sustainability

MDPI

Review

A Review of Subjective Assessments in Virtual Reality for Lighting Research

Michelangelo Scorpio [1,*], Davide Carleo [1], Martina Gargiulo [1], Pilar Chías Navarro [2], Yorgos Spanodimitriou [1], Parinaz Sabet [1], Massimiliano Masullo [1] and Giovanni Ciampi [1]

[1] Department of Architecture and Industrial Design, University of Campania Luigi Vanvitelli, 81031 Aversa, Italy; davide.carleo@unicampania.it (D.C.); martina.gargiulo@unicampania.it (M.G.); yorgos.spanodimitriou@unicampania.it (Y.S.); parinaz.sabet@unicampania.it (P.S.); massimiliano.masullo@unicampania.it (M.M.); giovanni.ciampi@unicampania.it (G.C.)

[2] Escuela de Arquitectura, Universidad de Alcalá, 28801 Alcalá de Henares, Spain; pilar.chias@uah.es

* Correspondence: michelangelo.scorpio@unicampania.it

Abstract: Immersive virtual reality allows showing people virtual environments with high levels of presence, realism, and "feeling of being", as if they were in the real world. With this aim, virtual environments must provide proper light distributions and elicit sensations similar to those seen in the real world. So far, experiments with human subjects are the most effective way to evaluate the accuracy of virtual reality in reproducing real spaces. This paper investigates the role of subjective assessments in lighting research using virtual reality. According to the review results, the investigations aimed at using immersive virtual reality for lighting can mainly be divided into three groups: (i) comparison between virtual and physical environments, (ii) analysis of different lighting scenarios, and (iii) investigation of users' interaction with the virtual model. On the one hand, the results show that immersive virtual reality is a useful tool for research and design in lighting. On the other hand, they highlight the limitations that still need to be overcome. Finally, the main findings and gaps concerning the subjective assessment were listed.

Keywords: virtual reality; subjective surveys; lighting; daylighting; human preferences; questionnaires; HDR; Radiance; tone mapping

check for updates

Citation: Scorpio, M.; Carleo, D.; Gargiulo, M.; Navarro, P.C.; Spanodimitriou, Y.; Sabet, P.; Masullo, M.; Ciampi, G. A Review of Subjective Assessments in Virtual Reality for Lighting Research. *Sustainability* **2023**, *15*, 7491. https://doi.org/10.3390/su15097491

Academic Editor: Ramchandra Pode

Received: 31 March 2023
Revised: 27 April 2023
Accepted: 1 May 2023
Published: 3 May 2023

1. Introduction

The changes in social and working habits and needs led people to spend more and more time inside buildings, making it mandatory to guarantee human-centered environments, especially considering that indoor conditions strongly affect users' satisfaction, health, and performance. Among the parameters influencing indoor quality, lighting and daylight can be regarded as the two most affecting people's moods. Much research has underlined the strong impact of light conditions on humans' well-being, task performance, and satisfaction [1–4]. In this scenario, investigation methods that allow considering both objective and subjective factors, placing human behavior at the core of the design process, are preferred [5–9]. This has prompted many researchers to investigate the effects of light on people, often using full-scale test rooms or living labs [10–12]. The use of physical spaces ensures the faithful reproduction of stimuli that people are used to perceiving and recognizing. Despite this, field investigations are costly and time-consuming [13]. For this reason, alternative, cheaper, less time-consuming, but equally reliable tools have been proposed in recent years.

In this scenario, immersive virtual reality (IVR) has been considered the most powerful tool offering immersive virtual spaces, allowing the end-users to have the perception of the first person in the scenes. As a result, IVR has sparked much interest in the scientific community for lighting design as well. Nevertheless, before IVR can be fully used for lighting design and research, virtual environments have to ensure a suitable reproduction of the real world. Validating virtual reality in its lighting applications requires achieving satisfying outcomes in the simulation of a visual world that is as accurate as the real one. However, such validation could be far-reaching, considering the aspects that still need to be thoroughly investigated, as well as limitations in the hardware and software for IVR [14–17]. Despite this, more and more researchers have highlighted the possibility of utilizing IVR for lighting design and research if the virtual model shown to participants provides a "suitable reproduction" of reality.

1.1. Virtual Reality for Lighting

According to the current literature, a virtual model for lighting can mainly be considered a suitable reproduction of the physical one if it is built: (i) to simulate light distribution in a photometrically correct way, (ii) to provoke perceptions that people experience in physical environments, and/or (iii) to ensure a plausible visual experience.

The choice to emphasize one or more aspects mainly depends on the research aims. For example, suppose the research object is to investigate the lighting influence on the impression of the perceived space. In that case, it is necessary to ensure that the light distribution is simulated correctly from the photometrical point of view [18].

With this aim, different methodologies have been proposed to guarantee the exhibition of virtual models by IVR with a correct light distribution to people. A first approach is used in [18], where the virtual environment in the game engine was realized through rendered images generated with validated physically based lighting simulation tools. In the second approach, calibrated high-dynamic range (HDR) pictures taken by a camera equipped with a fisheye lens in physical spaces were used [19,20] to create 360° HDR panoramas to be uploaded to a game engine and shown to users. The third approach is based on constructing the virtual environments and accurately setting the parameters related to the light sources directly in the game engine [21]. The verification of the light distribution can occur through (i) further physically based lighting simulation software [22], (ii) measurements in the real space [23,24], (iii) photometric data of real luminaires after a suitable calibration in the game engine [21], or (iv) subjective assessments [19]. In general, the efforts are focused on ensuring comparable luminance distributions between the virtual and real environments, as well as similar illuminance values at users' eyes between the Head-Mounted Display (HMD) and real scenes.

At the same time, the ability of immersive virtual reality to induce human reactions and experiences such as those felt in the real ones was investigated. Those aspects have been analyzed mainly by questionnaires to explore different perceptual aspects and visual tasks performed in real and/or virtual environments, allowing the evaluation of the light environment accuracy from users' points of view. In light of the above, subjective assessments play a crucial role in evaluating the quality of virtual environments, as well as in obtaining information about people's preferences, perceptions, and moods with lighting through IVR.

According to the literature (Figure 1), IVR has been used for lighting and daylighting investigations mainly to: (i) compare virtual and physical environments (VPC), (ii) display different light scenarios (DSC) or (iii) evaluate the humans' interaction with systems to investigate their lighting preferences (HII). Concerning the experimental design, the literature suggests two common factor analyses: (i) within-subjects design, where subjects are shown all scenarios; and (ii) between-subjects design, where participants are divided into different groups, and each group is shown a limited number of the available scenarios or participants are divided on the basis of some investigation variable.

A most recently investigated aspect is related to virtual reality's capability to reproduce colors correctly, since color is an essential component of people's visual sense. Indeed, people may quickly identify and divide objects into several categories, based on their color. To the authors' knowledge, only a few papers are focused on this topic. In particular, Rodríguez et al. [25] assessed the spectral properties of the display used for HMD in terms of luminance and chromaticity and the effects of the software used to control the HMD. A precise color calibration of the HMD in terms of luminance and chromaticity is needed to regulate the given stimuli properly. The vast field of vision of the HMD, the scene rendering software, or how the model is transferred from the software to the HMD may change how colored items look, implying a modification of the environments and objects perception.

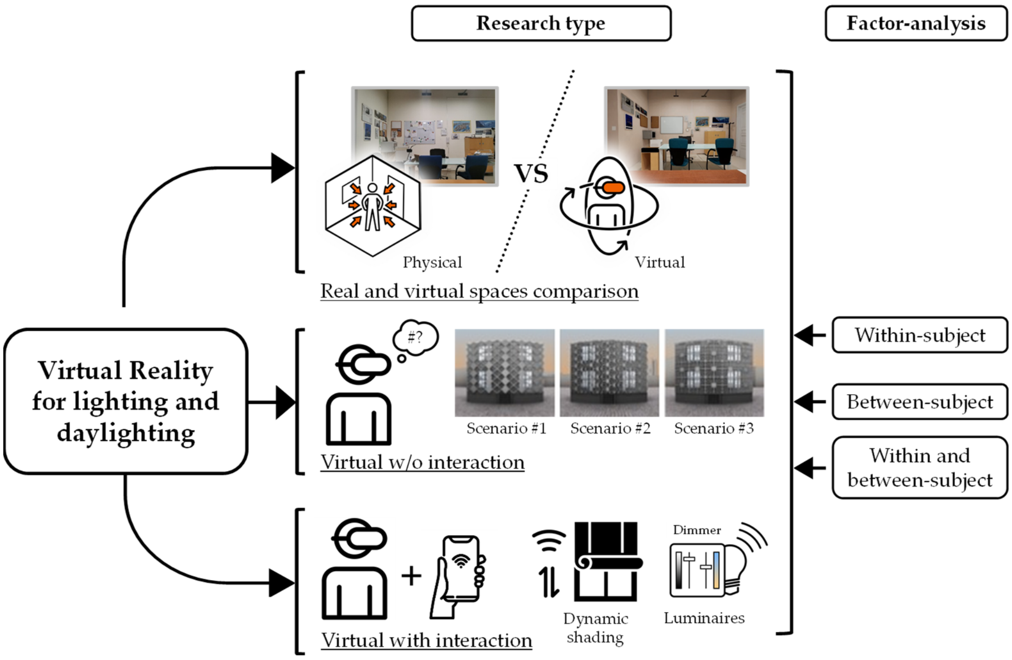

Figure 1. Scheme of investigation methods and factor analysis for lighting with IVR.

1.2. Purpose of the Research

The review underlines the prominent role of people's subjective responses and perceptions in understanding the potentiality of IVR for lighting and evaluating the accuracy of virtual environments in reproducing physical spaces.

This paper presents a literature review of studies where IVR is used to carry out lighting and daylighting investigations on people's lighting preferences and performances through subjective assessments. A specific focus on environmental and psychological factors, as well as tests performed, tools, and surveys used, are given. In particular, the literature review is focused on three different aspects associated with the usage of virtual reality to: (i) compare virtual and real light environments; (ii) compare different light scenarios; and (iii) evaluate the people's interaction with lighting and/or shading systems. In addition, the methods used for the statistical analysis and comparison of results are analyzed. Finally, the limitations and perspectives of using IVR for lighting were listed.

2. Papers Collection Methodology

The methodology followed to collect papers consists of three steps: (i) recognition of databases for publications in international journals, (ii) identification of databases for different types of publications, and (iii) definition of keywords.

The Scopus and Web of Science databases were used for papers in international journals, while PubMed, Google Scholar, and Research Gate were used for other types of contributions. Regarding the keyword definition, no periods were specified during paper collection, and three sets of words were taken into account:

- Lighting-related aspects, using the words: lighting, visual, daylighting, daylight, or view;
- Target keywords, using the words: virtual reality, virtual environment, immersive virtual reality, and immersive virtual environment;

- Human perspective, using the words: perception, subjective assessment, subjective response, and human psychological response.

The application of the above-described methodology returned a total of 72 papers; the last access to databases was in March 2023. The stored papers were filtered by evaluating the paper title and keywords (to remove duplicates and select papers perfectly matching the research topic), as well as reading the full text (to establish the eligibility of the remaining papers). At the end of the filtering process, 24 papers were included in the review and analyzed in detail. Figure 2 shows the preferred sources and the number of papers published in each, proving *Building and Environment*, *Lighting Research & Technology*, *Automation in Construction*, *LEUKOS*, and *Sustainability* as the most considered journals to publish contributions on the considered topic.

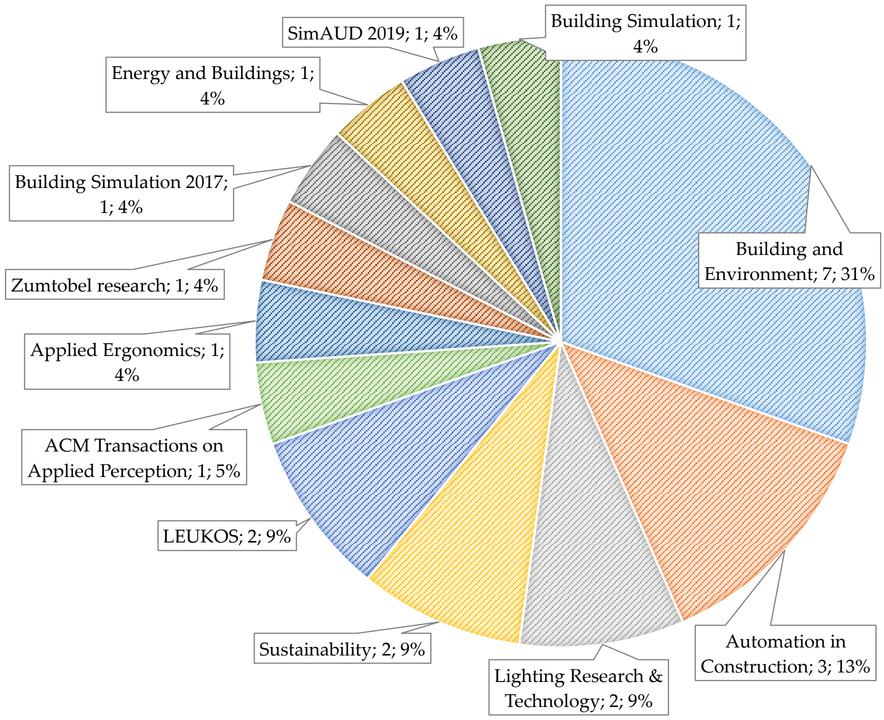

Figure 2. Most significant sources among the papers analyzed in the literature review.

3. Virtual Reality for Comparison between Virtual and Real Light Environments

3.1. Studies Based on Within-Subject Design

In studies comparing virtual and physical environments, a virtual model of the real space is built following different approaches, and participants are asked to judge one or more light scenarios in one or both environments. This research seeks to understand the potentiality of different methods to create virtual environments reproducing light distribution in various types of physical spaces.

Chamilothori et al. [18] offered a novel method for creating virtual reality scenes based on rendered images that accurately reproduce indoor daylight distribution from a photometrical perspective. Subjective assessments were used to evaluate the ability of the proposed method to ensure a precise perception of the environment, the sense of presence, physical symptoms before and after the HMD usage, as well as the effects of the presentation order. Twenty-nine participants were asked to see and judge physical and virtual daylit environments during experiments. The subjects' number was determined through a prior analysis, considering a statistical power of 0.80 and an effect size of 0.56. Since the one-sample Kolmogorov–Smirnov test [26] underlined the non-normality of data and the paired responses, the collected data were analyzed through Wilcoxon matched-pairs signed rank test (with $\alpha = 0.05$).

Another methodology to reproduce indoor artificial light distribution in virtual reality (VR) was proposed in [19]. Starting from images taken using a camera with a fisheye lens and luminance values measured in the physical environment, calibrated HDRs were obtained. Then, the calibrated HDR images were converted into Low Dynamic Range (LDRs) and stitched to create stereoscopic 360° images to be shown through the HMD to participants. The reliability of the proposed method in reproducing artificial light distribution, as well as people's light perception and performance in carrying out a task, were evaluated by asking them to perform two tasks: the achromatic characters contrast test and the Stroop test. In addition, a color discrimination task was considered to investigate the effects of resolution in VR and the conversion process from HDRs to LDRs. The two environments were presented to twenty volunteers and asked to perform visual tasks. Information about vision problems, demographic data, physical symptoms, and the Simulator Sickness Questionnaire (SSQ) [27] was requested. After completing tasks in each space, volunteers answered questionnaires about the perception of visual-quality, light appearance and room's impression, stress, as well as positive and negative affect schedule (PANAS) [28]. After the VR environment test, they also answered the SSQ. Finally, subjects completed a questionnaire on perceived presence. Different statistical analyses were used as a function of the distribution of data. Shapiro–Wilk and Kolmogorov–Smirnov [26] statistical tests, Levine's test, as well as Q-Q plots, were used to verify the homogeneity and normality of variance. When the assumptions were satisfied, the mean and standard deviation values were considered for data elaboration (time for characters contrast and sense of presence). Otherwise, the non-parametric Wilcoxon signed-rank test was used (color naming task, perception of visual-quality, light appearance and room's impression, perception of stress, PANAS, and SSQ).

The ability of different display media (photo, video, and VR) to create the feeling experienced by people under physical lighting conditions was explored in [13]. The virtual models of the room were made by acquiring a stereoscopic video as well as a conventional video and photo, upon varying light scenarios. The stereoscopic video was shown through a smartphone in the VR helmet. Forty subjects were asked to see each lighting environment and complete a questionnaire about the presentation, perceptual, emotional, and overall satisfaction attributes. The subjects' scores were analyzed by comparing the mean values and standard deviation for each item to evaluate the reliability of different media in reproducing the physical space. In addition, the correlation of Pearson, factor analysis, and dimension reduction analyses were calculated for presentation items to assess their correlation and the weight of items, as well as non-parametric tests (post hoc analyses and Friedman's test) were used to examine lighting and presentation attributes as well as compare real and virtual spaces. Finally, ANOVA analysis was computed for the overall satisfaction item.

The first experiment performed by Hong et al. [29] aimed to investigate subjects' reactions to a windowed office with its digital twin. The digital twin of the space was realized in SketchUp, then passed in 3ds Max and, finally, in Unreal Engine. The parameters characterizing how the software computes the light distribution of the indoor light sources (such as total energy emitted by the light source, location, number, length, and radius) were established according to the real artificial lighting system. The building's latitude and longitude, as well as the sun's azimuth and altitude, were used to simulate daylight. After a preliminary survey about personal information, the fifty volunteers were invited to see the physical and virtual environment, and complete questionnaires on satisfaction (to analyze the sense of visual comfort, privacy, inner space, and openness) for each of them. Finally, the volunteers accomplish the questionnaire on general and spatial presence, involvement, and realism in the virtual environment. The sense of presence was evaluated by calculating the mean values and standard deviation and comparing the values with those from other research. The items of the satisfaction survey in the physical and virtual environment were analyzed through a paired sample *t*-test.

The usability of stereoscopic images to assess the effects of daylight on the perception of indoor architectural quality was explored in [30]. The evaluation was carried out using two physical rooms (with room surfaces white or black painted and three different window size) and virtual models of each room. Virtual models were obtained starting from pictures taken inside rooms with a conventional photographic camera and shown by means of two full high-definition projectors. Twenty-six participants were divided into three groups; each group started the experiment by examining one of the two physical or virtual spaces. The precision of stereoscopic images for daylighting research was evaluated through a questionnaire on excitement, complexity, order, coherence, openness, legibility, spaciousness, and pleasantness; no numerical scale was used in the survey. People's responses were shifted into two groups according to the room's color and analyzed using the Bland–Altman method [31].

3.2. Studies Based on Between-Subject Design

Two studies [20,32] were based on the between-subject design. Rockcastle et al. [20] investigated how people perceive an illuminated space in VR, using the between-subject design to diminish the order bias effect and the session period. HDR pictures of physical space were taken under different lighting conditions. Before starting the experiment, participants were requested to provide demographic information (gender, age, and visual problems), then see the physical or virtual environment, and answer questions about pleasantness, brightness, evenness, visual comfort, glare perception, and contrast for each lighting scenario. In addition, the Landolt Rings test was presented to subjects, and the reading time was noted. A *t*-test was used to evaluate the effects of vision problems and gender on the rating of items and type of environment. Items answers were compared considering their distribution, as well as computing the mean values and frequency distribution. The normality of responses was evaluated through the Shapiro–Wilk test. A non-parametric Mann–Whitney test was considered to compare the items and reading speed for each lighting scenario.

The perception of daylit space simulated in the game engine was examined in [32]. The virtual model of a real space with different areas was modeled in SketchUp, passed to 3ds and imported in Unreal Engine 4. A directional light, for which the UE4 parameter named "light intensity" was set to 110,000 lx and correlated color temperature (CCT) to 5500 K, was used to simulate the sun and daylight. The real space was shown to the first group of 18 participants, while the virtual one was displayed to the second group of 18 people. The participants walked into the real or virtual scene, taking pictures of different points of the scene and rating the perceived brightness. In addition, the participants that observed the virtual scene responded to a questionnaire on the sense of presence. The participants' number was defined by performing a prior analysis with G*Power software, considering a statistical power of 0.80, large effects, and an effect size of 0.97. For data analysis, nine areas were identified in the space; the number of pictures taken and the rated perceived brightness for each area and space type were calculated and compared to evaluate the space perception. This information was condensed and displayed through Perceptual Light Maps. Finally, the participants' answers to the sense of presence questionnaire were compared in terms of the response distribution and median values.

3.3. Research Using Both within and Between-Subject Design

Both within-subject and between-subject designs were used in [33,34]. Heydarian et al. [33] evaluated the adequacy of immersive virtual reality (IVR) in reproducing real spaces, evaluating participants' visual task performances and responses to a questionnaire. Reading (considering speed and comprehension) and identification (identification and number of colored books in a bookcase) tasks were considered for users' performances. In addition, a questionnaire on the sense of presence and immersion was used to assess space and lighting perception. The virtual model of the real space was created in Revit 2013, modified in 3ds Max, and then imported into Architecture Interactive so that users could observe it. The illuminance levels acquired in the physical room were used to set the lighting levels in the virtual scenes. The 120 participants were randomly assigned to one of the four lighting scenarios and asked to perform the visual task and respond to the questionnaire. Results about the identification of books and reading speed were compared by performing the t-test, as well as in terms of average values and standard deviation. The average values and standard deviation were considered to compare the rating on the questionnaire's items.

The ability to reproduce users' experiences in physical spaces through different display media was investigated in [34]. Virtual environments of a real mock-up were realized following different methods: taking a traditional photo, acquiring a 360° panorama, and building a VR model into Unity3D; then shown to participants by HMD. A prior analysis suggests a number of subjects equal to 25 subjects per each environment (100 in total). The participants were first asked to examine and walk the real and VR spaces, and then to answer verbally a questionnaire on an evaluation of the environments, emotion, and presence. In addition to subjective measurements, physiological parameters (electrodermal function and heart rate) were also acquired. Responses on environment and emotion were analyzed by means of the non-parametric Mann–Whitney U test to underline differences and relationships among environments; data about presence were analyzed using the Slater, Usoh, and Steed method.

3.4. Main Findings

Table 1 summarizes the key design considerations, along with their labels and citations, from papers focused on the use of IVR between virtual and physical space. Fixture elements and design themes were separated into design parameters. The literature review suggests:

- Attributes linked to the sense of presence [18,19,29,32,34] and perception [13,18,19] are the most investigated;
- Investigations are generally carried out using the within-subject design, a 5-point scale questionnaire, and involving young people (average age around 30 years);
- Indoor spaces (especially offices) are usually considered when the virtual and real environments are compared;
- Results are primarily compared in terms of mean and standard deviation values.

Table 1. Main design parameters and measures for comparison between virtual and real light environments.

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
Factors investigated	[18]	Perceptual impressions Physical symptoms Reported presence	[29]	General presence Spatial presence Involvement Experienced realism Sense of visual comfort Sense of privacy Sense of inner space Sense of openness	[32]	Reported scene Sense of presence Perception of brightness
	[19]	Visual-quality perception Lighting appearance perception Perception of room's impressions Reported presence Simulator sickness questionnaire Stress Positive and negative affects	[30]	Unpleasant—Pleasant Dull—Exciting Chaotic—Ordered Simple—Complex Illegible—Legible Incoherent—Coherent Tight—Spacious Closed—Open Spatially Undefined—Defined	[33]	Focus Immersion and involvement Gaming Distraction factors Control factors and IVR interaction

Table 1. *Cont.*

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
	[13]	Presentation ability Perceptual attributes Emotional attributes Overall satisfaction	[20]	Visual comfort Pleasantness Evenness Brightness Contrast Glare perception Reading time of Landolt Rings test	[34]	Environmental evaluation Emotion Sense of presence
Number of subjects	[18]	29	[29]	50	[32]	36 (divided into two equals groups)
	[19]	20	[30]	26	[33]	112
	[13]	40	[20]	30 (physical space) 23 (virtual space)	[34]	100 (divided into four equals groups)
Average age {age range} year	[18]	28 {-}	[29]	25.1 {21–34}	[32]	- {21–25 (92%) 26–30 (8%)}
	[19]	26 {-}	[30]	32.7 {24–62}	[33]	21 {18–33}
	[13]	- {24 subjects aged 18–30 8 subjects aged 31–50 8 subjects aged over 51}	[20]	- {18–33 (80%) 34–58 (20%)}	[34]	32.7 {23–51}
Type of factor-analysis	[18]	Within-subjects	[29]	Within-subjects	[32]	Between-subjects
	[19]	Within-subjects	[30]	Between-subjects	[33]	Within-subjects Between-subjects
	[13]	Within-subjects	[20]	Between-subjects	[34]	Within-subjects Between-subjects
Space type	[18]	Office	[29]	Office	[32]	Multipurpose space with nine areas
	[19]	Office	[30]	Bedroom	[33]	Office
	[13]	Bedroom	[20]	Studio space	[34]	Shopping environment
Statistical power	[18]	0.8	[32]	0.8		
Effect size	[18]	0.56	[32]	0.97	[33]	>0.99
Questionnaire scale	[18]	5-point unipolar, with verbal anchors at the endpoints	[29]	7-point unipolar	[32]	5-point unipolar for the sense of presence 4-point unipolar for perceived brightness
	[19]	5-point unipolar, with verbal anchors at the endpoints	[30]	No numerical scale (a 7-point bipolar scale was considered in the elaboration phase)	[33]	3-point, 5-point, and 7-point unipolar
	[13]	6-point unipolar for Presentation-ability 7-point bipolar for the other	[20]	5-point unipolar	[34]	7-point unipolar
Data normality test	[18]	One-sample Kolmogorov–Smirnov	[19]	Shapiro–Wilk and Kolmogorov–Smirnov statistical tests, Levine's test, as well as Q-Q plots	[20]	Shapiro–Wilk test

Table 1. *Cont.*

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
Statistical test	[18]	Wilcoxon matched-pairs signed-ranks Cohen's d effect size	[29]	Mean values and standard deviation for the presence questionnaire Paired sample *t*-test for satisfaction questionnaire	[32]	Responses distribution and median values for the sense of presence questionnaire Number of pictures, rated perceived brightness and Perceptual Light Maps for space perception
	[19]	Mean values and standard deviation for normal and homogenous distribution Wilcoxon Signed-Rank Test for non-normal and homogeneous distribution	[30]	Bland–Altman method	[33]	*t*-test for identification of books and reading speed Mean and standard deviation values for visual tasks and items
	[13]	Mean values and standard deviation for all questionnaires items Correlation of Pearson, factor analysis, and dimension reduction analysis for presentation items Non-parametric Friedman tests and post hoc analyses for presentation and lighting items ANOVA for overall satisfaction item	[20]	*t*-test Answers rating distribution Answers frequency distribution Non-parametric Mann–Whitney test	[34]	Non-parametric Mann–Whitney U test for responses on environment and emotion Slater, Usoh and Steed method for the sense of presence

4. Virtual Reality for Comparison among Different Light Scenarios

4.1. Studies Based on Within-Subject Design

This study [35] describes an experiment utilizing a virtual reality headset to assess rendered daylit architectural scenes. The authors compared subjective perception ratings of architectural renderings to image-based measures of visual attraction by changing sky conditions and view directions from a fixed viewpoint. Virtual reality enabled head tracking data to reveal how participants viewed immersive situations. Eight architectural scenes were designed in Rhinoceros and rendered in Radiance to provide 360° HDR scenes across different periods of the year. Then, 180° HDR Radiance renderings from various view orientations were tone-mapped and utilized to create immersive virtual scenes of the daylit buildings. Oculus Rift CV1 and an Acer Predator 17-X laptop supported the VR headgear. The subjects were 18–50 years old, with a mean age of 29. Verbal surveys and head tracking were used to acquire subjective and objective data for the qualitative lighting study. Pearson's Correlation Coefficient analysis was used to compare subject responses (rating scales for each scene) to quantitative algorithms (contrast, visual interest, and brightness). Moreover, this research introduced a preliminary composite rating, "PIE," from a selection of attributes in the experiment. The match between subjective ratings and image-based algorithms designed to forecast them proves that immersive scenes may predict pleasant, interesting, and exciting sensations and that view direction affects such predictions.

Responses to energy-saving retrofit solutions in office buildings, reporting on the view assessment, and emotional reactions to ETFE double-skin facades were examined in [36]. Virtual reality and physics-based imaging techniques were used to assess the user experience of a window view in an office space with a pneumatic ETFE cushion serving as a second building skin. Evaluation criteria included view clarity, amount of view, view appearance, as well as emotional states of pleasure, arousal, and domination. Three double-skin facade scenarios with different ETFE cushions—clear, fritted, and switchable—were examined and compared to the original single-skin façade with double-glazed windows. The office space's physical and luminous conditions were reproduced in a virtual environment using a proven physically based imaging technique and presented to 22 volunteers using a virtual reality headset. Participants filled out a questionnaire about their views, perception, and emotions while in the virtual environment. The virtual environment was created from the physical one according to the procedure described in [19]. Shapiro–Wilk test was used to determine the type of data distribution, while Friedman's ANOVA was considered to state statistically significant differences. Questionnaire responses were analyzed with the non-parametric Wilcoxon signed-rank test, and Bonferroni corrections were considered for pair-wise comparisons. Finally, the Ferguson [37] classification was used to evaluate effect sizes.

Subjective and physiological evaluations were compared in [38] to investigate the perception of the view from various observing sites. This project examined visual perception using a physically based virtual 360-degree environment. Therefore, three research objectives were made: (i) making a replica in virtual reality based on the physical and lighting conditions at three viewing locations: close, middle, and far from the window in an office room, (ii) obtaining subjective responses on view quality parameters, such as view restorative ability, view content and size preferences, view valance/arousal, self-reported stress, and positive and negative effects, and (iii) measuring physiological markers, such as skin conductance (SC), heart rate variability (HRV), and heart rate (HR). The second and third objectives were used to analyze the differences in visual perception at various viewing points within a virtual office setting. Thirty-two volunteers were exposed to each of the three conditions, while subjective and physiological assessments were gathered. View perception was evaluated based on four factors: view vital capacity, view content, size preferences, and view valence/arousal. Two questions relating to the visual attraction and complexity of daylight were also utilized. Recovery from stress was studied using the positive and negative affect schedule (PANAS) [28]. After answering questions about demographic and sickness information, participants were asked to look at the virtual model (built according to the procedure in [19]), respond to the view questions, and perform the Stroop test, for each condition. Since different types of data were gathered, different statistical analyses were used: (i) the z-scores method for physiological data; while (ii) view, stress, and PANAS responses were analyzed through the ANOVA test (non-parametric Friedman's ANOVA test for not normal data distribution). The Kolmogorov–Smirnov [26] and Shapiro–Wilks [39] tests were considered for evaluating data normality. The Huynh-Feldt [40] and Bonferroni methods [41] were applied to correct data normality or experimental-wise error. The results demonstrated significant changes in subjective parameters and skin conductance based on the distance from the window.

A second experiment performed by Hong et al. [29] aimed to investigate the impact of four different window sizes on occupants' satisfaction, showing virtual environments built into Unreal Engine. As for the first experiment, people were asked to answer the satisfaction questionnaire. Repeated measures ANOVA was used to analyze responses, while the sphericity of data was tested by Mauchly's method.

In [42], the virtual model of urban space was used to investigate the psychological effects of different lighting scenarios through different alternative representation techniques, namely a Liquid Crystal Display (LCD) screen and HMD. Nineteen lighting scenarios were realized, upon varying lighting levels of the square, facades, and urban furniture. Twenty-one volunteers were asked to observe and judge each scene by means of ten adjective pairs. At the end of the test, an additional 5-point scale questionnaire was used for the investigation of virtual reality and immersion experience, as well as motion sickness. All the data were elaborated, computing the mean and standard deviation values.

An urban park in Aversa was modeled into Unreal Engine to inspect the impact of brightness and CCT on users' perception of lighting quality, fixation, and eyes' movement [43]. Nine lighting scenarios were obtained, combining three illuminance levels and CCT values. Twenty-six volunteers were asked to see each scenario and fill out the POLQ questionnaire [44]. In addition, participants' pupil diameter and gaze movement were achieved with the HMD-integrated eye-tracking system. A prior analysis with G*Power software confirmed that the recruited people were enough to ensure a test power greater than 0.95 with an effect size equal to 0.25. The sphericity of data was checked and Greenhouse–Geisser correction was employed. Firstly, the influence of each POLQ item on participants' preferences was evaluated by applying the two-way repeated measures ANOVA. Then, the mean and standard deviation values of two indices related to perceived strength and comfort quality were computed for each scenario. In addition to subjective assessment, physiologic parameters such as the maps of gaze, the diameter of pupils, and the IPA index [45] were also evaluated.

The link between participants' satisfaction with brightness and other major perceptual aspects of the scene was investigated in [46]. A hundred volunteers were immersed in a virtual office environment. The brightness level in all virtual scenarios remained constant, while the office shading system's design pattern, rendering materials, and furniture were altered to evaluate how different elements affect participants' pleasure with brightness. Virtual models of five versions of a typical workplace with a large south-facing window were obtained starting from images rendered in Radiance, combined into a 360° image, and shown through the Oculus Go headset. A verbal 11-point rating scale questionnaire was used to find out the participants' satisfaction with brightness, pleasantness, interest, complexity, and pleasure with the lookout for each scene. For statistical analysis, a linear mixed effects model was employed to account for the repeated measures design in which each participant was asked to evaluate many photographs. The linear mixed effects model describes the conditional correlations between brightness satisfaction and the five other perceptual qualities, including unknown participant attributes such as positivity. In addition, a composite index of satisfaction with selected features was created in R by averaging the results for pleasant, interesting, satisfaction with access to an outdoor view, and pleasure with ambiance.

4.2. Research Using Both within and Between-Subject Design

Moscoso et al. [47] investigated regional differences in the perception of spaces following a mixed approach: within-subjects design was used to investigate window sizes and space size, while the influence of sky type, spatial contexts, and the location was evaluated with between-subject design. The study was conducted in Norway, Switzerland, and Greece, using virtual reality to reproduce the same experiment in multiple locations. A total of 406 participants between 18–50 were surveyed, and their opinions on eight different spatial qualities were compiled using an 11-point Likert scale. The results showed that the participants' answers were dependent on where they lived. There were considerable differences in how pleasant and calm people saw the space, indicating even small differences in latitude. Regarding the methodology, the first scene was a blank screen with the logos of the two academic institutions participating. This was followed by a monochromatic scene for a minimum of 15 s. This scene corresponded to the mean RGB value of all the stimuli that were given to the participants in the experiment sequence, which was as follows: adaptation scenario, stimulus presentation, and verbal questionnaire. The results indicate that areas with particular fenestration characteristics might not induce the same response at various latitudes in Europe. Data were analyzed through mean and standard deviation values, as well as the Linear Mixed Model (LMM) [48,49] and pair-wise comparisons. For the study, a power of 0.80 was calculated through G*Power. In addition, Sidak [50] made adjustments to compare factors with significant effects, and plots of data to contrast ratings of stimuli were used.

The effects of urban park light sources, characterized by different intensities and CCT, on emotion, motivation of exploring, and feeling of safety were investigated by Masullo et al. [51]. The virtual model of the urban park was recreated in Unreal Engine, and combining three illuminance and CCT values, nine lighting scenarios were prepared and presented to 36 volunteers utilizing the HMD HTC Vive Pro Eye. A prior analysis confirms a test power of 0.8 and an effect size of 0.25 with the number of subjects considered. Before the test, PANAS questionnaire [52] was administered to collect volunteers' moods. For each scenario, the participants were asked to see the virtual environment and rate the scenario on a 9-point scale. In the first step, two repeated-measure ANOVAs were applied to data to explore the within-subjects variable, while in the second step, a between-subjects design was used, dividing volunteers into two groups according to the results on mood, motivation, and safety.

Chamilothori et al. [53] examined the influence of facade design and related lighting patterns on occupant subjective perception and physiological reactions using an innovative experimental approach that combines virtual office spaces obtained from physically based rendered images displayed in virtual reality with wearable biometric equipment. A total of 72 participants took part in a study that observed and compared three facade configurations with similar aspect ratios and varied space use situations (social or working environment). All of the variations of the facade were performed on an interior scenario with a clear sky and direct daylight access. After each exploration, the participants were asked to answer a questionnaire on interest, pleasantness, and excitement of the space verbally; at the same time, skin conductance and heart rate were recorded to evaluate the physiological reaction. The Kolmogorov–Smirnov test [26] was considered to verify the distribution normality of subjective responses. The influence of the geometry was evaluated according to the within-subject design, while results were gathered with Friedman's one-way ANOVA [54]. In addition, the Wilcoxon signed-ranks matched-pairs test was adopted for post hoc tests. The effects of context were evaluated using the between-subject design applying the Wilcoxon rank-sum test. Finally, Spearman's correlation was used to investigate the link between physiological and subjective responses.

The effects of different skylight solutions for a heritage building were investigated through IVR by Marzouk et al. [55]. The virtual models of the space with three different skylights were modeled by means of Rhino and Grasshopper, and then displayed in the HMD utilizing the application VR-Prospect. For each scenario, the 48 participants were asked to walk in the space, take off the HMD and complete a survey on their space perception. The data validity was checked with the Mann–Whitney test, grouping participants by gender and occupation. Regarding the data analysis, the investigation employed different types of statistical tests. In particular, it used: (i) mean and standard deviation values to quantify subjective responses to each attribute for each skylight solution, (ii) Wilcoxon signed-rank test to evaluate differences among attributes for the different skylight solutions, (iii) Kruskal–Wallis test [56] to compare attribute ratings within the three design solutions, (iv) Kendall's Tau correlation [57] to assess links among attributes within the three design solutions, and (v) Friedman test to contrast the most statistically significant attributes among the three design solutions.

4.3. Main Findings

Table 2 lists the key design considerations, along with their labels and citations, from papers focused on using IVR to compare different lighting scenarios. Fixture elements and design themes were separated into design parameters. The literature review suggests:

- Attributes linked to the interest [35,38,46,47,53], pleasantness [35,46,47,53] and brightness [35,46,47] are the most analyzed;
- Investigations are generally carried out using the within-subject design, questionnaire with different scales and involving young people (average age around 30 years);
- Usually, indoor office spaces (only three investigated outdoor environments) are considered when subjects are asked to compare different light scenarios;
- Results are mostly compared through repeated measures ANOVA or Friedman's ANOVA.

Table 2. Main design parameters and measures for comparison among different lighting scenarios.

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
Factors investigated	[29]	General presence Spatial presence Involvement Experienced realism Sense of visual comfort Sense of privacy Sense of inner space Sense of openness	[35]	Pleasantness Interesting Exciting Contrast Visual interest Brightness	[38]	View restorative ability View content View size View valence/arousal View interest and complexity
	[42]	Effect of light Virtual reality experience Immersion experience Motion sickness	[47]	Pleasantness Calmness Interest Excitement Complexity Spaciousness Amount of view Brightness	[46]	Pleasantness Interest Complexity Satisfaction with brightness Satisfaction with view out Satisfaction with ambiance
	[51]	Calmness Happiness Energy Tiredness Nervousness Sadness Feeling of Safety Motivation	[36]	Engagement View clarity View satisfaction View amount View appearance View perception View vividness Spatial pleasure Spatial arousal Spatial dominance	[53]	Pleasantness Interest Excitement
	[43]	Clear—Drab Strong—Weak Unfocused—Focused Subdued—Brilliant Dark—Light Mild—Sharp Hard—Soft Warm—Cool Glaring—Shaded Natural—Unnatural	[55]	Pleasantness Contrasting Brightness Distribution Visual Comfort Satisfaction		
Number of subjects	[29]	50	[35]	65, with a minimum of 15 subjects per space	[38]	32
	[42]	21	[47]	150 (Norway) 118 (Switzerland) 138 (Greece)	[46]	100
	[51]	36	[36]	22	[53]	71
	[43]	26	[55]	48		
Average age {age range} year	[29]	25.1 {21–34}	[35]	29 {18–50}	[38]	28 {-}
	[42]	- {21–53}	[47]	- {20–49 20–50 19–44}	[46]	- {>18}
	[51]	28.7 {20–57}	[36]	29 {-}	[53]	25.9 {18–32}
	[43]	29.7 {-}	[55]	- {>19}		
Type of factor-analyses	[29]	Within-subjects	[35]	Within-subjects	[38]	Within-subjects
	[42]	Within-subjects	[47]	Within-subjects Between-subjects	[46]	Within-subject
	[51]	Within-subjects Between-subjects	[36]	Within-subjects	[53]	Within-subjects Between-subjects

Table 2. *Cont.*

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
	[43]	Within-subjects	[55]	Within-subjects Between-subjects		
Space type	[29]	Office	[35]	Office	[38]	Office
	[42]	Urban space	[47]	Office	[46]	Office
	[51]	Urban Park	[36]	Office	[53]	Office
	[43]	Urban Park	[55]	Heritage building		
Statistical power	[51]	0.8	[43]	0.95	[47]	0.8
Effect size	[51]	0.25	[43]	0.25	[47]	Small
	[55]	0.28				
Questionnaire scale	[29]	7-point unipolar	[35]	10-point unipolar, with verbal anchors at the scale ends	[38]	continuous unipolar (“Not at all” (=0) to “Very much” (=10))
	[42]	5-point unipolar	[47]	11-point unipolar	[46]	11-point unipolar
	[51]	9-point unipolar	[36]	continuous unipolar (“Not at all” (=0) to “Very much” (=7))	[53]	10-point unipolar
	[43]	7-point bipolar	[55]	5-point unipolar		
Data normality test	[47]	Visual examination of the residuals’ and variables’ normal probability plots	[38]	Kolmogorov–Smirnov Shapiro–Wilks	[53]	Kolmogorov–Smirnov
	[36]	Shapiro–Wilk	[55]	Kolmogorov–Smirnov		
Statistical test	[29]	Repeated measures ANOVA	[35]	Subjects’ answers distribution Non-parametric Kruskal–Wallis pair-wise for influence of space and parameters Correlation of Pearson for relation between subjective and algorithms results Logistic Regression View direction frequency distribution for head tracking	[38]	Non-parametric Friedman’s ANOVA Pair-wise comparisons with Bonferroni corrections
	[42]	Mean and standard deviation values	[47]	Mean and standard deviation values Linear Mixed Model Pair-wise comparisons	[46]	Linear Mixed Model A specially created composite index
	[51]	Repeated measures ANOVA	[36]	Friedman’s ANOVA for answers to window condition Non-parametric Wilcoxon Signed-Rank Test	[53]	Wilcoxon Rank-Sum for between-subject design on context factors Friedman’s one-way ANOVA for within-subject design on geometry factors Wilcoxon Signed-Ranks Matched-Pairs for post-ho analysis on significant cases

Table 2. *Cont.*

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
	[43]	Repeated measures ANOVA	[55]	Mean and standard deviation values for subjective responses quantification Wilcoxon Signed-Rank test for the evaluation of attributes differences Kruskal–Wallis test for attributes rating comparison Kendall's Tau correlation for assessing attributes correlations Friedman test for contrasting the most statistically significant attributes		

5. Virtual Reality for People's Interaction with Lighting and/or Shading Systems

5.1. Studies Based on Within-Subject Design

Heydarian et al. [22] present an approach of collecting end-user lighting-related behavior in a virtual environment, integrating it with building performance simulation (BPS) tools, then utilizing user preference data to evaluate design choices that satisfy preferences and reduce energy consumption. The lighting preferences, skills (reading speed and comprehension), personality factors, and environmental views of 90 participants were gathered using IVRs to assess this method's usability. The virtual model of an office was modeled in Unity and equipped with shading systems on the three windows and twelve luminaires. In addition, lighting preferences were turned into illuminance value distributions with the BPS tool so that different designs could be compared and decisions about user-centered design could be made. Participants were divided into one of four groups based on the number of electric luminaires switched on in the scene. From a dark scenario, users could select their favorite lighting arrangement among thirty-two shade and electric lighting systems combinations to read a text. Meanwhile, reading speed and text comprehension were acquired. Finally, they filled out questionnaires on personality, space values, and views. The chi-square test was applied to evaluate users' preferences, while the *t*-test assessed the correlation between lighting preferences and reading speed.

Mahmoudzadeh et al. [58] evaluated the effects of personal control over the lighting system on lighting preferences, lighting comfort, and performance using the virtual model of an office. Three models of the office, characterized by different control systems for artificial lighting and shading systems, were displayed to thirty participants, asking them to adjust light conditions and read a text. In addition to lighting preferences, reading speed, and text comprehension, participants filled out a survey on visual comfort and light distribution based on [59], as well as cognitive load. After completing the experiment, people were asked to complete two further surveys to assess personality [60,61]. Regarding lighting control, Friedman's and Wilcoxon signed-rank tests with Bonferroni correction were applied to evaluate and determine differences among the scenarios, respectively. Scatter plot graphs and frequencies were considered to inspect the relationship between users' choices and personalities. Finally, acceptance of technology data consistency was evaluated by means of Cronbach's alpha value, while results were compared in terms of mean and standard deviation values.

5.2. Research Using Both within and Between-Subject Design

In [24], within- and between-subject designs were applied to assess how the location and type of light controls affected people's preferences. The virtual model of a real office was created in Revit, optimized in 3ds Max, and imported into Worldviz's Vizard to be presented to 114 volunteers, randomly allocated into four groups. Four different control strategies in the virtual model for controlling the artificial lighting and shading systems were implemented. The illuminance values acquired in the physical space were used to regulate the light level in the simulated space. Each group was assigned to one of the four scenarios, asking them to enter the virtual room, adjust the artificial lighting and shading systems, read a text, and answer a 7-point Likert scale questionnaire about interaction with IVR and environmental conscientiousness. The users' artificial lighting and shading system adjustments were collected in addition to the questionnaire. The chi-square test was used to analyze lighting preferences, while mean values and standard deviation were considered for the users' feedback.

5.3. Main Findings

Table 3 lists the key design considerations, along with their labels and citations, from papers aimed at investigating users' lighting preferences, allowing them to interact with the virtual environment. Fixture elements and design themes were separated into design parameters. The literature review suggests:

- IVR with interactive virtual environments are used to evaluate the lighting preferences of participants, as well as their interaction with control systems for artificial lighting and shadings;
- Reading task and text comprehension are the preferred activities in interactive models;
- Investigations are generally performed using the within-subject design, questionnaires with different scales and involving young people (average age around 30 years);
- Indoor environments are exclusively considered;
- Results are analyzed using different parametric and non-parametric statistical tests.

Table 3. Main design parameters and measures for IVR used to allow users' interaction with the virtual environment.

Design Parameters	Ref.	Label	Ref.	Label	Ref.	Label
Factors investigated	[24]	IVR interaction Environmental conscientiousness	[58]	Reading Task Text comprehension Subject's lighting preferences Cognitive load Visual comfort Light distribution Technology acceptance Subjects' personality	[22]	Reading task Text comprehension Subject's lighting preferences Space view Space value Subjects' personality
Number of subjects	[24]	114	[58]	30	[22]	89
Average age {age range} year	[24]	21 {18–33}	[58]	- {22–36}	[22]	22 {18–35}
Type of factor-analyses	[24]	Within-subjects Between-subjects	[58]	Within-subjects	[22]	Within-subjects
Space type	[24]	Office	[58]	Office	[22]	Office
Statistical power	[24]	0.8	[58]	0.95		
Questionnaire scale	[24]	7-point unipolar	[58]	Different unipolar scales		
Statistical test	[24]	Chi-square for lighting preferences Mean and standard deviation values for questionnaires	[58]	Friedman for data analysis Wilcoxon signed-rank, with Bonferroni correction, for difference identification Cronbach's alpha and Spearman's Rho for acceptance of technology	[22]	Chi-square goodness of fit for lighting preferences *t*-test Ordinary least square linear regression

6. Additional Investigations

6.1. Studies Based on Within-Subject Design

Based on the findings presented in [30], stereoscopic images were used as references to evaluate the usefulness of virtual environments shown through HMD for daylight investigations [62]. Six virtual bedroom models were modeled considering three different window sizes and two color of room surfaces (white and black). Models for HMD presentation were prepared according to the method proposed by Chamilothori et al. [18]. HDR rendered images were generated with Radiance and displayed using the system in [30], considering overcast sky conditions. A prior stereoscopic vision test was used to select participants. According to the within-subject design, twenty participants were asked to fill out a questionnaire on attributes linked to the spatial perception for each scenario. The Wilcoxon signed-rank test was applied for data analysis, and differences among attributes were quantified with the Bland–Altman method [31].

6.2. Main Findings

Table 4 lists the main design considerations, along with their labels and citations, from papers aimed at using IVR for research with different objectives. Fixture elements and design themes were separated into design parameters.

Table 4. Main design parameters and measures for IVR used research.

Design Parameters	Ref.	Label
Factors investigated	[62]	Pleasantness Calmness Interest Excitement Complexity Spaciousness Amount of View
Number of subjects	[62]	20
Average age {age range} year	[62]	25.9 {20–47}
Type of factor-analyses	[62]	Within-subjects
Space type	[62]	Bedroom
Questionnaire scale	[62]	10-point unipolar
Statistical test	[62]	Wilcoxon signed-rank Bland–Altman method for difference quantification

7. Discussion

During the past few years, interest in virtual reality has grown significantly because it provides a wonderful multisensory experience. This revolutionary technology's boundless potential has broadened its usage field beyond its first creative applications to comprehend complex and interconnected problems. At the same time, the potentialities and limitations of IVR for lighting have to be deeply investigated before it can be used as a substitute for conventional design systems. Due to its characteristics, IVR needs to be explored from two points of view simultaneously, namely from objective and subjective points of view. The first is related to the ability of IVR to reproduce the proper day-electric light distribution. In contrast, the second refers to the possibility of provoking perceptions comparable to those caused by physical spaces. This review aims to underline how different researchers faced this problem, focusing on the role of subjective assessments.

7.1. Discussion of the Findings

The literature review results suggest that subjective assessments are a powerful tool for evaluating IVR's accuracy and potentiality. In particular, different uses of IVR for lighting can be identified. The ability of IVR to reproduce physical spaces lit by daylight or electric light is understood by comparing the virtual model with the physical one. Even though different methods to obtain virtual environments with daylight or electric light luminance distributions more and more similar to the real ones have been proposed, some of them are time expensive, allow the evaluation of the virtual space from a fixed point of view, or can be used only if physical space is available. Despite this, all research agrees that subjective assessment is the main tool to understand the possibilities and limits of IVR for lighting. The benefit of using the IVR to compare different light scenes is more evident. The investigations indicated the IVR as the best tool for presenting virtual lit environments compared to other media, allowing more of a sense of immersion, users' satisfaction, and congruence with physical space. Although only a few papers have been carried out, the advantages of IVR for lighting to evaluate how users interact with shadings and artificial light control systems, as well as their lighting preferences, seem to be confirmed.

Figure 3 shows the link among the parameters and factors considered in the various research analyzed in the review grouped by the type of research and criteria considered during the design of the experiment.

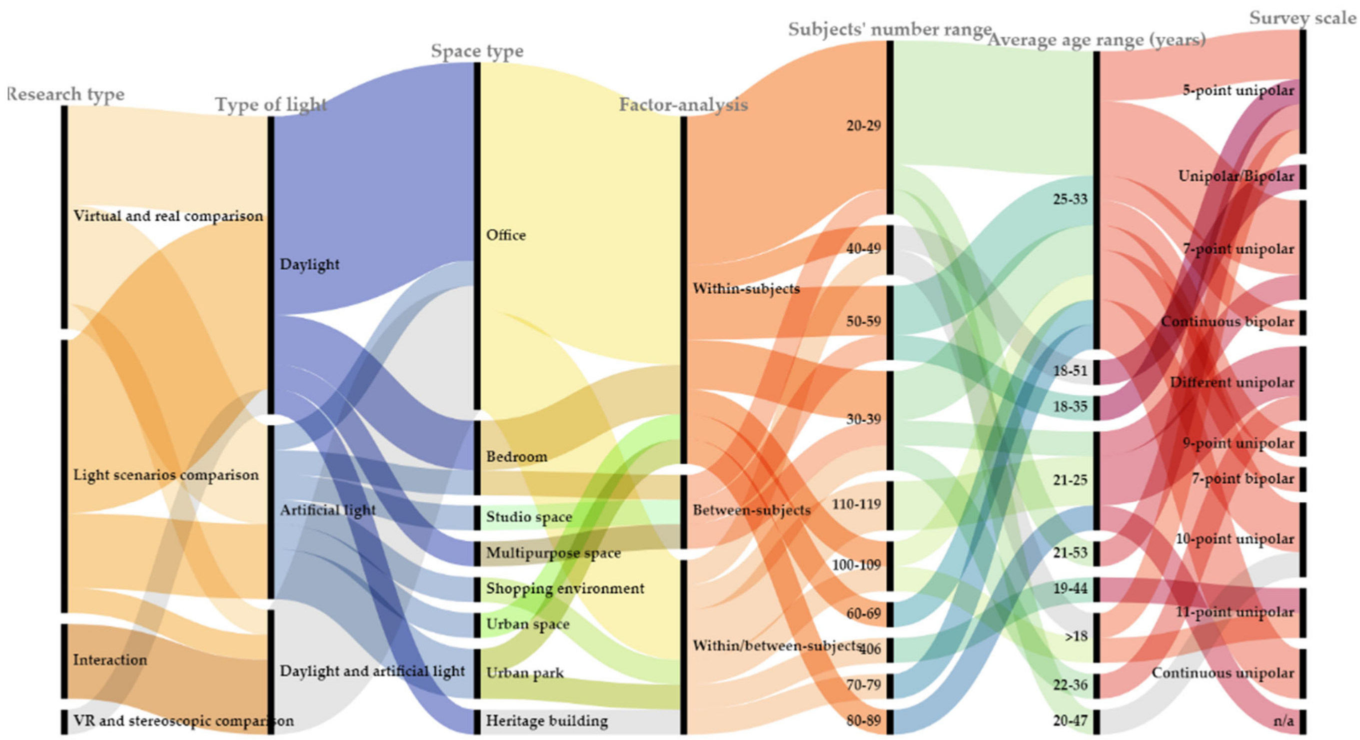

Figure 3. Alluvial diagram of the design parameters and measures involved in lighting research through IVR.

Among the 24 papers reviewed, nine evaluated the accuracy of virtual environments compared to the physical ones, eleven used IVR to compare various light conditions, three assessed the users' interaction with the environment, and one evaluated the effectiveness of IVR in comparison to other media. In addition, twelve of the twenty-four papers evaluated the effects of daylight, seven analyzed artificial light, and the remaining five articles considered both types of light. Concerning the space type, most of the research was conducted considering offices and only three investigated outdoor environments. Factor analysis was categorized into three groups, underlining that fourteen out of twenty-four studies were based on within-subject experimental design, and only three out of twenty-four on between-subject design. The review points out a significant variability in the number of people involved in testing through the studies analyzed. In order to simplify the readability of the graph, the number of participants was gathered in intervals of ten. From Figure 3, it is possible to notice that the number of participants in seven research ranged between 20 to 29, while only 5 out of 24 considered more than 100 participants. The average age ranges were obtained by grouping similar values, while the average age range is considered the same as the age range reported in Tables 1–3 when the participants' average age value is not provided in the reviewed paper. The results point out that in 18 out of 24 papers, the average age of recruited people ranges from 18 to 36 years. The participants' questionnaire responses are mainly measured on a 5-point unipolar scale (5 out of 24 research) or a 7-point unipolar scale (4 out of 24 papers).

Figure 4 displays the dendrogram of the measures and parameters examined in research based on IVR for lighting, stacked on the main research type. The graph stresses the connection between the research types and the other parameters considered: the type of light, survey scale, subjects' number range, average age range, factor analysis and space type. For each parameter, the values involved in the various studies were shown.

From Figures 3 and 4, it is possible to infer that:

- Daylight and artificial light are usually investigated separately in VPC and DSC investigations, while in HII research, both day and artificial light are considered;
- Regarding the survey scale, a 5-point unipolar scale, 5-point unipolar, and 7-point unipolar scales are used for VPC, DSC and HII analysis, respectively;
- For investigations using IVR to compare virtual and physical spaces or various lighting scenarios, the number of subjects is in the range of 20–29, while more than thirty people are involved in research evaluating lighting preferences and user interaction;
- On average, the age of participants ranges between 21 and 33 years;
- The most used factor analysis is the within-subject design;
- The office is the most investigated space, whatever the research type is.

A more complex question concerns the type of factors considered and the statistical tests used for data analysis.

In VPC research, the most considered factors are the sense of presence, presentation ability, and perceived brightness. Concerning the statistical analyses used for data interpretation, the results underline that the most-used parametric tests are the mean and standard deviation values (4 out of 9) or the *t*-test (3 out of 9). In contrast, the most-used non-parametric test is the Mann–Whitney test (2 out of 9). Often, two or more statistical tests are used for data elaboration in the same research.

In DSC investigations, Pleasantness, Brightness, Interest, Excitement, Calmness, and Complexity are the factors mainly investigated. In this case, the most used parametric tests are repeated ANOVA (3 out of 9), as well as mean and standard deviation values (3 out of 9). Friedman's ANOVA (4 papers) and Wilcoxon Signed-Rank (3 papers) are the main used non-parametric tests. Usually, more than one test is used in the same article.

In HII studies, the efforts are focused on evaluating the text comprehension, subjects' personalities, as well as space view and value. Two of the three studies used the parametric chi-squared test for data analysis. Friedman's and Wilcoxon signed-rank tests are considered if non-parametric analyses are required.

The literature outcomes underline a significant variability in experimental design as well as methodologies for data analysis and interpretation. This could make it challenging to design future investigations properly or compare results among different research. The overview provided by this review could help both understand the solutions used so far (in terms of experimental design, data analysis, experimental procedures, and results interpretation) and define benchmarks for future investigations using IVR for lighting. Based on the information about the critical analysis of the reviewed research and the identification of their key parameters summarized in the paper, standardized methodologies following a systematic investigation patch could be deduced. At the same time, the information reported in the study could be a reference for defining objectives or experimental strategies of new research based on IVR for lighting.

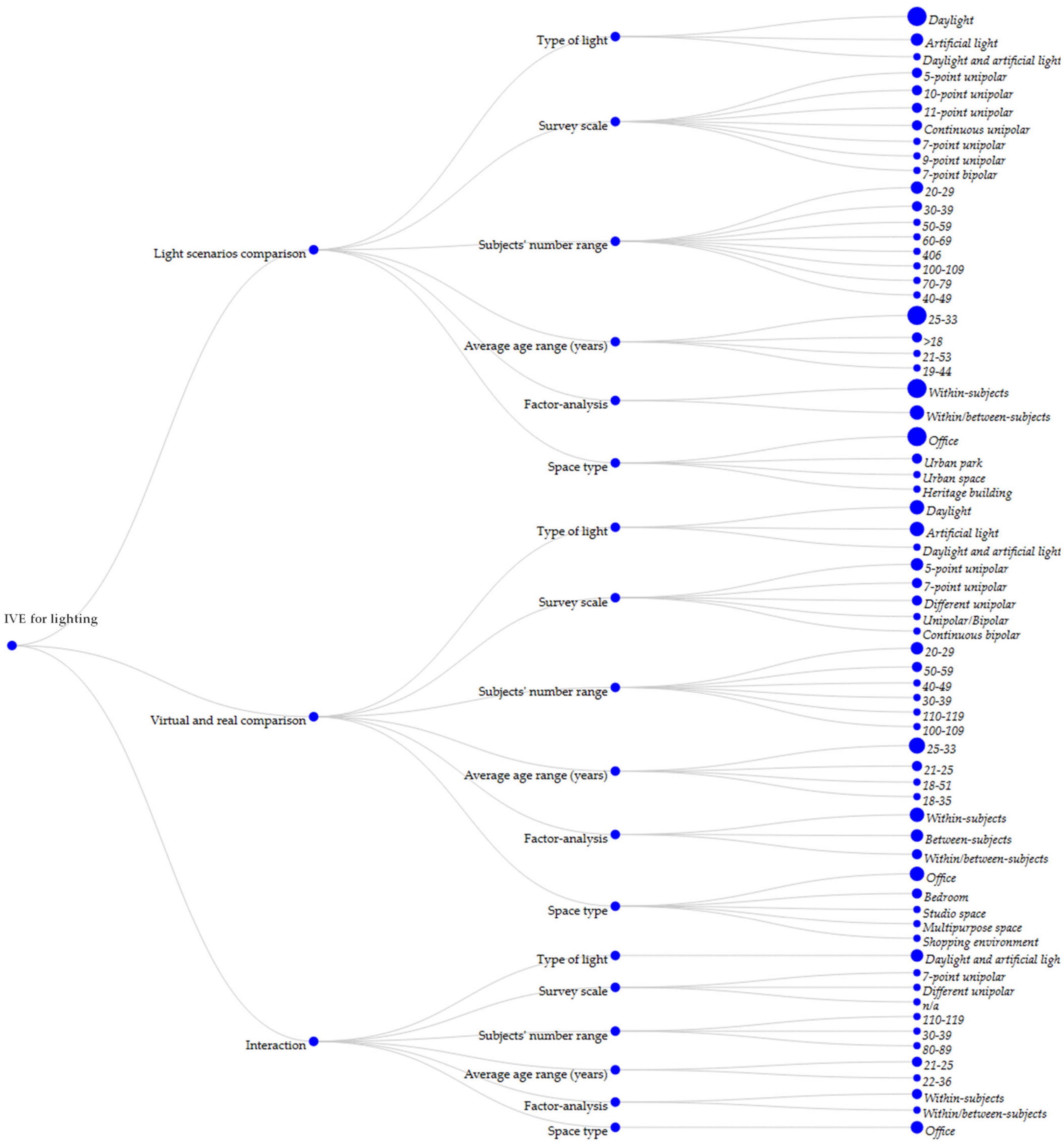

Figure 4. Design parameters and measures dendrogram stacked on the main research type.

7.2. Limitations

While the review results reinforce the value of IVR as a tool for researchers and designers, they also highlight gaps and limits that will need to be overcome before such technology is widely spread for lighting research and design. In particular, the review points out some main lacks or limitations of:

- A systematic approach(es) to subjective assessments and data analysis methodologies to simplify results comparison from different studies;
- Variability of the environments considered. So far, about 66% of the research considered office-like spaces;
- Involving old or young people in experimentations (participants' age mainly range from 21 to 33 years);
- Considering a larger sample size of the study to improve the power of the statistical analysis;
- Supporting the number of subjects considered for research with a specific analysis of statistical power and effect size;
- Performing analysis based on the comparison among different lighting scenarios or users' interaction with the virtual environment according to the between-subject design;
- Evaluating IVR's effectiveness in reproducing outdoor environments (to the best of the authors' knowledge, no research has compared an outdoor virtual reproduction with the physical environment);
- Using IVR to investigate people's responses to different lighting solutions for outdoor spaces.

The main limitations derived from the analyzed articles about the use of IVR for lighting are the:

- Limited HMD luminance range (making it challenging to evaluate glare);
- Limited HMD field of view;
- Limited resolution of HMD as well as of the devices used to create the virtual environment;
- Accuracy of the methodology used to develop the virtual environment;
- Tone mapping operator used to adapt the virtual model luminance values in the software to the characteristics of the HMD screens;
- Limited movement that users can have (usually only the rotation of the head is allowed);
- Accuracy of the virtual model in comparison to the physical one;
- Weather condition variability in the physical environment (mainly for investigation focused on daylight);
- Background stimuli (such as acoustic and thermal) during tests.

7.3. Main Perspective of Using IVR for Lighting and Future Investigations

Today's population spends a large portion of their time indoors, necessitating lighting solutions that support visual comfort and well-being as well as vision-related jobs. This has compelled scientists to consider how light affects people rather than just assessing the quantity and distribution of light. These studies are typically conducted in test rooms or living laboratories, but are costly and time-consuming.

Less-expensive and time-consuming inquiry techniques, such as virtual reality, can aid in analyzing various light settings and assist in discovering the best ones from a human perspective, if they can make people experience stimuli that are close to those in reality.

Indeed, IVR would allow obtaining feedback from people for design solutions that are challenging to achieve in the real world, such as the replacement of objects or parts of the scene, different boundary conditions, or geographic areas, as well as to overcome the constraints regarding stimuli control in the physical environments.

Finally, through IVR, people (e.g., citizens, buildings users) could be involved in the planning process (of lighting systems, shading systems, as well as control systems or strategies), asking them to select the preferred design solution (participatory design) or express a judgment about a design solution (to evaluate its social impact).

Future investigations based on the use of IVR for lighting should face the current limitations to overcome the gaps between the IVR environments and the physical one or identify the limits of this technology and the application fields. At the same time, the limitations call for the definition of a standardized methodology to harmonize the experimental design and statistical analysis of data.

8. Conclusions

One of the most promising technological advancements for creating lighting designs that meet users' needs and industry standards is immersive virtual reality. If less-expensive and time-consuming inquiry methods, such as virtual reality, can make individuals experience stimuli that are similar to those in reality, it will be easier to analyze different light conditions and determine which ones are ideal from a human perspective. In this scenario, subjective assessments are the most effective method to evaluate the accuracy of the virtual models.

This paper includes a review of the literature on research using IVR to investigate lighting and daylighting preferences and performance through subjective evaluations, mainly focused on environmental and psychological aspects, as well as the tests, methods, and surveys used.

The review results underline three main types of research performed with IVR, namely: (i) to evaluate the accuracy of virtual models in reproducing physical lit environment; (ii) to investigate people's preferences comparing various light scenarios, and (iii) to assess users' lighting preferences analyzing interaction with lighting and/or shading systems. Review results highlight that:

- Participants' age ranges from 21 and 33 years;
- Most experiments are planned according to the within-subject design, involving between 20 and 29 participants, and considering office-like spaces;
- The participants' number is chosen to ensure statistical power of not less than 0.8, whatever the research type is, as well as an effect size in the range of 0.25 to 0.28 for DSC, and greater than 0.50 for VPC and HII;
- The most-investigated factors are the sense of presence, pleasantness, and lighting preferences in VPC, DSC, and HII research, respectively;
- The most used statistical tests are the mean and standard deviation values, Friedman's ANOVA, and chi-squared test in VPC, DSC, and HII research, respectively.

At the same time, the review highlights a lack of standardization in the methods used for data analysis and subjective evaluations, making the comparison of the findings from various studies challenging. Finally, the review suggests gaps and proposals for future research to understand the boundaries of and use of the IVR for lighting, such as:

- Standardized procedures for experimental design and data analysis should be defined;
- Older and younger participants should be involved in tests;
- Sample size should be increased to reach higher power of the statistical analysis;
- The effectiveness of IVR in simulating lit scenarios should be evaluated for different types of space, as well as outdoors;
- The effects of walking and background stimuli on people's perception of IVR scenarios should be deeply investigated.

Author Contributions: Conceptualization, M.S., G.C., P.C.N. and M.M.; methodology, M.S., G.C., P.C.N. and M.M.; validation, M.S., G.C., P.C.N. and M.M.; formal analysis, M.S., G.C., P.S., Y.S., D.C. and M.G.; investigation, M.S., G.C., P.S., Y.S., D.C. and M.G.; data curation, M.S., D.C. and M.G.; writing—original draft preparation, M.S., G.C., P.S., Y.S., D.C. and M.G.; writing—review and editing, M.S., G.C., P.C.N. and M.M.; visualization, M.S., G.C., P.S., Y.S., D.C. and M.G.; supervision, M.S., G.C., P.C.N. and M.M.; project administration, M.S. and G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Konstantzos, I.; Sadeghi, S.A.; Kim, M.; Xiong, J.; Tzempelikos, A. The Effect of Lighting Environment on Task Performance in Buildings—A Review. *Energy Build.* **2020**, *226*, 110394. [CrossRef]
2. Houser, K.W.; Boyce, P.R.; Zeitzer, J.M.; Herf, M. Human-Centric Lighting: Myth, Magic or Metaphor? *Light. Res. Technol.* **2021**, *53*, 97–118. [CrossRef]
3. Vasquez, N.G.; Amorim, C.N.D.; Matusiak, B.; Kanno, J.; Sokol, N.; Martyniuk-Peczek, J.; Sibilio, S.; Scorpio, M.; Koga, Y. Lighting Conditions in Home Office and Occupant's Perception: Exploring Drivers of Satisfaction. *Energy Build.* **2022**, *261*, 111977. [CrossRef]
4. Gentile, N.; Lee, E.S.; Osterhaus, W.; Altomonte, S.; Naves David Amorim, C.; Ciampi, G.; Garcia-Hansen, V.; Maskarenj, M.; Scorpio, M.; Sibilio, S. Evaluation of Integrated Daylighting and Electric Lighting Design Projects: Lessons Learned from International Case Studies. *Energy Build.* **2022**, *268*, 112191. [CrossRef]
5. Scorpio, M.; Laffi, R.; Masullo, M.; Ciampi, G.; Rosato, A.; Maffei, L.; Sibilio, S. Virtual Reality for Smart Urban Lighting Design: Review, Applications and Opportunities. *Energies* **2020**, *13*, 3809. [CrossRef]
6. Narasimha, S.; Dixon, E.; Bertrand, J.W.; Chalil Madathil, K. An Empirical Study to Investigate the Efficacy of Collaborative Immersive Virtual Reality Systems for Designing Information Architecture of Software Systems. *Appl. Ergon.* **2019**, *80*, 175–186. [CrossRef]
7. Ruotolo, F.; Maffei, L.; di Gabriele, M.; Iachini, T.; Masullo, M.; Ruggiero, G.; Senese, V.P. Immersive Virtual Reality and Environmental Noise Assessment: An Innovative Audio-Visual Approach. *Environ. Impact Assess. Rev.* **2013**, *41*, 10–20. [CrossRef]
8. Zhu, Y.; Tang, G.; Liu, W.; Qi, R. How Post 90's Gesture Interact with Automobile Skylight. *Int. J. Hum. Comput. Interact.* **2022**, *38*, 395–405. [CrossRef]
9. Desmet, P.; Fokkinga, S. Beyond Maslow's Pyramid: Introducing a Typology of Thirteen Fundamental Needs for Human-Centered Design. *Multimodal Technol. Interact.* **2020**, *4*, 38. [CrossRef]
10. Gentile, N.; Goven, T.; Laike, T.; Sjoberg, K. A Field Study of Fluorescent and LED Classroom Lighting. *Light. Res. Technol.* **2018**, *50*, 631–650. [CrossRef]
11. Scorpio, M.; Ciampi, G.; Gentile, N.; Sibilio, S. Effectiveness of Low-Cost Non-Invasive Solutions for Daylight and Electric Lighting Integration to Improve Energy Efficiency in Historical Buildings. *Energy Build.* **2022**, *270*, 112281. [CrossRef]
12. Kong, Z.; Liu, Q.; Li, X.; Hou, K.; Xing, Q. Indoor Lighting Effects on Subjective Impressions and Mood States: A Critical Review. *Build. Environ.* **2022**, *224*, 109591. [CrossRef]
13. Chen, Y.; Cui, Z.; Hao, L. Virtual Reality in Lighting Research: Comparing Physical and Virtual Lighting Environments. *Light. Res. Technol.* **2019**, *51*, 820–837. [CrossRef]
14. Erickson, A.; Kim, K.; Bruder, G.; Welch, G.F. Exploring the Limitations of Environment Lighting on Optical See-Through Head-Mounted Displays. In Proceedings of the Proceedings—SUI 2020: ACM Symposium on Spatial User Interaction, Virtual Event Canada, Ottawa, ON, Canada, 31 October–1 November 2020; Association for Computing Machinery, Inc.: New York, NY, USA, 2020; pp. 1–8.
15. El Jamiy, F.; Marsh, R. Survey on Depth Perception in Head Mounted Displays: Distance Estimation in Virtual Reality, Augmented Reality, and Mixed Reality. *IET Image Process.* **2019**, *13*, 707–712. [CrossRef]
16. Vieri, C.; Lee, G.; Balram, N.; Jung, S.H.; Yang, J.Y.; Yoon, S.Y.; Kang, I.B. An 18 Megapixel 4.3″ 1443 Ppi 120 Hz OLED Display for Wide Field of View High Acuity Head Mounted Displays. *J. Soc. Inf. Disp.* **2018**, *26*, 314–324. [CrossRef]
17. Cheng, D.; Wang, Q.; Liu, Y.; Chen, H.; Ni, D.; Wang, X.; Yao, C.; Hou, Q.; Hou, W.; Luo, G.; et al. Design and Manufacture AR Head-Mounted Displays: A Review and Outlook. *Light. Adv. Manuf.* **2021**, *2*, 336. [CrossRef]
18. Chamilothori, K.; Wienold, J.; Andersen, M. Adequacy of Immersive Virtual Reality for the Perception of Daylit Spaces: Comparison of Real and Virtual Environments. *Leukos* **2018**, *15*, 203–226. [CrossRef]
19. Abd-Alhamid, F.; Kent, M.; Bennett, C.; Calautit, J.; Wu, Y. Developing an Innovative Method for Visual Perception Evaluation in a Physical-Based Virtual Environment. *Build. Environ.* **2019**, *162*, 106278. [CrossRef]
20. Rockcastle, S.; Danell, M.; Calabrese, E.; Sollom-Brotherton, G.; Mahic, A.; Van Den Wymelenberg, K.; Davis, R. Comparing Perceptions of a Dimmable LED Lighting System between a Real Space and a Virtual Reality Display. *Light. Res. Technol.* **2021**, *53*, 701–725. [CrossRef]
21. Scorpio, M.; Laffi, R.; Teimoorzadeh, A.; Ciampi, G.; Masullo, M.; Sibilio, S. A Calibration Methodology for Light Sources Aimed at Using Immersive Virtual Reality Game Engine as a Tool for Lighting Design in Buildings. *J. Build. Eng.* **2022**, *48*, 103998. [CrossRef]
22. Heydarian, A.; Pantazis, E.; Wang, A.; Gerber, D.; Becerik-Gerber, B. Towards User Centered Building Design: Identifying End-User Lighting Preferences via Immersive Virtual Environments. *Autom. Constr.* **2017**, *81*, 56–66. [CrossRef]
23. Wang, L. Use of Immersive Virtual Environment for Collecting Occupant Preference, Understanding the Impact of Wall Luminance and Uniformity on Preferred Task Illuminance. Master's Thesis, Hong Kong Polytechnic University, Hong Kong, April 2018. Available online: https://theses.lib.polyu.edu.hk/handle/200/9733 (accessed on 30 April 2023).
24. Heydarian, A.; Carneiro, J.P.; Gerber, D.; Becerik-Gerber, B. Immersive Virtual Environments, Understanding the Impact of Design Features and Occupant Choice upon Lighting for Building Performance. *Build. Environ.* **2015**, *89*, 217–228. [CrossRef]
25. Gil Rodríguez, R.; Bayer, F.; Toscani, M.; Guarnera, D.; Guarnera, G.C.; Gegenfurtner, K.R. Colour Calibration of a Head Mounted Display for Colour Vision Research Using Virtual Reality. *SN Comput. Sci.* **2022**, *3*, 22. [CrossRef]

26. Riffenburgh, R.H.; Gillen, D.L. Tests on Variability and Distributions. In *Statistics in Medicine*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 311–336.
27. Kennedy, R.S.; Lane, N.E.; Berbaum, K.S.; Lilienthal, M.G. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *Int. J. Aviat. Psychol.* **2009**, *3*, 203–220. [CrossRef]
28. Thompson, E.R. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *J. Cross Cult. Psychol.* **2007**, *38*, 227–242. [CrossRef]
29. Hong, T.; Lee, M.; Yeom, S.; Jeong, K. Occupant Responses on Satisfaction with Window Size in Physical and Virtual Built Environments. *Build. Environ.* **2019**, *166*, 106409. [CrossRef]
30. Moscoso, C.; Matusiak, B.; Svensson, U.P.; Orleanski, K. Analysis of Stereoscopic Images as a New Method for Daylighting Studies. *ACM Trans. Appl. Percept.* **2015**, *11*, 1–13. [CrossRef]
31. Martin Bland, J.; Altman, D.G. Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement. *Lancet* **1986**, *327*, 307–310. [CrossRef]
32. Hegazy, M.; Ichiriyama, K.; Yasufuku, K.; Abe, H. Comparing Daylight Brightness Perception in Real and Immersive Virtual Environments Using Perceptual Light Maps. *Autom. Constr.* **2021**, *131*, 103898. [CrossRef]
33. Heydarian, A.; Carneiro, J.P.; Gerber, D.; Becerik-Gerber, B.; Hayes, T.; Wood, W. Immersive Virtual Environments versus Physical Built Environments: A Benchmarking Study for Building Design and User-Built Environment Explorations. *Autom. Constr.* **2015**, *54*, 116–126. [CrossRef]
34. Higuera-Trujillo, J.L.; López-Tarruella Maldonado, J.; Llinares Millán, C. Psychological and Physiological Human Responses to Simulated and Real Environments: A Comparison between Photographs, 360° Panoramas, and Virtual Reality. *Appl. Ergon.* **2017**, *65*, 398–409. [CrossRef]
35. Rockcastle, S.F.; Chamilothori, K.; Andersen, M. An Experiment in Virtual Reality to Measure Daylight-Driven Interest in Rendered Architectural Scenes. In Proceedings of the Building Simulation 2017: 15th Conference of IBPSA 2017, San Francisco, CA, USA, 7–9 August 2017.
36. Flor, J.F.; Aburas, M.; Abd-AlHamid, F.; Wu, Y. Virtual Reality as a Tool for Evaluating User Acceptance of View Clarity through ETFE Double-Skin Façades. *Energy Build.* **2021**, *231*, 110554. [CrossRef]
37. Ferguson, C.J. An Effect Size Primer: A Guide for Clinicians and Researchers. *Prof. Psychol. Res. Pr.* **2009**, *40*, 532–538. [CrossRef]
38. Abd-Alhamid, F.; Kent, M.; Calautit, J.; Wu, Y. Evaluating the Impact of Viewing Location on View Perception Using a Virtual Environment. *Build. Environ.* **2020**, *180*, 106932. [CrossRef]
39. Shapiro, S.S.; Wilk, M.B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
40. Huynh, H.; Feldt, L.S. Estimation of the Box Correction for Degrees of Freedom from Sample Data in Randomized Block and Split-Plot Designs. *J. Educ. Stat.* **1976**, *1*, 69–82. [CrossRef]
41. Kitchin, J. Basic Statistical Inference. In *Methods in Experimental Physics*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 28, pp. 155–186.
42. Greule, R.; Braun, T. *Psychological Effect of Lights in An Urban Environment*; Zumtobel: Dornbirn, Austria, 2017.
43. Masullo, M.; Cioffi, F.; Li, J.; Maffei, L.; Ciampi, G.; Sibilio, S.; Scorpio, M. Urban Park Lighting Quality Perception: An Immersive Virtual Reality Experiment. *Sustainability* **2023**, *15*, 2069. [CrossRef]
44. Johansson, M.; Pedersen, E.; Maleetipwan-Mattsson, P.; Kuhn, L.; Laike, T. Perceived Outdoor Lighting Quality (POLQ): A Lighting Assessment Tool. *J. Environ. Psychol.* **2014**, *39*, 14–21. [CrossRef]
45. Duchowski, A.T.; Biele, C.; Niedzielska, A.; Krejtz, K.; Krejtz, I.; Kiefer, P.; Raubal, M.; Giannopoulos, I. The Index of Pupillary Activity: Measuring Cognitive Load Vis-à-Vis Task Difficulty with Pupil Oscillation. In Proceedings of the Conference on Human Factors in Computing Systems—Proceedings, Montreal, QC, Canada, 21–26 April 2018; Association for Computing Machinery: New York, NY, USA, 2018; Volume 1, pp. 1–13.
46. Sawyer, A.O.; Chamilothori, K. Influence of Subjective Impressions of a Space on Brightness Satisfaction: An Experimental Study in Virtual Reality. *Simul. Ser.* **2019**, *51*, 57–64.
47. Moscoso, C.; Chamilothori, K.; Wienold, J.; Andersen, M.; Matusiak, B. Regional Differences in the Perception of Daylit Scenes across Europe Using Virtual Reality. Part I: Effects of Window Size. *LEUKOS* **2021**, *18*, 294–315. [CrossRef]
48. McCulloch, C.E.; Searle, S.R.; Neuhaus, J.M. *Generalized, Linear, and Mixed Models*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2008; ISBN 978-0-470-07371-1.
49. Seltman, H.J. *Experimental Design and Analysis*; Carnegie Mellon University: Pittsburgh, PA, USA, 2018.
50. Šidák, Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *J. Am. Stat. Assoc.* **1967**, *62*, 626–633. [CrossRef]
51. Masullo, M.; Cioffi, F.; Li, J.; Maffei, L.; Scorpio, M.; Iachini, T.; Ruggiero, G.; Malferà, A.; Ruotolo, F. An Investigation of the Influence of the Night Lighting in a Urban Park on Individuals' Emotions. *Sustainability* **2022**, *14*, 8556. [CrossRef]
52. Watson, D.; Clark, L.A.; Tellegen, A. Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *J. Pers. Soc. Psychol.* **1988**, *54*, 1063–1070. [CrossRef]
53. Chamilothori, K.; Chinazzo, G.; Rodrigues, J.; Dan-Glauser, E.S.S.; Wienold, J.; Andersen, M. Subjective and Physiological Responses to Façade and Sunlight Pattern Geometry in Virtual Reality. *Build. Environ.* **2019**, *150*, 144–155. [CrossRef]
54. Siegel, S. Nonparametric Statistics. *Am. Stat.* **1957**, *11*, 13–19. [CrossRef]

55. Marzouk, M.; ElSharkawy, M.; Mahmoud, A. Analysing User Daylight Preferences in Heritage Buildings Using Virtual Reality. *Build Simul.* **2022**, *15*, 1561–1576. [CrossRef]
56. Kruskal, W.H.; Wallis, W.A. Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583. [CrossRef]
57. Newson, R. Parameters behind "Nonparametric" Statistics: Kendall's Tau, Somers' D and Median Differences. *Stata J. Promot. Commun. Stat. Stata* **2002**, *2*, 45–64. [CrossRef]
58. Mahmoudzadeh, P.; Afacan, Y.; Adi, M.N. Analyzing Occupants' Control over Lighting Systems in Office Settings Using Immersive Virtual Environments. *Build. Environ.* **2021**, *196*, 107823. [CrossRef]
59. Eklund, N.H.; Boyce, P.R. The Development of a Reliable, Valid, and Simple Office Lighting Survey. *J. Illum. Eng. Soc.* **1996**, *25*, 25–40. [CrossRef]
60. John, O.P.; Naumann, L.P.; Soto, C.J. Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In *Handbook of Personality: Theory and Research*; The Guilford Press: New York, NY, USA, 2008; pp. 114–158.
61. Balog, A.; Pribeanu, C. Developing a Measurement Scale for the Evaluation of AR-Based Educational Systems. *Stud. Inform. Control* **2009**, *18*, 137–148.
62. Moscoso, C.; Nazari, M.; Matusiak, B.S. Stereoscopic Images and Virtual Reality Techniques in Daylighting Research: A Method-Comparison Study. *Build. Environ.* **2022**, *214*, 108962. [CrossRef]