Human-Building Ecosyst., 2015, pp. 159–167, https://doi.org/10.1061/9780784479681.017.

[12] S. Saeidi, C. Chokwitthaya, Y. Zhu, M. Sun, Spatial-temporal event-driven modeling for occupant behavior studies using immersive virtual environments, Autom. ConStruct. 94 (2018) 371–382, https://doi.org/10.1016/j.autcon.2018.07.019.

[13] J.P. Carneiro, A. Aryal, B. Becerik-Gerber, Understanding the influence of orientation, time-of-day and blind use on user's lighting choices and energy consumption using immersive virtual environments, Adv. Build. Energy Res. (2019) 1–27 https://doi.org/10.1080/17512549.2019.1639075, 0.

[14] J.P. Carneiro, A. Aryal, B. Becerik-Gerber, Influencing occupant's choices by using spatiotemporal information visualization in Immersive Virtual Environments, Build. Environ. 150 (2019) 330–338, https://doi.org/10.1016/j.buildenv.2019.01.024.

[15] P. Mahmoudzadeh, Y. Afacan, M.N. Adi, Analyzing occupants' control over lighting systems in office settings using immersive virtual environments, Build. Environ. 196 (2021) 107823, https://doi.org/10.1016/j.buildenv.2021.107823.

[16] M. Hegazy, K. Ichiriyama, K. Yasufuku, H. Abe, Comparing daylight brightness perception in real and immersive virtual environments using perceptual light maps, Autom. ConStruct. 131 (2021) 103898, https://doi.org/10.1016/j.autcon.2021.103898.

[17] Y. Chen, Z. Cui, L. Hao, Virtual reality in lighting research: comparing physical and virtual lighting environments, Light. Res. Technol. 51 (2019) 820–837, https://doi.org/10.1177/1477153518825387.
[18] J.H. Ma, J.K. Lee, S.H. Cha, Effects of lighting CCT and illuminance on visual perception and task performance in immersive virtual environments, Build. Environ. 209 (2022) 108678, https://doi.org/10.1016/j.buildenv.2021.108678.
[19] F. Abd-Alhamid, M. Kent, J. Calautit, Y. Wu, Evaluating the impact of viewing location on view perception using a virtual environment, Build. Environ. 180 (2020) 106932, https://doi.org/10.1016/j.buildenv.2020.106932.
[20] A. Emamjomeh, Y. Zhu, M. Beck, The potential of applying immersive virtual environment to biophilic building design: a pilot study, J. Build. Eng. 32 (2020) 101481, https://doi.org/10.1016/j.jobe.2020.101481.
[21] K. Chamilothori, G. Chinazzo, J. Rodrigues, E.S. Dan-Glauser, J. Wienold, M. Andersen, Subjective and physiological responses to façade and sunlight pattern geometry in virtual reality, Build, Environ. Times 150 (2019) 144–155, https://doi.org/10.1016/j.buildenv.2019.01.009.
[22] T. Hong, M. Lee, S. Yeom, K. Jeong, Occupant responses on satisfaction with window size in physical and virtual built environments, Build. Environ. 166 (2019), https://doi.org/10.1016/j.buildenv.2019.106409.
[23] S. Yeom, H. Kim, T. Hong, M. Lee, Determining the optimal window size of office buildings considering the workers' task performance and the building's energy consumption, Build. Environ. 177 (2020) 106872, https://doi.org/10.1016/j.buildenv.2020.106872.
[24] C. Moscoso, K. Chamilothori, J. Wienold, M. Andersen, B. Matusiak, Window size effects on subjective impressions of daylit spaces: indoor studies at high latitudes using virtual reality, LEUKOS - J. Illum. Eng. Soc. North Am. 17 (2021) 242–264, https://doi.org/10.1080/15502724.2020.1726183.
[25] F. Vittori, I. Pigliautile, A.L. Pisello, Subjective thermal response driving indoor comfort perception: a novel experimental analysis coupling building information modelling and virtual reality, J. Build. Eng. (2021) 102368, https://doi.org/10.1016/j.jobe.2021.102368.
[26] M. Marzouk, M. ElSharkawy, A. Mahmoud, Analysing user daylight preferences in heritage buildings using virtual reality, Build. Simulat. 15 (2022) 1561–1576, https://doi.org/10.1007/s12273-021-0873-9.
[27] J. Li, Y. Jin, S. Lu, W. Wu, P. Wang, Building environment information and human perceptual feedback collected through a combined virtual reality (VR) and electroencephalogram (EEG) method, Energy Build. 224 (2020) 110259, https://doi.org/10.1016/j.enbuild.2020.110259.
[28] S. Yeom, H. Kim, T. Hong, Psychological and physiological effects of a green wall on occupants: a cross-over study in virtual reality, Build. Environ. 204 (2021) 108134, https://doi.org/10.1016/j.buildenv.2021.108134.
[29] J. Yin, J. Yuan, N. Arfaei, P.J. Catalano, J.G. Allen, J.D. Spengler, Effects of biophilic indoor environment on stress and anxiety recovery: a between-subjects experiment in virtual reality, Environ. Int. 136 (2020) 105427, https://doi.org/10.1016/j.envint.2019.105427.
[30] J. Yin, S. Zhu, P. MacNaughton, J.G. Allen, J.D. Spengler, Physiological and cognitive performance of exposure to biophilic indoor environment, Build. Environ. 132 (2018) 255–262, https://doi.org/10.1016/j.buildenv.2018.01.006.
[31] G. Ozcelik, B. Becerik-Gerber, Benchmarking thermoception in virtual environments to physical environments for understanding human-building interactions, Adv. Eng. Inf. 36 (2018) 254–263, https://doi.org/10.1016/j.aei.2018.04.008.
[32] D. Yeom, J.H. Choi, Y. Zhu, Investigation of physiological differences between immersive virtual environment and indoor environment in a building, Indoor Built Environ. 28 (2019) 46–62, https://doi.org/10.1177/1420326X17731945.
[33] S. Saeidi, G. Rentala, T. Rizzuto, T. Hong, N. Johannsen, Y. Zhu, Exploring thermal state in mixed immersive virtual environments, J. Build. Eng. 44 (2021) 102918, https://doi.org/10.1016/j.jobe.2021.102918.
[34] R. Doggett, E.J. Sander, J. Birt, M. Ottley, O. Baumann, Using virtual reality to evaluate the impact of room acoustics on cognitive performance and well-being, Front. Virtual Real. 2 (2021) 1–9, https://doi.org/10.3389/frvir.2021.620503.
[35] I. Muhammad, M. Vorländer, S.J. Schlittmeier, Audio-video virtual reality environments in building acoustics: an exemplary study reproducing performance results and subjective ratings of a laboratory listening experiment, J. Acoust. Soc. Am. 146 (2019) https://doi.org/10.1121/1.5126598, EL310–EL316.
[36] A. Latini, E. Di Giuseppe, M. D'Orazio, C. Di Perna, Exploring the use of immersive virtual reality to assess occupants' productivity and comfort in workplaces: an experimental study on the role of walls colour, Energy Build. 253 (2021) 111508, https://doi.org/10.1016/j.enbuild.2021.111508.
[37] G. Chinazzo, K. Chamilothori, J. Wienold, M. Andersen, Temperature–color interaction: subjective indoor environmental perception and physiological responses in virtual reality, Hum. Factors 63 (2021) 474–502, https://doi.org/10.1177/0018720819892383.
[38] S. Huang, G.W. Scurati, R. Etzi, F. Ferrise, S. Graziosi, L. Chiara Tagliabue, A. Gallace, M. Bordegoni, Effects of Coloured Ambient Light on Perceived Temperature for Energy Efficiency: A Preliminary Study in Virtual Reality, 2019. https://unity3d.com.
[39] J.Y. Jeon, H.I. Jo, B.B. Santika, H. Lee, Crossed effects of audio-visual environment on indoor soundscape perception for pleasant open-plan office environments, Build. Environ. 207 (2022) 108512, https://doi.org/10.1016/j.buildenv.2021.108512.
[40] ISO - International Organization for Standardization, ISO 10551:2019 Ergonomics of the Physical Environment -- Subjective Judgement Scales for Assessing Physical Environments, 2019.
[41] ASHRAE standard, ANSI/ASHRAE standard 55-2004: thermal environmental conditions for human occupancy, Am. Soc. Heating, Refrig. Air-Conditioning Eng. Inc. (2004) 1–34 2004.
[42] ISO - International Organization for Standardization, ISO 12913-1:2014 Acoustics — Soundscape — Part 1: Definition and Conceptual Framework, 2014.
[43] B.S. ITU-R, 1116-3, Methods for the subjective assessment of small impairments in audio systems BS Series Broadcasting service (sound), Int. Telecommun. Union. 3 (2015).
[44] F. Abd-Alhamid, M. Kent, C. Bennett, J. Calautit, Y. Wu, Developing an innovative method for visual perception evaluation in a physical-based virtual environment, Build. Environ. 162 (2019) 106278, https://doi.org/10.1016/j.buildenv.2019.106278.
[45] G. Ozcelik, B. Becerik-Gerber, R. Chugh, Understanding human-building interactions under multimodal discomfort, Build. Environ. 151 (2019) 280–290, https://doi.org/10.1016/j.buildenv.2018.12.046.
[46] F. Salamone, A. Bellazzi, L. Belussi, G. Damato, L. Danza, F. Dell'aquila, M. Ghellere, V. Megale, I. Meroni, W. Vitaletti, Evaluation of the visual stimuli on personal thermal comfort perception in real and virtual environments using machine learning approaches, Sensors 20 (2020), https://doi.org/10.3390/s20061627.
[47] J. Munafo, M. Diedrick, T.A. Stoffregen, The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects, Exp. Brain Res. 235 (2017) 889–901, https://doi.org/10.1007/s00221-016-4846-7.
[48] S. Niu, W. Pan, Y. Zhao, A virtual reality integrated design approach to improving occupancy information integrity for closing the building energy performance gap, Sustain. Cities Soc. 27 (2016) 275–286, https://doi.org/10.1016/j.scs.2016.03.010.
[49] S. Saeidi, C. Chokwitthaya, Y. Zhu, M. Sun, Spatial-temporal event-driven modeling for occupant behavior studies using immersive virtual environments, Autom. ConStruct. 94 (2018) 371–382, https://doi.org/10.1016/j.autcon.2018.07.019.
[50] M. Slater, M. Usoh, A. Steed, Depth of presence in virtual environments, Am. J. Manag. Care 10 (2004) 115–117.
[51] M. Usoh, E. Catena, S. Arman, M. Slater, Using presence questionnaires in virtual reality, Conf. Hum. Factors Comput. Syst. - Proc. (2019) 1–16, https://doi.org/10.1145/3290605.3300590.
[52] H. Akiduki, S. Nishiike, H. Watanabe, K. Matsuoka, T. Kubo, N. Takeda, Visual-vestibular conflict induced by virtual reality in humans, Neurosci, Lettres 340 (2003) 197–200, https://doi.org/10.1016/S0304-3940(03)00098-3.
[53] A.K.T. Ng, L.K.Y. Chan, H.Y.K. Lau, A study of cybersickness and sensory conflict theory using a motion-coupled virtual reality system, Displays 61 (2020), https://doi.org/10.1016/j.displa.2019.08.004.
[54] ASHRAE standard, J. June 2019 (2019).
[55] A. Miyake, N.P. Friedman, M.J. Emerson, A.H. Witzki, A. Howerter, T.D. Wager, The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis, Cognit. Psychol. 41 (2000) 49–100, https://doi.org/10.1006/cogp.1999.0734.
[56] A. Diamond, Executive functions, Annu. Rev. Psychol. 64 (2013) 135–168, https://doi.org/10.1146/annurev-psych-113011-143750.
[57] H.K. Mohajan, Two criteria for good measurements in research: validity and reliability, Ann. Spiru Haret Univ. 17 (2017) 59–82, https://doi.org/10.26458/1746.
[58] F.J. Gravetter, L-A.B Forzano, Research Method for Behavioural Science, 2012.

[59] K. Chamilothori, J. Wienold, M. Andersen, Adequacy of immersive virtual reality for the perception of daylit spaces: comparison of real and virtual environments, LEUKOS - J. Illum. Eng. Soc. North Am. 15 (2019) 203–226, https://doi.org/10.1080/15502724.2017.1404918.
[60] A. Heydarian, B. Becerik-Gerber, Use of immersive virtual environments for occupant behaviour monitoring and data collection, J. Build. Perform. Simul. 10 (2017) 484–498, https://doi.org/10.1080/19401493.2016.1267801.
[61] J.M. Jones, Validity of nutritional screening and assessment tools, Nutrition 20 (2004) 312–317, https://doi.org/10.1016/j.nut.2003.11.005.
[62] F.A. Faul, E. Erdfelder, A.-G. Lang, Buchner, G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, Behav. Res. Methods 35 (2007) 175–191. https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.
[63] T. Goto, J. Toftum, R. De Dear, P.O. Fanger, Thermal sensation and thermophysiological responses to metabolic step-changes, Int. J. Biometeorol. 50 (2006) 323–332, https://doi.org/10.1007/s00484-005-0016-5.
[64] Unity, https://unity.com, 2021. (Accessed 7 May 2021) accessed.
[65] SteamVR. Plugin, https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647, 2021.
[66] M. Wendt, S. Klein, T. Strobach, More than attentional tuning - investigating the mechanisms underlying practice gains and preparation in task switching, Front. Psychol. 8 (2017) 1–9, https://doi.org/10.3389/fpsyg.2017.00682.
[67] N. Unsworth, R.P. Heitz, J.C. Schrock, R.W. Engle, An automated version of the operation span task, Behav. Res. Methods 37 (2005) 498–505, https://doi.org/10.3758/BF03192720.
[68] J.R. Stroop, Studies of interference in serial verbal reactions, J. Exp. Psychol. 18 (1935) 643–662, https://doi.org/10.1037/h0054651.
[69] ISO - International Organization for Standardization, ISO 9920:2007 Ergonomics of the Thermal Environment — Estimation of Thermal Insulation and Water Vapour Resistance of a Clothing Ensemble, 2007.
[70] I. Ajzen, Constructing a theory of planned behavior questionnaire, Biofeedback and Selfregulation 17 (2010) 1–7, https://doi.org/10.1016/0749-5978(91)90020-T.
[71] H.K. Kim, J. Park, Y. Choi, M. Choe, Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment, Appl. Ergon. 69 (2018) 66–73, https://doi.org/10.1016/j.apergo.2017.12.016.
[72] R. Studio, https://www.rstudio.com, 2021. (Accessed 31 May 2021) accessed.
[73] G. Chinazzo, J. Wienold, M. Andersen, Influence of indoor temperature and daylight illuminance on visual perception, Light. Res. Technol. 52 (2020) 350–370, https://doi.org/10.1177/1477153519859609.