Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Effects of an embodied pedagogical agent on learning outcomes, presence, and cognitive load in immersive virtual reality

Yao Lu
Guangdong University of Science and Technology, Songshanhu High-tech Industrial Development Zone
Pohsun Wang
City University of Macau
Ni Kang
nikang@cityu.edu.mo
City University of Macau
Yuxuan Dong
Guangdong University of Science and Technology, Songshanhu High-tech Industrial Development Zone

Article

Keywords: Digital Transformation in education, Immersive Virtual Reality, Pedagogical agent, Presence, Learning outcomes, Electrodermal activity

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11676884/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Immersive virtual reality (IVR) is increasingly used to enrich online learning through spatial presence, embodied experience, and interactive engagement. However, the added value of embodied (PA) in fully immersive learning environments remains unclear. This study examined how PA affect learning performance, presence, cognitive load, and physiological arousal in an IVR-based college photography lesson. Ninety first-year undergraduates were randomly assigned to desktop video learning, IVR without a PA, or IVR with a PA (n = 30 each) for a 9-min lesson. Learning performance was assessed before and after instruction; presence, cognitive load, and physiological responses were measured in the two IVR conditions. Results showed a significant Group and Time interaction. The learning outcomes of the VR-PA group are significantly higher than those of the VR-NPA group and the DP group. The PA condition also reported higher spatial presence, involvement, and germane cognitive load, without increasing intrinsic or extraneous cognitive load. Exploratory analysis reveals that under the condition of VR integration into PA, there are correlations among involvement, germane cognitive load and post-test performance.

1. Introduction

Online learning has expanded access, especially since the COVID-19 pandemic, through platforms such as Coursera and large-scale MOOC systems in China (Buckner et al., 2025; Haleem et al., 2022; Sunardi et al., 2025; Qian et al., 2024). However, the convenience and scalability of desktop-based video online learning do not necessarily ensure effective learning outcomes. Limited real-time interaction, weak instructor presence, reduced peer awareness, high dropout rates, low completion, and learner distraction remain persistent challenges (Lv et al., 2025; Sarker et al., 2025; Xing et al., 2025). Although desktop video online learning broadens access, it often provides less social interaction and instructor presence than face-to-face teaching. Limited real-time interaction, weak peer awareness, and isolation may reduce engagement and performance, (Sarker et al., 2025). These problems help explain why high dropout rates, low course completion, and distraction continue to affect MOOCs and other asynchronous online courses (Lv et al., 2025; Xing et al., 2025). Instructor visibility further highlights the role of social support. Social presence, or the perceived presence of others in a learning environment, remains a relevant factor in engagement and learning (Poquet et al., 2018).

Immersive virtual reality (IVR) can provide spatial presence, embodiment, emotional engagement, and real-time interaction that are difficult to reproduce on conventional screens (Boffi et al., 2023; Lampropoulos & Kinshuk, 2024; W.-S. Wang et al., 2024). However, immersion does not guarantee better learning and may increase cognitive load when design is poorly aligned with learner needs (Ahn et al., 2022; Hartmann et al., 2023; Ratan et al., 2025). Pedagogical agent (PA), computer-generated characters that provide instruction, feedback, guidance, or social interaction, may add social and instructional cues, but their incremental value in fully immersive VR remains unclear (Hadjiliasi et al., 2024;Moreno et al., 2001;X. Wei et al., 2025).

Current studies have begun to confirm the learning effects of the combination of PA and IVR(G. Zhang et al., 2026), but there is still a lack of comprehensive evidence regarding how these effects manifest themselves at the experiential, cognitive, and time-resolved physiological levels. This study compared desktop video learning, IVR without a PA, and IVR with a PA. It examined learning outcomes across all three conditions and compared the two IVR conditions on presence, Cognitive Load, and physiological responses. Combining performance, subjective experience, cognitive responses, and physiological measures. Were able to assess how these effects of PA manifest themselves simultaneously at the experiential, cognitive, and temporal physiological.

2. Literature Review

2.1 Current Trends and Limitations of IVR-Based Learning

Immersive virtual reality (IVR) has rapidly developed and been increasingly adopted in educational settings (Lehrman, 2025; Sunardi et al., 2025). Since 2012, advances in computing resources and VR devices have enabled the development of unprecedented experiential learning environments (Lampropoulos & Kinshuk, 2024). In higher education, Wei & Yuan (2023) analysed 1,293 SSCI publications found that higher-education VR research was concentrated in health sciences, engineering design, and biology, with immersive interactive approaches dominant (Savytska et al., 2024).

From a constructivist perspective, learning involves actively constructing meaning through interaction with the environment and existing knowledge; therefore, VR-based learning environments may support knowledge construction through immersion and interactivity(Li et al., 2022). VR may support learning through presence, emotional engagement, embodiment, and interaction (Boffi et al., 2023). Through handheld controllers, users can interact with virtual environments in real time, fostering more engaging and experiential learning (W.-S. Wang et al., 2024). Stronger presence is often associated with engagement and learning outcomes (Jiawei & Mokmin, 2023). In addition, immersion and embodiment in digital learning environments can help learners retain meaningful information (Barrow et al., 2024), while IVR can foster presence and positive emotions among students (Lønne et al., 2023).

However, the educational benefits of IVR are not automatic. Although immersive VR can generate positive hedonic experiences such as enjoyment and immersion, these experiences do not always translate into improved learning outcomes and may increase cognitive load during prolonged use ((Ahn et al., 2022; Ratan et al., 2025). Ahn et al., (2022) argued that excessive spatial presence may consume cognitive resources and inhibit the encoding of semantic knowledge, particularly when the task itself imposes high cognitive load. Individual differences, such as visuospatial imagery

ability, can also moderate learning effectiveness in VR environments (Hartmann et al., 2023). Therefore, successful IVR learning is not only dependent on immersion itself, but also on the reasonable design of cognitive load and appropriate environmental guidance.

2.2 The Role of Pedagogical agent in IVR

PA（Pedagogical agent） are computer-generated characters that support learning through instruction, feedback, guidance, or social interaction. Research has focused mainly on desktop-based video multimedia, so their use in immersive VR remains underexplored, although emerging evidence suggests possible effects on learning outcomes (X. Wei et al., 2025).

Social Agency Theory (SAT) proposes that anthropomorphic agents activate social interaction rules similar to those used with people (Moreno et al., 2001). Social cues may enhance motivation, emotion, and cognitive involvement, while effectiveness depends on agent type, interaction mode (Herbert & Dołżycka, 2024).

PA may influence immersion, attention, and memory. Visual and auditory cues can guide attention in spatial tasks (Davis et al., 2025). Hasenbein et al., (2022) found that in VR teaching environments, PA were more effective than traditional screen prompts in guiding attention. Y. Wang et al., (2025) f further noted that PA capable of demonstrating reactive empathy can guide learners to focus on key learning content, thereby enhancing memory and transfer performance.

PA may also affect Cognitive Load. For example, G. Zhang et al., (2026) demonstrated that embodied PA in 360-degree panoramic video lectures can reduce cognitive load and enhance intrinsic motivation. In an augmented reality (AR) learning environment, embodied virtual agents significantly enhance students' motivation to learn(Cai et al., 2025). Their benefits are not guaranteed: appearance, feedback, and non-verbal behaviour must avoid distraction, overload, and reduced trust (Xiao et al., 2025; Y. Wang et al., 2023).

Taken together, PA may provide certain benefits for learning in VR environments. However, these benefits depend on appropriate instructional design. In particular, it remains unclear how PA influence learners' presence, cognitive load, physiological responses, and learning performance. Further research is therefore needed to clarify the additional effects of PA in IVR learning environments and to examine the experiential, cognitive, and physiological processes associated with these effects.

2.3 Research aims and Hypotheses

Although previous studies have demonstrated the educational potential of immersive virtual reality (VR) and pedagogical agent (PA), several gaps remain. First, VR and PA are often examined separately, leaving limited evidence on PA contributions within IVR. Second, relatively few studies have compared immersive VR with conventional desktop-based learning by integrating learning performance, subjective experience, cognitive responses, and physiological measures. Third, the processes linking PA to performance, especially the roles of presence, Cognitive Load, and arousal, remain unclear.

We therefore compared desktop-based video, IVR without a PA, and IVR with a PA. Learning outcomes were compared across all conditions; the two IVR conditions were also compared on presence, Cognitive Load, and physiological responses. This design addressed the incremental contribution of PA. Drawing on Social Agency Theory and prior research on immersive learning, the following hypotheses were proposed:

H1: Freshman year students in the group VR-PA will achieve higher learning outcomes than those in group VR-NPA and group DP .

H2: Students in the VR-PA condition will report higher levels of presence and exhibit greater physiological arousal than those in the VR-NPA condition.

H3: The introduction of PA into the VR learning environment will significantly affect learners' Cognitive Load, and Cognitive Load will be associated with learning performance.

This study aims to inform the design of socially enriched IVR learning environments.

3. Method

The study used a three-group between-participants design. Ninety first-year undergraduates were randomly assigned to DP (desktop-based video), VR-NPA, or VR-PA. We collected post-tests, presence questionnaires, Cognitive Load scales, EDA, and retrospective interviews, then integrating quantitative and qualitative analyses, we compared the learning outcomes and experience differences under different teaching conditions and we also explored associations among learning experience, cognitive load, and performance., as illustrated in Figure 1. This study was approved by the Ethics Review Committee of City University of Macau. All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki. Informed consent was obtained from all participants prior to data collection.

3.1 Experimental Design

3.1.1 Stimulus Materials

Materials came from a self-contained undergraduate Fundamentals of Photography module. A female teacher image served as the main instructional figure, informed by evidence that agent gender may influence learning experiences, with female teacher associated with more positive emotions and stronger social connections(Zhao et al., 2024; Zhao et al., 2025).

The same content was used in all conditions, the teacher's voice, the sequence of explanations, the direction of gestures, and the amount of information presented on the screen remained the same. Teachers shall maintain a highly consistent modeling image and a fully uniform vocal presentation. The two-dimensional course materials were processed and integrated into a Unity-based VR environment. The resulting headset experience preserved the online course content while adding a controlled three-dimensional presentation (see Figure 2). The DP group watched the original learning video on a computer. Under the two IVR conditions, content was presented in a 3D virtual environment: the VR-NPA group watched the original instructional video without embodied PA in the virtual environment, while the VR-PA group learned the original video with embodied PA in the virtual environment.

The VR environment was recreated in a studio-style laboratory with semi-transparent virtual companions. In VR-PA, the teacher model was based on the female instructor in the DP video and had a formal, professional appearance.

3.1.2 Participant Selection

First-year undergraduate students were selected because they are in a transitional stage of higher education and often face challenge of social support and learning engagement. This population was therefore considered appropriate for examining whether immersive learning environments can provide additional instructional and social support.

A total of 90 first-year students responded to recruitment advertisements for an immersive learning study. Participants came from Digital Media Arts, Product Design, and Visual Communication Design programmes. They were aged 19-21, had basic photography and videography knowledge but had not studied the target material, and had normal or corrected vision without colour deficiency or VR motion sickness.

3.1.3 Experimental Equipment

The DP group used a Windows computer equipped with a PC monitor, whilst the VR group used a Meta Quest 3S VR headset and Touch controllers; the headset had a resolution of 1832×1920 pixels and a field of view of 96°×90°; Additionally, the VR group normaliz the Ralovi S2 wrist-worn multimodal wrist-worn physiological sensor to collect EDA from participants, with a sampling rate of 4 Hz. E4 (4 Hz), in combination with multimodal devices, is commonly used to measure physiological arousal and In research on emotions and cognitive load (Dargahi Nobari & Bertram, 2024).

3.1.4 Experimental Procedure

The experiment took place in a VR laboratory with controlled lighting, temperature, and noise. Participants gave informed consent, and procedures followed institutional ethical guidelines. The experimental procedure is illustrated in Figure 3 and consisted of four stages.

In the first stage, participants completed the pre-test and experimental preparation. Upon arrival, participants completed a 5-min pre-test assessing their prior knowledge, followed by normalization instructions from the researcher. Participants assigned to the two IVR conditions were then fitted with a VR headset and a physiological wristband for electrodermal activity (EDA) recording. A 1-min adaptation period was provided to allow them to become familiar with the equipment and virtual environment.

In the second stage, participants completed the learning experiment. All participants received the same instructional content during a 9-min lesson. The DP group viewed the lesson on a computer screen, whereas both IVR groups learned in the same simulated classroom and could move within the virtual environment and inspect learning materials using handheld controllers. The VR-PA condition additionally included an embodied PA.

EDA was recorded only in the two IVR conditions (VR-PA and VR-NPA), while the DP group served as a psychological and behavioral baseline. This design avoided direct physiological comparisons across different display formats and allowed the physiological analysis to focus specifically on the additional effect of the PA under otherwise comparable immersive conditions.

In the third stage, participants completed the post-test and questionnaires. Following the learning session, participants underwent a 1-min recovery period and then completed a 5-min post-test covering the same knowledge domains as the pre-test, with the item order adjusted to reduce recall effects. Participants subsequently completed the Igroup Presence Questionnaire (IPQ) and Cognitive Load Scale (CLS).

In the final stage, retrospective interviews were conducted. Participants were randomly selected for brief retrospective interviews to share their learning experiences. The interview contents were recorded and used for thematic analysis.

3.2 Instruments

Validated measures were selected to assess learning outcomes and learning experience.

3.2.1 Learning Outcome Measure (Knowledge Test)

Two introductory-photography lecturers developed the knowledge test to align with the instructional content and objectives. The test evaluated participants' understanding of key concepts presented during the learning session. Both the pre-test and post-test covered the same knowledge domains, with item order adjusted in the post-test to normalize recall effects. The test comprised 13 items: 9 true/false (55%), 2 single-choice (15%), and 2 multiple-choice questions (30%), assessing factual knowledge, application in specific scenarios, and analytical understanding, respectively.

3.2.2 IPQ Scale

In this study, the IPQ was administered after the VR lesson to measure participants' subjective immersion and presence. The Igroup Presence Questionnaire (IPQ) scale (Schubert et al., 2001; Tran et al., 2024) was used to assess participants' perceived presence in the VR environment, defined as the subjective feeling of 'being there'. It comprises 14 items rated on a 7-point Likert scale and assesses four dimensions: Spatial Presence, Involvement, Realism, and Overall Presence. Spatial Presence reflects the perceived sense of being physically located in the virtual space; Involvement reflects attentional and psychological engagement; Realism concerns the perceived naturalness of the environment; and Overall Presence provides a global evaluation. This multidimensional structure permits separate assessment of distinct aspects of presence rather than a single undifferentiated immersion score.

3.2.3 CLS Scale

In this study, the CLS measured perceived cognitive processing demands following the VR learning session. The Cognitive Load Scale (Leppink et al., 2013), which is widely adopted in the fields of educational psychology and human computer interaction (Andersen & Makransky, 2021; Han et al., 2021; Makransky & Petersen, 2021; Choi & Lee, 2022). It was used to assess participants' subjective Cognitive Load after VR learning. It comprises 9 items rated on a 7-point Likert scale and distinguishes three components: Intrinsic Cognitive Load, reflecting task complexity; Extraneous Cognitive Load, reflecting demands introduced by the presentation; and Germane Cognitive Load, reflecting effort devoted to learning-related processing. This structure separates the demands of the task and its presentation from learners' reported investment in learning.

3.2.4 Electrodermal Activity

Electrodermal activity (EDA) was recorded continuously with a wrist-worn sensor during the VR lesson to assess peripheral physiological arousal. EDA includes skin conductance level (SCL), which reflects slower changes in skin conductance, and skin conductance responses (SCRs), which reflect brief changes associated with discrete events.

This study normalizes a wearable physiological sensing device (Ralovi S2 Bio-Wristwatch) was used to continuously monitor participants' physiological responses during the experiment. Raw EDA signals were denoised to remove high-frequency electromyographic noise and ultra-low-frequency drift. The analysis-period mean defined the SCL reference baseline, while individual normalization used a dynamic threshold (mean+0.5 SD); and outliers exceeding 3σ were removed using the z-score method. Rapid artifacts (>0.5 μS and <1 s) were replaced by the mean. A zero-phase fourth-order Butterworth low-PAs filter (0.5 Hz) was then applied to suppress high-frequency noise while retaining SCR-related activity. SCR events were identified by amplitude (>0.01 μS) and rise-time (<5 s) criteria, and their onset, peak amplitude, and peak latency were extracted. The resulting measures included SCL, SCR, HR, and HRV. And calculate RMSSD, which is one of the most used time-domain indicators in heart rate variability (HRV). It is calculated based on the RR intervals between consecutive heartbeats, mainly reflecting the degree of heart rate fluctuation over a short period, and is sensitive to parasympathetic nerve activity, especially vagus nerve activity.

3.3 Data Analysis method

Data analysis followed a three-step strategy, progressing from experimental effects to potential explanatory relationships and physiological responses. First, descriptive statistics were calculated, followed by a 3 × 2 mixed-design ANOVA to examine learning outcomes, with Group (DP, VR-NPA, VR-PA) as the between-participants factor and Time (pre-test, post-test) as the within-participants factor. The Group × Time interaction tested differences in learning gains. Planned contrasts compared the VR-PA with VR-NPA to assess the additional effect of the PA. For the two IVR groups, the internal consistency of the IPQ and Cognitive Load Scale was assessed using Cronbach's alpha, followed by between-group comparisons of presence and Cognitive Load.

Second, physiological data provided an objective complement to the subjective and performance measures. As physiological data were collected only in the two IVR conditions, analyses focused on the additional effect of the PA within the immersive environment. Skin conductance level (SCL),

heart-rate change (ΔHR), and root mean square of successive differences (RMSSD), a time-domain HRV indicator (Shaffer & Ginsberg, 2017), were compared between VR-PA and VR-NPA using independent-samples t-tests. Comparisons were conducted for both the full 9-min learning period and four event-aligned phases to examine whether PA-related physiological differences varied across instructional phases.

Finally, using partial correlation analysis and hierarchical multiple regression analysis, an exploratory analysis was conducted to describe exploratory associations among presence, cognitive load, and post-test performance.

4. Results

The sample comprised 90 participants: 34 males (37.78%) and 56 females (62.22%), aged 19-21 years (M = 19.58, SD = 0.669). All had normal or corrected vision, no colour deficiency or VR motion sickness, and provided informed consent.

4.1 Comparison of Learning Outcomes

To verify the baseline differences among the groups and the extent of improvement after the learning condition, a one-way ANOVA was conducted on the performance scores of the three groups. The comparison of pre-test scores (Table 1) showed that the p-values for comparisons among the three groups were all greater than 0.05, indicating no significant differences ($p > 0.05$). The post-test scores of the three groups differed significantly after receiving different instructional learning condition ($p < 0.05$). This suggests that the three groups had comparable levels of prior knowledge before the learning session, while post-test performance differed across instructional conditions. Gain scores (post-test minus pre-test) to compare learning improvements across groups. Teaching method had a significant effect on learning gains, F=14.555，$p < .001$.

Table 1: One-way ANOVA Results for Scores Among the Three Groups

	Group (mean ± SD)			***F(2,87)***	***p***
	DP(*n*=30)	**VR-NPA(*n*=30)**	**VR-PA(*n*=30)**		
Pre-test	30.00±14.56	28.83±7.39	30.33±10.33	0.149	0.861
Post-test	40.83±15.32	40.83±10.35	54.00±10.20	11.666	<0.001**
Score Gap	10.83±11.82	12.00±9.52	23.67±9.00	14.555	<0.001**

* p<0.05 ** p<0.01*** p<0.001

A 3 (Group: DP, VR-NPA, VR-PA) × 2 (Time: pre-test, post-test) mixed-design ANOVA showed a significant main effect of Time, $F(1, 87) = 208.30$, $p < .001$, $\eta p^2 = .705$, indicating an overall improvement from pre-test to post-test. The main effect of Group was also significant, $F(2, 87) = 4.51$, $p = .014$, $\eta p^2 = .094$. More importantly, the Group × Time interaction was significant, $F(2, 87) = 14.55$, $p < .001$, $\eta p^2 = .251$, indicating different learning gains. As shown in **Table 2**. The VR-PA group showed significantly greater improvement than the VR-NPA group, mean difference = 11.67, $t(87) = 4.43$, $p < .001$, 95% CI [6.44, 16.90], $\eta p^2 = .184$; the VR-PA group showed significantly greater improvement than the DP group, mean difference = 12.83, t(87) = 4.878, p < .001, 95% CI [7.60, 18.06], $\eta p^2 = .215$. Thus, the interaction reflected an additional gain associated with the PA, supporting H1.

Table 2　Mixed-Design ANOVA and Planned Contrasts for Learning Gains Across Teaching Conditions

Effect / Planned contrast	**Statistic**	**df**	*p*	**ηp²**	**Mean difference [95% CI]**
Mixed-design ANOVA					
Group	F = 4.51	2, 87	.014	.094	—
Time	F = 208.30	1, 87	< .001	.705	—
Group × Time	F = 14.55	2, 87	< .001	.251	—
Planned contrasts on change scores					
VR-PA vs. VR-NPA	t = 4.43	87	< .001	.184	11.67 [6.44, 16.90]
VR-PA vs. DP	*t* = 4.878	87	< .001	.215	12.83 [7.60, 18.06]

Note. Change scores were post-test minus pre-test. The Group x Time interaction tests whether change differed across DP, VR-NPA, and VR-PA. The IVR-versus-DP contrast compares the two IVR groups with DP; the PA contrast compares VR-PA with VR-NPA. Values are reported to two decimal places; p <= .05 was significant.

4.2 Between-Group Comparison of Presence and Cognitive Load

The IPQ Immersion Scale (14 items) showed good internal consistency, with Cronbach's α = 0.824 (n = 60). The detailed item-level statistics are provided in Supplementary Data. Independent-samples t tests showed higher Spatial Presence (p = .002) and Involvement (p < .001) in the VR-PA group than in the VR-NPA group, but no difference in Realism (*p* = .164), see Table 3. Overall Immersion was assessed by combining the independent item G1 with the three dimensions. When G1 was considered separately or combined with core items (Tran et al., 2024), the VR-PA group's G1 score was also significantly higher than that of the VR-NPA group.

Table 3: T tests analysis of the dimensions of Immersion

	Group (mean ± SD)		***t***	***p***
	VR-NPA(n=30)	**VR-PA(n=30)**		
sense of Spatial Presence	5.19±0.78	5.79±0.64	-3.275	0.002**
sense of Involvement	5.08±0.47	5.79±0.58	-5.259	0.000**
sense of Realness	4.78±0.87	5.09±0.84	-1.41	0.164
SP 1	5.50±0.82	5.97±0.85	-2.164	0.035*
SP 2	5.03±1.25	5.80±0.89	-2.747	0.008**
SP 3	5.27±1.23	5.67±0.76	-1.516	0.136
SP 4	5.13±1.22	5.67±1.09	-1.78	0.08
SP 5	5.03±0.96	5.87±0.78	-3.687	0.001**
INV 1	5.67±0.92	5.93±0.91	-1.129	0.264
INV 2	4.93±0.83	5.87±0.82	-4.39	0.000**
INV 3	4.23±1.14	5.50±1.11	-4.377	0.000**
INV 4	5.47±0.82	5.87±0.68	-2.056	0.044*
REAL 1_	4.70±1.47	4.50±1.85	0.464	0.645
REAL 2	4.27±1.70	5.10±1.03	-2.296	0.026*
REAL 3	5.37±1.13	5.67±0.99	-1.092	0.279
GP 1	4.13±1.33	5.27±0.94	-3.801	0.000**
G 1	4.80±0.89	5.50±0.90	-3.034	0.004**

* p<0.05 ** p<0.01*** p<0.001

SP= Spatial Presence INV= Involvement REAL= Realism

The CLS comprised Intrinsic, Extraneous, and Germane Cognitive Load. Reliability was high (Cronbach's α = 0.877), Intrinsic and Extraneous load did not differ significantly and were slightly lower in VR-PA. Germane load was higher in VR-PA(M = 5.7, SD = 0.71) than VR-NPA (M = 4.88, SD = 1.32), $p \leq 0.05$ (Table 4), Thus, the PA did not increase task or interface load but was associated with greater investment in learning-relevant processing.

T**able 4: T-test analysis of the dimensions of cognitive load**

	Group (mean ± SD)		t	p
	VR-NPA(n=30)	VR-PA(n=30)		
ICL	3.96±1.22	3.52±1.11	1.460	0.150
ECL	3.59±1.48	3.17±1.23	1.200	0.235
GCL	4.88±1.32	5.70±0.71	-2.996	0.004**
ICL-1	4.10±1.16	3.60±1.25	1.61	0.113
ICL-2	4.20±1.30	3.80±1.24	1.219	0.228
ICL-3	3.93±1.60	3.60±1.19	0.917	0.363
ECL-1	3.57±1.63	3.33±1.27	0.618	0.539
ECL-2	3.73±1.62	3.23±1.30	1.318	0.193
ECL-3	3.47±1.55	2.93±1.31	1.44	0.155
GCL-1	5.03±1.43	5.73±0.83	-2.325	0.024*
GCL-2	4.93±1.46	5.80±0.76	-2.882	0.006**
GCL-3	4.67±1.47	5.57±0.94	-2.829	0.007**

* p<0.05 ** p<0.01 ***p<0.001

ICL =Intrinsic Cognitive Load, ECL =Extraneous Cognitive Load, GCL =Germane Cognitive Load

4.3 Comparison of EDA Parameters:

A 1-min resting baseline was obtained before continuous time-domain recording of physiological signals. Independent-samples t tests compared responses between the VR-PA and VR-NPA conditions across the full recording period. Four participants were excluded from the physiological analyses because their recordings showed unstable sensor contact, missing signals, or other evident artifacts. leaving 30 participants in the PA group and 26 in the VR-NPA group with valid EDA, HR, and HRV data.

Across the full learning period, no significant between-group differences were observed in mean skin conductance level (SCL), skin conductance responses (SCR), heart rate (HR), or heart rate variability (HRV), indicating that the overall level of physiological arousal did not differ significantly between the PA and NPA conditions. However, a different temporal pattern emerged during the early phase of learning. The PA group showed relatively higher electrodermal activity at the beginning of the learning session, with a more pronounced difference between the two groups during the early stage. As learning progressed, the EDA trajectories of the two groups gradually stabilized and converged.

The 9-min session was aligned with instructional events and divided into four phases (P1–P4; see Figure 4). Phase-specific results are reported in **Table 5**.

P2 analysis showed the decrease in HR was smaller in the VR-PA group than in the VR-NPA group (−4.50 vs. −9.03 bpm, raw p = .049, d = .56, FDR-adjusted q = .160), however, after FDR correction, these inter-group differences no longer had statistical significance. There were no significant differences in SCL, ΔHR, and RMSSD at other stages between the two groups.

Overall, none of the phase-specific between-group comparisons remained statistically significant after Benjamini–Hochberg FDR correction. The phase-specific findings should therefore be interpreted as exploratory temporal patterns rather.

Table 5:Physiological responses across the four instructional phases in the VR-PA and VR-NPA groups

Phase	Time window (s)	SCL (μS)	*p*	FDR *q*	ΔHR (bpm)	*p*	FDR	Cohen's d	RMSSD (ms)	*p*	FDR
		VR-PA / VR-NPA			VR-PA / VR-NPA		*q*		VR-PA / VR-NPA		*q*
P1	0–25	0.016 / 0.021	0.709	0.768	-3.19 / -3.44	0.876	0.876	0.04	72.60 / 72.12	0.967	0.967
P2	25–180	0.035 / 0.030	0.768	0.768	-4.52 / -9.03	**0.049***	0.160	0.56	66.29 / 71.27	0.411	0.823
P3	181–480	0.026 / 0.041	0.281	0.561	-3.01 / -7.05	0.080	0.160	0.49	54.18 / 53.78	0.932	0.967
P4	481–540	0.015 / 0.000	0.227	0.561	-3.16 / -5.05	0.447	0.596	0.21	46.57 / 51.69	0.234	0.823

Note. PA = pedagogical agent; NO-PA = virtual reality condition without a pedagogical agent; SCL = skin conductance level; HR = heart rate; ΔHR = change in heart rate relative to the pre-learning baseline; RMSSD = root mean square of successive differences. Between-group comparisons were conducted using Welch's independent-samples t tests. Raw p values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure separately for SCL, ΔHR, and RMSSD across the four instructional phases. Bold raw p values indicate nominal significance at $p < .05$. None of the phase-specific comparisons remained significant after FDR correction ($q>.05$).

4.4 Exploratory Associations and Potential Pathways

Partial correlations used the two IVR groups while controlling for Group and Pre. The analysis examined links among presence, Cognitive Load, and performance; the correlation matrix is shown in Figure 5. INV correlated positively with GCL (partial $r = 0.255$, $p < 0.05$), GCL was significantly and positively correlated with post-test performance(partial $r = 0.291$, $p = 0.024$). SP, INV, and REAL showed positive but nonsignificant correlations with post-test performance.

To further examine the progressive predictive effects of different psychological and experiential factors on learning performance. Hierarchical regression used 60 participants from the two IVR conditions(see **Table 6**). Post-test performance (Post) was predicted in three steps: controls (Pre, Group), presence (SP, INV, REAL), then Cognitive Load (ICL, ECL, GCL). Model changes were evaluated with R^2, adjusted R^2, ΔR^2, and F-change. Table 6 showed, Adding SP, INV, and REAL increased explained variance by 2.5%, but not significantly ($\Delta R^2 = .025$, $\Delta F = 0.969$, $p = .414$). Adding ICL, ECL, and GCL increased explained variance to 60.2% ($R^2 = .602$, adjusted $R^2 = .540$), the 7.1% increase was significant ($\Delta R^2 = .071$, $\Delta F = 3.031$, $p = .038$). GCL was the only significant cognitive predictor ($\beta = .268$, $p = .015$).

To further explore the potential statistical associations learning experience and performance within the PA condition. Exploratory mediation used only VR-PA participants. Within VR-PA (n = 30), an exploratory model assessed the association of involvement with post-test performance through germane cognitive load, controlling for pre-test score.

The indirect effect was estimated using 5,000 bootstrap samples with a 95% confidence interval and was interpreted as exploratory rather than causal. The estimated indirect association was $\beta = 0.157$ (bootstrap 95% CI [0.020, 0.361]). With concurrent measures and n = 30, this result does not establish a causal mediation pathway, shall be regarded solely as an exploratory direction.

Table 6:Hierarchical multiple regression predicting post-test performance (n=60)

Predictor	Model 1		Model 2	Model 3		
	p	*β*	*p*	*β*	*p*	*β*
Pre-test	0.458	< .001	0.488	< .001	0.512	< .001
Group	0.507	< .001	0.408	< .01	0.314	0.01
SP	—	—	0.132	0.312	0.078	0.532
INV	—	—	0.083	0.579	−.011	0.937
REAL	—	—	−.011	0.93	0.123	0.327
ICL	—	—	—	—	−.327	0.15
ECL	—	—	—	—	0.117	0.589
GCL	—	—	—	—	0.268	0.015
R^2	0.506		0.531		0.602	
Adjusted R^2	0.489		0.488		0.54	
ΔR^2	—	—	0.025	0.414	0.071	0.038
F	29.193	< .001	12.239	< .001	9.649	< .001
ΔF	—	—	0.969	0.414	3.031	0.038

5. Discussion

This study investigated learning in VR environments. The results indicate that learners under the VR-PA condition demonstrated significantly higher learning outcomes compared to those under the VR-NPA and DP conditions, indicating an additional learning benefit associated with the inclusion of the PA within the immersive VR environment. The PA also significantly enhanced learners' presence, particularly Spatial Presence and Involvement, although no significant difference was observed in Realism. Physiological data indicate that certain phases may suggest an association between PA and cognitive patterns; however, no statistically significant phase-specific differences were observed across groups after FDR correction. Regarding Cognitive Load, the VR-PA group reported significantly higher Germane Cognitive Load, whereas Intrinsic Cognitive Load and Extraneous Cognitive Load did not differ significantly between the two VR conditions. Exploratory analyses further indicated that GCL was positively associated with learning performance and may represent a potential statistical associations Involvement with post-test performance within the VR-PA condition. Overall, H1 was supported, whereas H2 and H3 were partially supported.

5.1 Effects of PA on Learning Outcomes

Research findings demonstrate that the VR-PA learning environment is more effective in promoting learning outcomes than both VR-NPA and DP, further supporting the educational potential of PA in immersive virtual reality. This finding partially aligns with G. Zhang et al. (2026) who found that learning performance depended on the interaction between media environment and PA presence. Specifically, embodied PA were more effective in immersive VR than in desktop-based video environments, highlighting the importance of considering the interaction between the media environment and the instructional role of the PA. In the present study, however, the potential distraction-related learning costs of VR highlighted by Ratan et al. (2025) were not evident.

The present study extends this evidence to first-year university students. G. Zhang et al. (2026) showed that embodied PA can improve learning performance across different media environments, particularly by strengthening the advantage of VR over desktop-based learning. Our findings further showed that first-year students with limited prior knowledge and relatively little VR experience achieved significantly greater pre-to-post learning gains in the VR-PA condition than in the VR-NPA condition. This suggests that PA can provide additional learning benefits for novice learners in immersive VR environments.

From the perspective of Social Agency Theory (SAT), PA may function as sources of both social and instructional support. Learners may perceive instructional agents as socially meaningful partners, which can increase social engagement and learning commitment (Moreno et al., 2001). Such social responses may encourage learners to engage more actively with instructional content and thereby support the learning process. The present findings are consistent with Li et al., (2022), and previous studies have reported positive effects of PA on learning outcomes across different learning contexts (Makransky et al., n.d.; Xiao et al., 2025). In traditional multimedia learning environments, emotionally expressive PA can improve retention and transfer while also enhancing positive emotions and intrinsic motivation (Y. Wang et al., 2023). A meta-analysis of social cues delivered by AI-driven PA further showed that such cues can facilitate learning (S. Zhang et al., 2025). Together, these findings support for H1, that PA may support learning through social and instructional support.

5.2 Effects of PA on Students' Learning Experience

In addition to the significant learning gain observed in the VR-PA condition, the effects of the PA were clearly reflected in learners' experiential and cognitive responses. VR-PA participants reported greater Spatial Presence and Involvement, nominal early SCL differences that did not survive FDR correction, and different Cognitive Load, indicating effects across performance, experience, and processing.

Participants in the VR-PA condition reported significantly greater Spatial Presence and Involvement than those in the VR-NPA condition. This finding is consistent with Social Agency Theory, which proposes that learners respond socially to anthropomorphic instructional agents (Moreno et al., 2001). In the present study, the embodied presence, gestures, gaze, and instructional cues of the PA may have provided social and spatial reference points, strengthening learners' sense of "being there" and directing attention toward relevant instructional information. This interpretation is consistent with findings that PA-related visual cues can guide learners' selective attention and influence engagement during multimedia learning(King et al., 2023). It is also consistent with evidence that Spatial Presence reflects the perceived experience of being physically situated within a mediated environment (Ahn et al., 2022) and that avatar behaviours can influence visual attention in immersive classrooms (Hasenbein et al., 2022). The interview data provided further support for this interpretation; as some participant noted, "As soon as I entered the scene, I could see the PA; it captured my visual attention and prevented me from feeling aimless."

Constructivist theory views learning as active meaning construction through interaction with the environment and prior knowledge (Li et al., 2022). Learners actively explore the virtual environment and interact with PA, integrating new and prior knowledge to construct individualised cognitive schemas (Jaganov et al., 2024). The inclusion of PA may have created a more socially guided learning environment that supported learners' attentional orientation, exploration, and cognitive engagement. Retrospective interviews further corroborated this conclusion regarding students' active learning and exploration, with many students reporting a sense of curiosity within the scenarios and engaging in proactive exploratory learning. "I explored many corners", "I looked closely at the details of the PA ", "It felt as though I could reach out and touch the camera myself". These responses were associated with greater presence and deeper interaction, complementing quantitative differences in Spatial Presence and Involvement.

Despite these improvements in Spatial Presence and Involvement, no significant difference in Realism was found between the VR-PA and VR-NPA conditions. The inference is that this is due to the fact that the realism of the scene and the authenticity of the PA characters did not reach a sufficiently high level of realism. This may reflect individual differences in preferences for avatar appearance and realism. To minimise the potential uncanny valley effect associated with highly realistic avatars (Shen, 2025; Zibrek et al., 2018), the PA was designed with a moderate level of anthropomorphism. However, its interactions were primarily scripted rather than personalised or adaptive, which may have limited its contribution to Realism. This interpretation was also reflected in the interviews, with some participants commenting that aspects of the character modelling lacked realism. Thus, the effectiveness of PA may depend not only on their visual presence but also on the responsiveness and quality of their interactions (Herbert & Dołżycka, 2024). Future studies should further examine how avatar realism and adaptive interaction jointly shape learners' experiences.

The Cognitive Load findings further suggest that the effectiveness of PA may depend less on uniformly reducing cognitive demand and more on influencing how cognitive resources are allocated. Although immersive VR environments can enhance engagement and experiential quality, previous research has shown that increased immersion does not necessarily translate into improved learning outcomes if Cognitive Load is not effectively managed (H. Zhang et al., 2024). Germane Cognitive Load refers to the cognitive resources that learners actively invest in order to construct schemas and deepen their understanding (Klepsch & Seufert, 2020). In the present study, no significant between-group differences were observed in Intrinsic Cognitive Load or Extraneous Cognitive Load, whereas Germane Cognitive Load was significantly higher in the VR-PA condition. The absence of an increase in Extraneous Cognitive Load is also consistent with G. Zhang et al., (2026), who reported that adding an embodied PA in VR did not significantly increase Cognitive Load. The present results further suggest that the effect of the PA on Cognitive Load was dimension-specific rather than reflecting an overall reduction or increase in Cognitive Load. In immersive VR environments, learners often need to allocate cognitive resources to navigate the virtual space and identify relevant information. PA may reduce unnecessary orientation demands through gestures, gaze behaviours, and instructional guidance, enabling learners to devote more resources to meaningful processing of learning materials. This interpretation is consistent with Wang et al., (2025) and Ji & Zheng, (2025). Overall, the findings suggest that PA-supported learning was associated with greater germane processing without significant changes in intrinsic or extraneous cognitive load.

Physiological measures offered limited evidence of a PA effect across the learning period. Baseline-corrected SCL did not differ significantly between the VR-PA and VR-NPA groups in any phase. In P2, the VR-PA group also showed a smaller reduction in HR, whereas no significant between-group difference was observed in RMSSD. Cardiac autonomic activity has been shown to vary with cognitive task demands and sustained attention(Forte et al., 2019; Luque-Casado et al., 2016). Given the minimal decrease in HR, and no significant reduction in RMSSD, the physiological pattern observed in P2 is more appropriately interpreted as task-related physiological activation or sustained attention rather than a definitive stress response. (Charles & Nixon, 2019). In the P2 stage, the explanations and key reminders provided by PA may serve as social support and visual stimulation, consistent with the trend of a smaller decrease in ΔHR in the P2 stage, this explanation may be consistent with the research results, that is, PA can enhance the Germane Cognitive through significant visual and social cues (Davis et al., 2025).These findings provide partial rather than complete support for H2.

5.3 Possible Associations Underlying PA-Supported Immersive Learning

The partial correlation and hierarchical regression results suggest several possible associations among learning experience, cognitive processing, and performance. After controlling for Group and Pre, none of the three presence dimensions was significantly associated with Post. However, Involvement was positively associated with Germane Cognitive Load, and Germane Cognitive Load was positively associated with Post. This pattern

indicates a statistical association between presence-related experience and learning-relevant cognitive processing, as well as a statistical association between learning-relevant cognitive processing and learning performance.

Taken together, the findings indicate that the contribution of PA may extend beyond learning gains to learners' experiential, physiological, and cognitive responses. Higher Spatial Presence and Involvement, and higher Germane Cognitive Load provide converging evidence of several potentially related processes in PA-supported IVR learning. In particular, the exploratory analyses showed that Involvement was associated with Germane Cognitive Load, which was in turn associated with post-test performance. One possible interpretation is that the social and attentional cues provided by PA may support learner involvement and task-relevant cognitive processing. Early physiological arousal may also reflect initial attentional orientation to the PA. These relationships remain exploratory and should not be interpreted as a confirmed causal mechanism.

5.4 Limitations

First, the PA used in this study was constrained by the technical capabilities of the experimental platform. Its interactions were primarily pre-scripted rather than adaptive to learners' behaviours or cognitive states. Future studies could employ more advanced adaptive PA to examine whether personalised interaction produces stronger effects on involvement and learning.

Second, the physiological measures were limited to non-invasive peripheral indicators such as electrodermal and cardiac responses. Although these measures captured temporal changes in physiological arousal, they cannot directly identify the specific cognitive processes underlying these changes. Future research could incorporate multimodal physiological or neurocognitive measures to provide more direct evidence.

Thirdly, due to participants' time constraints, this study included only an immediate post-test, limiting the assessment of knowledge retention and long-term learning outcomes. Future research could employ repeated interventions and follow-up assessments to examine the long-term learning benefits of PA with IVR.

Third, the resource-intensive nature of the VR experiment limited the sample size and the extent to which the proposed mechanism could be tested causally. Therefore, the pathway linking PA, Involvement, Germane Cognitive Load, and learning performance should be regarded as preliminary rather than causal. Future studies should further test the causal relationships underlying this mechanism.

6. Conclusion

This controlled experiment compared desktop-based video learning, IVR without PA, and IVR with PA. VR with PA improved learning outcomes relative to desktop-based learning and VR without PA. The VR-PA condition also showed higher Spatial Presence, Involvement, and Germane Cognitive Load, together with exploratory early-stage differences in electrodermal activity were observed, but did not remain significant after FDR correction.

These findings indicate that PA can enhance immersive learning in immersive environments and are associated with greater learner involvement and germane cognitive processing. The exploratory association between INV and GCL suggests one possible pathway that warrants direct causal testing.

Future research should further test this mechanism using larger and more diverse samples, longer interventions, adaptive PA designs, and richer behavioural and physiological measures.

Abbreviations

Abbreviation	Definition
IVR	Immersive Virtual Reality
VR	Virtual Reality
PA	Pedagogical Agent
SAT	Social Agency Theory
DP	Desktop-based Video Learning
VR-NPA	Virtual Reality without a Pedagogical Agent
VR-PA	Virtual Reality with a Pedagogical Agent
IPQ	Igroup Presence Questionnaire
SP	Spatial Presence
INV	Involvement
REAL	Experienced Realism
CLS	Cognitive Load Scale
ICL	Intrinsic Cognitive Load
ECL	Extraneous Cognitive Load
GCL	Germane Cognitive Load
EDA	Electrodermal Activity
SCL	Skin Conductance Level
SCR	Skin Conductance Response
HR	Heart Rate
ΔHR	Change in Heart Rate from Baseline
HRV	Heart Rate Variability
RMSSD	Root Mean Square of Successive Differences

Declarations

Acknowledgements

The authors gratefully acknowledge the following research units for their research supported facilities: the Human Factors Engineering and Intelligent Interaction Research Center(Coloane), the Human Factors Engineering and Intelligent Interaction Research Center(Hengqin), and the Intelligent Perception and Human-Computer Interaction Laboratory.

This work was supported by the Research Project of Guangdong University of Science and Technology (Grant No. GKZLGC2024443).

Author Contributions

Conceptualization: Yao Lu

Methodology: Ni Kang, Pohsun Wang

Writing (original draft preparation): Yao Lu, Yuxuan Dong

Writing (review and editing): Pohsun Wang, Ni Kang

Experimental material: Yuxuan Dong

Authors and Affiliations

Yao Lu[1,2], Pohsun Wang[2], Ni Kang*[2], Yuxuan Dong[1]

Yao Lu,

U25092110269@cityu.edu.mo

Pohsun Wang

phwang@cityu.edu.mo

Ni Kang*Corresponding Author

nikang@cityu.edu.mo

Yuxuan Dong

dyx8172@gmail.com

[1] Guangdong University of Science and Technology, College of Creative Design, Songshanhu High-tech Industrial Development Zone, Dongguan, Guangdong 523668, China.

[2] City University of Macau, Fac Innovat & Design, Macau 999078, Peoples R China.

Ethics Declaration: Ethical guidelines were followed in line with Ethics Committee of City University of Macau, and approval was obtained from the University ethics committee. This study secured participants' consent prior to data collection.

Informed consent All the participating students signed the consent forms before participating in this research.

Consent to publish The authors agree to publish.

Competing interests Not applicable.

Data Availability Relevant data has been included in the supplementary information.

References

1. Ahn, S. J., Nowak, K. L. & Bailenson, J. N. (eds) (Grace),& Unintended consequences of spatial presence on learning in virtual reality. *Computers & Education*, *186*, 104532. (2022). https://doi.org/10.1016/j.compedu.2022.104532
2. Andersen, M. S. & Makransky, G. The Validation and Further Development of the Multidimensional Cognitive Load Scale for Physical and Online Lectures (MCLS-POL). *Front. Psychol.* **12**, 642084. https://doi.org/10.3389/fpsyg.2021.642084 (2021).
3. Barrow, J., Hurst, W., Edman, J., Ariesen, N. & Krampe, C. Virtual reality for biochemistry education: The cellular factory. *Educ. Inform. Technol.* **29** (2), 1647–1672. https://doi.org/10.1007/s10639-023-11826-1 (2024).
4. Boffi, P., Clerici, M., Gallace, A. & Lanzi, P. L. An educational experience in ancient Rome to evaluate the impact of virtual reality on human learning processes. *Computers Education: X Real.* **2**, 100014. https://doi.org/10.1016/j.cexr.2023.100014 (2023).
5. Buckner, E. et al. Charting the path after 2030: What should higher education's role be in the future of the sustainable development agenda? *Compare-a J. Comp. Int. Educ.* https://doi.org/10.1080/03057925.2025.2571462 (2025).
6. Cai, S., Huang, H., Jiao, X. & Zhang, Y. Generative AI + AR: A novel embodied conversational agent in an interactive learning environment. *Interact. Learn. Environ.* 1–15. https://doi.org/10.1080/10494820.2025.2589942 (2025).
7. Charles, R. L. & Nixon, J. Measuring mental workload using physiological measures: A systematic review. *Appl. Ergon.* **74**, 221–232. https://doi.org/10.1016/j.apergo.2018.08.028 (2019).
8. Choi, Y. & Lee, H. Psychometric Properties for Multidimensional Cognitive Load Scale in an E-Learning Environment. *Int. J. Environ. Res. Public Health*. **19** (10), 5822. https://doi.org/10.3390/ijerph19105822 (2022).
9. Dargahi Nobari, K. & Bertram, T. A multimodal driver monitoring benchmark dataset for driver modeling in assisted driving automation. *Sci. Data*. **11** (1), 327. https://doi.org/10.1038/s41597-024-03137-y (2024).
10. Davis, R. O., Vincent, J., Yang, E. B., Lee, Y. J. & Lee, J. H. The Exploration of Combining Hologram-like Images and Pedagogical Agent Gesturing. *Information* **16** (5), 350. https://doi.org/10.3390/info16050350 (2025).
11. Forte, G., Favieri, F. & Casagrande, M. Heart Rate Variability and Cognitive Function: A Systematic Review. *Front. NeuroSci.* **13**, 710. https://doi.org/10.3389/fnins.2019.00710 (2019).
12. Hadjiliasi, A., Nisiotis, L. & Polycarpou, I. A Comparative Assessment of Technology Acceptance and Learning Outcomes in Computer-based versus VR-based Pedagogical Agents. *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 513–516. (2024). https://doi.org/10.1109/ISMAR-Adjunct64951.2024.00148
13. Haleem, A., Javaid, M., Qadri, M. A. & Suman, R. Understanding the role of digital technologies in education: A review. *Sustainable Oper. Computers*. **3**, 275–285. https://doi.org/10.1016/j.susoc.2022.05.004 (2022).
14. Han, J., Zheng, Q. & Ding, Y. Lost in Virtual Reality? Cognitive Load in High Immersive VR Environments. *Journal of Advances in Information Technology*, *12*. (2021). https://doi.org/10.12720/jait.12.4.302-310

15. Hartmann, C., Orli-Idrissi, Y., Pflieger, L. C. J. & Bannert, M. Imagine & immerse yourself: Does visuospatial imagery moderate learning in virtual reality? *Comput. Educ.* **207**, 104909. https://doi.org/10.1016/j.compedu.2023.104909 (2023).
16. Hasenbein, L. et al. Learning with simulated virtual classmates: Effects of social-related configurations on students' visual attention and learning experiences in an immersive virtual reality classroom. *Comput. Hum. Behav.* **133**, 107282. https://doi.org/10.1016/j.chb.2022.107282 (2022).
17. Herbert, C. & Dołżycka, J. D. Teaching online with an artificial pedagogical agent as a teacher and visual avatars for self-other representation of the learners. Effects on the learning performance and the perception and satisfaction of the learners with online learning: Previous and new findings. *Front. Educ.* **9**, 1416033. https://doi.org/10.3389/feduc.2024.1416033 (2024).
18. Jaganov, T., Nnadi, L. C. & Watanobe, Y. The Learning Labyrinth: Integrating Learning Theories in VR. *2024 5th Asia Service Sciences and Software Engineering Conference, Asse 2024*, 158–165. (2024). https://doi.org/10.1145/3702138.3702156
19. Ji, X. & Zheng, X. Can pedagogical agents and visual cues enhance students' deep learning in a virtual reality environment? Insights from multiple data analysis. *Educ. Inform. Technol.* **30** (15), 22343–22365. https://doi.org/10.1007/s10639-025-13639-w (2025).
20. Jiawei, W. & Mokmin, N. A. M. Virtual reality technology in art education with visual communication design in higher education: A systematic literature review. *Educ. Inform. Technol.* **28** (11), 15125–15143. https://doi.org/10.1007/s10639-023-11845-y (2023).
21. King, J., Marcus, T. & Markant, J. Individual differences in selective attention and engagement shape students' learning from visual cues and instructor presence during online lessons. *Sci. Rep.* **13** (1), 5075. https://doi.org/10.1038/s41598-023-32069-7 (2023).
22. Klepsch, M. & Seufert, T. Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instr. Sci.* **48** (1), 45–77. https://doi.org/10.1007/s11251-020-09502-9 (2020).
23. Lampropoulos, G. & Kinshuk Virtual reality and gamification in education: A systematic review. *Etr&d-Educational Technol. Res. Dev.* **72** (3), 1691–1785. https://doi.org/10.1007/s11423-024-10351-3 (2024).
24. Lehrman, A. L. Embodied Learning Through Immersive Virtual Reality: Theoretical Perspectives for Art and Design Education. *Behav. Sci.* **15** (7), 917. https://doi.org/10.3390/bs15070917 (2025).
25. Leppink, J., Paas, F., Van Der Vleuten, C. P. M., Van Gog, T. & Van Merriënboer, J. J. G. Development of an instrument for measuring different types of cognitive load. *Behav. Res. Methods*. **45** (4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1 (2013).
26. Li, W., Wang, F., Mayer, R. E. & Liu, T. Animated pedagogical agents enhance learning outcomes and brain activity during learning. *J. Comput. Assist. Learn.* **38** (3), 621–637. https://doi.org/10.1111/jcal.12634 (2022).
27. Lønne, T. F., Karlsen, H. R., Langvik, E. & Saksvik-Lehouillier, I. The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: A randomized controlled study. *Heliyon* **9** (6), e17196. https://doi.org/10.1016/j.heliyon.2023.e17196 (2023).
28. Luque-Casado, A., Perales, J. C., Cárdenas, D. & Sanabria, D. Heart rate variability and cognitive processing: The autonomic response to task demands. *Biol. Psychol.* **113**, 83–90. https://doi.org/10.1016/j.biopsycho.2015.11.013 (2016).
29. Lv, P., Wang, X., Nie, Y., Wu, R. & Wang, L. Exploring the factors influencing college students' learning satisfaction in generative AI-supported MOOCs learning environment: A learning experience framework perspective. *Front. Psychol.* **16**, 1633686. https://doi.org/10.3389/fpsyg.2025.1633686 (2025).
30. Makransky, G. & Petersen, G. B. The Cognitive Affective Model of Immersive Learning (CAMIL): A Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychol. Rev.* **33** (3), 937–958. https://doi.org/10.1007/s10648-020-09586-2 (2021).
31. Makransky, G., Wismer, P. & Mayer, R. E. (n.d.). *A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. | EBSCOhost*. Retrieved 7 June from (2026). https://openurl.ebsco.com/EPDB%3Agcd%3A16%3A36908368/detailv2?sid=ebsco%3Aplink%3Agetftr&id=ebsco%3Adoi%3A10.1111%2Fjcal.12335&getft_integrator=clarivate&crl=f&link_origin=www.webofscience.com
32. Moreno, R., Mayer, R. E., Spires, H. A. & Lester, J. C. The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents? *Cognition Instruction*. **19** (2), 177–213. https://doi.org/10.1207/S1532690XCI1902_02 (2001).
33. Poquet, O. et al. (n.d.). *Social Presence in Massive Open Online Courses*.
34. Qian, J., Du, M., Yang, Y. & Qiu, Y. (n.d.). *Promoting educational equity and facilitating effective use of high-quality resources: An empirical study on improving massive open online courses completions*.
35. Ratan, R. et al. Time matters in VR: Students benefit from longer VR class duration, but certain outcomes decline after 45 minutes, with large individual variance. *Comput. Educ.* **235**, 105328. https://doi.org/10.1016/j.compedu.2025.105328 (2025).
36. Sarker, P. S. et al. Perceived key factors affecting online university classrooms. *Educ. Inform. Technol.* **30** (17), 24855–24886. https://doi.org/10.1007/s10639-025-13744-w (2025).
37. Savytska, I., Tryfonova, O., Feltsan, I., Alekseiko, V. & Ahiy, Y. Immersive technologies in the educational space. *Revista Amazonia Investiga*. **14** (75), 343–358. https://doi.org/10.34069/AI/2024.75.03.29 (2024).
38. Schubert, T., Friedmann, F. & Regenbrecht, H. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators Virtual Environ.* **10** (3), 266–281. https://doi.org/10.1162/105474601300343603 (2001).

39. Shaffer, F. & Ginsberg, J. P. An Overview of Heart Rate Variability Metrics and Norms. *Front. Public. Health*. **5**, 258. https://doi.org/10.3389/fpubh.2017.00258 (2017).
40. Shen, Z. How to develop student-preferred virtual teachers in AI-driven education? Exploring determinants and optimization strategies. *Interact. Learn. Environ.* 1–20. https://doi.org/10.1080/10494820.2025.2596061 (2025).
41. Sunardi, M., Hendric, S. W. H. L. & Rosmansyah, Y. Exploring Immersive Virtual Reality in Higher Education: Research Gap and Future Direction—A Scoping Review. *IEEE Access.* **13**, 76308–76321. https://doi.org/10.1109/ACCESS.2025.3565385 (2025).
42. Tran, T. Q., Langlotz, T., Young, J., Schubert, T. W. & Regenbrecht, H. Classifying Presence Scores: Insights and Analysis from Two Decades of the Igroup Presence Questionnaire (IPQ). *ACM Trans. Computer-Human Interact.* 3689046. https://doi.org/10.1145/3689046 (2024).
43. Wang, W. S. et al. Signaling feedback mechanisms to promoting self-regulated learning and motivation in virtual reality transferred to real-world hands-on tasks. *Interact. Learn. Environ.* **32** (10), 7661–7676. https://doi.org/10.1080/10494820.2024.2331151 (2024).
44. Wang, Y., Gong, S., Cao, Y. & Fan, W. The power of affective pedagogical agent and self-explanation in computer-based learning. *Comput. Educ.* **195**, 104723. https://doi.org/10.1016/j.compedu.2022.104723 (2023).
45. Wang, Y., Gong, S., Cao, Y. & Liu, Y. Parallel empathy or reactive empathy? The role of emotional support provided by affective pedagogical agent in online learning. *Internet High. Educ.* **67**, 101035. https://doi.org/10.1016/j.iheduc.2025.101035 (2025).
46. Wei, X. et al. Effects of immersive virtual reality on primary students' science performance in classroom settings: A generative AI pedagogical agents-enhanced 5E approach. *Interact. Learn. Environ.* https://doi.org/10.1080/10494820.2025.2514101 (2025).
47. Wei, Z. & Yuan, M. Research on the Current Situation and Future Development Trend of Immersive Virtual Reality in the Field of Education. *Sustainability* **15** (9), 7531. https://doi.org/10.3390/su15097531 (2023).
48. Xiao, M., Li, W., Han, L. & Zheng, S. Optimising Multimedia Learning: Effects of Pedagogical Agents' Appearance and Voice. *J. Comput. Assist. Learn.* **41** (4), e70101. https://doi.org/10.1111/jcal.70101 (2025).
49. Xing, W., Liu, Z., Kim, T. & Song, Y. Why do students leave instructional videos? Understanding students' in-video dropout behavior in a large online math learning platform. *Distance Educ.* https://doi.org/10.1080/01587919.2025.2579792 (2025).
50. Zhang, G., Ren, Z., Qi, B. & Huang, R. Learning With a Pedagogical Agent in Virtual Reality: Does the Impact of Environment Type and Agent Presence Matter? *J. Comput. Assist. Learn.* **42** (3), e70257. https://doi.org/10.1002/jcal.70257 (2026).
51. Zhang, H., Zhang, Y., Xu, T. & Zhou, Y. Effects of VR instructional approaches and textual cues on performance, cognitive load, and learning experience. *Etr&d-Educational Technology Research and Development*, *72*(2), 585–607. (2024). https://doi.org/10.1007/s11423-023-10313-1
52. Zhang, S., Li, Y., Gan, G., Stevany, Pang, S. & Kim, J. H. Effects of social cues of artificial intelligence-powered pedagogical agents: A multilevel meta-analysis. *Educational Res. Rev.* **49**, 100746. https://doi.org/10.1016/j.edurev.2025.100746 (2025).
53. Zhao, F. et al. Role of Race and Gender of Pedagogical Agents in Multimedia Learning. *J. Educational Comput. Res.* **63**, 1–40. https://doi.org/10.1177/07356331251322469 (2025).
54. Zhao, F. et al. Recognizing and Relating to the Race/Ethnicity and Gender of Animated Pedagogical Agents. *J. Educational Comput. Res.* **62** (3), 455–481. https://doi.org/10.1177/07356331231213932 (2024).
55. Zibrek, K., Kokkinara, E. & McDonnell, R. The Effect of Realistic Appearance of Virtual Characters in Immersive Environments—Does the Character's Personality Play a Role? *Ieee Transactions on Visualization and Computer Graphics*, *24*(4), 1681–1690. 25th IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR). (2018). https://doi.org/10.1109/TVCG.2018.2794638

Figures

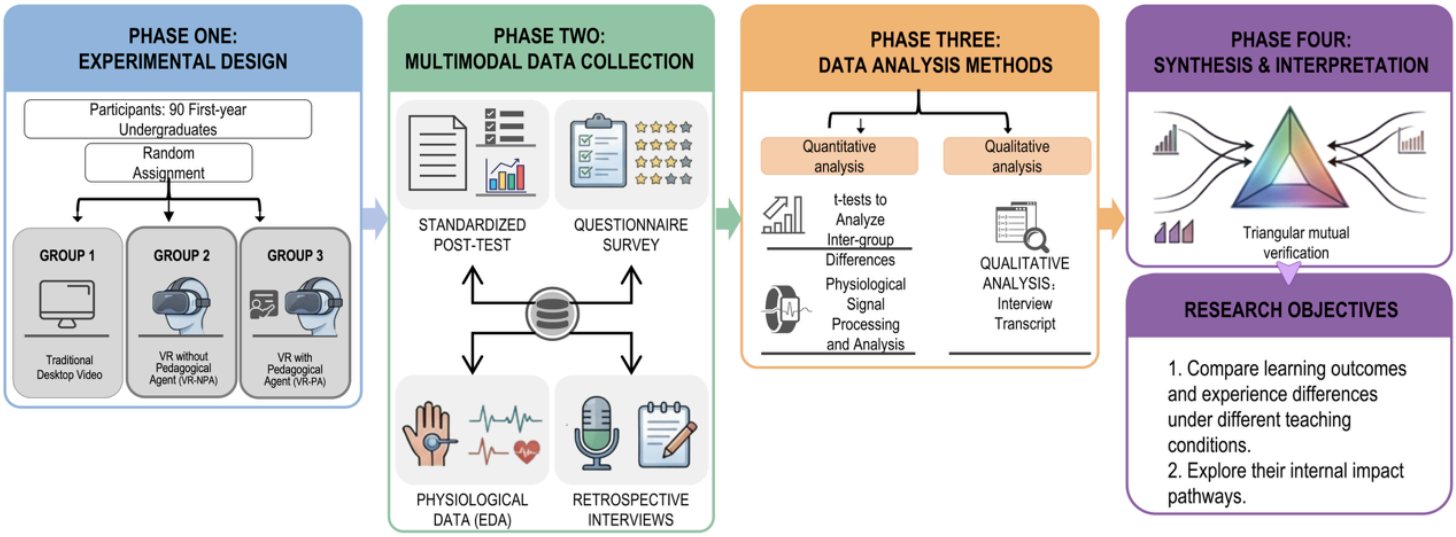

Figure 1

Research flowchart (created by authors)



Figure 2

A screenshot from the VR experiment：a: Student's subjective perspective；b：Simulated photography classroom scene；c：Pedagogical agent；

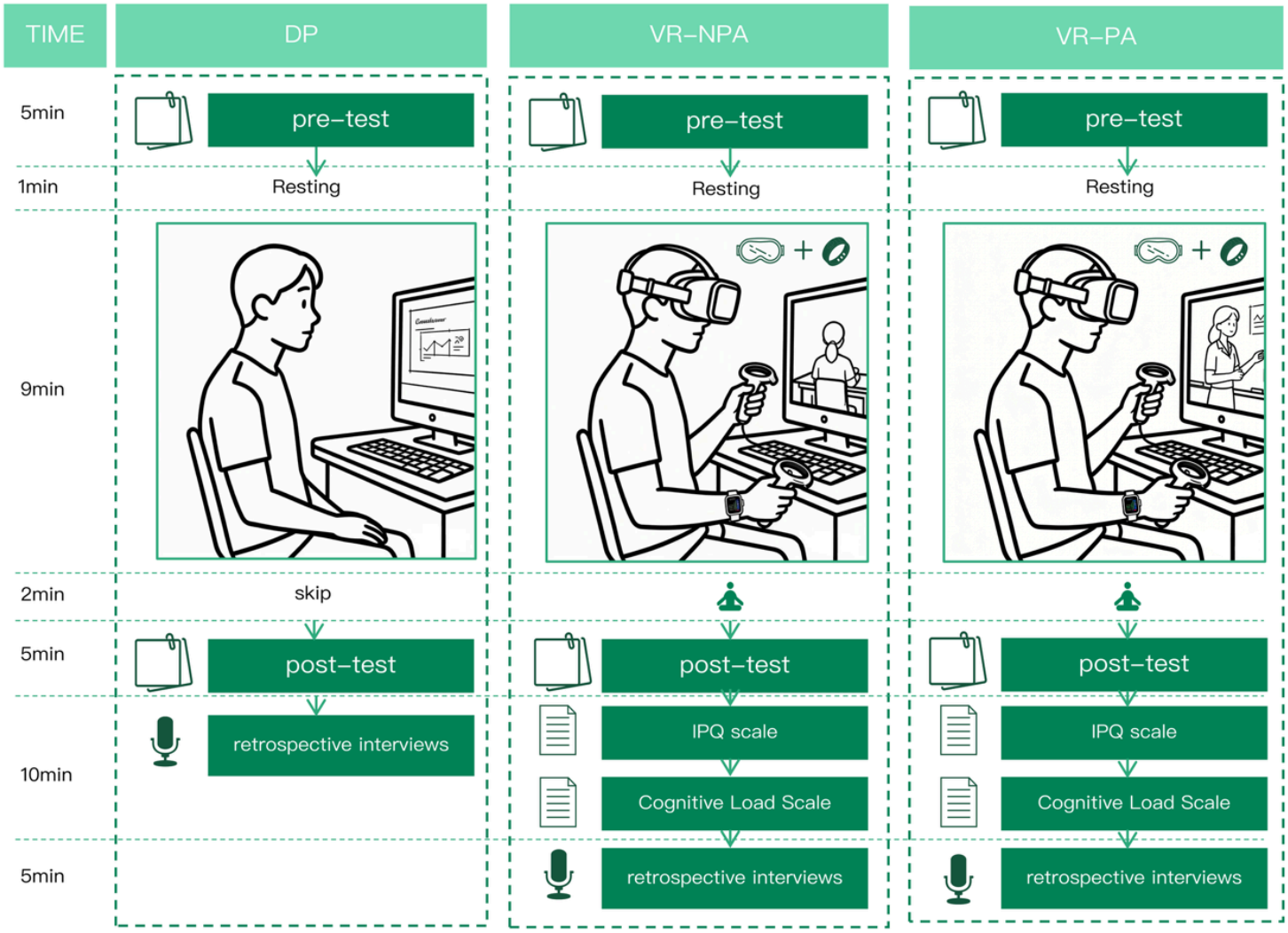

Figure 3

Experimental design (created by authors)

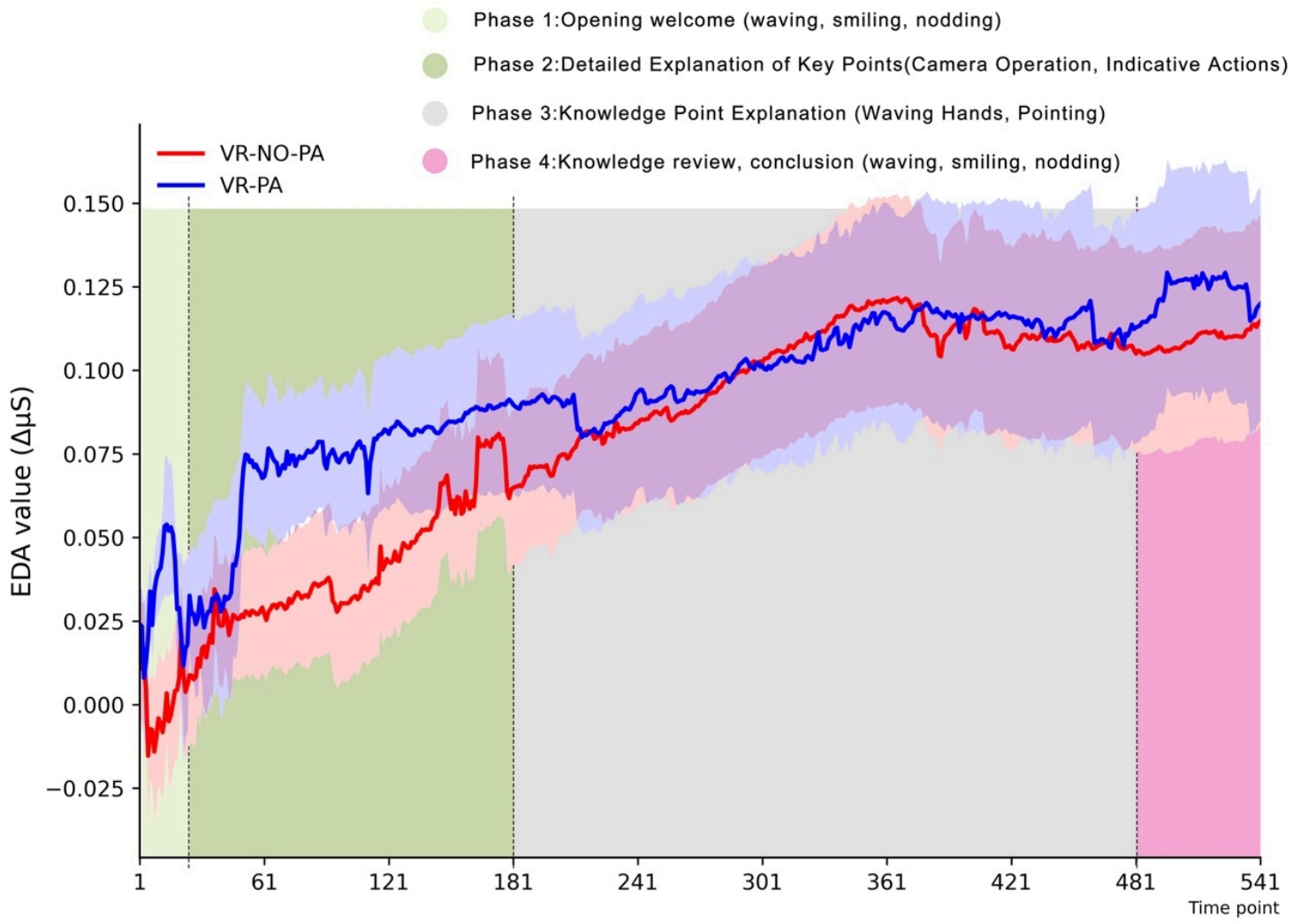

Figure 4

Multimodal Timeline Comparison Chart (created by authors)

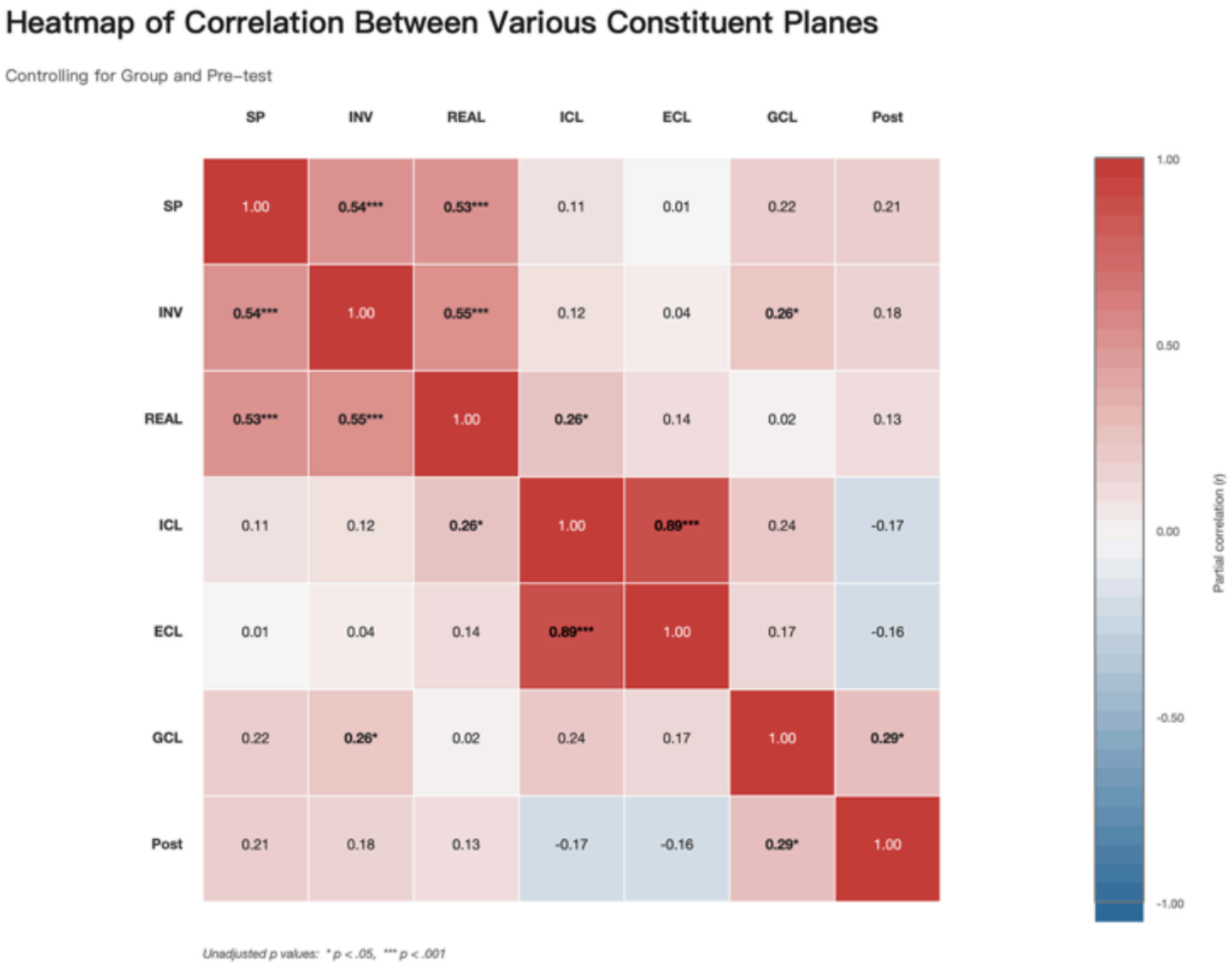

Figure 5

Heatmap of Correlation Between Various Constituent Planes (created by authors)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.zip