Systematic Literature Review of Virtual Reality Instruments for User Experience in Education Research Studies

Abstract

Virtual reality (VR) has integrated many facets of our daily lives. This virtual environment has also made its mark in the business and industry sectors. Many research studies have yielded valuable results that have galvanized this technology to the next level. However, the education sector is still lagging, and further research and analysis are needed to fully understand how to integrate VR. Hesitation may be due to the lack of a standardized protocol for virtual reality instruments testing user experience in the virtual environment. Instruments such as questionnaires and surveys are critical aspects of research studies. This paper presents a synthesis of VR studies and instruments that have been utilized. The results indicate a lack of existing instruments. Although existing instruments may be well studied, the prevalence of instruments used in assessing VR technology studies is researcher-made instruments with unknown qualities.

Introduction

In recent years, the prominent rise of a new generation of virtual reality (VR) systems has opened up new methodologies and interventions for researchers in a vast number of fields. As a new computing paradigm that redefines the interface between humans and computers by enabling them to experience reality at inconceivable levels. VR is not just an interactive multimedia tool but also a learning environment that extraordinarily emulates reality. As a result, it may significantly improve traditional learning. Therefore, various sectors range from physical exercise and physical training (Shaw et al., 2015; Yoo & Kay, 2016), high-fidelity response studies (Dahlquist et al., 2007; Gold et al., 2006), education (Lee, 2012; Bhayani & Andriole, 2005) among many others have incorporated this tool into their training, education, and way of teaching. Invariably responses on questionnaires are typically used to collect pre- and post-experience metrics. In addition, instruments in the form of surveys or questionnaires are commonly used in developing and assessing virtual reality experiences, whether the experience is for leisure or work. However, compared to other technological advances for instructional aids, there has been little empirical research on the utility of VR instruments.

Purposes

Currently, there is no standard virtual reality instrument(s) explicitly designed to test the VR experience as a technology integration tool into the existing curriculum. The motivation behind this paper is to investigate whether authors employ comparable terminology using a questionnaire presentation and response collection mechanism. Special attention is given to how authors use questionnaire presentation and response collection mechanisms. It is further explored whether authors use similar terminology and reflect their preferences. In its current state, it is currently difficult to make comparisons between VR user studies and other research studies

because VR user research has not yet developed well-defined administration procedures, classification schemes, or standard toolkits for presenting questionnaires. There is no standard for presenting questionnaires in VR user research studies to guide these considerations and facilitate comparisons with other studies. VR user research only has a narrow range of standard administration procedures, classification schemes, and toolkits for presenting questionnaires in VR studies. Therefore, this study aims to explore the various VR instruments utilized in case studies targeted toward VR technology specifically. The focus will be on literature from 2011 to 2021 and will be evaluating instruments for virtual reality technology in educational settings. For the purpose of this study, educational systems shall be defined as K-12 and higher education.

This work is based on the following research questions to build an understanding of the current instruments utilized in research studies:

RQ. 1 – Which VR instruments exist in the literature?

RQ. 2 – Is there a frequency of particular VR instruments used?

RQ. 3 – Is there an existing VR experience instrument in literature?

Theoretical Framework

In evidence-based research, questionnaires are crucial sources of information (Babbie, 1990; Field & Gournay, 2003; Lazar et al., 2017). Questionnaires represent self-reports and collect the participants' subjective experiences (Field & Gournay, 2003). In VR user research, questionnaires are typically used to investigate subjective responses to a distinctive experience. There are two formats in which questions can be displayed: structured and unstructured (Rosenthal & Rosnow, 1984). Unstructured or open-ended questions enable the subjects to respond freely, while structured questions allow for precise categorization (Saris & Gallhofer, 2014). Unstructured or open-ended questions assist the researcher in gaining a better insight into

the participant's ideas, sentiments, and emotions by enabling them to express themselves in their own words, which may necessitate greater effort on the side of the participant. Survey methods have acquired considerable attention in the literature, and their concessions and shortcomings have been extensively discussed (Baker et al., 2002; Bhattacherjee, 2012; Shaughnessy & Zechmeister, 1985). Questionnaires that are both reliable (measurement consistency) and validated (measuring the correct concept) are essential for reproducible and consistent research (Bhattacherjee, 2012). Choi and Pak (2004) classified potential biases into three categories: question design, questionnaire design, and administration. Question design deals with the ramifications of poor wording, such as jargon, double-barreled, or double-negative questions, negative phrasing, or language that pushes decisions (Choi & Pak, 2005; Lazar et al., 2017). Biases in questionnaire design are caused by the design and complexity of the surveys (Aday & Cornelius, 2006; Choi & Pak, 2005), as well as the duration and format of the questions (Bradburn & Sudman, 1979). In addition, context-dependent forgetfulness (Abernethy, 2010; Godden & Baddeley, 1975) caused by environmental change (Pohl & Phol, 2001) can cause biased responses. It is contended that consistent delivery of surveys can mitigate a sequence of random inaccuracies, notably in immersive contexts.

Methods

Webster and Watson's (2002) and Kitchenham et al. (2009) systematic literature reviews have been widely utilized in order to acquire comprehensive insight into a specific research field. The main characteristics of systematic literature reviews are:

1. Implicitly define the research question that the study will attempt to address.
2. Establish a precise and reproducible method for achieving the stated objectives.

3. Ensure that all relevant studies that meet the eligibility criteria are included in the search string.
4. Provide an analysis of the quality and validity of the specified studies.
5. The synthesis and presentation of data extracted from specified studies should be systematic.
6. Scientific and decision-making purposes should be served by making the study findings accessible.

Data Source

The following five databases were used to search for research studies utilizing questionnaires about virtual reality technology: Academic Search Complete, APA PsycInfo, Computer Source, ERIC, and Library & Information Science Source, as shown in Table 1. The keywords used in the search criteria for publications were virtual reality, VR, K-12, Education, higher education, or college or university, as shown in Table 1. The inclusion and exclusion criteria were determined in the next stage and implemented in the search, shown in Table 2. First, a preliminary filter was completed by reading the titles of publications relevant to VR applications in education, followed by a more extensive filter by skimming abstracts, findings, and conclusions. In the second stage of filtering, many publications related to VR applications in the medical field and education, systematic reviews, non-head mounted displays (HMD), dissertations, VR implementation projects, teacher training with VR, and preservice teacher training with VR were eliminated. Finally, secondary references were extensively evaluated, especially in the recent systematic review of papers with crucial findings concerning this review. The primary analysis of the selected articles was based on the higher education audience, if a VR

application was tested, the context in which the VR application was utilized, and whether a VR technology questionnaire was provided at the research study's conclusion.

Results

Fifty-nine studies were selected and analyzed, eight of which yielded instruments measuring VR technology for integration. Of the twenty-one analyzed, nine did not measure user experience through a form of an instrument, and four were not validated. Many instruments which were reviewed were adapted from other instruments. Instruments which were not created with virtual reality in mind. The last eight instruments which were evaluated, were adapted from existing instruments not generally used for a virtual environment. Table 3 provides a synopsis of each instrument.

Significance of Study

This systematic literature review's central focus was to identify instruments utilized in VR research studies assessing VR and technology within the classroom, specifically within K-12 and higher education. This study contributes to the literature by acknowledging a lack of an existing standard for an instrument in VR technology. Currently employed within existing research studies, instruments are not explicitly created for use in a VR environment. Therefore, it cannot be determined if they are validated for use with VR research studies. Many instruments being utilized in VR studies are adapted from a superfluity of instruments. The question remaining is sufficient for an analysis of a virtual environment. My analysis is that more measurements are needed to understand better the emotions, thoughts, and feelings participants exhibit towards VR as a technology and tool for integration with an educational setting.

Although some existing instruments may be well studied, the prevalence of instruments used in assessing the VR technology research case studies is researcher-made instruments with unknown qualities.

Implications For Practice and Future Research

Instruments constitute a significant source of data collecting in research projects. VR technologies must be evaluated like no other technology before them. The development and validation of an instrument(s) will contribute to the study of VR in education. Researchers will better understand the role of VR in education and the requirements of the participants as learners. It is recommended that future research expand the search to include other databases such as Google Scholar and alike. Such studies will be critical in providing more profound and more reliable insights into the context of VR instruments.

References

Abernethy, E. M. (2010). The Effect of Changed Environmental Conditions Upon the Results of College Examinations. *Http://Dx.Doi.Org/10.1080/00223980.1940.9917005*, *10*(2), 293–301. https://doi.org/10.1080/00223980.1940.9917005

Aday, L. A., & Cornelius, L. J. (2006). *Designing and conducting health surveys : a comprehensive guide* (3rd ed.). Jossey-Bass Publishers. https://books.google.com/books/about/Designing_and_Conducting_Health_Surveys.html?id=Xr-OCQAAQBAJ

Babbie, E. R. (1990). *Survey research methods* (2nd ed.). Wadsworth Publishing.

Baker, C., Pistrang, N., & Elliott, R. (2002). *Research Methods in Clinical Psychology: An Introduction for Students and ... - Chris Barker, Nancy Pistrang, Robert Elliott - Google Books* (2nd ed.). John Wiley & Sons Ltd. https://books.google.com/books?hl=en&lr=&id=goFsCgAAQBAJ&oi=fnd&pg=PP10&dq=Chris+Baker,+Nancy+Pistrang,+and+Robert+Elliott.+2002.+Research+Methods+in+Clinical+Psychology:+An+Introduction+for+Students+and+Practitioners+(second+edition+ed.).+John+Wiley+%26+Sons+Ltd.+&ots=cAOxnAjLGg&sig=Vyvodpf3bLhMya7GBvlPc9lzPyE#v=onepage&q&f=false

Bhattacherjee, A. (2012). Social Science Research: Principles, Methods, and Practices. *Textbooks Collection*. https://digitalcommons.usf.edu/oa_textbooks/3

Bhayani, S. B., & Andriole, G. L. (2005). Three-Dimensional (3D) Vision: Does It Improve Laparoscopic Skills? An Assessment of a 3D Head-Mounted Visualization System. *Reviews in Urology*, *7*(4), 211. /pmc/articles/PMC1477587/

Bradburn, N. M., & Sudman, S. (1979). *Improving Interview Methods and Questionnaire Design*

- *Response Effects to Threatening Questions in Survey Research*. Jossey-Bass Publishers. https://www.ojp.gov/ncjrs/virtual-library/abstracts/improving-interview-method-and-questionnaire-design-response

Choi, B. C. K., & Pak, A. W. P. (2005). A catalog of biases in questionnaires - PubMed. *PubMed*, *2*(1). https://pubmed.ncbi.nlm.nih.gov/15670466/

Dahlquist, L. M., McKenna, K. D., Jones, K. K., Dillinger, L., Weiss, K. E., & Ackerman, C. S. (2007). Active and Passive Distraction Using a Head-Mounted Display Helmet: Effects on Cold Pressor Pain in Children. *Health Psychology*, *26*(6), 794–801. https://doi.org/10.1037/0278-6133.26.6.794

Field, A., & Gournay, K. (2003). How to Design and Report Experiments Andy Field How to Design and Report Experiments and Graham Hole Sage No of pages: 400 £18.99 0761973834 0761973834 [Formula: see text]. *Nurse Researcher*, *11*(1), 83–84. https://doi.org/10.7748/NR.11.1.83.S13

Godden, D. R., & Baddeley, A. D. (1975). Context-Dependent MemoryIn Two Natural Environments: On Land and Underwater. *British Journal of Psychology*, *66*(3), 325–331. https://doi.org/10.1111/J.2044-8295.1975.TB01468.X

Gold, J. I., Kim, S. H., Kant, A. J., Joseph, M. H., & Rizzo, A. (2006). Effectiveness of Virtual Reality for Pediatric Pain Distraction during IV Placement. *Http://Www.Liebertpub.Com/Cpb*, *9*(2), 207–212. https://doi.org/10.1089/CPB.2006.9.207

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1), 7–15. https://doi.org/10.1016/J.INFSOF.2008.09.009

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). Research Methods in Human-Computer Interaction - Jonathan Lazar, Jinjuan Heidi Feng, Harry Hochheiser - Google Books. In *Research Methods in Human-Computer Interaction* (2nd ed.). Morgan Kaufmann. https://books.google.com/books?hl=en&lr=&id=hbkxDQAAQBAJ&oi=fnd&pg=PP1&dq=Jonathan+Lazar,+Jinjuan+Heidi+Feng,+and+Harry+Hochheiser.+2017.+Research+Methods+in+Human-Computer+Interaction+(2+ed.).+Morgan+Kaufmann,+Cambridge,+MA.&ots=Sq685gY3cN&sig=mhqJsWE87RAb1ykdcRtmthKgVmY#v=onepage&q=Jonathan Lazar%2C Jinjuan Heidi Feng%2C and Harry Hochheiser. 2017. Research Methods in Human-Computer Interaction (2 ed.). Morgan Kaufmann%2C Cambridge%2C MA.&f=false

Lee, K. (2012). Augmented Reality in Education and Training. *TechTrends 2012 56:2*, *56*(2), 13–21. https://doi.org/10.1007/S11528-012-0559-3

Pohl, R., & Phol, R. F. (2001). *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking ... - Google Books* (Rudiger. Pohl F. (Ed.)). Psychology Press. https://books.google.com/books?hl=en&lr=&id=k5gTes7yyWEC&oi=fnd&pg=PR3&dq=Riger+Pohl+and+Riger+F+.+Pohl.+2004.+Cognitive+Illusions:+A+Handbook+on+Fallacies+and+Biases+in+Thinking,+Judgement+and+Memory.+Psychology+Press.&ots=DYRxSeJiPy&sig=gh832T73nhmwafstCRqA41lCNsE#v=onepage&q=Riger Pohl and Riger F . Pohl. 2004. Cognitive Illusions%3A A Handbook on Fallacies and Biases in Thinking%2C Judgement and Memory. Psychology Press.&f=false

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of Behavioral Research: Methods and Data Analysis*. McGraw-Hill. https://dash.harvard.edu/handle/1/34945148

Saris, W. E., & Gallhofer, I. N. (2014). *Design, Evaluation, and Analysis of Questionnaires for*

Survey Research - Willem E. Saris, Irmtraud N. Gallhofer - Google Books. John Wiley & Sons. https://books.google.com/books?hl=en&lr=&id=zDOpAgAAQBAJ&oi=fnd&pg=PR13&dq=Willem+E.+Saris+and+Irmtraud+N.+Gallhofer.+2014.+Design,+Evaluation,+and+Analysis+of+Questionnaires+for+Survey+Research.+John+Wiley+%26+Sons.&ots=HnZvIMF-q3&sig=23JYjKYYoY2Cd1ZvXLuFHzqNJeI#v=onepage&q=Willem E. Saris and Irmtraud N. Gallhofer. 2014. Design%2C Evaluation%2C and Analysis of Questionnaires for Survey Research. John Wiley %26 Sons.&f=false

Shaughnessy, J. J., & Zechmeister, E. B. (1985). *Research methods in psychology*. *Alfred A. Knopf series in psychology.* Alfred A. Knopf. https://books.google.com/books/about/Research_Methods_in_Psychology.html?id=fgYQAQAAIAAJ

Shaw, L. A., Wünsche, B. C., Lutteroth, C., Marks, S., & Callies, R. (2015). *Challenges in virtual reality exergame design*. 61–68. https://openrepository.aut.ac.nz/handle/10292/8787

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review on JSTOR. *MIS Quarterly*, *26*(2), xiii–xxiii. https://www.jstor.org/stable/4132319

Yoo, S., & Kay, J. (2016). VRun: Running-in-place virtual reality exergame. *Proceedings of the 28th Australian Computer-Human Interaction Conference, OzCHI 2016*, 562–566. https://doi.org/10.1145/3010915.3010987

Table 1

Databases and key terms used in selected studies

Database	Search	String and Search Terms	No of articles
Academic Search Complete	Main terms utilized in search	"virtual reality" or "VR" AND "K-12" AND "Education" AND "higher education or college or university"	41
ERIC	Main terms utilized in search	"virtual reality" or "VR" AND "K-12" AND "Education" AND "higher education or college or university"	16
APA PsycInfo	Main terms utilized in search	"virtual reality" or "VR" AND "K-12" AND "Education" AND "higher education or college or university"	12
Library & Information Science Source	Main terms utilized in search	"virtual reality" or "VR" AND "K-12" AND "Education" AND "higher education or college or university"	10
Computer Science	Main terms utilized in search	"virtual reality" or "VR" AND "K-12" AND "Education" AND "higher education or college or university"	3

Table 2

Selection of literature for SLR studies utilizing inclusion and exclusion criteria

Criteria	Decision
Research Studies on VR in K-12 education	Inclusion
Research Studies on VR in higher education	Inclusion
Research study published in a scientific peer-reviewed journal	Inclusion
Research study must be written in the English language	Inclusion
All studies irrespective of design	Inclusion
Duplicated studies	Exclusion
Studies published prior to 2011	Exclusion
Studies published after 2021	Exclusion
Dissertations	Exclusion
Studies not utilizing VR head mounted display	Exclusion
Research studies not original research studies	Exclusion
Non-educational research studies	Exclusion
Evaluations of software studies	Exclusion
Literature reviews	Exclusion
Meta-analysis	Exclusion
Systematic reviews	Exclusion

Figure 1

PRISMA chart for study selection

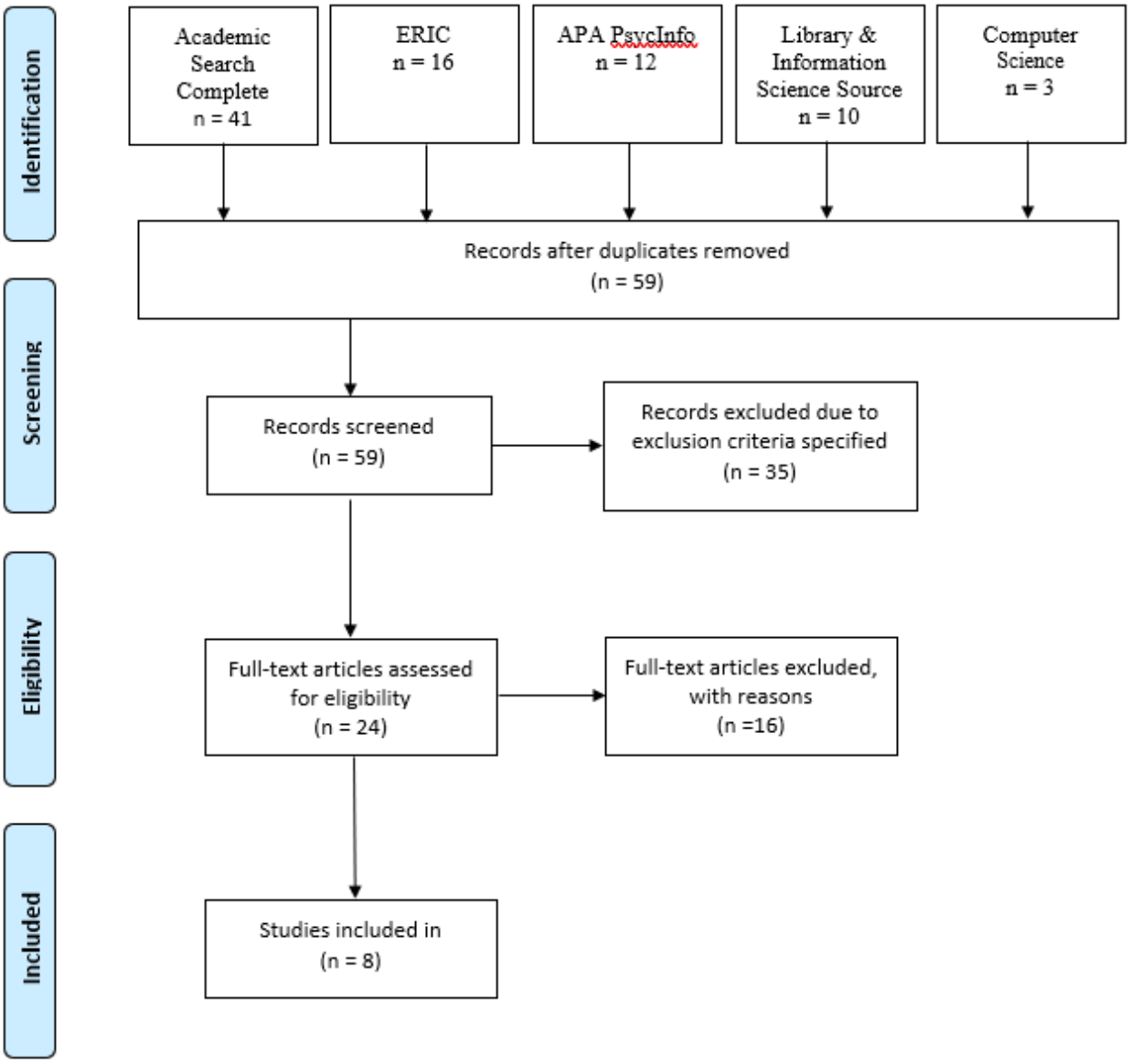

Table 3

A summary of the instruments examined

Instrument	Author/year	Objective	No. of items	Rating scale
Simulator Sickness Questionnaire (SSQ)	Kennedy et al. (1993)	Describe and assess simulator sickness.	16	Four point scale 0 (None), 1 (Slight), 2 (Moderate), 3 (Severe)
Technology Acceptance Model (TAM) Questionnaire	Davis (1989)	Understand more clearly of how users' attitudes related to their willingness to use a technology.	7	Likert scale - 5-point 1 (strongly disagree) to 5 (strongly agree)
The Engagement Scale	Wang et al. (2016)	Divided into four sections, it contained eight items for cognitive engagement, eight for behavioral engagement, 10 for emotional engagement and seven for social engagement.	33	Likert scale - 5-point 1 (strongly disagree) to 5 (strongly agree)
The Technology Acceptance Questionnaire	Chu et al. (2010)	It included six items for perceived usefulness and seven for perceived ease of use.	13 (out of 37)	Likert scale - 5-point 1 (strongly disagree) to 5 (strongly agree)
Interest/Enjoyment scale from the Intrinsic Motivation Inventory	Deci et al. (1994)	(IMI) is a multidimensional measurement device intended to assess participants subjective experience related to a target activity in laboratory experiments.	5 (out of 45)	Likert scale - 5-point 1 (strongly disagree) to 5 (strongly agree)
Motivated Strategies for Learning Questionnaire (MSLQ)	Pintrich et al. (1991)	Developed to measure the types of learning strategies and academic motivation used in higher education.	5 (out of 44)	Likert scale - 5-point 1 (strongly disagree) to 5 (strongly agree)
Perceived Usefulness and Perceived Ease of Use	Martocchio & Webster (1992)	Perceived usefulness and ease of use are people's subjective appraisal of performance and effort, respectively, and do not necessarily reflect objective reality.	14 (out of 20)	Likert scale - 7-point 1 (extremely likely) to 7 (extremely unlikely)
Presence Questionnaire	Witmer & Singer (1989)	This questionnaire attempts to cover as many aspects of presence as possible in order to provide a multifaceted presence measure.	8 (out of 29)	Likert scale - 7-point 1 (strongly disagree) to 5 (strongly agree)