Evaluation framework for BIM-based VR applications in design phase

Jung In Kim1, Sining Li1, Xingbin Chen2, Calvin Keung1, Minjae Suh3 and Tae Wan Kim4,*

1Department of Architecture and Civil Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon 999077, Hong Kong SAR; 2College of Water Conservancy and Civil Engineering, South China Agricultural University, Tianhe District, Guangzhou, Guangdong 510642, China; 3Department of Engineering Technology, Sam Houston State University, Huntsville, TX 77341, USA and 4Division of Architecture and Urban Design, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 22012, South Korea

*Corresponding author. E-mail: taewkim@inu.ac.kr http://orcid.org/0000-0002-7887-1055

Abstract

The integration of building information modeling (BIM) and virtual reality (VR) has attracted increasing attention in the architecture engineering and construction (AEC) industry. Despite the rapid emergence of BIM-based VR applications, no evaluation framework specialized for the technologies exists in the AEC industry. After extensively reviewing existing studies and interviewing experts, the research team proposed an evaluation framework for BIM-based VR applications, which consists of 3 stages, 5 areas, 14 criteria, and 29 metrics and focuses on the design phase of the projects. To assess the usefulness of the framework, the team applied it to five BIM-based VR applications using a BIM-based design project for an educational building in Hong Kong. The team also interviewed experts to discuss the comprehensiveness of the framework. The results show that the framework provides consistent results for comprehensive evaluation criteria and metrics in a quantitative and flexible manner. Further research that considers additional criteria and/or metrics related to other phases (e.g. construction phase) is required to extend the framework to the whole lifecycle.

Keywords: virtual reality (VR); building information modeling (BIM); performance evaluation framework; design phase; immersive virtual environment

1. Introduction

As building information modeling (BIM) is a digital representation of the geometric and nongeometric information of buildings, it supports visualizing the physical and functional properties of buildings and carrying out a variety of analyses (e.g. cost estimation, clash detection, and structural analysis; Herr & Fischer, 2019; Shin et al., 2020). In addition, as virtual reality (VR) technologies integrate information systems and immersive environments, the architecture engineering and construction (AEC) industry has accepted VR uses with BIM to enhance the visualization of a virtual world and interaction with the virtual space and its components.

BIM-based VR technologies allow stakeholders to walk through an artificial environment while observing a one-to-one-scale three-dimensional (3D) model. In addition, stakeholders who use the technologies can retrieve properties of parametric objects and manipulate BIM components (Sampaio, 2018). As some BIM-based VR applications provide a multiuser shared virtual environment without consideration of users’ locations, they...
enhance the cross-discipline understanding of projects, which ultimately contributes to improved collaboration among multiple stakeholders. For example, the technologies can efficiently support collaboration to adjust to conflicts in designs or construction plans (Shi et al., 2016; Du et al., 2017, 2018; Sun et al., 2019).

Due to these benefits, BIM-based immersive systems and interfaces have attracted more and more attention from researchers and industry professionals. With the rapid emergence of BIM-based VR applications, project teams need to make decisions about which applications they adopt based on their preferences (e.g., purpose) and the capability of the application. However, as software vendors do not provide the information about their applications in a comprehensive, structured, and quantitative manner, these teams have difficulty making an informed decision on the selection of the applications appropriate for their purpose in a consistent and flexible manner.

Several researchers have developed evaluation frameworks for the BIM implementation levels of projects, yet there is no existing evaluation framework for BIM-based VR technologies in the AEC industry (Sampaio, 2010). Thus, the authors also reviewed existing studies about evaluating VR technologies from other industries (Teizer et al., 2013). Although these studies support the evaluation of specific aspects (e.g., usability and collaboration) of VR technologies, they are not specialized to cover comprehensive aspects of BIM-based VR applications.

This paper aims to develop a framework for evaluating BIM-based VR applications. To support informed decision making on the application, the evaluation framework must provide consistent results based on comprehensive criteria in a quantitative and flexible manner. In addition, as the first step in developing the evaluation framework for BIM-based VR applications in the AEC industry, this evaluation framework focuses on the design phase of the projects because the utilization of the applications is relatively mature in design tasks (Calderon-Hernandez & Brioso, 2018).

After introducing existing studies about BIM-based VR technologies, evaluation frameworks for emerging technologies in the AEC industry, and evaluation frameworks for VR technologies in other industries, this paper describes an evaluation framework for BIM-based VR applications applicable for the design phase of projects that was formalized through a review of the literature and interviews with experts from multiple disciplines in the AEC industry. Finally, the paper discusses the results of applying the framework and interviewing experts to highlight the usefulness of the framework.

2. Points of Departure

This section describes existing studies about BIM-based VR technologies. After introducing the evaluation frameworks for emerging technologies (e.g., BIM) in the AEC industry and VR technologies in other industries, the authors discuss the limitations in the frameworks required to assess BIM-based VR technologies.

The increasing availability of VR technologies and the affordability of VR hardware have recently led to the emergence of VR-related applications in the AEC industry. The most advanced VR application areas in the AEC industry are the inspection of BIM models and collaboration in the virtual environment. The immersive virtual environment allows the improvement of architectural designs by identifying any issues in the designs in advance and finding better solutions (Teizer et al., 2013; Zaker & Coloma, 2018). In addition, collaboration in the virtual environment supports cross-team interactions among all participants, specifically benefiting the communication between designers and clients. As BIM-based VR applications serve as an immersive and interactive medium, they lower barriers to understanding complex designs developed by multiple stakeholders (Heydarian et al., 2015). In addition, BIM-based VR applications are increasingly used for construction safety identification and training. Users of such applications can recognize inherent risks and hazards in the construction process through four-dimensional simulations implemented in the immersive virtual environment (Germani et al., 2009, 2012). Moreover, these applications are used for construction engineering education because they present construction sites and construction processes to help inexperienced students comprehend the relevant concepts through authentic multisensory feelings (Heydarian et al., 2015; Martens, 2016; Zaker & Coloma, 2018).

As there is no assessment framework for BIM-based VR applications, the authors reviewed existing frameworks for emerging technologies in the AEC industry. Several researchers have developed frameworks to assess the BIM implementation levels of projects. For instance, the U.S. National Building Information Model Standard™ proposed a Capability Maturity Model to determine the maturity level of the BIM-based projects from 11 aspects (Sebastian & van Berlo, 2010). A BIM Proficiency Matrix was developed to measure the proficiency level of BIM models quantitatively using eight categories. The researchers also formulated a Virtual Design and Construction (VDC) Scorecard to comprehensively assess the implementation level of VDC technologies in the planning, adoption, technology, and performance areas (Kam et al., 2017). In addition to the evaluation frameworks for BIM, existing frameworks to assess emerging technologies have been developed to support informed decisions about (1) whether or not the project teams or companies will apply specific technologies and (2) which technologies they will apply for their projects. To do this, the assessment frameworks must consider comprehensive aspects of implementing specific technologies in a consistent, flexible, and quantitative manner (Succar, 2010).

The authors also reviewed the existing evaluation frameworks for virtual technologies from other industries. Grübel et al. (2017) developed a framework to evaluate the navigation performance in a virtual environment. Evaluation frameworks formalized by Sidani et al. (2019) and Li et al. (2012) assess visualization and representation performance of VR software. In addition, Sutcliffe and Kaur (2000) developed a walkthrough method that uses a goal-oriented task action, exploration, and navigation in the virtual environment to evaluate the usability of VR user interfaces. Meanwhile, Zhang and Zhao (2005) proposed a hierarchical framework for evaluating the collaboration in the virtual environment. After carrying out tasks required in a typical usage situation, they collected responses from a questionnaire they created based on the framework to determine the usefulness of VR collaboration. Some researchers have also built various assessment systems for industry-specific VR applications. For example, Krohn et al. (2020) developed a VR-check framework and its detailed workflow for evaluating the characteristics and quality of VR applications in the clinical neuropsychology field.

However, existing evaluation frameworks from other industries have limitations in directly applying BIM-based VR technologies. As they are not specialized to account for comprehensive aspects of BIM-based VR applications, they cannot support informed decision making on the applications. To overcome this limitation, the research proposed the development of an
evaluation framework tailored for BIM-based VR applications, focusing on the design phase.

3. Evaluation Framework of BIM-Based VR Applications

This section describes the framework developed in this research. To formalize the framework, the research team reviewed existing studies about evaluation criteria and metrics for VR applications and carried out interviews with experts to finalize the criteria, metrics, and their hierarchies. Based on the mutually exclusive and collectively exhaustive principle, the research team first classified entire VR implementation processes into three mutually exclusive stages: the commissioning, usage, and post-usage stages (Yu et al., 2015; Zakharchenko et al., 2016). The team also segmented the three stages into five areas (i.e. preparation, immersion, inspection, collaboration, and side effect) based on the main concerns of each stage. The research team identified candidate evaluation criteria and metrics for each area from existing studies and interviews with seven experts (i.e. three architects, one surveyor, one building service engineer, one project manager, and one structural engineer) who have worked on building design projects for more than 10 years (Bystrom et al., 1999; Carruzzo & Bergamasco, 2010; Seo et al., 2016; Mütterlein, 2018; Wienrich et al., 2018; Harms, 2019; Soler-Dominguez et al., 2020). Finally, the comprehensive criteria and metrics were tailored for the design phase of BIM-based VR applications in a collectively exhaustive manner.

The framework consists of three stages (Fig. 1). First, the focus of the commissioning stage is on the issues occurring before implementing the immersive virtual environment. VR applications present the information included in the BIM files to the VR devices, and the stakeholders in the AEC industry perceive realistic visual impressions of the designed buildings in the immersive virtual environment via the devices (Wang et al., 2019). Thus, in the commissioning stage, the framework is required to consider two aspects: the communication between the VR applications and devices (i.e. software and hardware) as well as the conversion of information from BIM files to the immersive virtual environment. Second, users take all benefits from implementing the BIM-based VR applications during the usage stage. For example, users perceive realistic impressions of the buildings and take the actions required to enhance the quality of building designs (Vila et al., 2003). Thus, the evaluation criteria considered in this stage must assess the usefulness of the VR applications. Third, the severity and frequency of side effects after using the VR applications have a substantial impact on usability. For instance, several scholars have reported that approximately 61% of users experienced symptoms of malaise, such as dizziness and fatigue, after 20 minutes in the immersive virtual environment (Wang & Dunston, 2006). Thus, the evaluation criteria considered in this stage need to assess the degree of side effects.

The framework contains one area for each of the commissioning and post-usage stages (i.e. preparation and side effect areas). The criteria and metrics related to the usage stage are included in the three areas (i.e. immersion, inspection, and collaboration areas). The five areas of the framework include 14 criteria and 29 metrics to assess BIM-based VR applications for the design phase in a comprehensive and consistent manner. The following subsections describe the assessment criteria and metrics included in each area in detail.

3.1 Preparation

The preparation area, which considers tasks carried out during the commissioning stage, focuses on cross-platform operability (Szpak et al., 2019). This area consists of two criteria: the VR equipment and the BIM file. As Fig. 2 illustrates, the VR equipment includes two detailed metrics (i.e. compatibility and convenience) while the BIM file includes three metrics (i.e. file format compatibility, model complexity, and convenience). As users prepare the VR devices for implementing the immersive virtual environment, the first criterion—the VR equipment—assesses the characteristics of the BIM-based VR applications related to the preparation of the VR devices. Each application accepts different VR devices; therefore, users are required to select the appropriate VR devices, in which the application operates steadily without any problems related to function limitations, forced termination, freezing, and frequent crashes. Thus, it is important to assess the performance of the application based on how many appropriate VR devices for a specific application exist and how many steps are required to appropriately install the devices. To this end, the VR equipment criterion is evaluated by two metrics: compatibility and convenience. Compatibility is proportional to the number of appropriate VR devices for the application, and convenience is inversely proportional to the complexity of merging the devices with the application. For example, if more steps are required to install the devices, the application will have a lower score on the convenience of the VR equipment criterion.

The second criterion in the preparation area is the BIM file, which contains geometry, materials, and manufacturer-specific
3.2 Immersion

Three areas—immersion, inspection, and collaboration—are proposed to grasp the outcomes that appear during the usage stage. Immersion describes the physical feeling of the realism of BIM models in the immersive virtual environment (Kato et al., 2000). The feeling of immersion involves several kinds of percipience (e.g. visual, auditory, and haptic sensations). However, as the application of VR in the AEC industry has a limited demand for auditory and haptic feedback, current BIM-based VR applications provide predominantly visual feedback. Thus, the immersion area includes the visual sensation as a single criterion, which indicates how close the visual output in the virtual environment is to real-world visual stimuli (Fig. 3).

To evaluate the criterion, the research team adopted two metrics (i.e. render quality and graphical quality), which are mainly used to assess VR-specific immersion quality (Yu et al., 2015; Zakharenko et al., 2016; Xu et al., 2018). The first metric, rendering quality, assesses the visual effect resulting from visible features (e.g. shading, texture-mapping, and reflection; Pardo et al., 2018). In addition, graphical quality evaluates the visual clarity, which depends on the quality of graphical outputs (e.g. resolution and frame rates; Blanchard et al., 1990; Nayyar et al., 2018).

3.3 Inspection

The inspection area evaluates the applicability of functions that enable users to view and interact with BIM objects. Using these functions, users can observe and manipulate parametric components in the virtual environment from multiple perspectives (Sampaio et al., 2012; EnscapeTM, 2016). The inspection area includes six criteria: walkthrough, navigation, object manipulation, annotation, acquisition of geometric information, and acquisition of nongeometric information (Fig. 4).

Each criterion consists of the same two metrics: capability and convenience.

The first criterion, walkthrough, assesses desirable characteristics of orienting oneself and roaming in the immersive virtual environment (Stappers et al., 2001; Thabet et al., 2002; Cherian, 2016; EnscapeTM, 2016). The capability metric assesses how extensively the BIM-based VR application covers the functions required for walkthrough. The second metric, convenience, is inversely proportional to the complexity and response time of operating the walkthrough-related functions (Jackson & Fagan, 2000).

The next criterion, navigation, evaluates the functions of the BIM-based VR application in searching for the destination when the subject has no prior knowledge of the whereabouts of the targeted destination (Carlsson & Hagsand, 1993). The difference between walkthrough and navigation is that walkthrough does not have a destination, but only moves through the immersive virtual environment.

The third criterion is object manipulation, which assesses the functions of interacting with all kinds of BIM objects and surrounding elements in the virtual environment. Users can adjust some default settings of the BIM model in the virtual environment, such as moving the object locations (Du et al., 2018). Adjustments to the weather or light in the virtual environment are also considered in this criterion (Shi et al., 2016). The next criterion, annotation, evaluates the judging, marking, and noting functions of the VR applications (Du et al., 2017). The last two
3.4 Collaboration

The realistic visualization of BIM-based designs in the immersive virtual environment enhances the understanding of multiple stakeholders, which supports the reduction of conflicts among them. As stakeholders often work remotely with each other, remote collaboration through the BIM-based VR application becomes increasingly important. The collaboration area consists of three criteria: multiple users, real-time communication, and real-time modification (Fig. 5). In addition, each criterion is assessed by the same two metrics as the inspection area (i.e. capability and convenience).

The first criterion about multiple users describes the function that provides a platform where participants can meet in a shared virtual immersive environment online. This function could accommodate users walking in and/or interacting with the virtual environment (Sampaio, 2018). Real-time communication is a remote communication method used within the BIM-based VR application (Du et al., 2017, 2018). Finally, the criterion for real-time modification assesses how extensively and conveniently the VR applications support the collaboration among multiple project participants to update the BIM model investigated in the VR environment. This criterion considers the functions of real-time modification directly within the VR applications as well as from the BIM authoring software synchronized with the VR applications (Heydarian et al., 2015; Du et al., 2018).
3.5 Side effects

Despite continued improvements in VR technologies, users still experience a few negative effects after using immersive VR applications (Lampton et al., 1994; Regan, 1995, 1997). To focus on this concern in the post-usage stage based on the mutually exclusive principle (Yu et al., 2015; Zakharchenko et al., 2016), the side effect area includes the criteria that directly refer to the symptoms of malaise after using the applications instead of causes (e.g. the time lag of visual information, unusual motion) of symptoms experienced while using the VR applications (Chang et al., 2020). The research team first collected a variety of symptoms by reviewing the existing literature and then determined two common side effects (i.e. dizziness and fatigue; Fig. 6) that are specifically important for BIM-based VR applications (LaViola, 2000; Chang et al., 2020). Dizziness means that the users feel dizzy, vertigo, and disorientation when experiencing the VR immersive condition. Meanwhile, fatigue indicates general discomfort, including headache, eyestrain, blurred vision, and difficulty concentrating. Each criterion has the same two metrics: occurrence time and severity. Occurrence time records the time a symptom occurs; severity is a self-reported measure to describe the degree of the side effects experienced (Sampaio, 2018).

4. Application of the Evaluation Framework

To demonstrate the procedures of evaluating BIM-based VR applications based on the framework and the usefulness of the results from the framework, the research team applied the framework for a case study, which is a design project developed for a three-story educational building in Hong Kong based on BIM (Fig. 7). This section describes the application procedures and results.

4.1 Application procedures

To evaluate the BIM-based VR applications based on the framework, the team designed four main task groups: (1) preparation for using VR applications, (2) design visualization, (3) design inspection, and (4) multiuser collaboration.

The first task group includes the experimental contents related to the preparation for using BIM-based VR applications. Participants are first required to connect the VR devices with the VR applications and deliver the information from BIM files to the VR applications. After that, they check if the controllers work correctly without any problems related to functional limitations and if the immersive virtual environment appropriately presents the information (e.g. color and style of components) included in the BIM file without any loss of content, data deficiencies, or data compression problems. During the preparation processes, participants assess the compatibility of BIM file formats, complexity of BIM files accepted, and convenience for the specific application. Figure 8 illustrates the experimental contents of the first task group.

The tasks included in the second group start with roaming in the immersive virtual environment to experience the

---

**Figure 6: The criteria and metrics in the collaboration area.**

**Figure 7: The three-floor BIM model.**
Evaluation framework for BIM-based VR applications

Figure 8: First task group: preparation of VR applications.

Figure 9: Second task group: navigation in the BIM model.

first-person movement or control the locomotion of the avatar. Through this experience, participants evaluate the capability and convenience of dynamic walking, which matches the criterion of walkthrough in the evaluation framework. During the experiment, the research team describes the starting and ending points (i.e. origin and destination) to the participants (Fig. 9). After being placed at the starting point, they find the routes to arrive at the ending point. During this process, participants assess navigation-supporting functions, such as plane view, map, and viewpoints.

The third group considers interaction with components in the immersive virtual environment and acquisition of information from parametric components. The adjustment of surrounding conditions in the immersive virtual environment allows users to assess their design under more realistic conditions. During the experiments, participants are required to adjust to daylight by considering the designated season and checking the building design. In addition, to assess the interaction-related functions in the BIM-based VR application, they manipulate components within the building model. Furthermore, as they can acquire both geometric and nongeometric information of the building designed through the application (Fig. 10), they evaluate their satisfaction by accessing information for default components and the difficulty of use based on their needs.

The research team designed the tasks in the last group to assess the collaboration capability of the VR application. Communication enhances the understanding of designs, which allows for the sharing of valuable feedback (e.g. identification of problems in design), thereby enhancing the collaboration among multiple stakeholders. Thus, participants communicate with others through auditory means and annotation in the immersive virtual environment and evaluate the functions related to real-time communication (Fig. 11). In addition, participants move, rotate, and hide building components to test functions of real-time modifications, enabling them to find appropriate design alternatives in a timely manner.

During the experiment, the research team records each participant’s immersion time and any symptoms of malaise experienced. This information is used to assess side effects. After participants complete the entire experimental procedure, they are
invited to fill out a questionnaire using five-point Likert-scale questions to evaluate different metrics of each application. The questions are closely connected to the predefined criteria in the evaluation framework, and scaled values are assigned to each option in the Likert questions to discriminate different levels of applicability.

### 4.2 Application results

To evaluate the BIM-based VR applications based on the framework, the research team conducted 15 experiments with 10 graduate students who had earned undergraduate degrees in architectural and civil engineering and 5 professionals with more than 5 years of work experience in the AEC industry. Each participant was asked to explore the functions and complete all experimental tasks for five different BIM-based VR applications. Based on the results collected from the questionnaire, the research team computed the average score in each area first and then calculated the average score of the five areas as an overall score. The five-point Likert-scale results were assigned to the scores along an interval from 0 to 10 (i.e., 0, 2.5, 5, 7.5, and 10), thereby enhancing the comparability between each area.
Table 1: Score of each VR application.

| Area           | Criteria                      | Metric                          | Application | 1          | 2          | 3          | 4          | 5          |
|----------------|-------------------------------|---------------------------------|-------------|-----------|-----------|-----------|-----------|-----------|
|                |                               |                                 |             | Mean      | Mean      | Mean      | Mean      | Mean      |
|                |                               |                                 |             | (S.D.)    | (S.D.)    | (S.D.)    | (S.D.)    | (S.D.)    |
| Preparation    | VR equipment                  | Compatibility                    | Application 1 | 9.2 (1.2) | 8(1.7)    | 7.2(1.9)  | 8.7(1.6)  | 7.0(2.2)  |
|                |                               | Compatibility                    | Application 2 | 9.0(1.3)  | 8.5(1.3)  | 7.2(2.1)  | 8.2(1.5)  | 6.7(2.2)  |
|                |                               | BIM format compatibility          | Application 3 | 8.5(1.6)  | 8.3(1.5)  | 7.8(1.6)  | 8.7(1.3)  | 6.5(1.8)  |
|                |                               | Model complexity                  | Application 4 | 9.2(1.2)  | 8.3(1.8)  | 8.3(1.2)  | 8.2(1.1)  | 2.2(1.9)  |
|                |                               | Conveniences                      | Application 5 | 8.7(1.6)  | 8.8(1.3)  | 7.7(2.0)  | 8.7(1.3)  | 6.7(1.8)  |
|                |                               | Final score of the preparation    | Application 1 | 8.9(0.8)  | 8.4(1.0)  | 7.6(1.3)  | 8.5(0.9)  | 5.8(1.5)  |
| Immersion      | Visual                         | Render quality                   | Application 2 | 7.8(1.3)  | 8.0(1.4)  | 6.8(2.2)  | 9.3(1.1)  | 3.5(2.3)  |
|                |                               | Graphic quality                  | Application 3 | 7.2(1.9)  | 6.2(1.9)  | 6.8(2.0)  | 9.8(0.6)  | 2.2(1.3)  |
|                |                               | Final score of the immersion      | Application 4 | 7.5(1.2)  | 7.1(1.3)  | 6.8(1.6)  | 9.6(0.6)  | 2.8(1.3)  |
| Inspection     | Wayfinding                     | Capability                       | Application 1 | 8.3(1.5)  | 7.8(2.1)  | 6.5(1.8)  | 6.5(2.3)  | 1.0(1.6)  |
|                |                               | Convenience                      | Application 2 | 8.2(2.2)  | 7.5(2.1)  | 6.3(1.9)  | 7.8(1.9)  | 1.2(1.6)  |
|                |                               | Capability                       | Application 3 | 8.0(1.4)  | 6.0(2.5)  | 5.8(1.5)  | 5.2(3.2)  | 1.0(1.6)  |
|                |                               | Convenience                      | Application 4 | 8.0(1.7)  | 5.5(2.4)  | 5.5(1.7)  | 5.7(3.3)  | 1.2(1.6)  |
|                |                               | Convenience                      | Application 5 | 6.2(2.3)  | 7.3(1.8)  | 3.3(2.4)  | 3.5(3.1)  | 0.7(1.1)  |
|                |                               | Final score of the immersion      | Application 1 | 7.5(1.2)  | 7.1(1.3)  | 6.8(1.6)  | 9.6(0.6)  | 2.8(1.3)  |
| Collaboration  | Real-time modification         | Capability                       | Application 2 | 7.2(2.5)  | 4.7(2.8)  | 4.7(3.0)  | 2.3(3.1)  | 1.5(1.8)  |
|                |                               | Convenience                      | Application 3 | 7.8(1.9)  | 5.3(2.5)  | 4.2(2.8)  | 2.5(3.0)  | 1.7(2.2)  |
|                |                               | Acquisition of geometric information | Capability | 6.0(2.3)  | 1.7(1.8)  | 4.2(3.2)  | 1.5(2.8)  | 0.5(1.0)  |
|                |                               | Convenience                      | Application 1 | 6.5(1.6)  | 2.2(2.3)  | 4.0(3.2)  | 1.7(2.8)  | 0.5(1.0)  |
|                |                               | Capability                       | Application 2 | 8.8(1.3)  | 6.7(2.4)  | 0.5(1.0)  | 0.3(0.9)  | 0.3(0.9)  |
|                |                               | Convenience                      | Application 3 | 8.8(1.9)  | 7.0(2.5)  | 0.3(0.9)  | 0.3(0.9)  | 0.2(0.6)  |
|                |                               | Final score of the inspection     | Application 4 | 7.6(0.8)  | 5.8(1.1)  | 4.1(1.6)  | 3.4(1.2)  | 0.9(0.9)  |
|                |                               | Capability                       | Application 1 | 2.2(1.6)  | 4.7(2.7)  | 1.0(1.6)  | 0.5(0.1)  | 0.0(0.0)  |
|                |                               | Convenience                      | Application 2 | 2.8(2.1)  | 4.3(3.2)  | 1.0(1.6)  | 0.7(1.1)  | 0.0(0.0)  |
|                |                               | Capability                       | Application 3 | 8.2(1.5)  | 0.8(1.5)  | 0.1(1.0)  | 0.7(1.5)  | 0.0(0.0)  |
|                |                               | Convenience                      | Application 4 | 8.5(1.6)  | 0.7(1.1)  | 0.3(0.9)  | 0.5(2.4)  | 0.0(0.0)  |
|                |                               | Capability                       | Application 5 | 7.5(1.6)  | 1.0(1.6)  | 0.3(1.0)  | 0.3(3.3)  | 0.0(0.0)  |
|                |                               | Convenience                      | Application 1 | 7.5(2.9)  | 2.0(1.2)  | 0.6(0.9)  | 0.5(3.1)  | 0.0(0.0)  |
|                |                               | Final score of the collaboration  | Application 2 | 8.2(2.4)  | 5.3(2.5)  | 7.3(3.1)  | 6.2(3.1)  | 6.5(4.3)  |
|                |                               | Occurrence frequency             | Application 3 | 8.3(1.5)  | 5.3(2.3)  | 8.0(1.9)  | 6.8(3.1)  | 9.8(0.6)  |
|                |                               | Severity                         | Application 4 | 7.7(2.6)  | 5.8(2.8)  | 7.5(3.3)  | 5.8(3.1)  | 7.0(3.9)  |
|                |                               | Occurrence frequency             | Application 5 | 8.0(2.4)  | 6.0(2.5)  | 8.3(2.0)  | 6.3(2.8)  | 9.3(1.1)  |
|                |                               | Severity                         | Application 1 | 8.0(2.1)  | 5.6(2.1)  | 7.8(2.1)  | 6.3(2.7)  | 8.2(2.0)  |
|                |                               | Final score of the side effects   | Application 2 | 7.6(0.6)  | 5.8(0.9)  | 5.4(0.8)  | 5.6(0.8)  | 3.5(0.7)  |

Final scores of the five areas and the total performance are presented in Table 1. Application 1 showed high scores in all areas. Although Application 4 received significantly higher scores than other applications in the immersion area, it received a low score in the VR inspection and collaboration areas. The assessment results of each area are described in detail next.

In the preparation area, Applications 1, 2, 3, and 4 similarly received high scores; only Application 5 had a distinctly lower rate. Applications 1, 2, 3, and 4 could identify and connect VR devices as well as render mainstream formats of BIM files automatically. In contrast, Application 5 was incompatible with VR devices and required users to manually convert the BIM files to rendered files in the application.

For the immersion evaluation, the results showed that Application 4 had the best performance because it was the only application to present real textures of BIM components and render detailed light and shadow effects to enhance visualization. The incompatibility problems of Application 5 caused data loss from the BIM files, which led to an extremely low user immersion level. The BIM model presentation of Applications 1, 2, and 3 allowed users to experience the immersive virtual environment to a certain degree.

In the inspection area, the five applications were divided into two levels based on their scores. Applications 1 and 2 had high scores, indicating that they could meet most participants’ basic needs. These two VR applications had their own unique strengths in the inspection area. Application 1 could transfer and present more comprehensive BIM information in the virtual environment. Participants could acquire various pieces of non-geometric information. Application 2 showed good performance for object manipulation, which provides numerous functions of interacting with BIM components (e.g. switching the lights) during the inspection process. In contrast, Applications 3, 4, and 5 could not offer practicable inspection functions for participants.

The collaboration area was the most immature area in the BIM-based VR applications. Only Application 1 created an acceptable platform for BIM collaboration. Multiple users could communicate in a virtual environment, yet the real-time...
modification was limited to highlighting or hiding components without further adjustment. Applications 2, 3, and 4 provided minimal collaborative functions. Finally, Application 5 did not provide collaboration capability in the virtual environment.

The last evaluation area is the side effects, including dizziness and fatigue. Most participants did not have obvious adverse reactions after testing Applications 1, 3, and 5. However, some participants indicated experiencing typical discomfort after testing Applications 2 and 4. Application 2 offered an unfamiliar mode of movement (i.e. flying). Furthermore, some motions (e.g. going through the doors) would conduct visual shaking. These are the reasons for users' unpleasant symptoms after 20 minutes of testing. For Application 4, several participants commented that the excessively abundant light and shadows during the observation caused dizziness. Figure 12 shows a spider diagram for comparison of scores for each area of the five applications.

5. Discussions

The research team proposed a framework to help users make informed decision when they choose a BIM-based VR application based on their preferences. To discuss the usefulness of the framework, the team assessed the consistency, comprehensiveness, and practical effectiveness of the evaluation framework for BIM-based VR applications.

First, the research team used the results of the 15 experiments detailed in the previous section to discuss the consistency of the evaluation framework. To select the most appropriate application, the consistent prioritization of applications is necessary. Thus, the team used Kendall’s coefficient of concordance (i.e. Kendall’s W) to assess the consistency of the framework in a quantitative manner (Zhao et al., 2012; Sánchez Calleja et al., 2016). This coefficient measures agreement among several participants for the rankings of evaluation objects. Therefore, the coefficient indicated the consistency in the rankings of the applications among the 15 participants. The null hypothesis of Kendall’s test was established.

H0: The evaluation results from the 15 participants are inconsistent.

The significance level was 0.05, and the results are presented in Table 2. The $P = 0.000$, which is lower than the significance level (0.05). Thus, the null hypothesis was rejected, and the hypothesis (i.e. the evaluation results from the 15 participants are consistent) was accepted. Furthermore, Kendall’s W is 0.788; normally, Kendall’s W ranges from 0.6 to 1 to indicate consistency (Zhao et al., 2012; Sánchez Calleja et al., 2016). Thus, the results from the 15 participants who used the evaluation framework are consistent in the rankings of these five VR applications. The team also computed Kendall’s W for students and professionals separately (Table 2). P values (0.001) for both groups are lower than the significance level (0.05). Kendall’s W values for 10 students and 5 professionals are 0.771 and 0.928, respectively. Although the results from both student and professional groups based on the evaluation framework indicate consistency in the ranking of the five VR applications, the results from the professional group are more consistent than those from the student group.

However, Table 1 shows that the standard deviations of scores in the framework range between 0 and 3.9. This means that inconsistency in the scores of different participants still exists with a variety of ranges even for the same applications. Specifically, metrics in the inspection and side effect areas were assessed with higher variability (i.e. less consistent) than those in the other areas.

To assess the comprehensiveness of the evaluation framework, the research team conducted three semistructured interviews with experts who have more than 5 years of experience working with BIM design and VR implementation. The interviewees rated the levels of comprehensiveness for each area in the framework on a scale ranging from 0 to 5. They indicated that
no additional criteria or metrics were necessary for the preparation, immersion, inspection, and side effects areas and, thus, the average scores for each of the four areas were 5.0. However, the average score for the collaboration area was 4.3. Two interviewees commented that, although the existing BIM-based VR applications lacked functions related to real-time modification, applications would increasingly enrich those functions in the near future and, thus, the framework would have to account for more extensive functions related to the real-time modification as its criteria. Thus, although the proposed evaluation framework was comprehensive for existing BIM-based VR applications, it will need to incorporate adaptive criteria for the future evaluation of advanced VR applications.

To address the practical effectiveness of the evaluation framework, the research team interviewed the five professionals who participated in the experiments. The five interviewees rated the levels of practical effectiveness of the framework on a scale ranging from 0 to 5, and the average score was 4.6. All interviewees stated that, because the framework provides the evaluation results in a comprehensive, structured, and quantitative manner, the framework would be able to support informed decision making on the BIM-based VR application based on their priority in a flexible manner. For example, although Application 4 received low scores for the inspection and collaboration areas, it obtained a significantly high score in the immersion area. As this score reflects how close the visual output in the virtual environment is to real-world visual stimuli, it would be a suitable choice for users whose focus of using the BIM-based VR application is on assessing a building design with a realistic visual effect (Kato et al., 2000; Nayyar et al., 2018; Pardo et al., 2018). In addition, as Application 1 obtained a high score in all areas, it is recommended for a project team with a variety of objectives for using the BIM-based VR application, including the retrieval of properties of objects and manipulation of BIM components (Sampiao, 2018).

However, the interviewees also addressed the limitations and suggestions of the framework to enhance its practical effectiveness. First, although the framework supports the provision of a certain degree of consistent assessment results, inconsistency in scores for each metric still exists. They mentioned that there might be two main causes of inconsistency: different rating rationales of participants and other different characteristics of participants (e.g. medical conditions). Thus, they suggested that the rating system (i.e. five-point Likert-scale) in the questionnaire be elaborated by providing reference values or examples for each score of the questions to improve the consistency of results. In addition, to consider other different characteristics of participants in a flexible manner, it is necessary to present adjusted results based on the characteristics. To this end, the questionnaire should include questions asking about experiment participants’ characteristics (e.g. medical conditions). Second, the experimental tasks were completed by a small number of students and professionals, so additional experiments with additional participants are required to ensure the validity of the experimental results and their consistency. The third limitation is that the team used only an educational building for the experiments. As users may have different expectations of using the BIM-based VR applications for different types of buildings, the consideration of additional types of buildings will enhance the validity of results from the experiments. The last limitation relates to the usage phase of this evaluation framework. The focus of the evaluation framework is on the design phase because the utilization of BIM-based VR technologies is relatively mature in design tasks (Du et al., 2017; Zou et al., 2017), yet BIM-based VR applications are increasingly being developed for other phases. For example, BIM-based VR applications can be used to recognize inherent risks and hazards in the construction process during the construction phase (Germani et al., 2009, 2012). Thus, the evaluation framework for the applications should be extended to cover the whole lifecycle.

6. Conclusion

The integration of BIM and VR technologies has a great potential to contribute to many aspects (e.g. design and education quality, safety) in the AEC industry. For instance, as inspecting design components of buildings in the immersive virtual environment enhances the understanding of designs, which allows for the identification of problems in advance, BIM-based VR technologies support the improvement of design quality. With the rapid emergence of these technologies, project teams and institutions must make informed decisions on which applications they will adopt depending on their preferences. However, as the existing evaluation frameworks for innovative technologies are not tailored for BIM-based VR applications, they have difficulty making informed decisions when selecting applications in a flexible manner.

Such informed decisions need consistent assessment results based on comprehensive criteria. Thus, the research team extensively reviewed existing studies to identify candidate evaluation criteria and metrics and conducted interviews with experts to finalize the areas, criteria, metrics, and their hierarchies. In addition, as the first step for developing the evaluation framework for BIM-based VR applications, the team focused on the design phase of the project for using the framework. Finally, the team proposed an assessment framework for BIM-based VR applications, which consists of 3 stages, 5 areas, 14 criteria, and 29 metrics.

To assess the usefulness of the framework, the team applied it to five BIM-based VR applications using a BIM-based design project for a three-story educational building in Hong Kong. After carrying out the tasks designed by the research team, the participants responded to the questionnaire developed based on the framework. The team also interviewed professionals to assess the comprehensiveness and practical effectiveness of the framework. The application results indicated that the framework provided consistent results for comprehensive evaluation criteria and metrics in a quantitative manner, which would support users in making informed decisions on the applications flexibly depending on users’ preferences. In particular, Kendall’s W for the participants was 0.788, which indicates good consistency in the rankings of the five BIM-based VR applications.

However, this research still has several limitations. First, the questionnaire’s rating system (i.e. five-point Likert scale) should improve the consistency of the results by providing reference values or examples for each score of the questions. Moreover, the research carried out the experiments with a small number of participants and used a case study about an educational building. Thus, extensive experiments with practitioners and other types of buildings are required to enhance the generality of the assessment results. In addition, the evaluation framework focused on the design phase when the utilization of the BIM-based VR applications is relatively mature. Consequently, the consideration of additional criteria and/or metrics pertaining to other phases (e.g. construction phase) when using the applications will contribute to extending the framework to the whole lifecycle.
Acknowledgment

This work was supported by a grant (Research Project No.: 6000703) from City University of Hong Kong.

Conflict of interest statement

None declared.

References

Aranda-Mena, G., & Wakefield, R. (2006). Interoperability of building information - Myth of reality. In Proceedings of the 6th European Conference on Product and Process Modelling - EWork and EBusiness in Architecture, Engineering and Construction (ECPPM 2006) (pp. 127–134).

Blanchard, C., Burgess, S., Harvill, Y., Lanier, J., Lasko, A., Oberman, M., & Teitel, M. (1990, February). Reality built for two: A virtual reality tool. In Proceedings of the 1990 Symposium on Interactive 3D Graphics (pp. 35–36).

Bystrom, K. E., Barfield, W., & Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. Teleoperators and Virtual Environments, 8(2), 241–244.

Calderon-Hernandez, C., & Brioso, X. (2018). Lean, BIM and augmented reality applied in the design and construction phase: A literature review. International Journal of Innovation, Management and Technology, 9(1), 60–63.

Carlsson, C., & Hagsand, O. (1993). DIVE—a multi-user virtual reality system. In Proceedings of IEEE Virtual Reality Annual International Symposium (pp. 397–400).

Carrozzino, M., & Bergamasco, M. (2010). Beyond virtual museums: Experiencing immersive virtual reality in real museums. Journal of Cultural Heritage, 11(4), 452–458.

Chang, E., Kim, H. T., & Yoo, B. (2020). Virtual reality sickness: a review of causes and measurements. International Journal of Human–Computer Interaction, 36(17), 1658–1682.

Cherian, S. (2016). Integrating BIM and Virtual Reality – The Next Big Thing. https://www.linkedin.com/pulse/integrating-bim-virtual-reality-next-big-thing-sarita-cherian. Accessed 14 December 2020.

Du, J., Zou, Z., Shi, Y., & Zhao, D. (2018). Zero latency: Real-time synchronization of BIM data in virtual reality for collaborative decision-making. Automation in Construction, 85, 51–64.

Du, J., Zou, Z., Shi, Y., & Zhao, D. (2017). Simultaneous data exchange between BIM and VR for collaborative decision making. In Congress on Computing in Civil Engineering, Proceedings (pp. 1–8).

EnscapeTM - Real-Time Rendering for Revit, SketchUp, Rhino & ArchiCad. (2016). https://enscape3d.com. Accessed 14 December 2020.

Germani, M., Mengoni, M., & Peruzzini, M. (2012). An approach to assessing virtual environments for synchronous and remote collaborative design. Advanced Engineering Informatics, 26(4), 793–813.

Germani, M., Mengoni, M., & Peruzzini, M. (2009). Metrics-based approach for VR technology evaluation in styling product design. In Proceedings of the ASME Design Engineering Technical Conference, 49026 (pp. 1325–1339).

Grübel, J., Thrash, T., Hölischer, C., & Schinazi, V. R. (2017). Evaluation of a conceptual framework for predicting navigation performance in virtual reality. Plos One, 12(9), e0184682.

Harms, P. (2019). Automated usability evaluation of virtual reality applications. ACM Transactions on Computer-Human Interaction, 26(3), 1–36.

Herr, C., & Fischer, T. (2019). BIM adoption across the Chinese AEC industries: An extended BIM adoption model. Journal of Computational Design and Engineering, 6(2), 173–178.

Heydarian, A., Carneiro, J. P., Gerber, D., Becerik-Gerber, B., Hayes, T., & Wood, W. (2015). Immersive virtual environments versus physical built environments: A benchmarking study for building design and user-built environment explorations. Automation in Construction, 54, 116–126.

Jackson, R. L., & Fagan, E. (2000). Collaboration and learning within immersive virtual reality. In Proceedings of the Third International Conference on Collaborative Virtual Environments (pp. 83–92).

Kam, C., Song, M. H., & Senaratna, D. (2017). VDC scorecard: Formulation, application, and validation. Journal of Construction Engineering and Management, 143(3), 04016100.

Kato, H., Billinghamurst, M., PoupvreY, I., ImamotO, K., & Tachibana, K. (2000). Virtual object manipulation on a table-top AR environment. In Proceedings - IEEE and ACM International Symposium on Augmented Reality, ISAR 2000 (pp. 111–119).

Krohn, S., Tromp, J., Quinque, E. M., Belger, J., Klotzsche, F., Rekers, S., Choejecki, P., de Mooij, J., Akbal, M., McCull, C., Villerger, A., Gaebler, M., Finke, C., & Thöne-Otto, A. (2020). Multidimensional evaluation of virtual reality paradigms in clinical neuropsychology: Application of the VR-check framework. Journal of Medical Internet Research, 22(4), e16724.

Lampton, D. R., Kolasaki, E. M., Knerr, B. W., Bliss, J. P., Bailey, J. H., & Witmer, B. G. (1994). Side effects and aftereffects of immersion in virtual environments. Proceedings of the Human Factors and Ergonomics Society, 38(18), 1154–1157.

LaViola, J., (2000, November). MSVT: A virtual reality-based multimodal scientific visualization tool. In Proceedings of the third IASTED international conference on computer graphics and imaging, (pp. 1–7).

Li, H., Chan, G., & Skitmore, M. (2012). Visualizing safety assessment by integrating the use of game technology. Automation in Construction, 22, 498–505.

Martens, D. (2016). Virtually usable: A review of virtual reality usability evaluation methods. https://danamartensmfadt.files.wordpress.com/2016/08/virtually usable.Pdf. Accessed 14 December 2020.

Müterlein, J. (2018). The three pillars of virtual reality? In Investigating the Roles of Immersion, Presence, and Interactivity. Proceedings of the 51st Hawaii International Conference on System Sciences.

Nayyar, A., Mahapatra, B., Le, D., & Suseendran, G. (2018). Virtual reality (VR) and augmented reality (AR) technologies for tourism and hospitality industry. International Journal of Engineering & Technology, 7(2.21), 156–160.

Pardo, P. J., SuerO, M. I., & Pérez, Á. L. (2018). Correlation between perception of color, shadows, and surface textures and the realism of a scene in virtual reality. JOSA A, 35(4), B130–B135.

Regan, C. (1995). An investigation into nausea and other side effects of head-coupled immersive virtual reality. Virtual Reality, 1(1), 17–31.

Regan, C. (1997). Some effects of using virtual reality technology. Virtual Reality, Training’s Future, 65(6), 77–83.

Sampaio, A. (2010). Virtual reality technology used in civil engineering education. The Open Virtual Reality Journal, 2(1), 18–25.

Sampaio, A. Z. (2018). Enhancing BIM methodology with VR technology. State of the Art Virtual Reality and Augmented Reality Knowhow, IntechOpen, London, UK, 59–79.

Sampaio, A. Z., Gomes, A. R., & Santos, J. P. (2012). Management of building supported on virtual interactive models:
Construction planning and preventive maintenance. Electronic Journal of Information Technology in Construction, 17, 121–133.

Sánchez Calleja, L., García Jiménez, E., & Rodríguez Gómez, G. (2016). Evaluation of the design of the aedem programme for emotional education in secondary education. Revista Electrónica de Investigación y Evaluación Educativa, 22(2), 1–10.

Sebastian, R., & van Berlo, L. (2010). Tool for benchmarking BIM performance of design, engineering and construction firms in the Netherlands. Architectural Engineering and Design Management, 6(Special Issue), 254–263.

Seo, N. J., Arun Kumar, J., Hur, F., Crocher, V., Motawar, B., & Lakshminarayanan, K. (2016). Usability evaluation of low-cost virtual reality hand and arm rehabilitation games. Journal of Rehabilitation Research and Development, 53(3), 321–334.

Shi, Y., Du, J., Lavy, S., & Zhao, D. (2016). A multiuser shared virtual environment for facility management. Procedia Engineering, 145, 120–127.

Shin, J., Rajabifard, A., Kalantari, M., & Atazadeh, B. (2020). Applying BIM to support dispute avoidance in managing multi-owned buildings. Journal of Computational Design and Engineering, 7(6), 788–802.

Sidani, A., Dinis, F. M., Sankudo, L., Duarte, J., Santos Baptista, J., Poças Martins, J., & Soeiro, A. (2019). Recent tools and techniques of BIM-based virtual reality: A systematic review. Archives of Computational Methods in Engineering, 28(2), 449–462.

Soler-Domínguez, J. L., de Juan, C., Contero, M., & Alcañiz, M. (2020). I walk, therefore I am: A multidimensional study on the influence of the locomotion method upon presence in virtual reality. Journal of Computational Design and Engineering, 7(5), 577–590.

Stappers, P. J., Saakes, D., & Adriaanse, J. (2001). On the narrative structure of Virtual Reality walkthroughs, Proceedings of the 9th International Conference on Computer Aided Architectural Design Futures, Eindhoven, Netherlands. 125–138.

Sucar, B. (2010). Building information modelling maturity matrix. Handbook of research on building information modelling and construction informatics: Concepts and technologies (pp. 65–103).

Sun, C., Hu, W., & Xu, D. (2019). Navigation modes, operation methods, observation scales and background options in UI design for high learning performance in VR-based architectural applications. Journal of Computational Design and Engineering, 6(2), 189–196.

Sutcliffe, A. G., & Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. Behaviour and Information Technology, Behaviour & Information Technology, 19(6), 415–426.

Szpak, A., Michalski, S. C., Saredakis, D., Chen, C. S., & Loetscher, T. (2019). Beyond feeling sick: The visual and cognitive aftereffects of virtual reality. IEEE Access, 7, 130883–130892.

Teizer, J., Cheng, T., & Fang, Y. (2013). Location tracking and data visualization technology to advance construction ironworkers’ education and training in safety and productivity. Automation in Construction, 35, 53–68.

Thabet, W., Shiratuddin, M. F., & Bowman, D. (2002). Virtual reality in construction: a review. Engineering computational technology. Civil-Comp press, 25–52.

Vila, J., Becue, B., & Anandikar, S. (2003). The gender factor in virtual reality navigation and wayfinding. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, HICSS 2003.

Wang, D., Guo, Y., Liu, S., Zhang, Y., Xu, W., & Xiao, J. (2019). Haptic display for virtual reality: progress and challenges. Virtual Reality & Intelligent Hardware, 1(2), 136–162.

Wang, X., & Dunston, P. S. (2006). Compatibility issues in augmented reality systems for AEC: An experimental prototype study. Automation in Construction, 15(3), 314–326.

Wienrich, C., Döllinger, N., Kock, S., Schindler, K., & Traupe, O. (2018). Assessing user experience in virtual reality – A comparison of different measurements. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)(pp. 573–589).

Xu, M., Li, C., Chen, Z., Wang, Z., & Guan, Z. (2018). Assessing visual quality of omnidirectional videos. IEEE Transactions on Circuits and Systems for Video Technology, 29(12), 3516–3530.

Yu, M., Lakshman, H., & Girod, B. (2015, September). A framework to evaluate omnidirectional video coding schemes. In 2015 IEEE International Symposium on Mixed and Augmented Reality(pp. 31–36).

Zaker, R., & Coloma, E. (2018). Virtual reality-integrated workflow in BIM-enabled projects collaboration and design review: A case study. Visualization in Engineering, 6(1), 1–15.

Zakharchenko, V., Choi, K. P., & Park, J. H. (2016, September). Quality metric for spherical panoramic video. In Optics and Photonics for Information Processing X (Vol. 9970, p. 99700C).

Zhang, H.-Y., & Zhao, Z.-X. (2005). A hierarchical framework for visualising and simulating supply chains in virtual environments. International Journal of Automation and Computing, 2(2), 144–154.

Zhao, X. G., Lin, J., Jiang, J. H., Wang, Q., & Ng, S. H. (2012). Validity and reliability of a method for assessment of cervical vertebral maturation. The Angle Orthodontist, 82(2), 229–234.

Zou, H., Li, N., & Cao, L. (2017). Emotional Response-Based approach for assessing the sense of presence of subjects in virtual building evacuation studies. Journal of Computing in Civil Engineering, 31(5), 04017028.