Exploring Immersive VR-based Emphatic Learning Using Interpersonal Reactivity Index

Jaziar Radianti
Department of Information Systems
University of Agder, Norway
jaziar.radianti@uia.no

Sofie Wass
Department of Information Systems
University of Agder, Norway
sofie.wass@uia.no

Marianne Hovet Steig
Department of Psychosocial Health
University of Agder, Norway
marianne.steig@uia.no

Mugula Chris Safari
Department of Psychosocial Health
University of Agder, Norway
chris.safari@uia.no

Abstract

In this study we explore the use of VR-based material as a tool for facilitating empathic reflection among bachelor students. We conducted a set of systematic experiments of 360-degree immersive materials using head-mounted Virtual Reality devices. Data was gathered from the second semester students through two experiment survey of Interpersonal Reactivity Index (IRI), covering pre- and post-tests before and after using VR. The results suggest some empathy improvement in some empathy subscales of IRI, although the amount of improvement is not so significant. The students show positive and supportive attitudes towards the use of immersive material for improving empathy skills.

Keywords: emphatic learning, VR, 360-degree immersive materials, *Interpersonal Reactivity Index*.

1. Introduction

Studies on the adoption and practices of virtual reality (VR) technology for higher education has been growing recently (Fromm et al., 2021; Majchrzak et al., 2022; Radianti et al., 2020). VR markets are estimated to reach around 1.25 USD by 2025, and VR has become more exciting (Cureton, 2023; Research-and-Market, 2023), and gradually shifting from only gaming and entertainment technologies to alternative usage in educational setting. This development has been supported by the emergence of Metaverse (Contreras et al., 2022) and advanced VR capabilities (Sanfilippo et al., 2022) that triggers further attraction toward tailoring VR to classroom settings (Jin et al., 2022). Extensive research has been conducted in various domains to explore the use of VR for educational purposes and develop VR materials for learning and training. One of the most widely discussed advantages of VR is the ability of VR to provide learners with immersive experiences of rare or risky situations, enabling them to make decisions without causing harm to themselves or their environments. Notably, the fields of medicine, nursing, and healthcare have witnessed a rising application of VR in educational settings.

However, there is a scarcity of VR studies in healthcare and nursing that specifically address the profession so-called "social educator" who work with individuals with intellectual disabilities. According to Carulla et al. (2011), intellectual disabilities are lifelong conditions characterized by the impairment of cognitive functions, which are associated with limitations in learning and adaptive behavior, and social and practical skills. The daily work setting of social educators presents numerous challenges, requiring them to adeptly navigate the delicate balance between exerting power and facilitating the autonomy and self-determination of people with intellectual disability. Consequently, the education and training of "social educators" underscore the development of empathetic skills and ethical judgment to effectively engage with and support people with intellectual disabilities. Thus, additional ethical challenges arise when incorporating observational learning practices to gain insights into how to effectively interact and support people with intellectual disabilities.

In this paper, we emphasize the significance of developing immersive materials that can be utilized in conjunction with VR technology, especially Head Mounted Display (HMD) variants. The objective is to uphold realism, presence, and engagement among students while they learn about interacting with and supporting individuals with intellectual disabilities. A set of 360-degree based immersive materials have been developed addressing day-to-day practical challenges in the interaction between social educators and people with intellectual disabilities. We use one of

URI: https://hdl.handle.net/10125/107022
978-0-9981331-7-1

the immersive scenarios for conducting experiments involving "social educator" bachelor students, with a particular focus on increasing empathy skills.

Empathy is a multifaceted socio-emotional phenomenon encompassing the ability to comprehend and respond to the emotions, thoughts, and experiences of others (Yalçın & DiPaola, 2020). This capacity is required to grasp and resonate with someone else's perspective. Many studies acknowledge that empathy plays a crucial role in shaping and nurturing social connections, as it aids in harmonizing actions, discerning the intentions of others, and promoting altruistic behavior among individuals (Yalçın & DiPaola, 2020). These features of emphatic skills are also required to be successful as a "social educator".

This paper addresses two research questions: *RQ1 How can the use of VR in combination with ethical group reflection increase bachelor students' empathy skills in? RQ2 Which empathic factors are inclined to change when exposed to immersive VR materials?*

To tackle these RQs, we conducted a comprehensive assessment of emphatic attitudes among social educator students. This involved measuring their attitudes before and after the experiments using the Interpersonal Reactivity Index measurement (Davis, 1980); see also Section 2.3. This paper presents a novel contribution by introducing the concept of using immersive materials for a more ethical approach to studying situations involving individuals with intellectual disabilities. Additionally, the paper introduces the measurement of empathy through VR-based experiments conducted with immersive materials. The case study was based on the *social educator* bachelor program (*Vernepleie*) in Norway.

2. Literature Review

2.1 VR-based Learning and Healthcare

Virtual Reality (VR) in the online Oxford dictionary is defined as a "... computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors." VR has three crucial elements: immersion, presence, and interactivity. Scholars deem this technology as promising, providing benefits beyond the entertainment purpose, and having positive attitude towards the use of VR for (higher) education (Freina & Ott, 2015; Kamińska et al., 2019). Furthermore, Jin et al. (2022) concretize on how VR will enter university classrooms by engaging relevant stakeholders (instructors, students).

VR also has inspired scholars to apply the technology as a means to enhance healthcare education. One promising healthcare area related to our study includes nurse education (e.g., Choi et al., 2022; Dean et al., 2020) and social work (e.g., Roberson & Baker, 2021). Indeed, we have noticed literature that scrutinizes VR with intellectual disabilities (Brown et al., 2016; Nabors et al., 2020). However, most of the focus of these papers are the skill developments or rehabilitation of people with intellectual disabilities such as increasing physical activity, improving literacy and understanding of different interventions for individuals with different types of impairments (Nabors et al., 2020), and life skill training (Brown et al., 2016). Existing studies do not so much scrutinize how to educate students to be prepared for various scenarios. Essentially, as a future social educator, the students should be able to handle different situations in an ethical manner, balancing their intervention approach between power and autonomy. Moreover, one of the skills required for this type of work is emphatic skill.

2.2 VR for Learning Empathy

Despite extensive history of empathy, there are inconsistent definitions of the term (Cuff et al., 2016). Empathy has been described as a concept that include cognitive and affective elements; the emotions of the target and observer; other stimuli such as imagination can evoke empathy; a self/other distinction. Empathy is influenced by both trait and state factors, and it is subject to automatic elicitation as well as top-down controlled processes. We use the empathy definition offered by Cuff et al (2016, p. 150):

> Empathy is an emotional response (affective), dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one's perception (directly experienced or imagined) and understanding (cognitive empathy) of the stimulus emotion, with recognition that the source of the emotion is not one's own

Moreover, the empathy concept has attracted VR researchers who consider that VR is an interesting instrument for studying empathy (Ventura et al., 2020). The authors argue that VR generates an emphatic behavior such as inducing helping behavior, reducing domestic violence, racial biases, and prejudice, as demonstrated by existing studies. Ventura et al (2020) pinpoint that eliciting empathy through VR is a research field that is grown quickly, and the results of the analysis provide evidence that VR tasks may be more effective at improving attitudes

toward specific social targets and motivating prosocial behavior compared with traditional interventions.

2.3 Measurements of empathy

One of the ways to explore empathy development is by measuring it. Measuring empathy has attracted researchers from time to time, as various measurement instruments have been proposed in the literature, especially in psychology and education/ counseling, with different emphasis and context concerning the applicability of these instruments. We found e.g., IRI Index or Interpersonal Reactivity index (Davis, 1980), Balanced Emotional Empathy Scale (Mehrabian, 1996), Emphatic Quotient (Baron-Cohen & Wheelwright, 2004), the Global Rating Scale (Haskard et al., 2008), Empathy Communication Coding Systems (Bylund & Makoul, 2005), Empathy Assessment Index (Lietz et al., 2011), and Roter´ Interaction Analysis Systems (Roter & Larson, 2002)—to name a few. The IRI, however, is one of the most widely used measures of empathy (Chrysikou & Thompson, 2016) and translated into multiple languages. The IRI is continuously compelling for scholars aiming at measuring empathy, and one of the reasons we adopt IRI in our study, besides its appropriateness to comprehend empathy.

The IRI is designed as a self-report questionnaire to measure dispositional facets of empathy, captured as a set of constructs (four-factor structure), i.e., perspective taking, empathic concern for others and personal distress. The IRI encompasses 28 questions Likert scale, evaluating the following factors: 1) *Perspective Taking (PT)* is about the tendency of someone to adopt the psychological point of view of others instinctively. It encompasses one's ability to project themselves into specific scenarios and respond in a more empathetic manner. 2) *Fantasy* deals with the respondents' tendencies to render the visions of fictitious characters in books, movies, plays, and in our case, the characters in the VR into the feelings and actions. 3) *Empathic Concern* is about is about being "the other-oriented" feelings of sympathy and concern for unfortunate others. 4) *Personal Distress* – the "self-oriented" feelings of personal anxiety and unease in tense interpersonal settings.

Note that sometimes the factor structure of IRI in literature often differs from Davis´s original four-factor model. Researchers have combined the subscale into a two-factor model, consisting of cognitive (recognizing emotions) and affective (responding to emotion dimensions). The cognitive empathy component measures the ability of someone to properly comprehend the emotions of others, as well as have an accurate emotional response in a given situation (Kaplan & Iacoboni, 2006; Smith, 2006). The perspective taking, fantasy, empathic concern are elements that capture cognitive elements that have been a part of the traditional notion of empathy (Pulos et al., 2004), and used in IRI index. Moreover, the affective empathy component constitutes unconscious processes involving the sharing of emotions, such as personal distress, affective responsiveness, and emotional contagion, and can be characterized as one's "gut reaction" to emotional stimuli. However, researchers also are critical on the simplification of the IRI subscales into a two-factor model (e.g., Chrysikou & Thompson, 2016; Murphy & Lilienfeld, 2019).

Chrysikou and Thompson (2016) claim that the IRI as predominantly used in the literature does not accurately measure cognitive and affective empathy and illuminate the advantages of assessing empathy by applying the original four-factor structure, which was also followed in this research.

Beyond research on empathy that emphasize the IRI Index, previous studies have explored the role of VR in enhancing empathy through experimental setups involving different groups (e.g., Hargrove et al., 2020; Schutte & Stilinović, 2017; Shin, 2018; Stavroulia & Lanitis, 2019; Stavroulia & Lanitis, 2023). Shin (2018) conducted an experiment using VR immersive storytelling materials that integrate presence, flow, empathy, and embodiment. Their experiment affirms that VR storytelling is crucial for fostering empathy and ultimately improving engagement with VR.

Schutte and Stilinović (2017) conducted an experiment exploring the connection between VR and empathy by comparing groups that used VR and those that did not. The authors report a positive relationship between exposure to VR experiences and higher levels of empathy. Stavroulia and Lanitis (2019) and Stavroulia and Lanitis (2023) run group experiments aim at enhancing teachers' empathic skills, a necessity for effective classroom management, especially in multicultural settings. In the former article, the authors employ a self-developed measurement tool, though the specific empathic aspects being measured were not specified. In the latter article, Stavroulia and Lanitis (2023) utilize Personal Taking (PT) dimension of the IRI Index in their experiments, highlighting the substantial value of using VR to change perspectives and strengthen empathetic behavior during educators' training sessions.

While some previous studies confirm the positive impact of VR on empathy, certain gaps are noticeable. For instance, these papers either do not measure detailed aspects of empathy or do not consider higher education contexts, where empathic skill is crucial for students who will be working with individuals with intellectual disabilities. Additionally, to the best of the

authors' knowledge, there is a lack of immersive materials designed to teach social educator students on various skills, including empathy, ethical considerations, and challenges specific to dealing with people with intellectual disabilities regularly within higher education settings.

Hence, combining immersive materials that simulate interactions between people with intellectual disabilities and caregivers within a group experimental framework, while employing the complete IRI Index to assess the effectiveness of VR in the context of social education, can be considered a novel approach.

3. Methodology

3.1. Context

This article is derived from experiments with Norwegian Social Educator students. The social educator program in Norway is a full-time undergraduate university program that spans three years and combines evidence-based and practice-focused health and social care (180 ECTS). Throughout the program, students are required to complete three mandatory placements in practice. Upon completion of the three-year program, graduates are eligible to apply for registration as social educators with the Norwegian Directorate of Health.

The social educator program aims to equip students with the necessary skills to provide health and care services to various groups, with a particular focus on individuals with intellectual and developmental disabilities. Social educators can work in different settings, including municipal care services, care homes for older adults, special education for individuals with intellectual disabilities in schools and kindergartens, and child and adult welfare centers.

The curriculum of the social educator program is governed by the National Curriculum for Social Education and encompasses specific knowledge, competencies, and skills that all students are required to achieve. Key subjects covered include social science, psychology, applied behavioral analysis, environmental/ milieu therapy, communication, pedagogy, health promotion, and rehabilitation and habilitation.

3.2. Experiment Setting

The experiment was conducted in a dedicated VR-lab that has three VR stations, thus allowing three people to simultaneously experience the 360 VR material (Figure 1). As the number of participants were more than the available devices, the experiments were conducted four times, with four groups. The VR scenario depicted a caretaker interacting with a person who has an intellectual disability. The interaction was intentionally designed to evoke empathetic learning among students, particularly when the caregiver attempted to discourage the individual from pouring a glass of soda that they had purchased together. This restriction on soda consumption is in place, as drinking soda is only permitted on Saturdays. The people with intellectual disability thought the soda should be poured out into the sink. Note that the VR scenario was developed by Ridderne (www.ridderne.no).

The VR scenario was followed by group reflections focusing on ethical aspects of the scenario and the interaction between the caretaker and the person with intellectual disabilities. Finally, focus group interviews were conducted, eliciting the students' feelings after viewing the scenario, and learning experience. However, this study only reports the part that including the pre- and post-experiment surveys, i.e., to answer the questions if learning from VR in combination with reflection groups can trigger changes in attitude of the students compared to their empathy predisposition prior to experiment.

Figure 1 Experiment Setting

3.3. Participants

Participants in this study were the second semester bachelor students of the Social Educator program attending the course called: "Practice: Functional impairment, environmental work, coping and welfare technology" (10 ECTS). The number of participants can be seen in Table 1.

Table 1 Participants completed the surveys.

Survey Category	# Participants	Completed survey
Compassion before test	21	14
Compassion after test	12	12
IRI Index before test	26	16
IRI index after test	17	16

Note that participation in this study was voluntary, and students had the option to withdraw their participation in accordance with the research ethics guidelines. As a result, there were more students who initially filled out the survey than those who completed it. Consequently, for the analysis in this

study, only completed surveys were included. Before the experiments, the students received an introduction course on "*Practice: Functional impairment, environmental work, coping and welfare technology*", and they were then verbally informed about this VR experiment and later through an announcement in their online university learning platform.

3.4. Data Collection Method: Survey

We asked the students to rate themselves regarding their empathic tendencies, based on IRI survey. The IRI was measured before and after the VR experiment. The IRI encompasses 28 questions using a Likert scale by evaluating the following factors: 1) Perspective Taking - the tendency of someone to adopt the psychological point of view of others instinctively, 2) Fantasy – respondents' tendencies to render the visions of fictitious characters in books, movies, plays, and in our case, the characters in the VR into the feelings and actions, 3) Empathic Concern – the "other-oriented" feelings of sympathy and concern for unfortunate others, and 4) Personal Distress – the "self-oriented" feelings of personal anxiety and unease in tense interpersonal settings (Davis, 1980). All measurement points of each factor were coded with numbers. For example, Perspective-taking (PT) scale aspects were coded as PT_1-PT_7; fantasy scale (FS) aspects were coded as FS_1-FS_7; empathic concern (EC) scale aspects coded as EC_1-EC_7, while personal distress (PD) scale aspects coded as PD_1-PD_7. All measured aspects of each factor are explained in Table 2-5 in Section 4.

To complement the emphatic measurement through the IRI Index, we included a survey on compassion. Compassion is conceived as a feeling of concern for another person's suffering which is accompanied by the motivation to help (Singer & Klimecki, 2014). The concept is slightly similar to empathy, but it is more about "other-related emotions" than "self-related emotions" towards e.g., other characters. In addition, we captured demographic data such as students' gender, age, previous experience of interacting with persons with intellectual disability, and previous experience of using VR technology. We also conducted systematic discussions with different groups of students to explore further impressions and experience of using VR, and to and to detect barriers and enablers for learning.

3.5. Data Analysis Method

We applied simple descriptive statistics to analyze the survey results. We started by assigning the score of each aspect of each factor in the IRI index. All aspects have five options ranging from "*describe me very well*" (scored as five) to "*does not describe me very well*" (scored as one). Note, some statements are marked with negative (-). In this case the scoring will be done in reverse fashion where "*describe me very well*" (scored as one) to "*does not describe me very well*" (scored as five). As the mean score approaches 0, the empathy measure decreases, whereas as the mean score moves closer to five, the empathy measure increases.

4. Results

4.1 Perception on Empathic Skills

Concerning the emphatic skills, we applied the IRI Index as has been described in the methodology section. The results of the measurement before and after the experiments can be seen in Figure 1-4, in line with the IRI factors included in this study: Position Taking (PT), fantasy (FS), empathic concern (EC) and Personal Distress (PD). Definition of each aspect being measured is described in Table 2-5. These factors or subscale should be used separately as the instrument is not for measuring global empathy (Konrath, 2013).

4.2 Position Taking

The position taking captures empathy skill in the sense of how easily a person can imagine themselves in other peoples' situations. Subscales measured under the position taking can be seen in Table 2.

Table 2 Position Taking measured in IRI Index

ID	Description
PT_1	Difficulty seeing things from others perspective (-)
PT_2	Seeing disagreement before making decision
PT_3	To understand friends and their perspectives
PT_4	Won't waste time for something that one is unsure (-)
PT_5	Two sides to every question
PT_6	Put in other "shoes" when angry
PT_7	Thinking other's place before criticizing

Figure 2 shows the average scores of our experiment participants. The results are presented as a spiderweb diagram to observe more easily the changes of the score in the pre-test, i.e., prior to the experiments, signified with solid (blue) line. The orange dash-line indicates the post-test score results. This division valid for the results of all four-factors of an empathy. In all the pre-test and post-test, N= 17 and 16 respectively.

Figure 2 shows that mean values of the different subscales of PT factor fall between 3.1 and 3.81 (pre-test, solid line) and between 3.29 and 4.14 (posttest, dashed line). In several PT subscales, the mean scores increase, especially PT_1, PT_3, PT_4 and PT_7. In overall impression, the dashed line area expands beyond the

pre-test, except for PT_2 (unchanged) and PT_5 (slightly lower mean score).

In overall impressions, most of students had a relatively good initial position taking, especially based on the mean score of their self-assessment, where the average mean score is above three. There are two possible explanations for these tendencies: *First*, the participants are social educator students who have have learned to respect other people's opinions early enough, and thereby indirectly being exposed to the "personal taking concept." *Second*, the subscales of PT closely align with the societal culture of the students. Therefore, when exposed to immersive material in VR, it becomes easier for them to absorb and take on the position and develop empathy after observing how a caretaker treats a person with intellectual disabilities. For example, after seeing perceived disrespect treatment to a person with disabilities as seen in the VR, the students expressed the feeling concerning this situation as: "irritated, frustrated, anger, overwhelmed, skeptical, and disappointed" addressed to the caretaker.

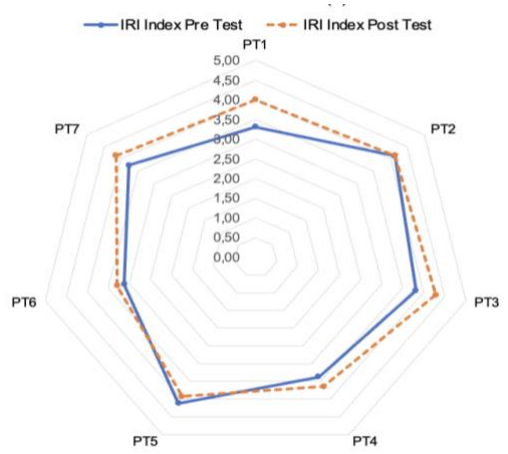

Figure 2 The Results of Position Taking

4.3 Empathic Concern Results

Recall that emphatic concern is about someone's ability to sympathize for the misfortunes that occur to others. It represents an emotional component of empathy. The definition of aspects/ subscales used for Empathic Concern can be seen in Table 3, while the results are presented as Figure 3.

Figure 3 illustrates the mean values of various subscales within the EC factor. The pre-test data, represented by a solid line, indicates that the mean values range from 3.56 to 4.75. On the other hand, the post-test results, depicted by a dashed line, show mean values between 3.64 and 4.29. Notably, there is a significant increase in the mean score of one particular EC subscale, particularly EC_2.

On contrary, mean score for EC_5 decreases from 4.75 to 4.07. Two subscales, namely EC_3 and EC_6 also show a slight decrease after experiment, while two subscales remain stable, i.e., EC_4 and EC_7. Moreover, EC_5 demonstrates a better mean value in the post-test than pre-test.

Table 3 Emphatic Concern (EC) measured in IRI

ID	Description
EC_1	Tender feelings for less fortunate people
EC_2	Do not feel sorry for other people when they are having problem (-).
EC_3	Feel protective toward someone being taken advantage of.
EC_4	Do not disturb with other people's misfortunes (-)
EC_5	Do not feel pity to see someone being treated unfairly (-)
EC_6	Touched by things that happen
EC_7	Describe him/ herself as a soft-hearted person.

However, we also observe a counterintuitive result, i.e., EC_5 that specifically measures whether one feels pity when witnessing someone being treated unfairly. This EC factor involved subscale with negative statements and thus the question could be tricky. The respondents might have mistakenly mixed up their answer either in the pre- or post-test, identifying themselves as "*not describing themselves well*", by choosing the lower scale when rating the question. In the focus group discussions after the students seemed to develop sympathy to the person with intellectual disability and expressed disappointment to the caretaker in the scenario. Therefore, in the context of EC5, the statement "Do not feel pity to see someone being treated unfairly" appears contradictory. The negative construction of the question can potentially mislead the rating process.

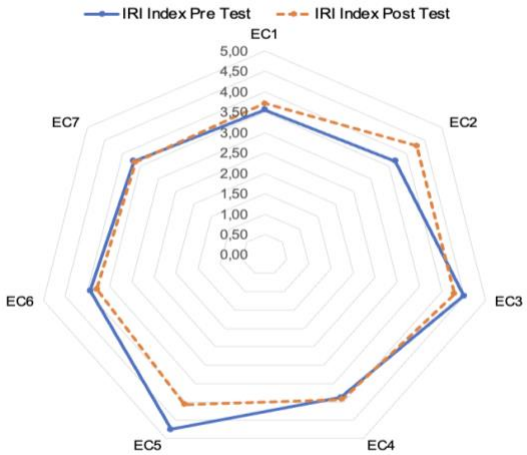

Figure 3 The Results of Empathic Concern

4.4 Fantasy Results

We also provide the definition of each aspect/subscale within the fantasy (FS) factor, as

elaborated in Table 4. This table includes the original definition of the measures. It is important to note that the fantasy scale items evaluate empathy towards fictional stories or characters in a specific situation. Therefore, although fantasy is typically associated with events or characters in books, plays, or movies, it is also applicable in a VR setting. The immersive materials presented to the participants in this study were based on a 360-degree video featuring a couple of characters and an event scenario.

Table 4 Fantasy (FS) measured in IRI

ID	Description
FS_1	Dream on regularity that might happen
FS_2	Get involved with the feelings of the characters in a novel.
FS_3	Objective when watch a movie or play, and not caught up in it (-)
FS_4	Rarely becoming extremely involved in a good book or movie (-)
FS_5	Felt one of the characters after seeing play/movie
FS_6	Can put him/herself in the place of a leading character in a movie
FS_7	Imagining events in the story were happening to him/ her when reading book

Figure 4 shows that mean values of the different subscales of FS factor fall between 2.62 and 3.94 (pre-test, solid line) and between 2.86 and 3.64 (post-test, dashed line). Generally, the mean scores decrease in most FS subscales, with the exception for FS_3 and FS_6 (unchanged) and FS_5 which shows a slightly higher mean score. The overall results indicate a lower mean score and a post-test area that slightly shifts towards the inner region of the pre-test area.

A lower mean value can be attributed to several reasons. Firstly, interpreting questions related to fantasy is relatively challenging. Secondly, although not reported in this article, a series of discussions, emotion elicitation, and reflection sessions were conducted, which might have had a greater impact on students' perception than the survey can capture in the post-test. Thirdly, the VR materials were considered too short, lacking sufficient actions, making it difficult to assess and evoke fantasy towards the characters.

4.5 Personal Distress Results

Personal Distress (PD) is the last factor measured in IRI, which also consists of seven subscales. It measures the negative feelings that may appear when encountering distress. It is basically an affective component of empathy. The PD measures can be seen in Table 5

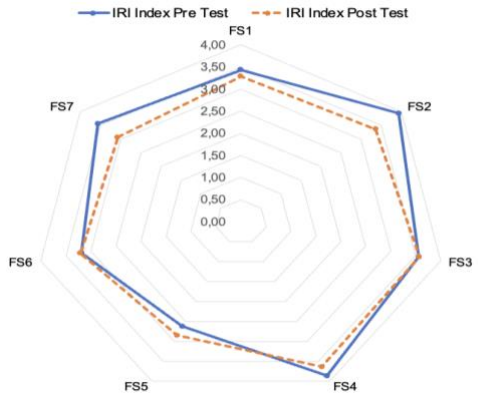

Figure 4 The Results of Fantasy

Figure 5 illustrates that mean values of the different subscales of PD factor fall between 1.93 and 2.87 (pre-test, solid line) and between 1.93 and 2.86 (posttest, dashed line). In most of PD subscales, both the score of empathic scale and overall region of posttest marked with dashed line were getting lower value, with exception of PD_7 (unchanged) and PD_1 (slightly increases).

Table 5 Personal Distress measured in IRI

ID	Description Personal Distress (PD)
PD_1	Feel apprehensive in emergency situations
PD_2	Feel helpless in the middle of a very emotional situation.
PD_3	Remain calm when seeing someone get hurt (-)
PD_4	Scare of being in a tense emotional situation (-)
PD_5	Effective in dealing with emergencies. (-)
PD_6	Tend to lose control during emergencies.
PD_7	Go into pieces when seeing someone who badly needs help in an emergency.

This is also another example of challenging questions to be applied to students who learn to be a social educator, to elicit negative feelings when encountering distress and assess their feeling. However, the tendency to get lower results in PD has been mentioned in the literature (e.g., Konrath, 2013) when applying IR index. Given counterintuitive results on this subscale, Bradford et al. (2019) argue that the personal distress component of empathy itself, as a measurable component, could be flawed. Other studies have concluded that personal distress is a measure reflecting emotional maturity and cognitive development as individuals transition from childhood to adulthood. In fact, it is argued to be more indicative of sympathy, wherein one identifies with another person but still focuses on self-protection. As individuals age or gain experience in a specific context, this component decreases owing to increased emotional control and diminished fears based on

personal experiences. Consequently, there is a shift towards developing sympathy for victims and engaging in pro-social behaviors (Bradford et al., 2019). Our participants are young students who may have little experience with personal distress and its consequences. Furthermore, it may be even more challenging to relate this personal distress component to a scenario seen in virtual reality (VR) that was relatively simple and lacked distressing situations.

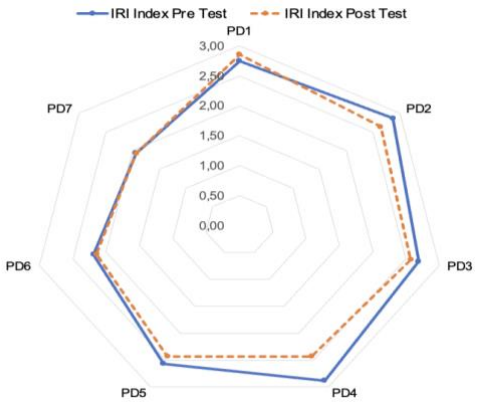

Figure 5 Personal Distress Results

Recall that in this paper the first RQ asks about "How can the use of VR in combination with ethical group reflection increase bachelor students' empathy skills in?" Through a set of experiments with the same scenario with different groups we have arrived at the results presenting in Figure 2-5. The overall qualitative results of the survey, we can observe that combined VR and reflection groups help students to improve partly their emphatic skills. However, there may be discrepancy between qualitative interpretation of the score and quantitative results (See Section 5).

The RQ in this study explores which emphatic factors tend to change when exposed to immersive VR materials. Qualitatively, we can assess from Figure 2-5 that the following subscales show higher tendencies to positive changes: PT_1, $PT_{3\text{-}4}$, PT_7, EC_2, accounting for a total of five subscales out of the 28 subscales.

5 Discussion: Do We Make Improvements?

After observing the four-factor empathy component and the scores of IRI qualitatively, this article presents the results using the box and whisker plots, that are efficient to summarize all key descriptive statistic in a single chart (Figure 6). It can capture the variations in the datasets of the mean score of subscales under each factor, pre- and post-test. The left figure represents the results before the VR experiments while the right represents the results after experiments. PT is represented by yellow, FS by orange, EC by green, and PD by brown. In each plot, the box represents the interquartile range (IQR), the line inside the box represents the median, and the whiskers extend to the minimum and maximum values within 1.5 times the IQR. For all four factors, the median values are similar to the VR pre- and post-test, i.e., *3.48*. In each emphatic factor, the IQR value appears to be slightly lower after the intervention, indicating decreased variability. No outliers are observed in each box and whisker plot.

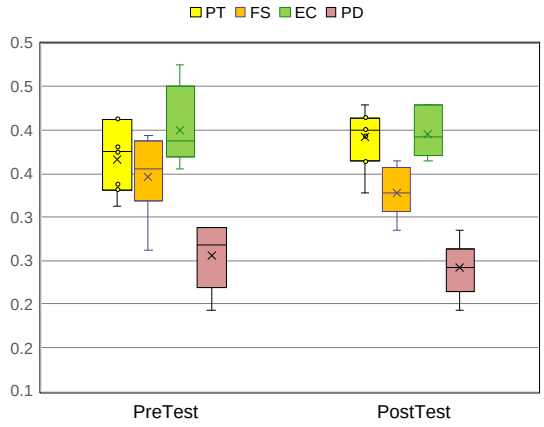

Figure 6 Four-factors empathy: pre - post test

Overall, the box and whisker plots suggest that there are no significant differences in the central tendencies (medians) and variability (range) of the factors PT, FS, EC, and PD before and after the study. However, there may be slight shifts in the median values for some factors, indicating potential changes.

We also conducted a paired t-test to compare the means of each variable, using the mean and standard deviation (SD), to determine if there are significant differences based on our findings. The null hypothesis (H_0) states that the mean difference is zero, while the alternative hypothesis (H_1) suggests that the mean difference is not zero. Our analysis used a significance level of 0.05 and degrees of freedom (df)=6. The critical t-value for a two-tailed test at this significance level is approximately 2.447 (Table 6). Upon comparing the calculated t-value with the critical t-value, our interpretation indicates that there is no significant difference in the means of *PT*, *FS*, *EC*, and *PD* before and after the study.

Table 6. Paired t-test for Means.

	Comparison (1)	Decision (2)	Meaning (3)
PT	2.132 < 2.447	Accept Ho	No significant differences
FS	0.665< 2.447		
EC	0.761 < 2.447		
PD	1.753 < 2.447		

However, despite the lack of a statistically significant difference, our discussions with the

students provided valuable insights into the potential use of for their learning and fostering empathy (e.g., how to support people with disability in a more ethical manner). These discussion results below may partly explain why there is no significant difference before and after experiments:

- **Preparation:** The use of VR for education necessitates more preparation beyond what we initially deemed adequate. For instance, several students reported that felt that they were not adequately informed beforehand. They also mentioned the audio quality was poor. Thus, we recommend other researchers to offer opportunities for students to try out the VR technology before watching the scenario. Hence, researchers can avoid technical disturbance that may impede learning for the students.
- **Expectations regarding content:** The focus of the pre-survey made some of the students believe that the scenario would concern a difficult empathic and/or challenging situation. As a result, the anticipated more advanced elements such as acting and challenging behavior (e.g., aggressivity and coercion) and a longer scenario.
- **Scenario setting:** The students mentioned that an immersive material showing a less daily dilemma and with more interactive possibilities would trigger more emotions. Some students disagreed with this, but for others, it was believed to trigger a stronger desire to help. However, this assumption raised by the students' needs further testing.

One point raised by the students was that the immersive material facilitated discussion and reflections among the students. They pointed out the importance of making these reflections in advance of going into practice or work. Moreover, VR was deemed a useful way to learn, as expressed as follow:

"*Very good to see it through VR glasses, so being so close. It sparked more emotions watching it like that rather than just watching it on film*." – Student 1

"*Genius way [to learn] in social education*" – Student 2

In short, in some of the students' comments, they acknowledged the learning potential using VR.

The limitations of this study come from several aspects: 1) limited numbers of respondents to be able to discuss the results of this study further with statistical analysis. 2) We only used one scenario, giving limited time which prevented longer experiments. 3) Variations of immersive materials were also limited.

6. Conclusions

This article has raised two issues. First, we explore possible use of combined VR and reflection groups to help students develop empathy. Second, we explore which empathic factors are inclined to change when exposed to immersive VR materials. We have conducted VR experiments with social educator students working with persons with intellectual disabilities. Our findings based on the survey shows no difference (statistically) in empathic learning before and after conducting VR testing. However, we have added explanations derived from the dialog and evaluation with the students. The discussion suggested that technical difficulties may prevent maximizing the benefits of VR. Thus, in future research, to circumvent the technical issues that may occur, we suggest offering students a possibility to repeat after a session in case some technical issues occur and vary the student participants. In sum, the students suggested that VR as a possible means of learning and a were positive to use VR in their education. Methodologically, a sample size with different segments might help to show the evidence in favor of VR for Social Educators, and thus can be a part of future research.

7. References

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, *34*, 163-175.

Bradford, L., Chun, K. P., Bonli, R., & Strickert, G. (2019). Does engagement build empathy for shared water resources? Results from the use of the interpersonal reactivity index during a mobile water allocation experimental decision laboratory. *Water*, *11*(6), 1259.

Brown, R., Sitbon, L., Fell, L., Koplick, S., Beaumont, C., & Brereton, M. (2016). Design insights into embedding virtual reality content into life skills training for people with intellectual disability. Proceedings of the 28th Australian Conference on Computer-Human Interaction,

Bylund, C. L., & Makoul, G. (2005). Examining empathy in medical encounters: an observational study using the empathic communication coding system. *Health communication*, *18*(2), 123-140.

Carulla, L. S., Reed, G. M., Vaez-Azizi, L. M., Cooper, S.-A., Leal, R. M., Bertelli, M., Adnams, C., Cooray, S., Deb, S., & Dirani, L. A. (2011). Intellectual developmental disorders: towards a new name, definition and framework for "mental retardation/intellectual disability" in ICD-11. *World Psychiatry*, *10*(3), 175.

Choi, J., Thompson, C. E., Choi, J., Waddill, C. B., & Choi, S. (2022). Effectiveness of immersive virtual reality in nursing education: systematic review. *Nurse Educator*, *47*(3), E57-E61.

Chrysikou, E. G., & Thompson, W. J. (2016). Assessing cognitive and affective empathy through the interpersonal

reactivity index: An argument against a two-factor model. *Assessment*, *23*(6), 769-777.
Contreras, G. S., González, A. H., Fernández, M. I. S., Martínez, C. B., Cepa, J., & Escobar, Z. (2022). The importance of the application of the metaverse in education. *Modern Applied Science*, *16*(3), 1-34.
Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion review*, *8*(2), 144-153.
Cureton, D. (2023). Virtual Reality Statistics to Know in 2023. *XR Today*.
Davis, M. H. (1980). Interpersonal reactivity index.
Dean, S., Halpern, J., McAllister, M., & Lazenby, M. (2020). Nursing education, virtual reality and empathy? *Nursing open*, *7*(6), 2056-2059.
Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives. The international scientific conference elearning and software for education,
Fromm, J., Radianti, J., Wehking, C., Stieglitz, S., Majchrzak, T. A., & vom Brocke, J. (2021). More than experience? - On the unique opportunities of virtual reality to afford a holistic experiential learning cycle. *The Internet and Higher Education*, *50*, 100804.
Hargrove, A., Sommer, J. M., & Jones, J. J. (2020). Virtual reality and embodied experience induce similar levels of empathy change: Experimental evidence. *Computers in Human Behavior Reports*, *2*, 100038.
Haskard, K. B., Williams, S. L., DiMatteo, M. R., Rosenthal, R., White, M. K., & Goldstein, M. G. (2008). Physician and patient communication training in primary care: effects on participation and satisfaction. *Health Psychology*, *27*(5), 513.
Jin, Q., Liu, Y., Yarosh, S., Han, B., & Qian, F. (2022). *How Will VR Enter University Classrooms? Multi-stakeholders Investigation of VR in Higher Education* Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA.
Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R. E., Avots, E., Helmi, A., Ozcinar, C., & Anbarjafari, G. (2019). Virtual reality and its applications in education: Survey. *Information*, *10*(10), 318.
Kaplan, J. T., & Iacoboni, M. (2006). Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social neuroscience*, *1*(3-4), 175-183.
Konrath, S. H. (2013). Critical synthesis package: interpersonal reactivity index (IRI).
Lietz, C. A., Gerdes, K. E., Sun, F., Geiger, J. M., Wagaman, M. A., & Segal, E. A. (2011). The Empathy Assessment Index (EAI): A confirmatory factor analysis of a multidimensional model of empathy. *Journal of the Society for Social Work and Research*, *2*(2), 104-124.
Majchrzak, T. A., Radianti, J., Fromm, J., & Gau, M. (2022). Towards Routinely Using Virtual Reality in Higher Education. HICSS,
Mehrabian, A. (1996). Manual for the Balanced Emocional Empathy Scale (BEES), Alta Mesa. *Monterrey, CA: Mehrabian*.
Murphy, B. A., & Lilienfeld, S. O. (2019). Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance. *Psychological Assessment*, *31*(8), 1062.
Nabors, L., Monnin, J., & Jimenez, S. (2020). A scoping review of studies on virtual reality for individuals with intellectual disabilities. *Advances in Neurodevelopmental Disorders*, *4*, 344-356.
Pulos, S., Elison, J., & Lennon, R. (2004). The hierarchical structure of the interpersonal reactivity index. *Social Behavior & Personality: an international journal*, *32*(4).
Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.
Research-and-Market. (2023). Global Virtual Reality (VR) in Education Market Report 2023: Increasing Internet Penetration Drives Growth. *GlobeNewswire*.
Roberson, C. J., & Baker, L. R. (2021). Designing and Implementing the Use of VR in Graduate Social Work Education for Clinical Practice. *Journal of Technology in Human Services*, *39*(3), 260-274.
Roter, D., & Larson, S. (2002). The Roter interaction analysis system (RIAS): utility and flexibility for analysis of medical interactions. *Patient education and counseling*, *46*(4), 243-251.
Sanfilippo, F., Blazauskas, T., Salvietti, G., Ramos, I., Vert, S., Radianti, J., Majchrzak, T. A., & Oliveira, D. (2022). A perspective review on integrating VR/AR with haptics into stem education for multi-sensory learning. *Robotics*, *11*(2), 41.
Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, *41*(6), 708-712.
Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, *78*, 64-73.
Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, *24*(18), R875-R878.
Smith, A. (2006). Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record*, *56*(1), 3-21.
Stavroulia, K.-E., & Lanitis, A. (2019). Enhancing Reflection and Empathy Skills via Using a Virtual Reality Based Learning Framework. *International journal of emerging technologies in learning*, *14*(7).
Stavroulia, K. E., & Lanitis, A. (2023). The role of perspective-taking on empowering the empathetic behavior of educators in VR-based training sessions: An experimental evaluation. *Computers & Education*, *197*, 104739.
Ventura, S., Badenes-Ribera, L., Herrero, R., Cebolla, A., Galiana, L., & Baños, R. (2020). Virtual reality as a medium to elicit empathy: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, *23*(10), 667-676.
Yalçın, Ö. N., & DiPaola, S. (2020). Modeling empathy: building a link between affective and cognitive processes. *Artificial Intelligence Review*, *53*(4), 2983-3006.