



2

Aprile 2026

**Immersive technologies and educational responsibility in teacher training:
empirical evidence from the MuSEd VR project**

**Tecnologie immersive e responsabilità educativa nella formazione docente:
evidenze empiriche dal progetto MuSEd VR**

Claudio Palazzi

Università degli Studi Roma Tre

claudio.palazzi@uniroma3.it

Doi: <https://>

ABSTRACT

The study investigates the conditions under which the integration of VR and GenAI may constitute a pedagogically grounded instrument of didactic innovation. The research comparatively evaluates MuSEd VR, an immersive environment featuring GenAI-based conversational avatars, against the traditional MuSEd "Mauro Laeng" museum. The mixed-methods study involves 2,524 participants: an experimental group (N = 1,443, VR experience) and a control group (N = 1,081, physical museum visit, non-randomized access). The results demonstrate the superiority of the VR experience across all evaluative dimensions (Cohen's d effect sizes ranging from 0.73 to 3.40; $p < .001$). The findings are discussed through the original AIM model (Immersive Mediated Learning) and their implications for both pre-service and in-service teacher education.

Keywords: teacher education, MuSEd, virtual reality, generative artificial intelligence, Immersive Mediated Learning.

RIASSUNTO

Lo studio indaga le condizioni in cui l'integrazione di realtà virtuale e intelligenza artificiale generativa possano costituire uno strumento di innovazione didattica pedagogicamente fondato. La ricerca valuta comparativamente il museo tradizionale MuSEd "Mauro Laeng" e il MuSEd VR, ambiente immersivo con avatar conversazionali basati su GenAI. Lo studio mixed-methods coinvolge 2.524 partecipanti: gruppo sperimentale (N = 1,443, esperienza VR) e gruppo di controllo (N = 1,081, visita al museo fisico, accesso non randomizzato). I risultati evidenziano una superiorità dell'esperienza VR su tutte le dimensioni valutative (dimensioni dell'effetto d di Cohen da 0,73 a 3,40; $p < 0,001$). I risultati sono discussi attraverso l'originale modello AIM (Apprendimento Immersivo Mediato) nelle loro implicazioni per la formazione iniziale e continua degli insegnanti.

Parole chiave: formazione docente, MuSEd, realtà virtuale educativa, intelligenza artificiale generativa, Apprendimento Immersivo Mediato.

1. INTRODUZIONE

La scuola e i sistemi formativi contemporanei sono attraversati da trasformazioni profonde che interrogano il significato dell'essere insegnanti oggi, rendendo evidente l'insufficienza di una concezione dell'insegnamento ridotta a funzione tecnica. In questo scenario, l'ingresso pervasivo dell'intelligenza artificiale e delle tecnologie immersive nei contesti educativi impone una riflessione critica sul rapporto tra tecnologie digitali, relazione educativa e responsabilità pedagogica (Agrusti, 2023), sostenendo soggettività professionali capaci di riconoscere la complessità dell'agire educativo in contesti tecnologicamente mediati (Nirchi *et al.*, 2025) ed evitando sia l'entusiasmo acritico che il rifiuto pregiudiziale delle innovazioni digitali (Isma'il & Ibrahim, 2025).

Il presente contributo indaga le condizioni sotto le quali le tecnologie immersive integrate con l'intelligenza artificiale generativa possano costituire uno strumento di innovazione didattica autenticamente pedagogico. La ricerca si fonda sull'implementazione del sistema MuSEd VR, un museo virtuale immersivo che ricrea digitalmente gli spazi del Museo della Scuola e dell'Educazione "Mauro Laeng" dell'Università Roma Tre, integrando *avatar* conversazionali che incarnano figure storiche della pedagogia italiana attraverso un'architettura di intelligenza artificiale generativa basata su corpus documentali di comprovata attendibilità storica. L'analisi comparativa tra l'esperienza immersiva e la visita al museo fisico, condotta su un campione di 2.524 partecipanti, offre evidenze empiriche significative per ripensare la formazione docente nell'era dell'intelligenza artificiale.

2. QUADRO DI RIFERIMENTO TEORICO

2.1 La professionalità docente tra innovazione tecnologica e responsabilità pedagogica

La trasformazione indotta dall'integrazione delle tecnologie intelligenti impone una revisione strutturale della professionalità insegnante, richiedendo competenze che trascendono la mera alfabetizzazione digitale (Nirchi *et al.*, 2016), configurandosi come sintesi complessa di saperi pedagogici, contenutistici e tecnologici da orchestrare in modo situato (Nascimento *et al.*, 2024). Il modello del Technological Pedagogical Content Knowledge (TPACK), sviluppato da Mishra e Koehler (Mishra & Koehler, 2006) a partire dal *framework* del Pedagogical Content Knowledge originariamente proposto da Shulman (Shulman, 1986), costituisce un *framework* essenziale per comprendere tale competenza professionale, postulando che l'efficacia dell'integrazione tecnologica dipenda dalla capacità di operare nell'intersezione tra conoscenza disciplinare, metodologie pedagogiche e *affordance* delle tecnologie educative (Su, 2023). Le tecnologie di intelligenza artificiale generativa e di realtà virtuale immersiva pongono sfide inedite alla riflessività pedagogica, chiamando l'insegnante a operare come progettista consapevole di esperienze di apprendimento e modello di pensiero critico.

2.2 Il modello dell'Apprendimento Immersivo Mediato (AIM) come framework per la formazione docente

Il modello dell'Apprendimento Immersivo Mediato (AIM) costituisce un contributo teorico sviluppato per rispondere alle sfide poste dalla digitalizzazione avanzata dei contesti educativi e culturali, integrando sistemicamente tre dimensioni fondamentali la cui interazione sinergica determina la qualità complessiva dell'esperienza educativa.

La prima dimensione, l'*immersione tecnologica*, si riferisce al grado di coinvolgimento sensoriale e cognitivo generato dall'ambiente virtuale. Radicandosi nella teoria della presenza di Lombard e Ditton (Lombard & Ditton, 1997), questa dimensione esplora come il senso di coinvolgimento spaziale faciliti l'*engagement* e riduca le distrazioni, creando le condizioni per un apprendimento profondo.

La seconda dimensione, la *mediazione pedagogica*, rappresenta il cuore del modello AIM. Ispirandosi ai principi vygotskijani della zona di sviluppo prossimale e al concetto di *scaffolding* (Fani & Ghaemi, 2011), questa dimensione concettualizza gli agenti di intelligenza artificiale come mediatori esperti che guidano l'apprendimento in modo dialogico e personalizzato attraverso uno *scaffolding* digitale (Chen *et al.*, 2023) adattivo. Nel progetto MuSed VR, tale mediazione si realizza attraverso *avatar* conversazionali che incarnano figure storiche della pedagogia italiana, alimentati da corpus documentali autentici mediante un'architettura RAG (*Retrieval-Augmented Generation*) che garantisce coerenza storica e fedeltà al pensiero pedagogico originale.

La terza dimensione, l'*esperienza culturale*, postula che la tecnologia debba mantenersi costantemente al servizio dei contenuti culturali e dei valori pedagogici, ponendosi in antitesi rispetto alla visione tecnocentrica che privilegia il "come" tecnologico rispetto al "perché" pedagogico. La tecnologia raggiunge la sua massima efficacia quando diventa sufficientemente trasparente da permettere un incontro diretto con il contenuto culturale.

L'aspetto più innovativo del modello AIM è la teorizzazione dell'interdipendenza dinamica delle tre dimensioni: l'immersione crea le precondizioni per un coinvolgimento profondo, la mediazione pedagogica struttura l'esperienza, l'esperienza culturale ne fornisce la sostanza e il senso.

2.3 Tecnologie, intelligenza artificiale e il diritto all'educazione: verso la democratizzazione culturale

Una delle questioni più rilevanti che emerge dal rapporto tra formazione docente e tecnologie immersive riguarda il potenziale democratizzante di questi strumenti nell'accesso all'educazione e alla cultura. Le istituzioni culturali tradizionali perpetuano di fatto molteplici forme di esclusione legate al *background* educativo, alla distanza geografica, alle condizioni di disabilità e alle difficoltà psicologiche.

Tuttavia, questa potenzialità deve essere valutata criticamente: le medesime tecnologie, se non accompagnate da una formazione adeguata, possono generare nuove forme di esclusione digitale. Come evidenziato dall'UNESCO (UNESCO, 2022), la questione è come formare i docenti affinché le utilizzino in modo pedagogicamente fondato ed eticamente responsabile. Il rischio della visione tecnocentrica risiede nella sottovalutazione del ruolo mediatore del docente, la cui competenza professionale e la cui "vigilanza epistemologica" rimangono il fattore determinante nell'efficacia di qualsiasi intervento educativo tecnologicamente mediato.

3. METODOLOGIA

3.1 Contesto della ricerca e sistema MuSed VR

La presente ricerca si fonda sulla progettazione, implementazione e valutazione del sistema MuSed VR, un'inedita applicazione di realtà virtuale immersiva che ricrea digitalmente gli spazi e i contenuti del Museo della Scuola e dell'Educazione "Mauro Laeng" dell'Università Roma Tre (Cantatore, 2022). Il sistema articola l'esperienza virtuale in quattro sale tematiche dedicate rispettivamente alla scuola contadina dell'Agro romano, all'aula montessoriana, alla scuola popolare di Giuseppe Lombardo Radice e al mondo della letteratura per l'infanzia, integrando *avatar* intelligenti che incarnano le figure storiche di Maria Montessori, Giuseppe Lombardo Radice, Duilio Cambellotti e Pinocchio. Ciascun *avatar* è stato alimentato mediante un'architettura RAG basata su documenti storicamente fondati comprendenti scritti originali, le trascrizioni di conferenze e altri materiali d'archivio delle rispettive figure storiche, garantendo che le risposte generate dall'intelligenza artificiale mantengano rigorosa coerenza storica, accuratezza concettuale e fedeltà al pensiero pedagogico autentico.

3.2 Disegno di ricerca e campione

Lo studio adotta un disegno quasi-sperimentale comparativo con *mixed-method*, coinvolgendo un campione complessivo di 2.524 partecipanti, suddivisi in un gruppo sperimentale, composto da 1.443 soggetti che hanno fruito dell'esperienza VR immersiva, e un gruppo di controllo, composto da 1.081

soggetti che hanno effettuato la visita tradizionale al museo fisico con guida umana. Entrambe le condizioni hanno avuto una durata standardizzata di 15 minuti per controllare l'effetto del tempo sull'elaborazione cognitiva. Il campione è stato stratificato per genere, fascia d'età e livello di istruzione. Nel gruppo VR, la distribuzione per genere comprende 768 maschi, 669 femmine e 6 soggetti che non hanno specificato il genere, mentre nel gruppo del museo fisico si registrano 887 maschi, 87 femmine e 107 soggetti non specificati.

Va precisato che i due gruppi presentano distribuzioni demografiche significativamente asimmetriche: il genere ($\chi^2(2) = 505,35$, $p < 0,001$) e il livello di istruzione ($\chi^2(4) = 1435,03$, $p < 0,001$) risultano non equivalenti, con una quota femminile del 46,3% nel gruppo VR contro l'8,0% nel gruppo fisico, e una prevalenza di diplomati e laureati magistrali nel VR rispetto alla dominanza di laureati triennali nel gruppo fisico (80,6%). Tuttavia, le analisi intra-gruppo confermano l'assenza di effetti sistematici di genere e istruzione sulle variabili di *outcome* (**es. interesse per genere nel VR: $M_M = 8,04$ vs $M_F = 8,11$, $t = -0,90$, $p = 0,367$; variazione per istruzione: 7,56–8,12), escludendo un *bias* sostanziale nei confronti principali. Questa limitazione è discussa nella sezione 6.**

I gruppi non sono stati formati per randomizzazione bensì per flusso naturale di accesso: disegno quasi-sperimentale (cfr. sezione 6).

3.3 Strumenti di raccolta dati

La raccolta dati è stata condotta mediante due questionari strutturati, somministrati post-visita, rispettivamente al gruppo sperimentale e al gruppo di controllo, ciascuno comprendente tre sezioni distinte. La prima sezione ha rilevato le informazioni anagrafiche e demografiche dei partecipanti, includendo età, genere e livello di istruzione. La seconda sezione ha valutato l'esperienza tramite scale *Likert* a 10 punti su sei dimensioni: interesse percepito, chiarezza delle spiegazioni, difficoltà attentiva, efficacia della presentazione, coinvolgimento emotivo e utilità percepita della modalità VR rispetto a quella tradizionale; inclusa una domanda sulla frequenza d'uso del medium per valutare la familiarità pregressa. La terza sezione comprendeva un test di apprendimento composto da 9 domande a scelta multipla calibrate secondo la tassonomia di Bloom, coprendo diversi livelli cognitivi dalla conoscenza fattuale alla comprensione concettuale, unitamente a due domande aperte relative alle funzionalità desiderate e agli aspetti percepiti come più utili per l'apprendimento.

La coerenza interna del test di apprendimento è stata valutata mediante il coefficiente KR-20 (Kuder-Richardson Formula 20), appropriato per item dicotomici. Nel gruppo VR il valore ottenuto è KR-20 = 0,577, indicativo di un'affidabilità moderata adeguata per strumenti esplorativi in contesti applicativi (Maggino et al, 2007). Nel gruppo del museo fisico il KR-20 risulta prossimo a zero, consistentemente con l'ipotesi che le risposte riflettano prevalentemente una scelta casuale in assenza di apprendimento sistematico dei contenuti specifici esposti, come discusso nella sezione 4.6.

3.4 Analisi dei dati

L'analisi quantitativa è stata condotta attraverso l'applicazione di test parametrici e non parametrici. Per il confronto tra i due gruppi dei punteggi su scala *Likert* sono stati impiegati il test t di Student per campioni indipendenti con correzione di Welch, adottata a seguito della verifica dell'assunto di omoschedasticità mediante il test di Levene e il test U di Mann-Whitney a conferma dei risultati. La

dimensione dell'effetto è stata quantificata mediante il d di Cohen. Le correlazioni tra variabili sono state analizzate attraverso il coefficiente r di Pearson. Tutte le analisi statistiche sono state condotte con il *software* Jamovi. Il livello di significatività adottato è $\alpha = 0,05$.

4. RISULTATI

4.1 Equivalenza delle condizioni di partenza

L'analisi preliminare delle condizioni di partenza dei partecipanti ha evidenziato un elemento di particolare rilevanza metodologica. La familiarità con il medium è stata operazionalizzata in modo necessariamente differente nei due gruppi: nel gruppo VR come frequenza d'uso di tecnologie di realtà virtuale ($M = 3,29$, $DS = 2,24$), nel gruppo museo fisico come frequenza di visita museale ($M = 3,40$, $DS = 1,50$). Sebbene i valori medi siano numericamente simili ($t = -1,398$, $p = 0,162$), i due *item* misurano costrutti distinti e non possono essere assimilati a una misura unica di familiarità con il proprio medium. Questa limitazione rende parziale il controllo delle condizioni di partenza rispetto all'esperienza pregressa con ciascuna modalità.

L'analisi comparativa delle caratteristiche demografiche ha evidenziato differenze statisticamente significative tra i gruppi per età ($t = 2,791$, $p = .005$, $d = 0,11$), genere ($\chi^2(2) = 505,35$, $p < 0,001$) e livello di istruzione ($\chi^2(4) = 1435,03$, $p < 0,001$). Nonostante la significatività statistica, la dimensione dell'effetto per l'età risulta trascurabile ($d = 0,11$), e le analisi intra-gruppo confermano l'assenza di correlazioni sostanziali tra queste variabili demografiche e gli *outcome* considerati. Tali differenze sono riconducibili alle caratteristiche strutturali dei rispettivi flussi di utenti e sono discusse come limitazione nella sezione 6.

4.2 Confronto sulle variabili Likert: interesse, chiarezza ed efficacia

L'analisi delle statistiche descrittive e inferenziali relative alle variabili Likert, le cui sintesi sono riportate nella Tabella 1, rivela una superiorità statisticamente significativa e praticamente rilevante dell'esperienza VR immersiva rispetto alla visita museale tradizionale su tutte le dimensioni valutative considerate.

Per l'interesse percepito, il gruppo VR ($M = 8,08$, $DS = 1,38$) ha registrato una differenza eccezionalmente elevata rispetto al gruppo fisico ($M = 2,63$, $DS = 1,79$), $t(1965,6) = 83,091$, $p < 0,001$, $d = 3,40$, con conferma non parametrica (Tabella 1).

Variabile	M (VR)	DS (VR)	M (Fisico)	DS (Fisico)	t	p	d di Cohen
Interesse percepito	8,08	1,38	2,63	1,79	83,09	< ,001	3,40
Chiarezza spiegazioni	7,99	1,41	2,68	1,88	77,79	< ,001	3,26
Difficoltà attenzione	6,93	2,73	1,85	2,14	52,39	< ,001	2,04
Efficacia presentazione	6,97	2,67	2,69	1,90	46,97	< ,001	1,80
Coinvolgimento emotivo	7,45	2,40	5,52	2,86	18,01	< ,001	0,74
Utilità VR vs tradizionale	8,53	1,11	5,51	2,89	32,66	< ,001	1,46

Tab. 1: Statistiche descrittive e inferenziali per le variabili Likert (scala 1-10), confronto tra il gruppo VR immersivo (N = 1.443) e il gruppo museo fisico (N = 1.081)

Nota: M = media aritmetica; DS = deviazione standard; t = valore del test t di Welch; p = livello di significatività statistica; d di Cohen = dimensione dell'effetto.

Per la variabile Difficoltà percepita nel mantenimento dell'attenzione, punteggi più elevati indicano maggiore difficoltà; la direzione della scala è quindi invertita rispetto alle altre variabili, per le quali punteggi più elevati indicano un esito positivo.

La chiarezza percepita ha evidenziato un *pattern* analogo (VR: M = 7,99, DS = 1,41 vs fisico: M = 2,68, DS = 1,88), $t(1924,7) = 77,793$, $p < 0,001$, $d = 3,26$. L'efficacia della presentazione ha mostrato una differenza egualmente marcata (VR: M = 6,97, DS = 2,67 vs fisico: M = 2,69, DS = 1,90), $t(2515,6) = 46,971$, $p < 0,001$, $d = 1,80$ (Tabella. 2, Tabella. 3, Figura 1-6).

Variabili		Statistiche	gdl	p		<i>Dimensione dell'Effetto</i>
Interesse	t di Student	86,2	2522	<.,001	d di Cohen	3,466
	t di Welch	83,1	1966	<.,001	d di Cohen	3,403
	U di Mann-Whitney	65216		<.,001	Correlazione biseriale di rango	-0,916
Chiarezza	t di Student	81,0	2522	<.,001	d di Cohen	3,258
	t di Welch	77,8	1925	<.,001	d di Cohen	3,192
	U di Mann-Whitney	78235		<.,001	Correlazione biseriale di rango	-0,900
Difficolta_Attenzione	t di Student	50,6 ^a	2522	<.,001	d di Cohen	2,037
	t di Welch	52,4	2517	<.,001	d di Cohen	2,071
	U di Mann-Whitney	145244		<.,001	Correlazione biseriale di rango	-0,814
Efficacia	t di Student	44,8 ^a	2522	<.,001	d di Cohen	1,803
	t di Welch	47,0	2516	<.,001	d di Cohen	1,844
	U di Mann-Whitney	203875		<.,001	Correlazione biseriale di rango	-0,739
Coinvolgimento_Em otivo	t di Student	18,5 ^a	2522	<.,001	d di Cohen	0,743
	t di Welch	18,0	2086	<.,001	d di Cohen	0,733
	U di Mann-Whitney	475319		<.,001	Correlazione biseriale di rango	-0,391
Utilita_VR	t di Student	36,3 ^a	2522	<.,001	d di Cohen	1,462
	t di Welch	32,7	1503	<.,001	d di Cohen	1,433

Tab. 2: Confronto gruppo VR e gruppo museo tradizionale sulle variabili esperienziali: risultati t-test di Student, t-test di Welch e test U di Mann-Whitney con relative misure della dimensione dell'effetto

Nota. gdl = gradi di libertà. Per il t-test di Welch i gradi di libertà in virgola mobile riflettono la correzione per varianze diseguali. La dimensione dell'effetto è espressa tramite d di Cohen per i test parametrici e correlazione biseriale di rango per il test non parametrico

Variabili	Gruppo	N	Media	Mediana	SD	SE
Interesse	VR	1443	8,08	8,00	1,38	0,0364
	tradizionale	1081	2,63	2,00	1,79	0,0545
Chiarezza	VR	1443	7,99	8,00	1,41	0,0372
	tradizionale	1081	2,68	2,00	1,88	0,0573
Difficolta_Attenzione	VR	1443	6,93	8,00	2,73	0,0718
	tradizionale	1081	1,85	1,00	2,14	0,0651
Efficacia	VR	1443	6,97	8,00	2,67	0,0704
	tradizionale	1081	2,69	2,00	1,90	0,0579
Coinvolgimento_Emotivo	VR	1443	7,45	8,00	2,40	0,0632
	tradizionale	1081	5,52	6,00	2,86	0,0869
Utilità VR	VR	1443	8,53	9,00	1,11	0,0293
	tradizionale	1081	5,51	6,00	2,89	0,0878

Tab.3: Statistiche descrittive per le variabili esperienziali nei gruppi VR e museo tradizionale

Nota. N = numerosità del campione; SD = deviazione standard; SE = errore standard della media. I valori riportati si riferiscono a scale di risposta da 1 a 10

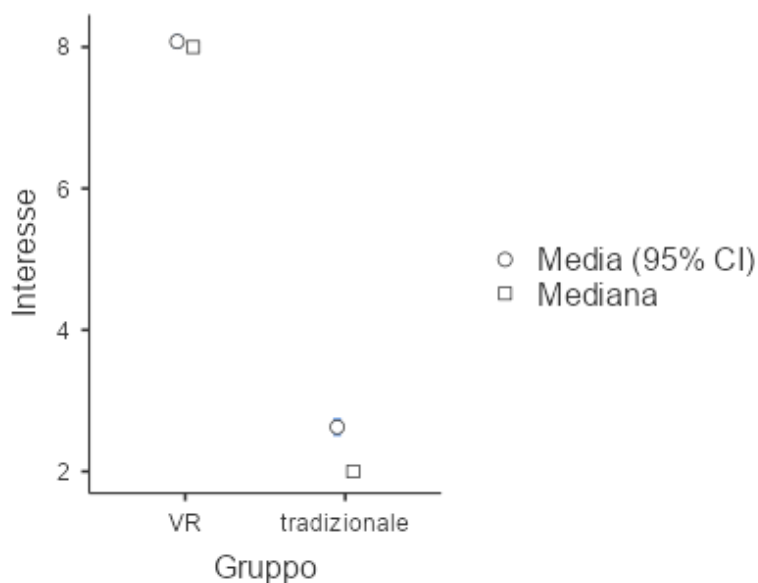


Fig. 1: Interesse

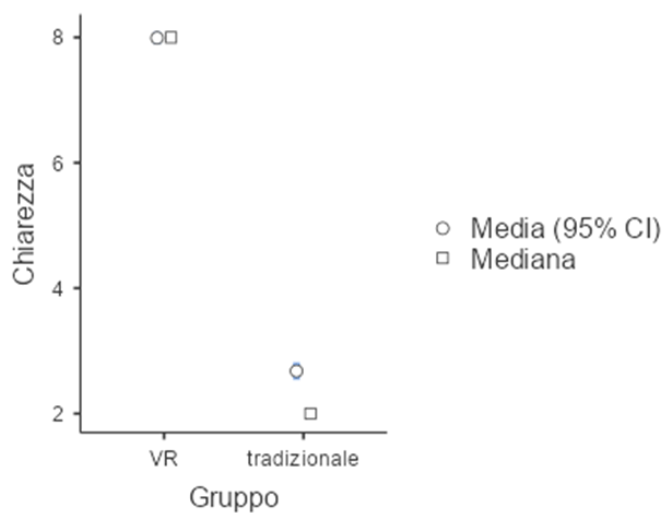


Fig. 2: Chiarezza

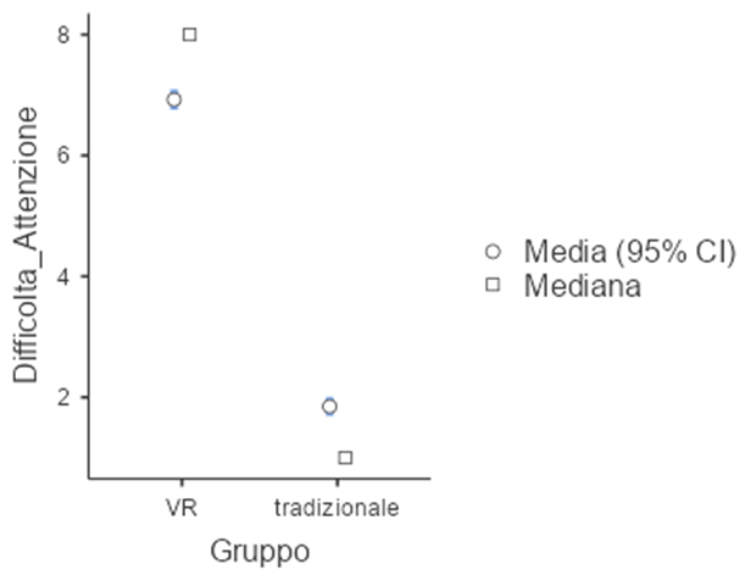


Fig. 3: Difficoltà attenzione

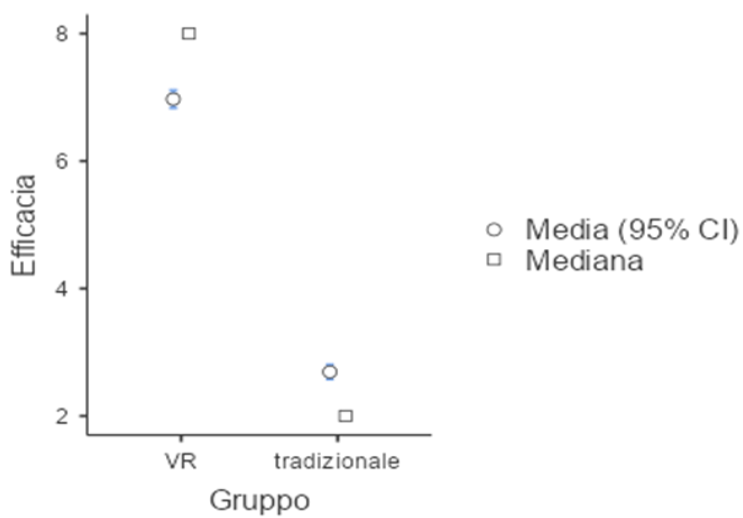


Fig. 4: Efficacia

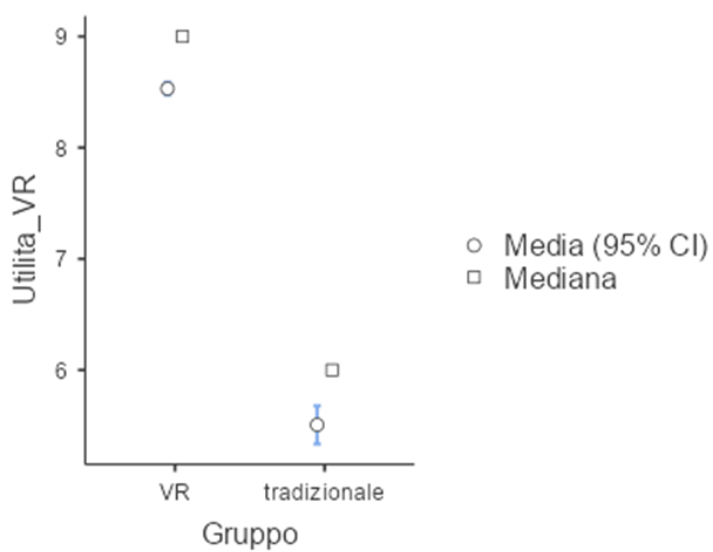


Fig. 5: Utilità VR

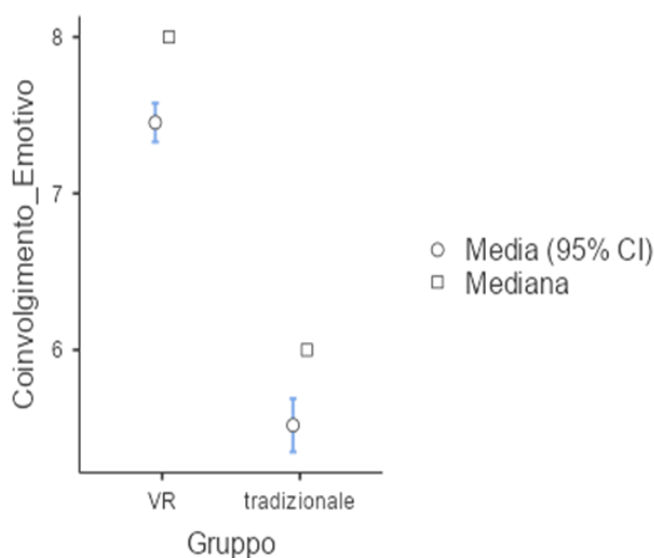


Fig. 6: Coinvolgimento emotivo

4.3 Coinvolgimento emotivo e difficoltà attentiva

Il coinvolgimento emotivo è la variabile con la dimensione dell'effetto più contenuta: VR ($M = 7,45$, $DS = 2,40$) vs fisico ($M = 5,52$, $DS = 2,86$), $t(2086,2) = 18,009$, $p < 0,001$, $d = 0,74$. Questo suggerisce che il museo fisico conserva una capacità specifica di risonanza emotiva, attraverso il contatto con gli oggetti autentici e la presenza della guida, non pienamente replicata dal contesto immersivo.

Il gruppo VR ha inoltre riportato una maggiore difficoltà percepita nel mantenimento dell'attenzione ($M = 6,93$, $DS = 2,73$) rispetto al gruppo del museo fisico ($M = 1,85$, $DS = 2,14$), con $t(2516,7) = 52,388$, $p < ,001$, $d = 2,04$ (Tabella 4, Tabella 5, Figura 7a, Figura 7b). Questo dato, apparentemente controintuitivo, risulta interpretabile alla luce della teoria del carico cognitivo come indicatore di un più intenso impegno cognitivo associato all'esperienza immersiva: la difficoltà attentiva percepita riflette la consapevolezza dell'intensità dello sforzo cognitivo richiesto da un'esperienza che stimola simultaneamente molteplici canali sensoriali e richiede un'elaborazione attiva delle informazioni presentate dagli *avatar* conversazionali. Questa lettura è coerente con la teoria del carico cognitivo, ma deve essere corroborata da misure più dirette, quali strumenti di *dual-task* o scale validate di *cognitive load*, che costituiscono una direzione prioritaria per le ricerche future (Sweller, 2020).

Variabili		Statistiche	gdl	p	Differenza media	Differenza SE		Dimensione dell'Effetto
Coinvolgimento_Emotivo	t di Student	18,5 ^a	2522	<0,001	1,94	0,105	d di Cohen	0,743
	t di Welch	18,0	2086	<0,001	1,94	0,1075	d di Cohen	0,733
Difficolta_Attenzione	t di Student	50,6 ^a	2522	<0,001	5,08	0,100	d di Cohen	2,037
	t di Welch	52,4	2517	<0,001	5,08	0,0969	d di Cohen	2,071

Tab. 4: Risultati del t-test di Welch per campioni indipendenti: confronto tra gruppo VR e gruppo museo fisico sul coinvolgimento emotivo e sulla difficoltà percepita nel mantenimento dell'attenzione

Nota. Il test di Levene è significativo ($p < 0,05$), suggerendo una violazione dell'assunzione di varianze uguali

Variabili	Gruppo	N	Media	Mediana	SD	SE
Coinvolgimento_Emotivo	VR	1443	7,45	8,00	2,40	0,0632
	tradizionale	1081	5,52	6,00	2,86	0,0869
Difficolta_Attenzione	VR	1443	6,93	8,00	2,73	0,0718
	tradizionale	1081	1,85	1,00	2,14	0,0651

Tab. 5: Statistiche descrittive per il coinvolgimento emotivo e la difficoltà percepita nel mantenimento dell'attenzione nei gruppi VR e museo fisico

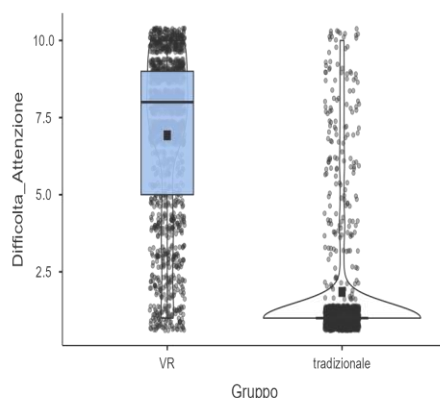


Fig. 7a: Confronto tra gruppo VR e gruppo museo fisico sul coinvolgimento emotivo e sulla difficoltà percepita nel mantenimento dell'attenzione: medie con intervalli di confidenza al 95%

Nota. Le barre rappresentano i valori medi; i segmenti indicano gli intervalli di confidenza al 95%. Le differenze tra i gruppi sono statisticamente significative per entrambe le variabili ($p < ,001$)

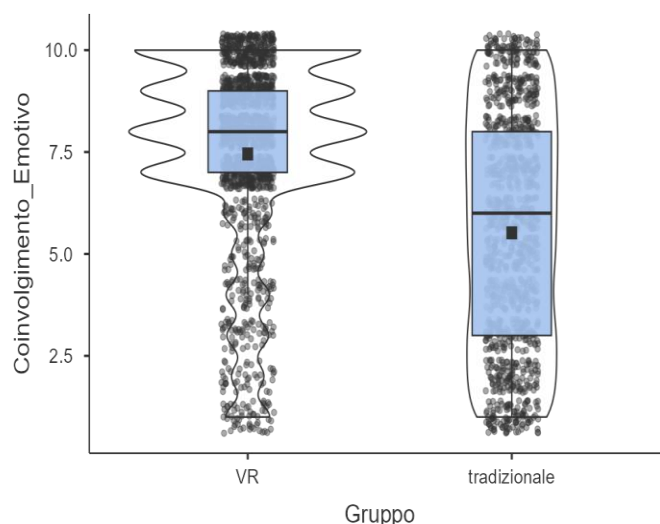


Fig. 7b: distribuzione dei punteggi di coinvolgimento emotivo e di difficoltà percepita nel mantenimento dell'attenzione nei gruppi VR e museo fisico

Nota. I box plot: mediana, range interquartile e valori anomali; i violin plot sovrapposti rappresentano la densità della distribuzione empirica. I punti singoli corrispondono alle osservazioni individuali (jittered)

4.4 Assenza di digital divide generazionale

L'analisi delle correlazioni tra l'età dei partecipanti e le variabili esperienziali nel gruppo VR costituisce uno dei risultati più rilevanti dal punto di vista delle implicazioni per la formazione docente. Le correlazioni risultano uniformemente prossime allo zero per tutte le dimensioni considerate: $r = 0,017$ per l'interesse percepito ($p = 0,527$), $r = 0,038$ per la chiarezza delle spiegazioni ($p = 0,149$), $r = -0,014$ per la difficoltà attentiva ($p = 0,597$), $r = -0,006$ per l'efficacia della presentazione ($p = 0,807$) e $r = 0,012$ per il coinvolgimento emotivo ($p = 0,652$) (Tabella 6, Figura 8). Nessuna correlazione raggiunge la significatività statistica, e le dimensioni delle correlazioni sono tutte trascurabili secondo i criteri convenzionali.

Variabili		Interesse	Chiarezza	Difficoltà attenzione	Efficacia	Coinvolgimento emotivo	Età
Interesse	r di Pearson	—					
	gdl	—					
	valore p	—					
	N	—					

Chiarezza	r di Pearson	0,002	—				
	gdl	1441	—				
	valore p	0,935	—				
	N	1443	—				
Difficoltà Attenzione	r di Pearson	0,016	-0,027	—			
	gdl	1441	1441	—			
	valore p	0,533	0,308	—			
	N	1443	1443	—			
Efficacia	r di Pearson	-0,064*	-0,022	-0,020	—		
	gdl	1441	1441	1441	—		
	valore p	0,014	0,405	0,444	—		
	N	1443	1443	1443	—		
Coinvolgimento Emotivo	r di Pearson	0,040	-0,028	0,034	-0,007	—	
	gdl	1441	1441	1441	1441	—	
	valore p	0,131	0,287	0,197	0,785	—	
	N	1443	1443	1443	1443	—	
Eta	r di Pearson	0,017	0,038	-0,014	-0,006	0,012	—
	gdl	1440	1440	1440	1440	1440	—
	valore p	0,527	0,149	0,597	0,807	0,652	—
	N	1442	1442	1442	1442	1442	—

Tab. 6: Correlazioni di Pearson tra l'età dei partecipanti e le variabili esperienziali nel gruppo VR

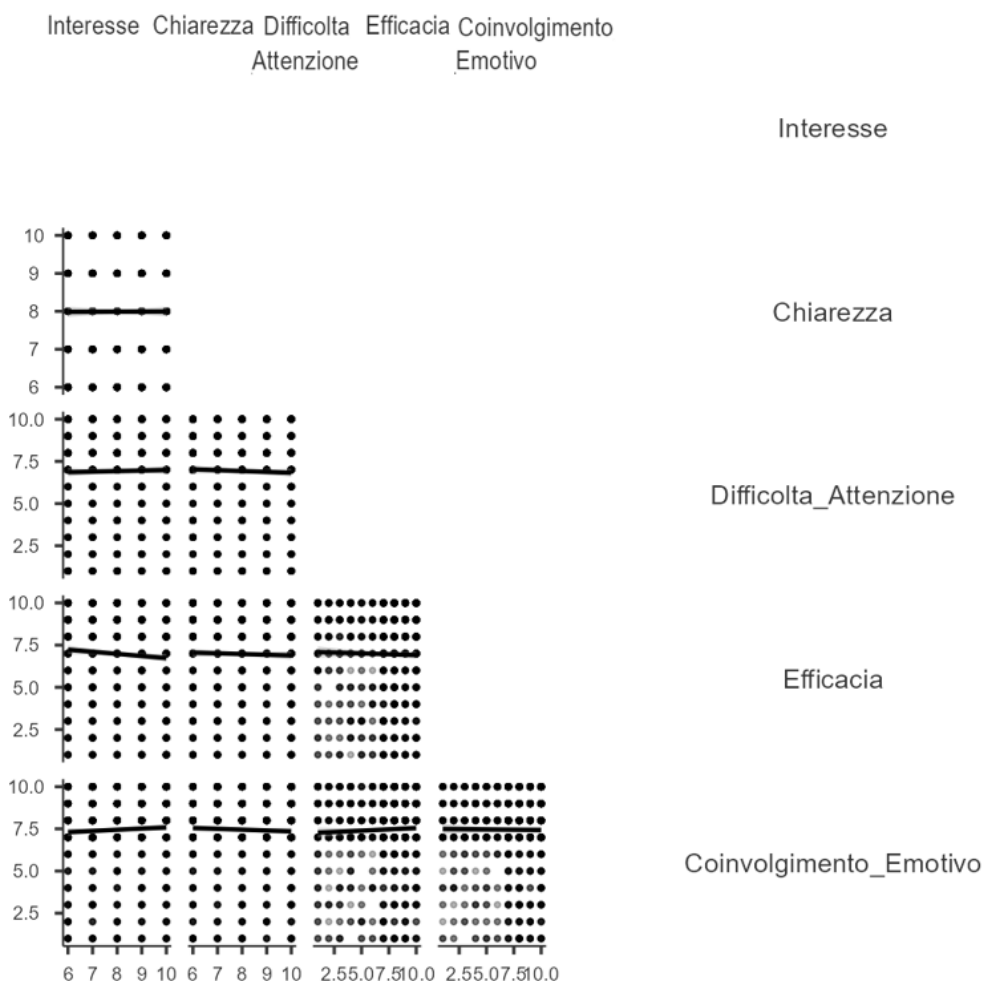


Fig. 8: Coefficienti di correlazione di Pearson tra l'età dei partecipanti e le variabili esperienziali nel gruppo VR

Fa eccezione la correlazione tra Interesse percepito ed Efficacia della presentazione ($r = -0,064$, $p = 0,014$), statisticamente significativa ma di entità trascurabile, che non modifica l'interpretazione complessiva.

4.5 Indipendenza delle dimensioni valutative

L'analisi della matrice di correlazione tra le variabili Likert nei due gruppi sperimentali rivela un *pattern* strutturale di notevole interesse teorico. Nel gruppo VR, le correlazioni tra le dimensioni valutative risultano uniformemente prossime allo zero, con valori compresi tra $r = -0,064$ e $r = 0,040$, indicando che ciascuna variabile misura un aspetto indipendente e distinto dell'esperienza immersiva. Nel gruppo del museo fisico, per contro, si osservano correlazioni deboli ma positive tra interesse, chiarezza, attenzione ed efficacia (r compreso tra 0,077 e 0,142), mentre il coinvolgimento emotivo mantiene la propria indipendenza dalle altre dimensioni. Questo *pattern* suggerisce che l'esperienza VR attivi dimensioni percettive e cognitive effettivamente distinte grazie alla sua ricchezza

multisensoriale, mentre nell'esperienza tradizionale le percezioni tendono a covariare moderatamente, configurando un'esperienza più unitaria.

4.6 Risultati del test di apprendimento

I risultati del test di apprendimento, composto da 9 domande a scelta multipla calibrate secondo la tassonomia di Bloom, evidenziano una differenza di eccezionale rilevanza pedagogica tra i due gruppi, le cui sintesi sono riportate nella Tabella 7. Il gruppo VR ha raggiunto una percentuale media di risposte corrette del 70,2% contro il 25,6% del gruppo del museo fisico, con una differenza complessiva di 44,6 punti percentuali che configura un divario di apprendimento sostanziale e sistematico.

È opportuno precisare che il test di apprendimento dei due gruppi valuta contenuti specificamente diversi (le figure e i temi presentati dagli *avatar* nel caso VR; i contenuti della visita guidata nel caso del museo fisico), rendendo i due strumenti non direttamente comparabili in termini di difficoltà assoluta. La prossimità del punteggio del gruppo fisico (25,3%) al valore atteso per risposta casuale su quattro alternative (25,0%), confermata dal $KR-20 \approx 0$, suggerisce che le domande del test fossero calibrate su contenuti non adeguatamente trasmessi dalla visita tradizionale, piuttosto che su una difficoltà intrinseca degli item. Questa limitazione deve essere considerata nell'interpretazione del divario di apprendimento.

4.7 Analisi qualitativa delle risposte aperte

Le risposte alle domande aperte (Q12 e Q13) sono state analizzate mediante analisi tematica induttiva (Braun & Clarke, 2006). Nel gruppo VR le richieste più frequenti riguardavano guide virtuali interattive, gamification e modalità di esplorazione autonoma; gli aspetti più utili per l'apprendimento erano le interazioni con gli *avatar* e la ricostruzione ambientale. Nel gruppo museo fisico prevalevano le richieste di percorsi personalizzati e supporti digitali integrati, mentre gli aspetti apprezzati si concentravano sul contatto con gli oggetti autentici e sulla relazione con la guida umana.

Domanda	VR (%)	Fisico (%)	Δ (pp)
Q1	71,3	26,8	+44,5
Q2	72,1	25,3	+46,8
Q3	69,7	25,0	+44,7
Q4	69,1	26,5	+42,6
Q5	70,1	25,1	+45,0
Q6	68,7	26,4	+42,3

Q7	70,8	24,9	+45,9
Q8	70,1	26,5	+43,6
Q9	69,5	23,8	+45,7
Media	70,2	25,6	+44,6

Tab. 7: Percentuale di risposte corrette per ciascuna delle 9 domande del test di apprendimento, confronto tra il gruppo VR immersivo ($N = 1.443$) e il gruppo museo fisico ($N = 1.081$)

Nota: Δ = differenza in punti percentuali tra il gruppo VR e il gruppo del museo fisico.

Un elemento che merita particolare attenzione riguarda la consistenza delle performance nel gruppo VR: le percentuali di risposte corrette risultano uniformemente comprese tra il 68,7% e il 72,1% su tutte le 9 domande, con una varianza minima che suggerisce un apprendimento omogeneo e trasversale all'intera esperienza museale immersiva. Il sistema MuSed VR si dimostra sistematicamente più efficace per tutti i contenuti trattati, indipendentemente dalla sala tematica di riferimento e dalla figura storica incarnata dall'*avatar* conversazionale. Nel gruppo del museo fisico, per contro, le percentuali di risposte corrette oscillano tra il 23,8% e il 26,8%, collocandosi in prossimità del valore atteso in caso di risposta casuale su un test a 4 alternative (25%), un dato che suggerisce un apprendimento fattuale assai limitato dalla visita tradizionale con guida umana.

Questo risultato assume implicazioni dirette e rilevanti per la formazione dei docenti. Se un sistema educativo fondato sulla realtà virtuale e sull'intelligenza artificiale generativa, progettato secondo i principi del modello AIM, produce un apprendimento fattuale significativamente superiore rispetto alla modalità tradizionale con un margine di tale ampiezza, è necessario interrogarsi sulla capacità dei percorsi di formazione insegnante di preparare professionisti in grado di progettare, orchestrare e valutare criticamente questo tipo di esperienze educative tecnologicamente mediate. La questione riguarda l'integrazione consapevole di strumenti che, quando pedagogicamente fondati, dimostrano un'efficacia empiricamente documentata.

5. DISCUSSIONE

5.1 Implicazioni per la formazione docente

Le dimensioni dell'effetto osservate ($d = 0,73-3,40$) si collocano nella categoria "grande" o "molto grande" secondo le convenzioni di Cohen, superando le medie riportate nelle principali meta-analisi sulla VR educativa ($d = 0,50$) (Merchant et al., 2014; Radianti et al., 2020). L'interpretazione di queste evidenze richiede tuttavia cautela, poiché la superiorità quantitativa dell'esperienza immersiva suggerisce che l'integrazione consapevole delle tecnologie possa amplificare l'efficacia dei processi di insegnamento-apprendimento senza tuttavia esaurirne la complessità pedagogica.

In primo luogo, i dati evidenziano che le tecnologie immersive integrate con l'intelligenza artificiale generativa possiedono un potenziale trasformativo eccezionale quando sono progettate secondo principi pedagogici rigorosi, come teorizzato dal modello AIM. Gli *avatar* conversazionali del

MuSEd VR realizzano una mediazione pedagogica che dimostra come l'intelligenza artificiale possa fungere da *partner* educativo capace di personalizzare l'esperienza e generare risonanza emotiva con i contenuti culturali.

In secondo luogo, l'assenza di *digital divide* generazionale demolisce l'assunto che tali strumenti sarebbero efficaci soltanto per le generazioni più giovani. Questa evidenza assume particolare rilevanza per la formazione in servizio degli insegnanti di ogni fascia d'età, suggerendo che la variabile determinante sia la qualità della progettazione pedagogica.

5.2 Responsabilità educativa e trasparenza tecnologica

Il dato relativo al coinvolgimento emotivo merita una riflessione specifica nel contesto della responsabilità educativa. La circostanza che questa sia la variabile nella quale la differenza tra le due modalità risulta meno marcata ($d = 0,74$), pur rimanendo statisticamente significativa, suggerisce che il museo fisico preserva una capacità specifica e non facilmente replicabile di generare risonanza emotiva attraverso il contatto con gli oggetti autentici, la fisicità degli spazi e la presenza umana della guida. Questa osservazione rafforza il principio cardine del modello AIM secondo cui la tecnologia deve rimanere trasparente, ovvero al servizio dell'esperienza culturale piuttosto che fine a sé stessa e che l'obiettivo educativo consiste nell'arricchire le modalità tradizionali e nel renderle accessibili a pubblici più ampi.

La responsabilità educativa del docente si configura come competenza composita: valutare quando la tecnologia arricchisce o impoverisce l'esperienza, riconoscere il valore della relazione umana e progettare percorsi ibridi centrati sul benessere degli studenti. La maggiore difficoltà attentiva nel gruppo VR richiama la necessità di una progettazione attenta dei tempi dell'esperienza immersiva, che tenga conto dei limiti della memoria di lavoro e del rischio di sovraccarico cognitivo.

5.3 Verso una formazione docente critica e trasformativa

Alla luce delle evidenze empiriche discusse, la formazione degli insegnanti nell'era dell'intelligenza artificiale e delle tecnologie immersive, oltre alla trasmissione di competenze tecniche nell'uso di strumenti, deve configurarsi come pratica critica e trasformativa capace di sviluppare quella che nel modello AIM viene definita "vigilanza epistemologica" rispetto all'utilizzo delle tecnologie digitali nei contesti educativi. I docenti devono essere educati a porsi costantemente la domanda fondamentale: l'integrazione tecnologica è subordinata agli obiettivi educativi e culturali, o è invece l'agire pedagogico a conformarsi alle caratteristiche e ai limiti della tecnologia?

Il modello AIM offre un *framework* operativo per questa formazione, educando i docenti a riconoscere che l'efficacia educativa risiede nell'integrazione sistemica tra immersione tecnologica, mediazione pedagogica ed esperienza culturale.

6. LIMITAZIONI

Le limitazioni principali includono la natura trasversale del disegno, la non completa randomizzazione del campione, l'assenza di misure longitudinali e la specificità museale, culturale e linguistica del contesto italiano, che richiede cautela nella generalizzazione ad altri Paesi. Le direzioni future includono studi longitudinali, la validazione del modello AIM in contesti diversificati, la

replica cross-culturale del protocollo in contesti museali internazionali, l'integrazione di misure fisiologiche del coinvolgimento e la sperimentazione di percorsi formativi che integrino esperienza diretta con le tecnologie immersive e riflessione critica sulle loro implicazioni pedagogiche.

Una limitazione rilevante riguarda lo squilibrio nella distribuzione di genere tra i gruppi, già documentato nel paragrafo 3.2. Sebbene le analisi intra-gruppo escludano un effetto confondente del genere sulle variabili di *outcome*, la non equivalenza demografica dei gruppi suggerisce cautela nell'attribuire le differenze osservate esclusivamente alla modalità di fruizione.

Ulteriore limitazione la non coincidenza temporale dei periodi di raccolta dati nei due gruppi (museo fisico: novembre 2024–ottobre 2025; gruppo VR: marzo 2025–marzo 2026). La sfasatura non consente di escludere completamente effetti di stagionalità, aggiornamenti del sistema VR nel corso del periodo, o variazioni nella composizione dei flussi di utenti.

Infine il possibile effetto *novelty effect*: i partecipanti del gruppo VR potrebbero aver espresso valutazioni più elevate in risposta alla natura insolita dell'esperienza immersiva, indipendentemente dalla sua efficacia pedagogica intrinseca. Studi longitudinali su partecipanti con diversa esposizione pregressa al sistema permetterebbero di isolare l'effetto della qualità pedagogica da quello della novità tecnologica.

7. CONCLUSIONI

Il presente studio fornisce evidenze empiriche su larga scala circa l'efficacia delle tecnologie immersive integrate con l'intelligenza artificiale generativa nella mediazione culturale. Il confronto tra esperienza VR e visita museale tradizionale, condotto su 2.524 partecipanti, documenta una superiorità statisticamente significativa dell'esperienza immersiva sulle dimensioni esperienziali valutate, con dimensioni dell'effetto grandi o molto grandi e assenza di *digital divide* generazionale. Il contributo principale risiede nelle implicazioni per la formazione docente. Il modello dell'Apprendimento Immersivo Mediato (AIM) proposto offre un *framework* teorico per comprendere che l'efficacia delle tecnologie educative immersive deriva da una progettazione pedagogica consapevole che integra immersione sensoriale, mediazione pedagogica e profondità culturale. La replicabilità del protocollo rende inoltre il caso MuSEd VR trasferibile a istituzioni museali di altri Paesi, aprendo alla verifica cross-culturale del modello AIM.

RIFERIMENTI BIBLIOGRAFICI

- Agrusti, F. (2023). L'AI literacy per una educazione attenta agli algoritmi. *Educazione e Intelligenza artificiale*, 19–32.
- UNESCO. (2022). *International Forum on AI and Education: Ensuring AI as a Common Good to Transform Education*, 7-8 December 2021; *synthesis report*. <https://unesdoc.unesco.org/ark:/48223/pf0000381226>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Cantatore, L. (2022). Storia e vita del MuSEd – Museo della Scuola e dell'Educazione dell'Università Roma Tre. *Educació i Història: Revista d'Història de l'Educació*, (39), 139–159.

- Chen, C.-H., Law, V., & Huang, K. (2023). Adaptive scaffolding and engagement in digital game-based learning. *Educational Technology Research and Development*, 71(4), 1785–1798.
- Fani, T., & Ghaemi, F. (2011). Implications of Vygotsky's Zone of Proximal Development (ZPD) in Teacher Education: ZPTD and Self-scaffolding. *Procedia - Social and Behavioral Sciences*, 29, 1549–1554. <https://doi.org/10.1016/j.sbspro.2011.11.396>
- Isma'il, A., & Ibrahim, H. B. (2025). Teachers' professional identity in the era of artificial intelligence: A phenomenological study. *International Journal of Professional Development, Learners and Learning*, 7(2), e2517.
- Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2). <https://doi.org/10.1111/j.1083-6101.1997.tb00072.x>
- Maggino, F. et al. (2007). *Rilevazione e analisi statistica del dato soggettivo*.
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & education*, 70, 29–40.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record*, 108(6), 1017–1054.
- Nascimento, L., Correia, M. F., & Califf, C. B. (2024). Towards a bright side of technostress in higher education teachers: Identifying several antecedents and outcomes of techno-eustress. *Technology in Society*, 76, 102428.
- Nirchi, S., Mangione, G. R. J., De Vincenzo, C., & Pettenati, M. C. (2025). Indagine esplorativa sulla percezione dei docenti neoassunti circa l'impiego dell'intelligenza artificiale nella didattica: Punti di forza, ostacoli e prospettive. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (30), 151–180.
- Nirchi, S. et al. (2016). Le competenze digitali dei docenti. Un'indagine esplorativa sull'uso delle ICT a scuola. *Formazione & insegnamento*, 14(3), 179–188.
- 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.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, 15(2), 4–14.
- Su, J. (2023). Preservice teachers' technological pedagogical content knowledge development: A bibliometric review. *Frontiers in education*, 7, 1033895.
- Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, 68(1), 1–16.

Copyright (©) Claudio Palazzi



This work is licensed under a Creative Commons Attribution NonCommercial-NoDerivatives 4.0 International License.

How to cite this paper: Palazzi, C. (2026). Tecnologie immersive e responsabilità educativa nella formazione docente: evidenze empiriche dal progetto MuSEd VR [Immersive technologies and educational responsibility in teacher training: empirical evidence from the MuSEd VR project]. *QTimes webmagazine*, anno XVIII, n. 2, 522-543, Doi: <https://doi.org> ...