VIRTUAL REALITY'S ROLE IN INCREASING EMPATHY LEVELS IN MEDICAL STUDENTS: A SYSTEMATIC REVIEW OF QUANTITATIVE STUDIES

Oleh

Dimas Ardi Bramundito[1], Diana Listy Nababan[2], Arif Rachman[3], Sofiana[4]
[1,2]Military Medicine Study Program, Faculty of Medicine, Indonesian Defense University
[3]Department of Cell Biology and Biomolecular, Faculty of Medicine, Indonesian Defense University
[4]Department of Forensic and Medicolegal, Faculty of Medicine, Indonesian Defense University
E-mail: [1]dimasbramundito@gmail.com, [2]diananababan2004@gmail.com, [3]arif.rachman@idu.ac.id

Abstract: Empathy is a cornerstone of effective medical practice, yet numerous studies indicate a decline in empathy as students advance through medical training. Virtual Reality (VR) offers immersive, experiential learning that can potentially counteract this decline by fostering perspective-taking and emotional engagement. This systematic review analyzed quantitative studies published between 2015 and 2025 assessing VR interventions designed to enhance empathy among medical students. Following PRISMA 2020 guidelines, nine eligible studies (N = 553) were identified from four databases. The included designs comprised five randomized controlled trials (RCTs) and four single-group pre–post studies, primarily utilizing immersive head-mounted VR simulations. Across studies, all reported post-intervention empathy increases, with six demonstrating statistically significant improvements (p < 0.05). The strongest effects were noted in VR simulations portraying stigmatized or vulnerable patient perspectives, such as psychosis, depression, and chronic illness. Methodological appraisal using the MERSQI tool indicated moderate study quality (scores 10–13.5/18). Overall, immersive VR demonstrates measurable effectiveness in improving empathy among medical students, surpassing traditional didactic approaches. However, further large-scale, longitudinal research is required to confirm sustainability and translate self-reported gains into clinical behavioral outcomes.

INTRODUCTION

Empathy is an integral component of medical professionalism, it forms the cornerstone of the fundamental and dynamic physician-patient relationship. The capacity for empathy is vital for providing high quality patient-centered care which minimizes communication barriers and fosters a strong relational alliance built on trust. Given this foundational importance, the cultivation and maintenance of empathy is regarded as a core professional

competency that must be explicitly addressed within medical education curriculum.

At the psychosocial level, empathic engagement fosters a trusting and open therapeutic relationship which diminishes communication constraints and encourages patients to disclose their full illness narrative without concealment. This enhanced rapport leads to a more accurate diagnosis and crucially, better patient compliance with treatment plans, and self care adherence which ultimately results in optimized healthcare outcomes and increased patient satisfaction.

Despite its critical importance, empirical evidence consistently highlights a challenging phenomenon within medical education which is the tendency for clinical empathy levels to erode as students progress through their medical training. This decline represents a failure of traditional curriculum to sustain a core humanistic quality against the pressures of professional socialization.

Longitudinal studies utilizing validated instruments such as the JSE have confirmed and identified a particularly vulnerable period in the curriculum. The most pronounced decline typically occurs during the critical transition to the clinical phase of medical school, often corresponding to the third year. This timing is highly paradoxical as this stage marks the student's increased exposure to direct patient contact, making empathy most needed at the moment it tends to diminish. This phenomenon is sometimes dramatically summarized as, "The devil is in the third year," reflecting the intense challenges and psychological pressures inherent to early clinical rotations. Cross-sectional studies corroborate this trend. Research shows that while medical students may start with high levels of empathy, there is a consistent and measurable drop in mean empathy scores with increasing years of education even if not always reaching statistical significance in every specific cohort study.

Virtual Reality (VR) technology presents a potent answer to the challenges facing traditional medical education by offering a significant shift toward scalable and standardized experiential learning. VR utilizes computer generated three dimensional (3D) environments to create highly immersive and interactive simulated experiences. These simulations are often delivered through Head Mounted Displays (HMD) which provide a high sense of presence and allow users to engage with virtual objects, characters, and complex scenes which compels the individual to assume that they are experiencing a real life scenario.

The primary functional advantages of immersive VR compared to physical simulation include the remarkable accessibility and replicability of the learning experience, minimal setup time, potential for remote and collaborative learning (multiplayer VR), and a long term reduction in costs and resource allocation compared to perpetual reliance on costly physical facilities and standardized actors. This controlled environment ensures that critical and high stakes clinical experiences are delivered in a safe and secure environment which allows for critical decision making practice without posing risk to actual patients.

Despite VR's clear technical capacity to address the core pedagogical needs of modern medical education, analysis of contemporary VR research shows an overwhelming emphasis on the development of technical and psychomotor competencies over humanistic skills. This substantial bias towards technical application leaves the domain of non psychomotor skills which includes empathy, communication, and clinical management, underrepresented.

This underrepresentation highlights the crucial absence of statistically significant and definitive validation by suggesting that the current body of quantitative literature dedicated

to non-technical skills, including empathy, suffers from several weaknesses which is a low volume of studies, methodological heterogeneity, and insufficient sample sizes. This validates the necessity for a systematic review that is highly focused exclusively on quantitative studies related to VR and empathy which ensures that the existing and scattered data is rigorously consolidated and analyzed to determine a definitive magnitude of effect.

1. **Empathy**

    Empathy is a complex and multidimensional construct that is integral to effective social and clinical interactions. It is generally defined as the ability to understand, be aware of, be sensitive to, and vicariously experience the feelings, thoughts, and experiences of another person. In the clinical setting, empathy is often described as the act of correctly acknowledging the emotional state of another. Multiple instruments have been developed to measure levels of empathy, among them are :

    - Jefferson Scale of Empathy
      A 20-item scale measuring cognitive empathy in medical students focusing on understanding and compassionate care towards patients. Uses a 7-point Likert scale, is widely validated, and specific to healthcare contexts.
    - Interpersonal Reactivity Index
      A general empathy measure with 28 items and four subscales: Perspective Taking (adopting others' viewpoints), Empathic Concern (sympathy for others), Personal Distress (self anxiety in tense settings), and Fantasy (imaginative transposition into fictional situations). Uses a 5-point Likert scale and measures multidimensional empathy not specific to medicine.
    - Toronto Empathy Questionnaire
      A 16-item scale assessing behavioral, emotional, cognitive, and physiological aspects of empathy with good reliability. Used broadly in research beyond medicine.

2. **Virtual Reality**

    Virtual Reality (VR) is defined as a computer generated three dimensional environment that users can explore and interact with, often through the use of specialized equipment such as head mounted displays (HMDs) such as HTC Vive and Meta Quest devices that enable immersive experiences. VR technology tricks the brain into perceiving these virtual environments as real by simulating multiple senses, primarily vision and hearing. Levels of immersion in virtual reality are divided into two which are:

    - Immersive
      Involves HMDs that provide a fully immersive experience by isolating users from the physical environment and offering a 360 degree interactive virtual space.
    - Semi-Immersive
      Utilizes setups like desktop VR or Cave Automatic Virtual Environment (CAVE) systems where users interact with a virtual environment but remain partly aware of the physical surroundings. Examples include flight simulators and virtual labs.

RESEARCH METHOD

This mixed methods systematic review was conducted and reported following the guidelines established in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Statement.

The inclusion and exclusion criteria were defined using the PICO framework. P (Population) included medical students across both pre-clinical and clinical years, excluding those in dentistry and nursing programs, I (Intervention) was strictly defined as the use of immersive Virtual Reality (VR) requiring a Head Mounted Display (HMD) for empathy or communication skills training. Studies using non-immersive digital methods such as 2D video, desktop simulations, or non-HMD virtual classrooms were excluded. C (Comparison) required the presence of a control or comparison group including standard lecture, simulation, or no intervention, as well as a pre-intervention baseline for single group studies. O (Outcome) was limited to studies reporting quantitative results using a standardized and validated empathy scale such as JSE or IRI. We included Randomized Controlled Trials (RCTs), Non-Randomized Controlled Trials, and Controlled Pre-Post studies. Exclusions included other review articles, qualitative studies, case reports, editorials, and abstracts without full text data. Only articles published between January 2015 and October 2025 (the last 10 years) in English or Indonesian were considered.

The literature search was performed across four major electronic databases: PubMed, Google Scholar, Semantic Scholar, and OpenAlex. The core search strategy utilized a combination of key terms and Boolean operators (AND/OR) to maximize sensitivity, focusing on the population, intervention, and outcome. Relevant keywords were utilized in the respective databases such as "Medical Students", "Empathy" Training" OR "Empathy Education" AND "VR" OR "Virtual Reality" OR "Immersive" AND "Empathy Score" OR "Empathy Level". Reference lists of all included articles and relevant systematic reviews were also hand searched to identify additional sources.

The screening process involved three stages conducted independently by two reviewers to minimize bias. First, all search results were aggregated and duplicates were removed which was followed by an independent screening of titles and abstracts. Second, all potentially relevant articles underwent a full text screening against the detailed eligibility criteria. Any disagreements between the two independent reviewers at the full text stage were resolved through discussion to reach a consensus. The entire screening process was documented in a PRISMA Flowchart.**Error! Bookmark not defined.**
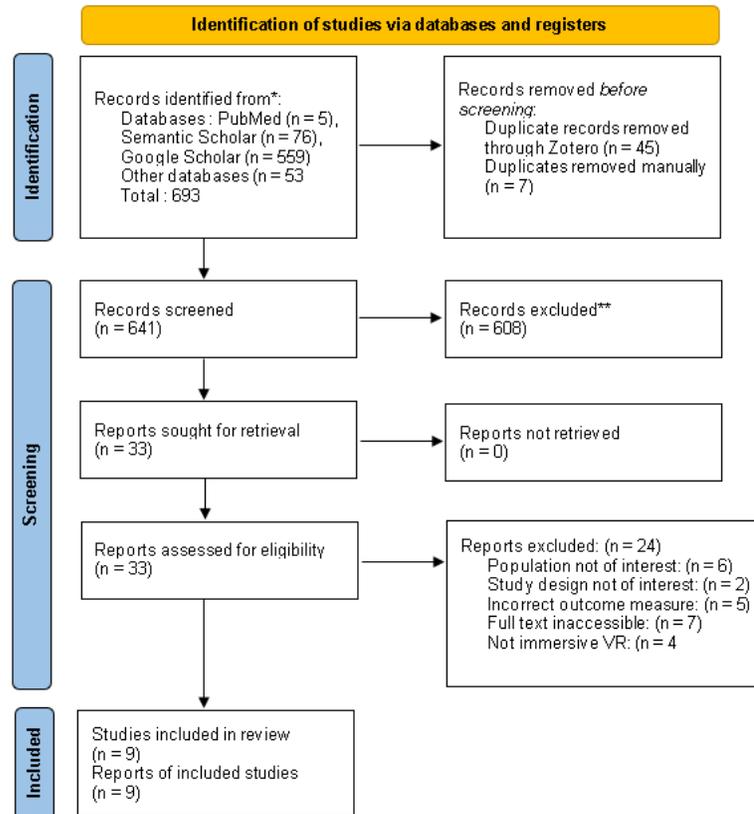
Figure 1. PRISMA Systematic Review Flow Chart

Ultimately, 9 articles met all established eligibility criteria and were retained for data extraction and synthesis. The retained literature represents a heterogeneous collection of interventional designs which provide a comprehensive quantitative perspective on the effect of VR on medical student empathy. The methodological quality of the nine studies was appraised using the Medical Education Research Study Quality Instrument (MERSQI) which is a validated tool designed to assess the methodological quality of research studies in medical education. The instrument assesses studies across six domains, particularly key features of studies such as study design, sampling, type of data, validity of the evaluation instruments used, data analysis, and the outcomes measured until yielding a total score with a maximum possible score of 18.

Table 1. MERSQI Appraisal of Included Studies

Author (Year)	Study Design	Sampling	Type of Data	Validity of Evaluation Instruments' Scores	Data Analysis	Outcome	MERSQI Total Score
Mircheraghi, Khadem-Rezaiyan, and Zeraati (2025)	3	2.5	1	3	3	1	13.5
Zare-Bidaki	3	1	1	3	3	1	12

et al. (2022)							
Lin et al. (2024)	3	2.5	1	3	3	1	13.5
Tomy et al. (2023)	3	1	1	2	3	1	11
Al-Basri (2019)	3	1	1	3	3	1	12
Nagasundaram et al. (2025)	1.5	1	1	3	3	1	10.5
Rehl et al. (2024)	1.5	1	1	3	3	1	10.5
Alieldin et al. (2024)	1.5	1	1	3	3	1	10.5
Kohn et al. (2025)	1.5	1.5	1	3	2	1	10

Source: Scoring from appraisal of included studies using MERSQI

The MERSQI (Medical Education Research Study Quality Instrument) scores for the nine included studies reflected a moderate standard of methodological quality across the evaluated dimensions. The majority of studies achieved the maximum score in categories such as study design, validity of evaluation instruments' scores, and data analysis which suggests strong adherence to rigorous research design and robust analytic approaches. Some studies were rated highly in sampling with a few studies scoring 2.5 indicating well executed but not fully optimal sampling strategies. The mean total MERSQI scores ranged from 10 to 13.5 out of a possible maximum with a median score of 11, highlighting a moderate quality although some studies scored lower mainly due to reduced scores in type of data or sampling. The studies by Mircheraghi et al. (2025)**Error! Bookmark not defined.** and Lin et al. (2024)**Error! Bookmark not defined.** obtained the highest total MERSQI scores of 13.5 which reflect a higher methodological rigor among the studies. Meanwhile, lower total scores such as the study by Kohn et al. (2025)**Error! Bookmark not defined.** was largely attributable to lower scores in data analysis suggesting some limitations in these areas.

RESULTS AND DISCUSSION
1. **General Results**

A total of 9 articles were included in the quantitative synthesis, comprising a diverse array of interventional designs to address the effect of VR on empathy in medical students. These studies involved a total population of ±1813 medical students and a final participant count or sample size of 553. The majority of the evidence was comprised of interventional designs including 5 Randomized Controlled Trials (RCTs) and 4 Single-Group Pre-Post Studies. All studies focused on undergraduate medical students across pre-clinical and clinical phases and used validated scales with the Jefferson Scale of Empathy (JSE) being the predominant measurement tool.

Table 2. Characteristics of Included Studies

No.	Author (Year)	Title	Study Design	Population & Sample Size (N_{total}/N_{sample})	VR Technology Utilization	Comparison Methodology	Empathy Scale Used
1	Mircheraghi, Khadem-Rezaiyan, and Zeraati (2025)**Error! Bookmark not defined.**	Enhancing Compassionate Care: Virtual Reality Boosts Medical Students' Empathy Toward Hypertensive Patients	RCT	Medical students, internal medicine residents, and fellows experiencing a clinical educational course (198/181)	VR scenario of audiovisual symptoms experienced by hypertensive patients.	Exposure to a written pamphlet about visual and auditory complications of hypertension.	Jefferson Scale of Empathy (JSE)
2	Zare-Bidaki et al. (2022)**Error! Bookmark not defined.**	Evaluating the Effects of Experiencing Virtual Reality Simulation of Psychosis on Mental Illness Stigma, Empathy, and Knowledge in Medical Students	RCT	2nd and 3rd year medical students (NR/144)	One session of a Virtual Reality Simulation of Psychosis (VRSP).	Routine educational program, which was visiting patients under supervision in the psychiatric ward	Jefferson Scale of Empathy (JSE) - Student Version
3	Lin et al. (2024)**Error! Bookmar**	Can virtual reality technology be used for	RCT	Fifth-year and sixth-year	VR scenario simulating the daily	VR scenario simulating the daily	Jefferson Scale of Empathy (JSE)

	k not defined.	empathy education in medical students: a randomized case-control study		medical students (59/59)	life of a depressed medical student	life of a general (normal) medical student	
4	Tomy et al. (2023)**Error! Bookmark not defined.**	Immersive and screen-based virtual reality simulations enhance empathy	RCT	Medical students in pre-medical and medical phases (42/36)	Immersive VR scenario focused on social isolation utilizing HTC Vive HMD.	Screen based VR scenario focused on social isolation utilizing Apple iMac and headphones	Jefferson Scale of Empathy (JSE) - Student Version
5	Al-Basri (2019)**Error! Bookmark not defined.**	Learning Empathy Through Virtual Reality: A Mixed Methods Study	RCT with pre/post test	Second to fourth year clinical year medical students (34/34)	Supplementary VR-based educational intervention comprising of patient POV and debrief.	"Education as usual" (medical school communication skills sessions and clinical interactions).	Toronto Empathy Questionnaire
6	Nagasundaram et al. (2025)**Error! Bookmark not defined.**	Measuring the effects and feasibility of virtual reality in developing empathy among medical students in	Single Arm Pre/Post Study	Doctor of Medicine medical students on child and adolescent health placement in teaching	Short clinical scenario in VR from the point of view of a child admitted with asthma.	Internal pre-interventional baseline	Jefferson Scale of Empathy (JSE)

		a paediatric setting		hospital (200/22)			
7	Rehl et al. (2024)**Error! Bookmark not defined.**	Feasibility of a cinematic-virtual reality training program about opioid use disorder for osteopathic medical students: a single-arm pre–post study	Single Arm Pre-post Study	Osteopathic medical students from three campuses (1011/48)	Storyline based cine-VR episodes designed to educate medical students about Opioid Use Disorder (OUD), social determinants of health, and implicit bias.	Internal pre-interventional baseline	Jefferson Scale of Empathy (JSE)
8	Alieldin et al. (2024)**Error! Bookmark not defined.**	Efectiveness of immersive virtual reality in teaching empathy to medical students: a mixed methods study	Single Arm Pre-post Study	First-year medical students (105/19)	VR scenario that focused on loneliness and social isolation in older adults and their efect on their health outcomes was selected	Internal pre-interventional baseline	Jefferson Scale of Empathy (JSE)
9	Kohn et al. (2025)**Error! Bookmark not defined.**	Virtual reality communication training in pain medicine: effects on	Within-Subjects Pre-post Study	Third- and fourth-year medical students (20/10)	Immersive virtual reality simulation using a HMD utilizing a human-	Internal pre-interventional baseline	Jefferson Scale of Empathy (JSE)

medical students' racial bias, empathy, and interview performance with virtual patients	operated Virtual Patient (VP) avatar of different races (black & white) interacting with participants in clinical setting.

Source: General findings from included studies

Geographically, the research demonstrates a global effort with studies originating from North America**Error! Bookmark not defined.Error! Bookmark not defined.Error! Bookmark not defined.Error! Bookmark not defined.**, the Middle East**Error! Bookmark not defined.Error! Bookmark not defined.Error! Bookmark not defined.**, Taiwan**Error! Bookmark not defined.**, and Australia**Error! Bookmark not defined.**, confirming the universal relevance of this educational challenge. The combined sample sizes, while individually modest, represent a collective exposure of several hundred medical trainees to VR empathy training.

The core of the intervention across the nine studies heavily relies on Immersive Virtual Reality (IVR) typically delivered via Head Mounted Displays (HMDs) which are employed to maximize the sense of presence and embodiment. This modality is favored by researchers as it is theorized to be the most effective mechanism for stimulating genuine perspective taking.**Error! Bookmark not defined.Error! Bookmark not defined.** A smaller subset utilized less immersive approaches such as Cinematic VR (Cine-VR)**Error! Bookmark not defined.** or a comparison between immersive and screen-based VR**Error! Bookmark not defined.** which allow for a future assessment of immersion level as a moderating factor. The specific patient scenarios are diverse and strategically target populations often facing stigma or complex communication barriers. These range from simulating the experience of psychosis**Error! Bookmark not defined.**, depression**Error! Bookmark not defined.**, and Opioid Use Disorder (OUD)**Error! Bookmark not defined.**, to general situations like the perspective of a child in a paediatric setting**Error! Bookmark not defined.**, social isolation in older adults**Error! Bookmark not defined.**, and patients with hypertension**Error! Bookmark not defined.**.

The primary outcome measure is highly standardized which lend strong comparability to the results. Nearly all studies utilized one of the versions of the Jefferson Scale of Empathy (JSE), either the student (JSE-S) or health professional student (JSE-HPS) edition which provides a consistent quantitative metric of self-reported empathy with one study utilizing the Toronto Empathy Questionnaire (TEQ).

2. **Quantitative Results**

....
Journal of Innovation Research and Knowledge

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

Of the five comparative studies included in the review which consisted of RCTs, three studies (60%) reported a statistically significant increase in empathy scores favoring the VR intervention group over the control or comparison group while two studies (40%) found no statistically significant difference in empathy change between the VR and control groups based on the post-test scores

Table 3. Descriptive Statistics of RCTs

Author (Year)	Empathy Scale Used	Sample Size	Intervention Group Size ($N_{VR)}$	Comparison Group Size ($N_{Comp)}$	Mean ±SD (VR Pre-Test)	Mean ±SD (VR Post-Test)	Mean ±SD (Comp Pre-Test)	Mean ±SD (Comp Post-Test)	P-value for change
Al-Basri (2019)	Toronto Empathy Questionnaire	34	17	17	47.11	52.18	47.53	48.94	p = 0.01
Zare-Bidaki et al. (2022)	Jefferson Scale of Empathy (JSE) – Student Version	144	72	72	118.81 ± 12.50	128.76 ± 9.40	115.75 ± 13.49	121.99 ± 12.39	p < 0.000
Mircheraghi, Khadem-Rezaiyan, and Zeraati (2025)	Jefferson Scale of Empathy (JSE)	181	93	88	107.26± 12.69	112.46± 12.67	106.77± 13.38	108.17± 12.59	p = 0.024
Lin et al. (2024)	Jefferson Scale of Empathy (JSE)	59	30	29	5.485 ± 0.422	5.558 ± 0.509	5.597 ± 0.468	5.691 ± 0.490	p = 0.39
Tomy et al. (2023)	Jefferson Scale of Empathy (JSE) – Student Version	42	32	10	113.50	122.31	115.70	119.80	p = 0.142

Source: Quantiative results from RCTs included

Al-Basri (2019)**Error! Bookmark not defined.** reported one of the strongest findings, confirming a statistically significant difference in the mean change of TEQ scores that favored the VR group. The VR based curriculum group achieved a mean post-test score of 52.18

compared to a pre-test score of 47.11, representing an average increase of 5.07 points. The reported probability value of p = 0.01 provides a high degree of confidence that the observed difference was due to the VR intervention rather than chance. In addition, the effect size of the VR intervention was measured by the author, suggesting a large effect size with a Cohen's d value of 0.89. Zare-Bidaki et al. (2022)**Error! Bookmark not defined.** utilized a VR simulation of psychosis to reduce mental health stigma and enhance empathy involving 144 medical students. the VR Simulation of Psychosis (VRSP) group demonstrated a significant increase in total empathy, reporting a mean post-intervention score of 128.76±9.40 compared to a pre-intervention baseline of 118.81±12.50. This study reported a statistically significant difference in empathy scores between the VR group and the control group (p < 0.000). Mircheraghi et al. (2025)**Error! Bookmark not defined.** which focused on empathy toward hypertensive patients in a VR simulation of audiovisual symptoms, reported a statistically significant mean difference in the JSE scores favoring the VR group as well as between pre-test and post-test scores in the VR group. The VR intervention group saw their mean Jefferson Scale of Empathy (JSE) score rise from a pre-intervention mean of 107.26±12.69 to a post-intervention mean of 112.46±12.67. The reported p-value for pre-test and post-test results in the intervention group was p<0.001 indicating a significance in the result scores. The reported between-group p-value was p = 0.024 which confirms the superiority of the VR intervention at the standard level of significance.

Lin et al. (2024)**Error! Bookmark not defined.** which examined a VR simulation of depression, failed to find a statistically significant difference in the change of JSE scores between the VR intervention group and the control group (General Medical Student Scenario) despite a slight positive mean change between pre-intervention and post-intervention groups (+0.073). The total JSE score mean increased slightly from 5.485±0.422 to 5.558±0.509 in the intervention group. The authors reported a p-value of p = 0.39 which is well above the conventional threshold for statistical significance (p < 0.05). Tomy et al. (2023)**Error! Bookmark not defined.** compared both immersive VR and screen-based VR against a passive control. Although the immersive VR groups showed significant within-group pre-post gains in empathy of +8.81 with a p-value of p < 0.001, the critical comparative analysis demonstrated no statistically significant difference when the immersive VR was directly contrasted with the control condition (p = 0.142).

Table 4. Descriptive Statistics of Single Group Pre-post Design

Author (Year)	Empathy Scale Used	Sample (N_{VR})	Mean $\pm$ SD (Pre-Test)	Mean±SD (Post-Test)	Mean Difference (Post - Pre)	P-value for change
Nagasundaram et al. (2025)	Jefferson Scale of Empathy (JSE)	22	113.9±7.6	122.5±8.1	+8.6	p < 0.0001
Rehl et al. (2024)	Jefferson Scale of Empathy	36	88.7±14.6	93.8±16.0	+5.1	0.023

	(JSE)						
Alieldin et al. (2024)	Jefferson Scale of Empathy (JSE)	19	121.52±7.36	127.47±6.23	+5.94	p < 0.01	
Kohn et al. (2025)	Jefferson Scale of Empathy (JSE)	10	118.6±7.1	121.4±7.8	+2.8	NR	

Source: Quantitative results from pre-post studies included

Of the four studies utilizing a single-group pre-post design, three (75%) reported a statistically significant increase in the mean empathy scores when comparing pre-intervention (baseline) results to immediate post-intervention results. 1 study was unable to report statistical inferences, stating a methodological limitation in sample size, however the study may portray a trend while unable to provide statistical evidence.

Nagasundaram et al. (2025)**Error! Bookmark not defined.** reported the strongest statistical outcome while focusing on a VR paediatric clinical scenario. Mean empathy scores for the intervention group increased from a pre-VR score of 113.9±7.6 to a post-VR score of 122.5±8.1, representing an average gain of +8.6 points. This increase in the mean empathy scores was observed and deemed highly significant with a probability value of $p < 0.0001$. This result indicates a negligible probability that the observed empathy gain occurred by chance. In addition, this study also reported a Cohen's d value of 1.1, suggesting a large effect by VR intervention on post-intervention empathy scores. Rehl et al. (2024)**Error! Bookmark not defined.** utilized a cinematic-VR program to address Opioid Use Disorder (OUD) and demonstrated a statistically significant improvement in empathy scores after the online cine-VR training. The mean JSE score for the VR group rose from 88.7±14.6 (baseline) to 93.8±16.0 (post-training). This finding was supported by a p-value of $p = 0.023$, establishing a significant measurable positive shift in attitudes toward a stigmatized patient population. The effect size of the intervention on post-test scores was also calculated and reported resulting in a Cohen's d value of 0.40 which suggests a small effect. Alieldin et al. (2024)**Error! Bookmark not defined.** focused on VR interventions focused on experiences of older adults struggling with social isolation. First-year medical students (N=19) undergoing an immersive VR scenario recorded an average gain of +5.94 points with mean scores rising from a pre-test average of 121.52±7.36 to a post-test average of 127.47±6.23. A paired sample t-test on the pre- and post-test JSE scores confirmed a statistically significant improvement in these empathy levels with a reported p-value of $p < 0.01$. Kohn et al. (2025)**Error! Bookmark not defined.**, while lacking statistical support, provided raw data which showed a directional change consistent with the other studies. Empathy scores increased from a mean of 118.6±7.1 (Pre-trial) to 121.4±7.8 (Post-trial). This small increase (+2.8) suggests a potential positive effect, however as no statistical inference was made, it cannot contribute as statistical evidence.

A complete and thorough analysis of the numerical data reveals a nuanced picture, encompassing not only studies showing significant success but also those whose findings fell short of statistical significance. Out of 9 studies, all of them reported an increase in empathy

....

ISSN 2798-3471 (Cetak)   **Journal of Innovation Research and Knowledge**
ISSN 2798-3641 (Online)

levels after exposure to VR intervention, however only 6 reported a statistical significance to improvements, 2 reported a lack of statistical significance, while 1 failed to report any inferencial data related to empathy levels.

VR's ability to trigger affective emotional responses that mimic genuine patient reactions is a key differentiator from traditional didactic or even many simulation techniques.**Error! Bookmark not defined.** Empirical evidence supports a correlation between subjective engagement and measured empathy gains. Specifically, correlation analyses revealed that a higher perceived sense of presence ($r = 0.415$, $p = 0.01$) and a greater sense of immersion ($r = 0.308$, $p = 0.049$) both demonstrated a statistically significant positive correlation with total JSE scores.**Error! Bookmark not defined.** This indicates that the more realistic and engaging the simulation feels to the trainee, the stronger the resultant empathy gain is likely to be. VR also consistently demonstrates effectiveness in boossting measurable components of empathy. Specifically the cognitive component of Perspective Taking which measures the intellectual capacity to understand another's viewpoint showed statistically significant increases in both the depression scenario ($p = 0.03$)**Error! Bookmark not defined.** and the pediatric scenario ($p < 0.0035$). Similarly, Compassionate Care which represents the affective/behavioral commitment to help and acknowledge patient emotions also improved significantly in depression ($p = 0.01$)**Error! Bookmark not defined.** and pediatric scenario ($p = 0.011$).**Error! Bookmark not defined.**

Overall, VR supports a strong affective and emotional component by triggering emotional responses identical to those of patients such as feelings of vulnerability, isolation, and loss of agency as well as effectively targets conditions susceptible to provider stigma such as Opioid Use Disorder (OUD) and psychosis leading to significant measurable empathy gains. This makes VR a valuable tool for cultivating humanism in healthcare and building emotional intelligence by increasing self-awareness of emotional implications and implicit biases.

While the preliminary quantitative findings are overwhelmingly positive, this study is still limited by methodological restraints based upon the quality of the evidence base, instruments used, as well as flaws in the sustainability and longetivity of the data. A core constraint is the pervasive reliance on small and convenient samples (N=19**Error! Bookmark not defined.**, N=22**Error! Bookmark not defined.**, etc.) often derived from single academic centers. This limits the generalizability of findings across diverse medical school settings. Another critical gap in the literature is the scarcity of data concerning the longitudinal persistence of empathy gains. The reported significant increases rely heavily on immediate post-intervention testing. Only one study included a prolonged follow-up which provided only self-reported qualitative data on the endurance of the experience after six months.**Error! Bookmark not defined.** The heavy reliance on self-reported instruments (JSE and TEQ) introduces vulnerability to social desirability and subject bias which potentially inflates the measured empathy scores.

CONCLUSION

The quantitative evidence strongly supports the effectiveness of Virtual Reality simulations in significantly enhancing empathy levels among medical students which demonstrates superior gains over traditional didactic and passive comparative methods.

This efficacy is driven by the unique capability of Immersive VR to induce affective emotional responses and facilitate deep cognitive perspective taking, therefore converting the intellectual abstraction of patient experience into tangible and personalized reality.

This study notions that the efficacy of VR translates directly into actionable applications in the medical education curriculum. VR interventions present substantial logistical advantages over traditional empathy training modalities like standardized patient encounters or long term workshops which are often time consuming and resource intensive. The virtual environment provides a non-judgmental and safe space for learners to make mistakes and process strong emotional reactions such as feelings of vulnerability or moral injury without real world patient consequences. A longitudinal curricular approach is recommended as VR exposure can be delivered early in pre-clinical years to prime humanistic values and later in clinical years to provide a psychologically safe space for students to reflect on and cope with difficult real world clinical experiences.

Further research must address the common methodological limitations of the included studies such as small sample sizes, the reliance on self-reported outcome measures such as JSE and TEQ, as well as the sustainablity of the data in the long term. Future investigations should seek to generalize these findings across diverse institutional settings and integrate objective behavioral metrics such as structured clinical observation during OSCEs to verify that the self-reported total empathy gains translate into demonstrable and improved patient-centered care. Since empathy is known to naturally decline over time and under systemic pressure, future large scale studies should incorporate quantitative longitudinal follow-up measures (at 3, 6, and 12 months, etc.) to establish the sustained impact and determine the necessary "dose" required to prevent the observed gains in empathy from fading.

REFERENCES

[1] Voultsos, P., A. Gavra, and G.N. Andreou. 2024. "The Role of Empathy in Medical Education and Clinical Practice". *Mental Health & Human Resilience International Journal.* 8(1): 1-8. https://doi.org/10.23880/mhrij-16000246.

[2] Tan. L., M.K. Le, C.C. Yu, S.Y. Liaw, T. Tierney, Y.Y. Ho, E. Lim, D. Lim, R. Ng, C. Ngeow, and J. Low. 2021. "Defining Clinical Empathy: A Grounded Theory Approach from the Perspective of Healthcare Workers and Patients in a Multicultural Setting". *BMJ Open.* 11: 1-9. https://doi.org/10.1136/bmjopen-2020-045224.

[3] Wang, Y., Q. Wu, Y. Wang, and P. Wang. 2022. "The Effects of Physicians' Communication and Empathy Ability on Physician–Patient Relationship from Physicians' and Patients' Perspectives". *Journal of Clinical Psychology in Medical Settings.* 29: 849–860, https://doi.org/10.1007/s10880-022-09844-1.

[4] Hojat, M., S.C. Shannon, J. DeSantis, M.R. Speicher, L. Bragan, dan L.H. Calabrese. 2020. "Does Empathy Decline in the Clinical Phase of Medical Education? A Nationwide, Multi-Institutional, Cross-Sectional Study of Students at DO-Granting Medical Schools". *Academic Medicine.* 95(6): 911–18. https://doi.org/10.1097/ACM.0000000000003175.

[5] Sihombing, N.R., I. Armyanti, and A. Wicaksono. 2021. "Tingkat Empati Mahasiswa Program Studi Kedokteran Fakultas Kedokteran Universitas Tanjungpura, Pontianak,

Indonesia". *Cermin Dunia Kedokteran.* 48(4): 197–207.

[6] Kokorelias, K.M., M. Chiu, S. Paul, L. Zhu, N. Choudhury, C.G. Craven, A. Dubrowski, T. Redublo, B. Kapralos, M.S.D Smith, A. Shnall, J. Sadavoy, and A. Burhan. 2024. "Use of Virtual Reality and Augmented Reality Technologies to Support Resilience and Skill-*Building* in Caregivers of Persons With Dementia: A Scoping Review". *Cureus*. 16(7): 1-13. https://api.semanticscholar.org/CorpusID:271083969.

[7] Aguirre-Calan. G.M., A. Mejía-Jálabe, J.A. Romero-Torres, and P.A. Morillo-Andrade. 2024. "Assessment of the Impact of Virtual Reality on Empathy and Learning in University Students". *International Journal of Religion.* 5(11): 6209-6225. https://api.semanticscholar.org/CorpusID:272196861.

[8] Collier, R., R. Darling, and K. Browne. 2025. "An Immersive Virtual Reality Simulation Scenario to Improve Empathy in Nursing Students". *CIN: Computers, Informatics, Nursing.* 43(5). https://api.semanticscholar.org/CorpusID:276093720.

[9] Dhar, E., U. Upadhyay, Y. Huang, M. Uddin, G. Manias, D. Kyriazis, U. Wajid, H. Al Shawaf, and S.S. Abdul. 2023. "A Scoping Review to Assess the Effects of Virtual Reality in Medical Education and Clinical Care". *Digital Health*. 9: 1-18. https://api.semanticscholar.org/CorpusID:257245656.

[10] Moudatsou, M., A. Stavropoulou, A. Philalithis, and S. Koukouli. 2020. "The Role of Empathy in Health and Social Care Professionals". *Healthcare*. 8(1): 1-9. https://doi.org/10.3390/healthcare8010026.

[11] Hojat, M., J. DeSantis, S.C. Shannon, L.H. Mortensen, M.R. Speicher, L. Bragan, M. LaNoue, and L.H. Calabrese. 2018. "The Jefferson Scale of Empathy: A Nationwide Study of Measurement Properties, Underlying Components, Latent Variable Structure, and National Norms in Medical Students". *Advances in Health Sciences Education*. 23: 899-920. https://doi.org/10.1007/s10459-018-9839-9.

[12] Akgün, Ö, M. Akdeniz, E. Kavukcu, and H.A. Avci. 2020."Medical Students' Empathy Level Differences by Medical Year, Gender, and Specialty Interest in Akdeniz University". 7: 1-11. https://doi.org/10.1177/2382120520940658.

[13] Liao, C.H. 2025. "The Role of Virtual Reality in Enhancing Behavioral Empathy: Exploring Cognitive Absorption, Engagement, and Emotional Moderation Using Multivariate Methods". *Humanities and Social Sciences Communications*. 12: 995. https://doi.org/10.1057/s41599-025-05334-5.

[14] Hamad, A. and B. Jia. 2022. "How Virtual Reality Technology Has Changed Our Lives: An Overview of the Current and Potential Applications and Limitations". *International Journal of Environmental Research and Public Health.* 19: 1-14. https://doi.org/10.3390/ijerph191811278.

[15] Page, M.J., J.E. Mackenzie, P.M. Bossuyt, I. Boutron, T.C Hoffman, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E Brennan, R. Chou, J. Glanville, J.M. Grimshaw, A. Hróbjartsson, M.M. Malu, T. Li, E.W. Loder, E. Mayo-Wilson, S. McDonald, L.A. McGuinness, L.A. Stewart, J. Thomas, A.C. Tricco, V.A. Welch, P. Whiting, and D. Moher. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews". *BMJ*. https://doi.org/10.1136/bmj.n71.

[16] Jaros, S. and G.B. Dallaghan. 2024. "Medical Education Research Study Quality Instrument: An Objective Instrument Susceptible to Subjectivity". *Medical Education*

Online. 29(1). https://pmc.ncbi.nlm.nih.gov/articles/PMC10810632/.

[17] Mircheraghi, S.F., M. Khadem-Rezaiyan, and A.A Zeraati. 2025. "Enhancing Compassionate Care: Virtual Reality Boosts Medical Students' Empathy Toward Hypertensive Patients". *Journal of Research and Health*. 15(5): 493-503. https://jrh.gmu.ac.ir/browse.php?a_id=2696\&slc_lang=en\&sid=1\&ftxt=1\&html=1.

[18] Zare-Bidaki, M., A. Ehteshampour, M. Reisaliakbarighomi, R. Mazinani, M.R.K. Ardakani, A. Mirabzadeh, R. Alikhani, M. Noroozi, F. Momeni, A.D. Samani, M.M.M. Tavana, A. Esmaeili, and S.B. Mousavi. 2022. "Evaluating the Effects of Experiencing Virtual Reality Simulation of Psychosis on Mental Illness Stigma, Empathy, and Knowledge in Medical Students". *Frontiers in Psychiatry*. 13: 1-9. https://api.semanticscholar.org/CorpusID:248815669.

[19] Lin, H.L., Y.C. Wang, M.L. Huang, N.W. Yu, I. Tang, Y.C. Hsu, and Y.S. Huang. 2024. "Can Virtual Reality Technology Be Used for Empathy Education in Medical Students: A Randomized Case-Control Study". *BMC Medical Education*. 24. https://api.semanticscholar.org/CorpusID:273818081.

[20] Tomy, T., C. Neumann, J. Vognsen, and S. Major. 2023. "Immersive and Screen-Based Virtual Reality Simulations Enhance Empathy". *International Journal of Healthcare Simulation.* 2(2): 58-60. https://www.researchgate.net/profile/Stella-Major/publication/369419357_Immersive_and_screenbased_virtual_reality_simulatons_enhance_empathy/links/642e502e609c170a13f9b21f/Immersive-and-screen-based-virtual-reality-simulations-enhance-empathy.pdf.

[21] Albasri, M., 2020. Learning Empathy Through Virtual Reality: A Mixed Methods Study. *Thesis,* Cambridge: Master of Medical Sciences in Medical Education. https://api.semanticscholar.org/CorpusID:235136406.

[22] Nagasundaram, N., C. van Heerden, S. Porashemi, and A. Gray. 2025. "Measuring the Effects and Feasibility of Virtual Reality in Developing Empathy among Medical Students in a Paediatric Setting". *Focus On Health Professional Education,.* 26(2): 21-32. https://doi.org/10.3316/informit.T2025072100009701494115898.

[23] Rehl, D., M. Mangapora, M. Love, C. Love, K. Shaw, J. McCarthy, and E.A. Beverly. 2024. "Feasibility of a Cinematic-Virtual Reality Training Program about Opioid Use Disorder for Osteopathic Medical Students: A Single-Arm Pre–Post Study". *Journal of Osteopathic Medicine.* 124(11): 509-516. https://doi.org/10.1515/jom-2023-0188.

[24] Alieldin, R., S. Peyre, A. Nofziger, and R. Borasi. 2024. "Effectiveness of Immersive Virtual Reality in Teaching Empathy to Medical Students: A Mixed Methods Study". *Virtual Reality*. 28: 1-13.

[25] Kohn, J.N., E.A. Troyer, K.L. Wilson, R. Reddy, C. Vieten, E. Viirre, W. Joshi, J. Unger, T. Williams, D.J. Gonzales, and S. Hong. 2025. "Virtual Reality Communication Training in Pain Medicine: Effects on Medical Students' Racial Bias, Empathy, and Interview Performance with Virtual Patients". *Pain Medicine*. 26: 120-123. https://academic.oup.com/painmedicine/article-abstract/26/2/120/7815733.

HALAMAN INI SENGAJA DIKOSONGKAN