Does Screen-Based VR Improve Learning Performance in Higher Education? A Meta-Analysis of Recent Research

Gulmira Yerlanova[1], Gulnar Shynatay[2], Nurzhan Shyndaliyev[2], and Gulmira Nurbekova[2,*]

[1]Department of Information and Technical Science, Alikhan Bokeikhan University, Semey, Kazakhstan
[2]Department of Computer Science, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
Email: erlanovagulmira@gmail.com (G.Y.); Goolnar0989@gmail.com (G.S.); nurzansyndaliev385@gmail.com (N.S.); nurbekovagulmira48@gmail.com (G.N.)
*Corresponding author

Manuscript received May 30, 2024; revised July 25, 2024; accepted August 26, 2024; published November 15, 2024

***Abstract*—Virtual reality has revolutionised the concept on learning throughout the world, and the higher education sector is no exception. Despite the emergence of high-fidelity simulation technologies, display-based learning environments have still been involved in educational research. But even now, it is under question whether low-immersion virtual learning really enhances undergraduate students' learning performance. As an attempt to produce pooled evidence on the topic, this review meta-analytically combined post-test effect sizes extracted from 16 relevant empirical studies comprising 2,508 participants. Overall, the virtual conditions had a large positive effect (Hedge's g = 0.98) on learning outcomes, but these findings are injured by substantial between-study variance. To understand differences in the effects, subgroup analyses were performed across research design (experimental versus quasi-experimental) and outcome type (practical skills measured or not). However, all subgroups were heterogeneous, so the total picture remains somewhat equivocal. Thus, the results imply screen-based virtual environments can boost post-secondary students' educational gains, but the extent has yet to be discovered. The contribution of this paper to extant research is that it highlights how much has not been clarified yet about the advantageousness of applying virtual environments in education. Academia is therefore urged to continue experimenting with synthesised instructional scenarios.**

***Keywords*—academic achievement, learning outcomes, meta-analysis, simulation, undergraduate students, virtual reality**

I. Introduction

Virtual Reality (VR) technology is slowly gaining a firm foothold in the field of education and training, allegedly because of its considerable influence on various learning effectiveness parameters, including skill transfer [1, 2]. VR simulations have been exploited for decades in sectors such as fire safety exercises, psychological therapy, aviation training, and foreign language teaching in order to abstractly represent certain aspects of real work situations in training scenarios and gradually adapt them to complex reality [3, 4].

From a pedagogical standpoint, the virtualization of practicum activities is supposed to amplify the comprehensiveness of the learning experience, ensure left-field teaching content, deepen students' feelings, and upturn their learning engagement, thus leading to elevated academic achievement. Moreover, virtual worlds can help overcome the limitations of facilities and infrastructure [5]. An educational virtual environment can be understood as a scene that emulates real or imaginary situations and objects actively manipulated by learners to carry out relevant activities in an imitative manner, thus obtaining the required skills and knowledge [6]. In the case of screen-based learning simulations, this fictitious environment is generated through a display device so that the individual is not entirely doused in the replica and navigates exocentrically [7].

VR technology in its low-immersive form offers a middle ground between traditional non-VR learning methods and high-immersion VR, making it an attractive option for educators seeking to improve student outcomes while working within budget constraints [8]. This balance is crucial in higher education, where resource constraints often limit the adoption of more sophisticated, high-immersive VR systems. Low-immersive VR experiences, characterized by their use of readily available technology like computer screens and tablet/mobile phone displays, provide an accessible means for educational institutions to incorporate virtual environments into their teaching methodologies [9]. For example, a student in a nursing program could use a screen-based simulation to practice administering medication or responding to a patient's needs in a virtual hospital setting. This approach offers a valuable alternative to expensive high-immersion setups while still providing an engaging and interactive learning experience. Such environments can simulate real-world settings and scenarios at a fraction of the cost and complexity associated with high-immersive VR, which often requires specialized equipment [10].

II. Literature Review

In a recent scrupulous meta-analysis [11], the overwhelming majority of reviewed studies used high-immersive VR technology as an experimental condition, whereas low-immersive studies were mostly published in the 2000s. It follows that the desktop-based virtualization of learning content has become obsolete, but the last few years have seen some new research dealing with display simulation modalities, which might be partly attributed to the fact that high-end VR devices such as headset helmets and glasses are still unaffordable for a large number of education facilities. Moreover, Mulders *et al.* [12] believe that applying highly immersive VR systems can be substantively flawed given possible distraction, working-memory capacity abatement, cognitive overload, health and safety risks, issues related to integration into designed scenarios, as well as possible student's discomfort associated with wearing a head-mounted display. As explained by Eisenlauer [13], personal mobile devices could not be introduced as an alternative since they do

doi: 10.18178/ijiet.2024.14.11.2190

not imply any input options like a mouse, controller, or keyboard; therefore, interactions with the learning materials are largely limited to passive watching, and such simulation scenarios exert a little realistic effect on the user. Therefore, display-based VR learning tools could be considered an optimal way to mimic real-world settings and cases in an educational context given their familiarity to individuals and relative inexpensiveness.

Looking back at related reviews published over the last several years, one can see the dearth of conclusive evidence from analysts who attempted to investigate low-immersive VR interventions implemented in post-secondary education. Despite the potential benefits and practicality of low-immersive VR, previous reviews have covered extensive topics about the effects of VR while often overlooking the principal outcome of VR use expected by educators, which is its impact on academic performance. Beck and Perkins [14] revised 127 empirical works dealing with desktop virtual world environments to categorise them by research methodologies utilised, type of data collected, academic field, data analysis method, and so on. However, that is a purely qualitative analysis aimed at exploring connections between the data and discovering gaps in types of research and methodologies. Generally, not so much can be deduced from the findings. Mulyadi *et al.* [15] undertook a systematic literature review that pooled learners' post-test learning outcomes and affective variables measured within simulated technology–based learning experiments, and the investigation yielded promising results in terms of knowledge acquisition pointing to a significant advantage of virtual learning modes over normative methods of instruction. Unfortunately, the scope of the analysis covered only nursing students. Furthermore, beyond just computer-generated simulations, the review involved interventions that employed high-fidelity manikins, and the latter had a better effect on knowledge acquisition as proven by subgroup analysis. Finally, a third of the primary studies looked at affective domain outcomes solely. This all raises the problem of how to generalise the findings. Similarly, Shorey and Ng [16] completed a systematic review of how immersive and non-immersive simulations affected diverse learning outcomes in registered nurses and nursing students. However, the analysis embraces 17 individual studies, 13 of which dealt with desktop-based VR interventions, and 4 out of 13 were published before 2015 when VR technology was at a relatively early stage of development. A systematic review by Chan *et al.* [17] in turn focuses on virtual chemical laboratories, and again, only one-fifth of included studies reported cognitive domain outcomes obtained by students subjected to a three-dimensional desktop VR condition. A recent review by Flavin and Bhandari [18] inspects research pertaining to virtual learning environments in tertiary education that were published between 2014 and 2018, but the work is descriptive in nature rather than analytical and lacks explicit takeaways.

Thus, it appears that academia still faces the challenge of figuring out whether a low-immersive VR experience can advance learning performance in post-secondary students. Trying to fill this gap, the current review seeks to find the overall effectiveness of applying screen-based virtual learning environments in higher education on the academic achievement of students as reported over the past five years. An additional objective is to examine if pre-determined moderator variables influence the effects of the interventions on learning performance. Therefore, the following questions guided this research:

1) What is the effectiveness of fostering learning success through display-based VR technology in higher education as pooled from relevant empirical studies?
2) What kinds of moderator variables affect screen-based VR learning outcomes?

III. Methods

To arrive at the research objective, it was agreed to utilise a meta-analysis, which refers to a statistical procedure that enables amalgamating findings from a number of empirical studies on a specific problem in the form of integrated quantitative evidence as an attempt to yield conclusive results when individual studies are ambiguous or conflicting [19].

A. Search Strategy and Quality Criteria

Using EBSCOhost, Science Direct, and Google Scholar electronic databases, a comprehensive search of journal publications was performed through the following search terms: {'virtual' OR 'VR'} AND {'learning' OR 'reality' OR 'simulation'} AND {'student' OR 'undergraduate'}, AND {'learning performance' OR 'outcome' OR 'knowledge' OR 'competence' OR 'skill' OR 'achievement'}.

The databases were selected for their comprehensive coverage of relevant research in education and technology. EBSCOhost provides access to a wide range of peer-reviewed journals in education and related fields. Science Direct offers a strong focus on scientific and technical literature, including computer science and engineering, which are pertinent to virtual reality research. Google Scholar, while broader in scope, serves as a valuable tool for identifying potentially relevant studies across various disciplines and for uncovering grey literature. This combination of databases ensured a robust and balanced search strategy, capturing both specialized and interdisciplinary research on virtual reality applications in education.

Only papers published from 2017 onwards were included in the meta-dataset. Focusing on more recent publications allowed us to capture the latest research incorporating newer VR technologies and pedagogical approaches. While earlier studies are valuable for understanding the historical development of virtual reality in education, our primary focus was on evaluating the effectiveness of more recent VR applications in enhancing undergraduate learning outcomes. To screen all downloaded articles, the following selection criteria were applied:

1) Studies must have evaluated the impact of screen-based VR learning (meaning that the virtual environment and objects represented three-dimensional graphics running on electronic device displays) on higher education students, with simulation-assisted learning being the primary intervention.
2) Studies must have adopted either a randomised or quasi-experimental research design involving at least one no-treatment control group.

3) Studies must have drawn upon objective tools measuring cognitive learning outcomes, i.e., skills and knowledge rather than self-evaluation reports.
4) Studies must have reported outcomes sufficient to compute an effect size.
5) Studies must have been published in English and peer-reviewed.

Studies meeting the following criteria were discarded: a) papers irrelevant to the current study objective; b) dissertations, conceptual papers, and conference materials; c) papers unavailable as full-text documents; d) given that a critical point in virtual learning is the opportunity for participants to interact with the simulated environment, studies in which non-manipulable digital content was offered to students such as lectures delivered through Zoom [20] or just video clips [17] were not included in this review, although the authors called the interventions virtual ones.

B. Literature Screening

The literature search results were screened against the quality criteria between January and April 2023. Upon the initial search, a total of 411 potentially relevant records were returned, which dropped to 356 after removing duplicates. Upon removing inapposite articles based on title and abstract primary screening, 62 full-text documents were downloaded and thoroughly examined for eligibility. In addition, 75 sources derived from earlier reviews were scrutinised. Eventually, a total of 16 studies comprising 2,508 students were considered for the final quantitative synthesis.

C. Data Synthesis

Hedge's g with a 95% confidence interval was used to estimate the mean weighted effect size to compare the effectiveness between virtual learning and non-virtual learning using the R programming language. To account for a possible high variance in effect sizes across the primary studies, the post-test effect sizes of the interventions were integrated using a random-effects model. Data were averaged in the event of two and more learning outcomes pertaining to the cognitive domain being presented within one investigation. The effect size was regarded as small (0.2), medium (0.5), and large (0.8) [21]. Z-statistics at the p <0.05 level were considered significant. Heterogeneity was appraised by virtue of the I2 parameter, with more than 75% indicating considerable heterogeneity between the studies [22]. Besides, the Q statistic was employed. A Baujat plot [23] was drawn to inquire into the contribution of each study to the overall heterogeneity. Using Meta-Essentials [24], studies were coded dichotomously within subgroup analyses for the potential moderating effect of study design, i.e., whether participants were assigned into groups randomly or conventionally, as well as outcome type measured, i.e., skills or knowledge. Publication bias was assessed using the Begg and Mazumdar rank test, along with Egger's regression, to determine whether works in the field were published selectively, favoring positive evidence. Additionally, a normal quantile plot showing the distribution of the data against the expected normal distribution was generated to visually elucidate its asymmetry, which would point to the publication bias of the sample data.

IV. RESULTS

The current integrative review was conducted on seven experimental studies (43.75%) and nine quasi-experimental studies (56.25%). Of them, two studies were carried out in European countries, including France and Finland (12.5%); seven in Asia, namely Taiwan, China, Singapore, Israel, and Turkey (43.75%); and seven in the Americas, namely Canada, the United States of America, and Brazil (43.75%). The most frequent learning domain was nursing education [25–34], followed by chemistry education [35–37], while one study each focused on medical education [38] and dental education [39], and another one involved undergraduate students from various subject fields [40]. As for learning application type, the largest proportion of the studies used a simulation as the treatment (68.75%) and the remainder utilised serious games (31.25%). Regarding implementation modality, only investigations by Lebdai *et al.* [38] and Chang *et al.* [39] applied the experimental tools in a ubiquitous manner, while in the remaining studies participants were required to use the synthetic environment in a formal research setting. The weighted mean effect of the screen-based VR interventions on the students' learning performance is shown in Fig. 1.

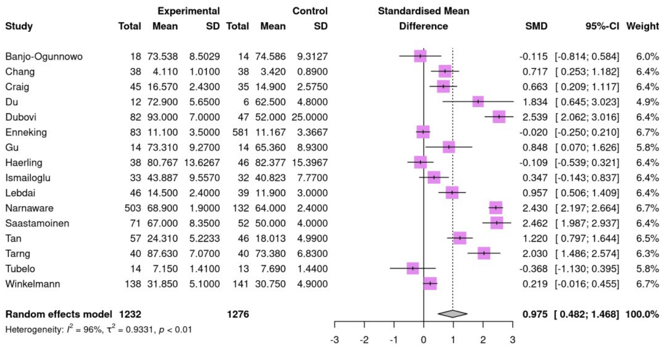

Fig. 1. Pooled estimate of the reviewed studies. (SD: Standard Deviation; SMD: Standardised Mean Difference; CI: Confidence Interval).

The overall effect size using the random-effects model is 0.975 which is commonly regarded as a large promotion effect. It was statistically significant at Z-value = 3.918, $p < 0.01$. It is therefore likely that the screen-based VR groups generally fared better than the comparison groups. As regards publication bias, the normal quantile plot for the selected studies is presented in Fig. 2.

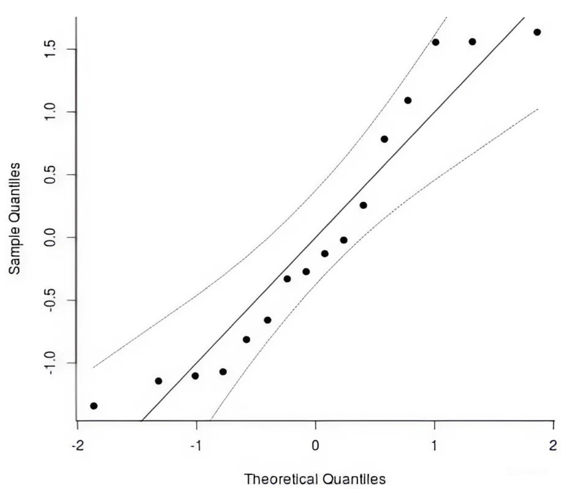

Fig. 2. Normal quantile plot for the reviewed studies.

In the diagnostic plot, all the dots representing effect, sizes calculated for the studies fall basically near the central line and form a mild curvature, which informs us that the investigations in the evaluated area were hardly published selectively. Furthermore, both the Begg and Mazumdar test (Kendall's tau = 0.12; $p = 0.528$) and the Egger regression test ($p = 0.859$) indicated no evidence of the data asymmetry. However, a statistically significant Q value (Q(df = 15) = 413.5111, p <0.001) along with the I2 criterion (96%) indicated the presence of heterogeneity across the studies, so the between-study variance is highly likely to be attributable to the discordance in the effect sizes, not to sampling error.

In spite of considerable between-study heterogeneity, there were no effect sizes three standard deviations above or below the combined estimate, so the three-sigma rule of thumb [41] was not violated and effect-size-based outliers were not detected. To assess sources of heterogeneity in the meta-analytic data visually, a Baujat plot was generated (Fig. 3).

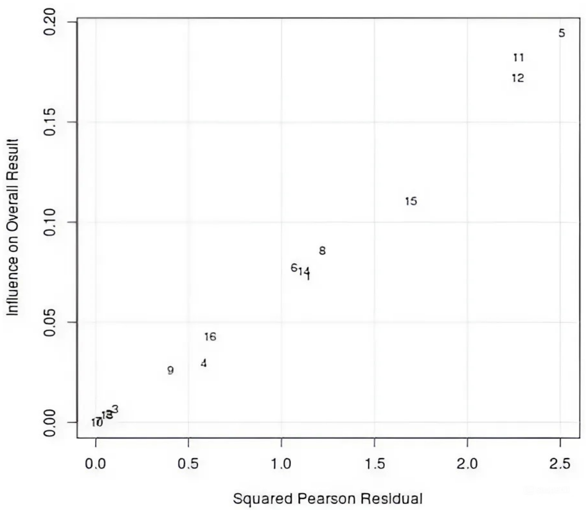

Fig. 3. Baujat plot for the reviewed studies.

As depicted in Fig. 3, half of the primary studies have a high overall heterogeneity contribution, and removing them from further analysis would be destructive to the meta-analysis quality. It was therefore decided to perform a subgroup analysis in an attempt to explain the conflicting effects by different experimental designs of the included studies (Fig. 4).

Study name / Subgroup name	Hedges' g	CI Lower limit	CI Upper limit	Weight	Q	p_Q	I^2	T^2	T	PI Lower limit	PI Upper limit
Banjo-Ogunnowo 2022	-0,12	-0,83	0,59	10,58%							
Chang 2021	0,72	0,25	1,18	11,06%							
Craig 2021	0,66	0,21	1,12	11,07%							
Dubovi 2017	2,54	2,06	3,02	11,03%							
Enneking 2017	-0,02	-0,25	0,21	11,37%							
Haerling 2018	-0,11	-0,54	0,32	11,11%							
Narnaware 2021	2,43	2,20	2,66	11,37%							
Saastamoinen 2022	2,46	1,99	2,94	11,04%							
Winkelmann 2020	0,22	-0,02	0,46	11,37%							
Non-randomised	**0,98**	**0,22**	**1,74**	**38,30%**	**378,93**	**0,000**	**97,89%**	**1,42**	**1,19**	**-1,92**	**3,87**
Du 2020	1,83	0,63	3,03	10,32%							
Gu 2017	0,85	0,06	1,64	13,34%							
Ismailoglu 2018	0,35	-0,15	0,84	15,63%							
Lebdai 2021	0,96	0,50	1,41	15,92%							
Tan 2017	1,22	0,80	1,65	16,11%							
Tarng 2017	2,03	1,49	2,57	15,21%							
Tubelo 2019	-0,37	-1,14	0,41	13,46%							
Randomised	**0,97**	**0,36**	**1,57**	**61,70%**	**37,76**	**0,000**	**84,11%**	**0,44**	**0,67**	**-0,83**	**2,76**
Combined Effect Size	**0,97**	**-1,16**	**3,10**		**416,88**	**0,000**	**96,40%**	**1,13**	**1,06**	**-1,16**	**3,10**

Fig. 4. Subgroup analysis based on research design. (CI: Confidence Interval; PI: Prediction Interval).

As seen in Fig. 4, neither the subgroup of randomised research nor the one representing quasi-experimental studies can be treated as a homogeneous population. Additionally, a subgroup analysis was undertaken to determine if the effect sizes differ depending on whether or not the measured outcomes included practical skills (Fig. 5). Unfortunately, both of the coded subgroups were markedly heterogeneous again and further examination of this moderator must be ruled out.

Study name / Subgroup name	Hedges' g	CI Lower limit	CI Upper limit	Weight	Q	p_Q	I^2	T^2	T	PI Lower limit	PI Upper limit
Du 2020	1,83	0,63	3,03	12,69%							
Haerling 2018	-0,11	-0,54	0,32	14,63%							
Narnaware 2021	2,43	2,20	2,66	14,91%							
Saastamoinen 2022	2,46	1,99	2,94	14,55%							
Tarng 2017	2,03	1,49	2,57	14,41%							
Tubelo 2019	-0,37	-1,14	0,41	13,90%							
Winkelmann 2020	0,22	-0,02	0,46	14,91%							
Knowledge	**1,21**	**0,28**	**2,14**	**23,67%**	**264,38**	**0,000**	**97,73%**	**1,74**	**1,32**	**-2,21**	**4,64**
Banjo-Ogunnowo 2022	-0,12	-0,83	0,59	10,29%							
Chang 2021	0,72	0,25	1,18	11,28%							
Craig 2021	0,66	0,21	1,12	11,32%							
Dubovi 2017	2,54	2,06	3,02	11,23%							
Enneking 2017	-0,02	-0,25	0,21	12,01%							
Gu 2017	0,85	0,06	1,64	9,93%							
Ismailoglu 2018	0,35	-0,15	0,84	11,18%							
Lebdai 2021	0,96	0,50	1,41	11,33%							
Tan 2017	1,22	0,80	1,65	11,43%							
Skills	**0,80**	**0,28**	**1,31**	**76,33%**	**109,52**	**0,000**	**92,70%**	**0,63**	**0,79**	**-1,13**	**2,72**
Combined Effect Size	**0,90**	**-1,24**	**3,03**		**416,88**	**0,000**	**96,40%**	**1,13**	**1,06**	**-1,24**	**3,03**

Fig. 5. Subgroup analysis based on intervention outcome. (CI: Confidence Interval; PI: Prediction Interval).

It is evident from the descriptive results outlined earlier that the reviewed works mostly share common application-related and contextual characteristics, which undermines the use of other moderators except for research design and assessed outcome type. Concerning the duration of the interventions, it is unspecified in nine of the 16 studies, so this analysis is not possible. The learner's gender does not seem to be a promising adjustment variable to the authors of this review.

V. Discussion

The primary aim of this study was to determine the effectiveness of screen-based VR simulations in improving the learning performance of higher education students. After calculating the weighted mean effect of relevant interventions on learners' knowledge and skills, we cannot definitively state whether a low-immersion virtual learning environment is beneficial for higher education students.

Radianti *et al.* [42] promote the idea that VR interventions with educational purposes should rest on learning theories, such as behaviourism, cognitivism, constructivism, and connectivism, since it is presumed that VR learning activities grounded in a solid theoretical framework are more likely to bring about corresponding learning achievement. Having critically reviewed 56 empirical works dedicated to desktop virtual worlds for education, Loke [43] inferred that the studies were supported by several theories that explain to some extent how mental operations, reflection, verbal interactions, and vicarious experiences that individuals undergo in a computer-generated learning environment are somewhat analogous to those occurring in a non-simulated mode of learning while affording participants the opportunity to break down psychological and physical barriers, as well as broaden their scope of action, thus instigating learning gains.

As for the 16 primary studies overviewed herein, 12 are reportedly not underpinned by solid learning theories. Particularly, Banjo-Ogunnowo and Chisholm [25] indicated that they employed the model "Thinking like a nurse" [44] as the theoretical framework for their study. Nevertheless, the intervention is described in general terms only, so it is uncertain how exactly the virtual learning modality was driven by the model. In the empirical work by Haerling [29], it is reported that the "study uses the National League for Nursing/Jeffries Simulation Theory as a guiding framework" [45]. In fact, the author just referred to the theory in order to demonstrate several variables on which control and treatment groups differed. Tarng *et al.* [36] claim that their virtual laboratory was developed based on the situated learning theory elaborated by Brown *et al.* [46], which insists that learning is incorporated in a certain socio-physical context. The theory was applied to the intervention in the way that the virtual laboratory designed by the authors implied learning under realistic situations by replicating hands-on operations. Saastamoinen *et al.* [32] mention experiential learning theory [47] as the theoretical framework for their serious-game-based study since they believe that simulation games enable experiential learning. Fromm and co-authors [48] clarify that the virtual environment imparts a copious and captivating education context in which individuals can learn by doing and reflect on the experiences to obtain new skills or ways of thinking. In a similar vein, constructivist learning theory [49] revolves around the notion that learning is effective as long as students construct knowledge through learning-by-doing [50]. In this respect, synthetic reality provides a controlled, exploratory environment for learning through experimentation. Meanwhile, social constructivist learning theory [51] approaches learning as collaborative in nature. Within this context, Loke [43] proposes narrowing the social interactions to verbal interaction since not all virtual actions are fully equal to those in non-emulated reality. In 14 of the 16 studies included in the current review, it is not clear from the article text whether the VR environment implied verbal interaction with virtual agents or real-world actors. It is pointed out in the paper by Enneking *et al.* [35] that "students typically worked in pairs," but this cooperation cannot actually be considered a VR-mediated verbal interaction. Regarding Tarng *et al.* [36], it is written in the article that learners could "follow the verbal instructions to select the virtual instruments for conducting the experiment," but this does not imply any verbal response from the participants, so it is a one-way process rather than an interaction. It therefore does not appear feasible to assess whether the embodiment principle was employed in the selected investigations. The principle stems from social agency theory [52] and holds that learners are more assiduous when virtual agents exhibit human-like facial expressions, gestures, and so forth, which increases students' social presence [53]. However, none of the observed theories provide an understanding of how the virtual world experience promotes the enhancement of real-world knowledge and skills given the lack of a firm connection between the direct experience and the virtual one.

When the primary studies were looked at in terms of whether a given study was value-added research or just adopted media comparison, that is whether it included a juxtaposition of different design components of the same simulation (for example, a virtual scenario either accompanied by narration or not) apart from a negative control group [52], only two of the reviewed studies provide a value-added comparison: in the research conducted by Du e*t al.* [40] either a single-player or a multiple-player version of a virtual gaming system was proposed for participants, whereas Craig *et al.* [25] evaluated low-fidelity and high-fidelity simulations beyond the treatment-control pair. This picture is in line with the findings of the scoping review by Chan *et al.* [54]. At the same time, in contrast to findings in the latter research, the current meta-analytical work shows that most of the selected studies look into participants' practice-anchored skills, probably because the mentioned review encompassed exclusively virtual chemical laboratories. On the other hand, it scoped school-aged students and all types of educational virtualisation rather than those restricted to the desktop category. Paradoxically, this indicates that display-based learning simulations have been utilised for the purpose of honing the professional skills of tertiary education students more extensively compared to natural user interfaces equipped with advanced tracking sensors.

In terms of learning effectiveness, our findings align with similar studies that suggest low-immersion VR can bolster learning outcomes [55, 56]. The overall positive effect indicates that virtual conditions indeed boost post-secondary students' educational gains. However, the substantial between-study variance implies that the effectiveness is not uniform and may depend on various factors, which brings us to the consideration of moderator variables. The moderator analyses performed here across research design (experimental versus quasi-experimental) and outcome type (practical skills measured or not) revealed that these variables did not account for the heterogeneity in the results. This suggests that other factors, such as the subject matter taught, might play a more significant role in determining the learning outcomes. Unfortunately, the number of relevant quantitative studies must be several times higher than it currently is to provide subgroups large enough for adequate comparison.

VI. Conclusion

As for limitations of this review, its most vulnerable point is the between-study variance, which reduces the statistical power of the work. The effect sizes were found to be too heterogeneous to allow for the estimation of a combined outcome, so the first research question remains unanswered. We attempted to address the issue by exploring potential moderators in the form of different research designs and intervention outcomes, but to no avail. Thus, the results imply that screen-based virtual environments can boost post-secondary students' educational gains, but the extent has yet to be discovered. Further experimental implementation of VR technology and teaching scenarios is encouraged for higher education. Despite not having yielded any concrete conclusions regarding the effectiveness of screen-based educational VR environments in post-secondary education as compared with conventional learning, this paper sheds light on how ambivalent the subject is: seemingly evaluated throughout, it still has blind spots to overcome. Moreover, this investigation extends and complements extant research on the overall impact of virtual learning on students' cognitive learning outcomes. In line with the call for a balanced integration of physical and virtual interaction, we recommend future research to explore the practical harmonization of virtual learning environments with contemporary educational paradigms, such as flipped or fuzzy-based learning approaches.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Gulmira Yerlanova and Gulnar Shynatay conducted the literature search, analyzed the data and wrote the paper. Nurzhan Shyndaliyev and Gulmira Nurbekova visualized the data and reviewed the manuscript. All authors have accepted the final version of the manuscript.

REFERENCES

[1] R. P. Antonio and R. R. Castro, "Effectiveness of virtual simulations in improving secondary students' achievement in physics: A meta-analysis," *International Journal of Instruction*, vol. 16, no. 2, pp. 533–556, April 2023.

[2] P. Yang and Z. Liu, "The influence of Immersive Virtual Reality (IVR) on skill transfer of learners: The moderating effects of learning engagement," *International Journal of Emerging Technologies in Learning*, vol. 17, no. 10, pp. 62–73, May 2020.

[3] M. Prange, "Virtual reality und augmented reality in der Bildung—Ein Überblick zum Thema," in *Digitale Bildung für Lehramtsstudierende*, D. Graf, N. Graulich, K. Lengnink, H. Martinez, and C. Schreiber Eds. Springer VS, 2021, pp. 235–246. https://doi.org/10.1007/978-3-658-32344-8_32

[4] R. R. Setyowati, S. Rochmat, Aman, and A. N. P. Nugroho, "Virtual reality on contextual learning during COVID-19 to improve students' learning outcomes and participation," *International Journal of Instruction*, vol. 16, no. 1, pp. 173–190, Jan. 2023.

[5] I. Yusuf and S. W. Widyaningsih, "Implementing e-learning-based virtual laboratory media to students' metacognitive skills," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 5, pp. 63–74, March 2020.

[6] S. M. Reeves, K. J. Crippen, and E. D. McCray, "The varied experience of undergraduate students learning chemistry in virtual reality laboratories," *Computers & Education*, vol. 175, 104320, Dec. 2021.

[7] M. Kozhevnikov, J. Gurlitt, and M. Kozhevnikov, "Learning relative motion concepts in immersive and non-immersive virtual environments," *Journal of Science Education and Technology*, vol. 22, no. 6, pp. 952–962, Feb. 2013.

[8] N. O. Onele, "The role of desktop virtual reality as an accessible and equitable strategy to improve career opportunities for women in technology," *Journal of Computer Assisted Learning*, vol. 39, no. 1, pp. 20–33, Feb. 2023.

[9] G. M. N. Avellar, M. L. Fioravanti, W. S. Deus, K. R. L. J. C. Branco, and E. F. Barbosa, "SSPOT-VR: An immersive and affordable mobile application for supporting K-12 students in learning programming concepts," *Education and Information Technologies,* 2024. https://doi.org/10.1007/s10639-024-12499-0

[10] D. Moreno-Lumbreras, R. Minelli, A. Villaverde, J. M. Gonzalez-Barahona, and M. Lanza, "CodeCity: A comparison of on-screen and virtual reality," *Information and Software Technology*, vol. 153, 107064, Jan. 2023.

[11] D. Angel-Urdinola, C. C. Castro and A. Hoyos, "Meta-analysis assessing the effects of virtual reality training on student learning and skills development," *The World Bank*, July 2023.

[12] M. Mulders, J. Buchner, and M. Kerres, "A framework for the use of immersive virtual reality in learning environments," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 24, pp. 208–224, Dec. 2020.

[13] V. Eisenlauer, "Digital literacies in virtual reality learning contexts," in *Augmented Reality and Virtual Reality,* T. Jung, M. tom Dieck, and P. Rauschnabel, Eds., Springer, 2020, pp. 269–281.

[14] D. Beck and R. Perkins, "Review of educational research methods in desktop virtual world environments: Framing the past to provide future direction," *Journal of Virtual Worlds Research*, vol. 7, no. 1, pp. 1–27, Jan. 2014.

[15] M. Mulyadi, S. I. Tonapa, S. S. J. Rompas, R. H. Wang, and B. O. Lee, "Effects of simulation technology-based learning on nursing students' learning outcomes: A systematic review and meta-analysis of experimental studies," *Nurse Education Today*, vol. 107, 105127, 2021.

[16] S. Shorey and E. D. Ng, "The use of virtual reality simulation among nursing students and registered nurses: A systematic review," *Nurse Education Today*, vol. 98, 104662, March 2021.

[17] P. Chan, T. van Gerven, J. L. Dubois, and K. Bernaerts, "Virtual chemical laboratories: A systematic literature review of research, technologies and instructional design," *Computers and Education Open*, vol. 2, 100053, 2021.

[18] M. Flavin and A. Bhandari, "What we talk about when we talk about virtual learning environments," *International Review of Research in Open and Distributed Learning*, vol. 22, *no.* 4, pp. 164–193, 2021.

[19] Y. Ma, J. Wang, J. Wu, C. Tong, and T. Zhang, "Meta-analysis of cellular toxicity for graphene via data-mining the literature and machine learning," *Science of the Total Environment*, vol. 793, 148532, Nov. 2021.

[20] J. X. Lee and I. Ahmed, "Advancing neurosurgery education in junior doctors and medical students—A neurosurgery virtual lecture series," *Annals of Medicine and Surgery*, vol. 68, 102578, August 2021.

[21] S. J. Peterson and S. Foley, "Clinician's guide to understanding effect size, alpha level, power, and sample size," *Nutrition in Clinical Practice*, vol. 36, no. 3, pp. 598–605, May 2021.

[22] M. Saqr and S. Lopez-Pernas, "The curious case of centrality measures: A large-scale empirical investigation," *Journal of Learning Analytics*, vol. 9, no. 1, pp. 13–31, March 2022.

[23] B. Baujat, C. Mahé, P. J. Pignon, and C. A. Hill, "Graphical method for exploring heterogeneity in meta-analyses: Application to a meta-analysis of 65 trials," *Statistics in Medicine*, vol. 21, no. 18, pp. 2641–2652, Sept. 2002.

[24] R. Suurmond, H. Rhee, and N. Hak, "Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis," *Research Synthesis Methods*, vol. 8, no. 4, pp. 537–553, August 2017.

[25] S. M. Banjo-Ogunnowo and L. J. Chisholm, "Virtual versus traditional learning during COVID-19: Quantitative comparison of outcomes for two articulating ADN cohorts," *Teaching and Learning in Nursing*, vol. 17, no. 3, pp. 272–276, 2022.

[26] S. J. Craig, J. C. Kastello, B. J. Cieslowski, and V. Rovnyak, "Simulation strategies to increase nursing student clinical competence in safe medication administration practices: A quasi-experimental study," *Nurse Education Today*, vol. 96, 104605, 2021.

[27] I. Dubovi, S. T. Levy, and E. Dagan, "Now I know how! The learning process of medication administration among nursing students with non-immersive desktop virtual reality simulation," *Computers & Education*, vol. 113, pp. 16–27, 2017.

[28] Y. Gu, Z. Zou, and X. Chen, "The effects of vSIM for NursingTM as a teaching strategy on fundamentals of nursing education in undergraduates," *Clinical Simulation in Nursing*, vol. 13, no 4, pp. 194–197, 2017.

[29] K. A. Haerling, "Cost-utility analysis of virtual and mannequin-based simulation," *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 13, no. 1, pp. 33–40, 2018.

[30] E. G. Ismailoglu and A. Zaybak, "Comparison of the effectiveness of a virtual simulator with a plastic arm model in teaching intravenous catheter insertion skills," *CIN: Computers, Informatics, Nursing*, vol. 36, no. 2, pp. 98–105, 2018.

[31] Y. R. Narnaware and M. Neumeier, "Use of a virtual human cadaver to improve knowledge of human anatomy in nursing students," *Teaching and Learning in Nursing*, vol. 16, no. 4, pp. 309–314, Oct. 2021.

[32] T. Saastamoinen, M. Harkanen, K. Vehvilainen-Julkunen, and A. Naslindh-Ylispangar, "Impact of 3D simulation game as a method to learn medication administration process: Intervention research for nursing students," *Clinical Simulation in Nursing*, vol. 66, pp. 25–43, May 2022.

[33] A. J. Q. Tan, C. C. S. Lee, P. Y. Lin, S. Cooper, L. S.T. Lau, W. L. Chua, and S. Y. Liaw, "Designing and evaluating the effectiveness of a serious game for safe administration of blood transfusion: A randomized controlled trial," *Nurse Education Today*, vol. 55, pp. 38–44, August 2017.

[34] R. A. Tubelo, F. F. Portella, M. A. Gelain, M. M. C. de Oliveira, A. E. F. de Oliveira, A. Dahmer, and M. E. B. Pinto, "Serious game is an effective learning method for primary health care education of medical students: A randomized controlled trial," *International Journal of Medical Informatics*, vol. 130, 103944, Oct. 2019.

[35] K. M. Enneking, G. R. Breitenstein, A. F. Coleman, J. H. Reeves, Y. Wang, and N. P. Grove, "The evaluation of a hybrid, general chemistry laboratory curriculum: Impact on students' cognitive, affective, and

psychomotor learning," *Journal of Chemical Education*, vol. 96, no. 6, pp. 1058–1067, 2019.
[36] W. Tarng, C. C. Hsie, C. M. Lin, and C. Y. Lee, "Development and application of a virtual laboratory for synthesizing and analyzing nanogold particles," *Journal of Computers*, vol. 12, no. 3, pp. 270–283, May 2017. https://doi.org/10.17706/jcp.12.3.270-283
[37] K. Winkelmann, W. Keeney-Kennicutt, D. Fowler, M. Lazo Macik, P. Perez Guarda, and J. C. Ahlborn, "Learning gains and attitudes of students performing chemistry experiments in an immersive virtual world," *Interactive Learning Environments*, vol. 28, no. 5, pp. 620–634, May 2019. https://doi.org/10.1080/10494820.2019.1696844
[38] S. Lebdai, M. Mauget, P. Cousseau, J. C. Granry, and L. Martin, "Improving academic performance in medical students using immersive virtual patient simulation: A randomized controlled trial," *Journal of Surgical Education*, vol. 78, no. 2, pp. 478–484, March–April 2021. https://doi.org/10.1016/j.jsurg.2020.08.031
[39] J. Y. F. Chang, T. C. Lin, L. H. Wang, F. C. Cheng, and C. P. Chiang, "Comparison of virtual microscopy and real microscopy for learning oral pathology laboratory course among dental students," *Journal of Dental Sciences*, vol. 16, no. 3, pp. 840–845, 2021.
[40] Y. C. Du, S. C. Fan, and L. C. Yang, "The impact of multi-person virtual reality competitive learning on anatomy education: A randomized controlled study," *BMC Medical Education*, vol. 20, p. 343, 2020.
[41] Y. Cui, H. Yan, Y. Su, L. Wang, T. Lu, D. Zhang, and W. Yue, "CYP2D6 genotype-based dose recommendations for risperidone in Asian people," *Frontiers in Pharmacology*, vol. 11, pp. 1–10, 2020.
[42] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Computers & Education*, vol. 147, 103778, April 2020.
[43] S. K. Loke, "How do virtual world experiences bring about learning? A critical review of theories," *Australasian Journal of Educational Technology*, vol. 31, no. 1, pp. 112–122, March 2015.
[44] C. A. Tanner, "Thinking like a nurse: A research-based model of clinical judgment in nursing," *Journal of Nursing Education*, vol. 45, no. 6, pp. 204–211, June 2006.
[45] P. R. Jeffries, B. Rodgers, and K. Adamson, "NLN Jeffries simulation theory: Brief narrative description," *Nursing Education Perspectives*, vol. 36, no. 5, pp. 292–293, 2015.
[46] J. S. Brown, A. Collins, and P. Duguid, "Situated cognition and the culture of learning," *Educational Researcher*, vol. 18, no. 1, pp. 32–42,1989.
[47] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Pearson FT Press, 2014.
[48] J. Fromm, J. Radianti, C. Wehking, S. Stieglitz, T. A. Majchrzak, and J. Brocke, "More than experience?—On the unique opportunities of virtual reality to afford a holistic experiential learning cycle," *Internet and Higher Education*, vol. 50, 100804, 2021.
[49] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation,* Cambridge University Press, 1991.
[50] T. Leung, F. Zulkernine, and H. Isah, "The use of virtual reality in enhancing interdisciplinary research and education," arXiv preprint, arXiv:1809.08585, Sept. 2018.
[51] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*, Harvard University Press*,* 1978.
[52] R. E Mayer, and L. Fiorella, *The Cambridge Handbook of Multimedia Learning*, 3rd ed. Cambridge University Press, 2022.
[53] G. Makransky and G. B. Petersen, "The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality," *Educational Psychology Review*, vol. 33*,* no. 3, pp. 937–958, Jan. 2021.
[54] H. Chan, H. Chang, and T. Huang, "Virtual reality teaching in chemotherapy administration: Randomised controlled trial," *Journal of Clinical Nursing*, vol. 30, no. 13–14, pp. 1874–1883, Feb. 2021.
[55] J. G. Cromley, R. Chen, and L. Lawrence, "Meta-analysis of STEM learning using virtual reality: Benefits across the board," *Journal of Science Education and Technology*, vol. 32, pp. 355–364, March 2023.
[56] C. Yang, J. Zhang, Y. Hu, X. Yang, M. Chen, M. Shan, and L. Li, "The impact of virtual reality on practical skills for students in science and engineering education: A meta-analysis," *International Journal of STEM Education*, vol. 11, no. 28, June 2024.