Asia Pacific Academy of Science Pte. Ltd.

Metaverse 2026, 7(4), m260066.
https://doi.org/10.54517/m260066

Article

AIGC and Metaverse-Enabling Technologies for Art and Design Education: A Source-Bounded Multi-Source Evidence Framework Linking GenAI Learner Data, VR Readiness, and 3D Resources

Kaijie Yu[1], Albina Volkotrubova[2], Liang Wang[3,*]

[1] Department of Preschool Education and Pedagogy, Kyrgyz State University named after I. Arabaev, Bishkek, 720026, Kyrgyzstan.
[2] Department of Linguistics, International University of Kyrgyzstan, Bishkek, 720007, Kyrgyzstan.
[3] School of Art and Design, Qilu University of Technology, Jinan, 250300, China.
*** Corresponding author:** Liang Wang, lxh@qlu.edu.cn

CITATION

Yu K, Volkotrubova A, Wang L. AIGC and Metaverse-Enabling Technologies for Art and Design Education: A Source-Bounded Multi-Source Evidence Framework Linking GenAI Learner Data, VR Readiness, and 3D Resources. Metaverse. 2026; 7(4): m260066. https://doi.org/10.54517/m260066

COPYRIGHT

Copyright © 2026 by author(s). *Metaverse* is published by Asia Pacific Academy of Science Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. https://creativecommons.org/licenses/by/4.0/

Abstract: Generative AI (GenAI) and immersive three-dimensional (3D) technologies are entering art and design education through separate evidence streams. This study proposes a source-bounded, multi-source evidence framework that integrates three publicly available sources at the construct level without merging samples. Learner-level evidence came from an open dataset of 519 university students who all used GenAI; art and design students could not be identified in it. Domain-specific evidence came from published item statistics for 226 Chinese art and design students evaluating VR-supported still-life sketching. Technical evidence came from open 3D-FRONT toolbox metadata (9,990 model records). In adjusted robust regression, Intrinsic Motivation was associated with Learning Engagement ($\beta = .617$, 95% CI .538 to .696), whereas the linear GenAI-frequency term ($\beta = .051$) and AI Learning Support ($\beta = -.064$) were not. A bootstrap decomposition showed a statistical indirect association via Intrinsic Motivation (.086, 95% CI .030 to .145), which does not establish causal mediation. Under nested cross-validation, tree-based models reached $R^2 \approx .42$–$.43$, and SHAP again ranked Intrinsic Motivation first. Art and design students reported positive VR willingness alongside low reported equipment use, a pattern we term a readiness access gap. Because access barriers were not measured directly, this gap is descriptive rather than explanatory. The framework specifies what each evidence layer can and cannot support and yields testable propositions for single-population studies.

Keywords: Generative AI (AIGC); Art and design education; Metaverse-enabling technologies; Learning engagement; Intrinsic motivation; Virtual reality; Explainable machine learning; Three-dimensional scenes

1. Introduction

Generative AI is not just a new tool; it is a part of academic life. Large language models can provide explanations, help to draft, give examples, give feedback, and provide alternative representations of a problem. These capacities not only open up new prospects for students and teachers, but have also led to concerns regarding reliability, academic writing, grading and quality of learning [1-5]. When it comes to the use of GenAI tools in the educational context the question is no longer 'if' students can use GenAI tools, but 'how' and 'when'. The open question is if there is an association between using GenAI and core learning processes, and if these associations are causal, or if they are confounded. Association is not causation.

Learners in higher education institutions feel a variety of emotions towards GenAI, including usefulness, ability to explain, personalisation and support, but

also confused about the absence of clear institutional guidelines, academic integrity, dependence and accuracy [3,6]. According to policy oriented work [7], this should be addressed openly, rather than blanket adoption or prohibition. Now students and teachers need to be able to interact with AI, understand what it can do and how it can be directed with prompts [8-10]. These skills are most relevant to the subject areas in the studio. Art and design education is based on creating artworks, making judgements, critiquing and developing a personal voice. These processes can be supported by a tool that generates ideas, but if they are not organized properly by teaching then this tool can be the one that can replace these processes.

The second strand of research explores the role of AI in the wider digital learning context. The literature on using AI in education recommends adapting the use of the tools to both learning practices and learners' capacities and needs [11–13]. The study on sustainable and adaptive learning also focuses on the control of the learner, feedback, designing tasks and integrating technologies [14, 15]. A meta-analysis of educational chatbots found that they can improve learning outcomes, although the effect varied with chatbot design and context [16]. Motivation is also very important. Research on AI adoption for language learning and student use of AI indicates students' intentions to use AI for the future is not solely determined by the perceived usefulness [17,18]. This helps to consider the use frequency as well as motivation and engagement separately.

A directly related and separate body of evidence – that of metaverse and immersive technologies. Typically, the term educational metaverse is used for a number of technical elements, such as: immersive user interfaces (VR/AR), 3D virtual environments and assets, AI-generated content and persistent, multiplayer social presence through avatars, although sometimes it is used more extensively than it is appropriate [19,20]. In the present study, we will concentrate on the first three building blocks, which we call metaverse-enabling technologies: immersive-interface readiness, re-usable 3D spatial content and AI-assisted learning. This scope does not cover persistence, avatars, multi-user co-presence, or virtual economies, and we do not claim to assess a full metaverse learning environment. Each of these enabling components needs to be evaluated individually, as there is no direct relationship between the technical immersion and the educational value [21–26].

In art and design, virtual environments can support spatial composition, perspective, the appearance of materials in space, and iterative visual exploration. Wen et al. studied VR-supported still-life sketching among Chinese art and design students and found positive attitudes toward its teaching value but low equipment-use rates [27]. Another study of 90 middle-school students reported pre/post changes in cognitive, behavioural, and affective engagement after VR experiences [28]. This is only used as contextual evidence that immersive learning engagement is multi-dimensional; not used as evidence of art and design outcomes.

Both of these parallel literatures share a methodological gap. The attitude, academic integrity or learning outcomes of GenAI use are usually the focus of research, but not the broader teaching ecosystem. Studies in the metaverse and VR are directed towards acceptance or immersion. In Technical 3D, the emphasis is on the representation and rendering. Pooling these sources is not appropriate because they differ in samples, measures, and units. A more defensible approach is to keep source boundaries intact and integrate evidence at the construct level. The same care must be taken with computational analysis: explainable AI is able to detect non-linear predictive patterns, which may not be causal explanations [29–32].

This study therefore proposes a multi-source evidence framework for the digital teaching ecosystem in art and design education and answers four research questions. In the learner data set, what are the relationships between: frequency of the use of GenAI, learning support offered with AI, Intrinsic Motivation and Learning Engagement? Second, will the pattern of the statistical model adjusted for the new data be similar to the explainable machine learning model? Thirdly, what are the signs of the readiness of immersive technology, according to the Chinese source of art and design VR? Fourth, what technical possibilities do the accessible 3D-FRONT resources offer and what can they not support? The study does not make any assumptions about the groups of the GenAI and VR learners and does not assume the GenAI sample is made up of art and design students. Instead, it constructs a hierarchy of sources to a body of evidence, such that each source answers only those questions to which its variables are relevant.

Existing approaches address only part of this problem. The technology-acceptance models [21] describe intention to use a single technology, the frameworks describe the components conceptually in relation to the metaverse in education [19,20] and the policy frameworks and reviews on GenAI use, academic integrity and AI literacy [3,7,8] describe different aspects of the framework. These tend to be either a single platform and/or a single population of the model, or are abstract and conceptual. The present framework differs in three ways: (1) Evidence status is added to the framework; (2) Each construct is coded by source (directly measured, contextual, technical or not measured), so that it can be analyzed consistently with public data; and (3) The present framework results in falsifiable propositions for a single-population study. The contribution is thus a clear way of synthesizing evidence from several sources, and not a tried and tested instructional practice.

2. Materials and methods

2.1. Study design and evidence sources

The study used a secondary, multi-source analytical design (**Figure 1**). The unit of integration was the construct, not the respondent. Source-specific analyses were completed first and then mapped to a common framework comprising three evidence layers and one contextual source: learner evidence, art and design immersive-readiness evidence, contextual VR engagement evidence and technical 3D infrastructure evidence. Independent samples were not merged row by row, and no effect sizes were pooled across sources with different populations or measures; the design is a structured triangulation, rather than a meta-analysis. Constructs that no source measured were explicitly coded as evidence gaps; for example, there was no source that measured engagement with art and design tasks using GenAI or VR.

The main quantitative data source was the publicly available de-identified dataset A Study on the Influence of GenAI Use on Learning Engagement, which is deposited on Mendeley Data with its questionnaire under CC BY 4.0 [33]. The data collection procedure was documented by the data owner, which involved anonymous online questionnaires with a total of 519 complete and valid questionnaires from students in three public universities in Fujian and Henan provinces in China. No sampling period or sampling procedure is reported and no rates of return are available; the data are therefore considered to be a non-probability convenience sample not from which estimates are generalized. Moreover, the documentation indicates that the collection of the data was geared towards the research about the role of Intrinsic Motivation and

AI Learning Support while the present study is a secondary study of data collection with a different purpose, estimation (robust regression and bootstrap indirect associations as non-causal estimates), and scope (integration of the learner data with art and design VR and 3D evidence). The 15 rows of coding notes (removed) were followed by one blank row and the worksheet rows filled with respondents 1–519. No missing values, no reverse coded items or respondents were excluded. There was no variance in the binary adoption variable since all respondents reported using GenAI before. The frequency of GenAI use, however, varied considerably and served as the main measure of GenAI use. Sixty-seven (12.9%) respondents answered all 17 engagement items, and 17 (3.3%) respondents answered all 27 Likert items the same. When the 67 straight-lining respondents were excluded (n = 452), the pattern was unchanged (Intrinsic Motivation β = .560, 95% CI .471 to .650; GenAI frequency β = .052, p = .255; AI Learning Support β = −.035, p = .488).

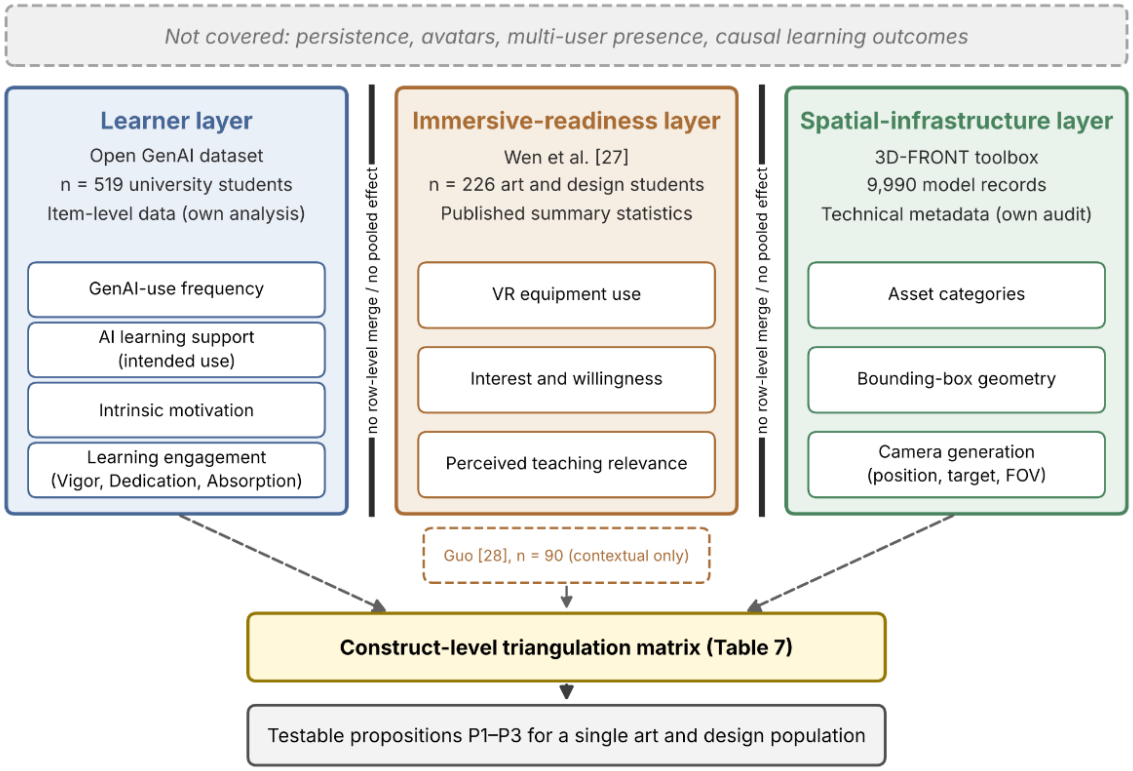

Figure 1. Source-bounded multi-source evidence framework. Solid boundaries indicate that sources are not merged; dashed arrows indicate construct-level triangulation only.

The questionnaire identified the commonly used GenAI tools in China, including DeepSeek, ERNIE Bot, Kimi, ChatGPT, etc., which were used as the definition of the learner constructed. Frequency was rated on a 5 point scale from 1 – rarely (more than 6 months ago) to 5 – very frequently (several times a day). Three engagement dimensions, Vigor (six items), Dedication (five items), and Absorption (six items) were included. AI Learning Support is represented by four items: Classroom question answering, Assignment support, Access to information on a course and AI-based evaluation or feedback, which are all formulated as intentions (I will use generative AI to…), and thus represent intended use of generative AI in a course, rather than the type of support that was sought or received. Intrinsic Motivation had six items on professional growth, interest, goal clarity, self-development and curiosity. Control

variables included were gender, age, grade, and general field of study. As the field variable only identifies Natural Sciences as opposed to Humanities and Social Sciences, it is not possible to identify art and design students. The GenAI dataset is therefore used instead as a general learner-process layer therefore offering measurable constructs and testable propositions but not as properties of art and design students.

The domain-specific source was the study by Wen et al. [27], which surveyed 226 first- to third-year Chinese university students of art and design about VR-assisted still-life sketching. The article reports means and standard deviations for ten items; respondent-level data were not available, so these statistics were used only for description and triangulation. Guo [28] reported a study on the engagement of 90 middle school mathematics students in Ganzhou, China, and reported the pre/post engagement means. It was only kept to demonstrate that the engagement in immersive learning may be multidimensional and the outcomes did not provide evidence for the effectiveness of art and design.

The publicly available 3D-FRONT-ToolBox repository, accompanying the datasets 3D-FRONT and 3D-FUTURE [34,35] was used for the technical layer. A toolbox README is provided with the description of scene JSON as input and camera JSON as output. It reads the layout, floor and furniture, and creates camera position, target and field of view parameters. There were 9,990 model-information records and 16,562 category-mapping rows in the metadata. Such resources were not viewed as evidence of Learner Data, but rather evidence of spatial content infrastructure. The 3D-FRONT-FUTURE project website was accessed for information only, and the entire 3D-FRONT scene corpus was not studied, so room level educational simulation was not claimed. Both repositories do not specify a license; only files which are openly available in the repositories were used, no files were redistributed, and no access to the full 3D-FRONT scene corpus, which is distributed separately by its creators, was made.

The sources were chosen according to the following criteria: (i) public availability: the sources should be available to the public so that they can be reproduced for further analysis; (ii) documented measurement: the sources should have a form of measurement, e.g. a questionnaire, published item statistics, or technical documentation; (iii) coverage of a distinct ecosystem layer: the sources should cover a different ecosystem layer, e.g. learner processes during the use of GenAI, art and design readiness for immersive VR, or reusable 3D spatial content; and (iv) relevance to art and design education either directly [27] or through a function in the studio, e.g. scene composition and viewpoint generation. All results cited as 'our analysis' are based on calculations done in this study with the GenAI dataset and the metadata for 3D-FRONT; the numbers reported by Wen et al. [27] and Guo [28] are summary statistics and were not recalculated. The features, breakdowns, and evidence roles are summarized in **Table 1**.

Table 1. Evidence sources and permitted analytical roles

Source	Analytical unit	Variables or evidence	Permitted role
GenAI dataset [33]	519 valid respondents	GenAI frequency; Vigor; Dedication; Absorption; AI Learning Support; Intrinsic Motivation; demographics	Primary learner-level quantitative evidence
GenAI questionnaire [33]	Instrument documentation	Item wording and coding	Construct-definition source

Continuation Table:

Source	Analytical unit	Variables or evidence	Permitted role
Chinese art and design VR study [27]	226 art and design students	Ten reported VR willingness and teaching-relevance items	Domain-specific descriptive triangulation
Ganzhou immersive-VR study [28]	90 mathematics students	Pre/post cognitive, behavioral, affective, overall engagement means	Contextual engagement evidence only
3D-FRONT resources [34,35]	9,990 model records; 16,562 mapping rows	Category metadata; bounding boxes; scene/camera processing	Technical enabling evidence; not learning outcomes

2.2. Variable construction and measurement checks

Prior to analysis, the scale scores were checked against the scale items. The Vigor, Dedication, and Absorption scores were the arithmetic average of their items, and Overall Learning Engagement was the arithmetic average of the three subscale scores (M = 3.23; the unweighted 17 item mean was 3.22). The AI Learning Support and Intrinsic Motivation scores were likewise reproduced from their items. The precomputed GenAI use score was the same as GenAI frequency and was therefore not entered as a separate variable. This avoided entering redundant information into the association and predictive models.

Descriptive analysis included frequency distributions for categorical variables, and means, standard deviations, medians, and distribution checks for continuous scores. Internal consistency of the scales was evaluated with Cronbach's α and McDonald's ω. Confirmatory factor analysis was then used as a measurement diagnostic. Vigor, Dedication, and Absorption were represented by a three-factor engagement model, and AI Learning Support and Intrinsic Motivation were added in a five-factor model. Model assessment included the use of the chi-square, CFI, TLI, RMSEA and SRMR. Additionally, average variance extracted and standardized loading ranges were calculated. The discriminant validity of the engagement dimensions was analyzed by using the HTMT, the Fornell–Larcker criterion, and the one factor comparison model. Common method variance was assessed with Harman's single-factor test.

CFA was used as a measurement diagnostic rather than as a test of a new latent structure: its purpose was to verify that the published items functioned as coherent, distinguishable subscales in this sample before composite scores were entered into subsequent models. Because the items were measured on five-point Likert scales, the diagnostic models were estimated with maximum likelihood on Pearson correlations; an ordinal estimator such as WLSMV is recommended for replication. The primary outcome used was Overall Learning Engagement while the secondary outcomes were Vigor, Dedication and Absorption. This choice reflects the fact that the three dimensions are highly correlated, and it reduces the risk of overinterpreting subtle differences among them.

Standardized continuous predictors and dummy coding of categorical controls were performed. Overall Learning Engagement was regressed against the following variables in the principal model: GenAI frequency, AI Learning Support, Intrinsic Motivation, gender, age, grade, and broad field of study. HC3 robust standard errors were used for all ordinary least squares (OLS) estimates. The same specification was used for Vigor, Dedication, Absorption and the Benjamini–Hochberg false discovery

rate correction was applied to the family of dimension-specific tests. The frequency was measured in five ordered categories, so it was entered both as a continuous and as a categorical variable. Results are reported with effect sizes and confidence intervals, not p values alone.

2.3. Indirect associations, predictive robustness, and explainability

A parallel path decomposition was conducted to determine if the cross-sectional relationship between GenAI frequency and Learning Engagement was statistically decomposable into shared and unique components with AI Learning Support and Intrinsic Motivation. This was done using 5,000 bootstrap resamples. The procedure is not called causal mediation, rather it is called indirect-association analysis. An indirect coefficient does not imply a temporal precedence, as cross-sectional data cannot provide this information, meaning that an indirect coefficient does not imply that use of GenAI leads to a change in a mediator which in turn leads to a change in engagement.

A second analytical track evaluated the robustness of predictions. GenAI frequency, AI Learning Support, Intrinsic Motivation, and dummy-coded gender, age, grade, and field (12 features) were used to predict Learning Engagement with Elastic Net, Random Forest, XGBoost and LightGBM. Nested cross validation was used to estimate performance: the outer loop was five-fold cross validation repeated five times (25 test folds; seed 42) and within each outer training set, hyperparameters were determined by a five-fold inner cross validation which minimized RMSE (seed 7) with test folds never used for hyperparameter tuning. Search spaces were: Elastic Net (standardized features), l1_ratio ∈ {0.1, 0.3, 0.5, 0.7, 0.9, 1.0} with a 100-value α path; Random Forest (300 trees), max_depth ∈ {None, 4, 8}, min_samples_leaf ∈ {1, 5, 10}, max_features ∈ {1.0, 0.5, √p}; XGBoost and LightGBM (300 trees, subsample and colsample = 0.8), max_depth ∈ {2, 3, 4} or num_leaves ∈ {4, 8, 15}, learning rate ∈ {0.03, 0.1}, and min_child_weight ∈ {1, 5} or min_child_samples ∈ {10, 30}. The outer folds were summarized using R^2, RMSE and MAE. The goal was to verify if the primary statistical pattern remained valid under the nonlinear prediction and not to be able to rank the algorithms.

The hyperparameters of the Random Forest used to compute SHAP values were optimized on the full sample using an inner five-fold cross-validation (with seed 7) and the number of trees was set to 300; the selected settings were max_depth = None, min_samples_leaf = 5, and max_features = 0.5. This model was used as a representative tree-based model, rather than a 'best' model. SHAP was used to explain the impact of the values of observed features on the model's predictions, but was not employed to give causal importance. This separation is similar to the recent XAI guidance which distinguishes between post hoc explanation and mechanism identification [29,30] and emphasises communicating uncertainty and dependence on the fitted model [29–32]. SHAP magnitudes are thus reported together with regression estimates and cross-validated performance and convergence is considered supportive while divergence would have been reported, but not resolved by selecting the better method.

In the exploratory heterogeneity tests, interactions of GenAI frequency with gender, age, grade and field of study were added, also with FDR correction. Interpretations were carried out with caution for the small categories, especially for the youngest and oldest age categories. Since the field variable did not isolate the art and design students, there was no specification for art and design students.

The 3D technical audit counted the following: model-information records,

number of category mapping rows, unique detailed categories, unique super-categories, joined bounding-box records, and median x/y/z dimensions. Rows that did not have a detailed category label were not counted as a category, but instead were mapped as rows. The distributions of the bounding box dimensions were shown on a scatterplot with a logarithmic scale, and the distributions of the categories were shown separately. These measures describe asset composition and geometric variation; they are not markers of immersion, presence, creativity, achievement or learning.

The sources were then integrated in a convergent evidence matrix. An educational construct which is directly measured in a source was coded as direct, one from another educational population as contextual, and one which describes infrastructure, not a learning outcome, as technical. There was no source used for a question that it did not measure, and the coding linked each interpretation to the framework. If a construct was measured by just one source, that source is represented in the matrix as single source evidence and not as convergent evidence.

3. Results

3.1. Learner profile and measurement quality

There were 519 valid respondents in the GenAI sample. All participants had previously used GenAI (adoption), but frequency of use varied among respondents ($M = 3.58$, $SD = 0.93$ on a five category scale) with 3.1% using GenAI rarely, 6.7% occasionally, 34.9% regularly, 39.9% frequently, and 15.4% very frequently. Of the sample, 72.1% were female and most respondents were 19–21 years old (77.1%). By grade, 18.9% were freshmen, 34.5% sophomores, 27.0% juniors, 8.5% seniors, and 11.2% postgraduates. The percentages were 82.5% for Natural Sciences and 17.5% for Humanities and Social Sciences. These are sample proportions, not population proportions.

Internal consistency of the learner scales was high. Vigor had $\alpha = .891$ and $\omega = .893$. Dedication had $\alpha = .905$ and $\omega = .906$. Absorption had $\alpha = .904$ and $\omega = .905$. AI Learning Support had $\alpha = .830$ and $\omega = .835$, and Intrinsic Motivation had $\alpha = .886$ and $\omega = .889$. The 17-item overall engagement score had $\alpha = .954$ and $\omega = .955$. The score distributions are displayed in **Figure 2**. Each construct showed sufficient variance for association analysis.

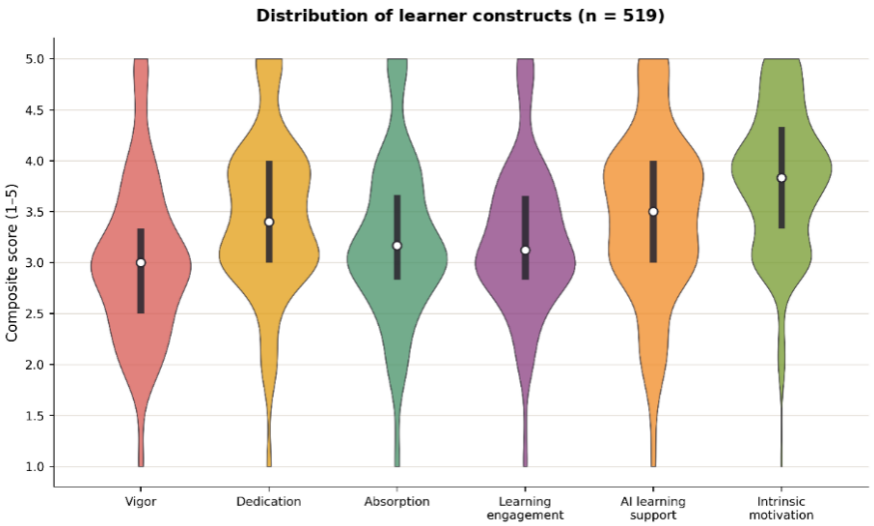

Figure 2. Distribution of learner constructs in the GenAI sample (n = 519). Violin widths show empirical density; thick vertical segments show interquartile ranges and white points show medians.

The diagnostic factor models showed mixed fit. The three-factor engagement model produced $\chi^2(116) = 512.85$, CFI = .939, TLI = .929, SRMR = .043, and RMSEA = .081 (90% CI .074–.089); SRMR indicated good fit, CFI and TLI were close to .95, and RMSEA was in the mediocre range, probably reflecting residual covariance among similarly worded items and ML estimation on ordinal items. No post-hoc error covariances were added. A one-factor model fitted clearly worse (CFI = .883, RMSEA = .111; $\Delta\chi^2(3) = 368.11$, $p < .001$), so the dimensions are statistically distinguishable. The five-factor model produced CFI = .924, TLI = .915, RMSEA = .067, and SRMR = .058. Composite reliability ranged from .835 to .906, AVE from .561 to .656, and standardized loadings from .612 to .885. Discriminant validity among engagement dimensions was nonetheless limited: latent correlations were .856–.890, squared correlations (.733–.792) exceeded the AVEs (.583–.656), failing the Fornell–Larcker criterion, and HTMT values (.858–.877) fell between the .85 and .90 thresholds. Overall Learning Engagement was therefore the primary outcome, and dimension-level differences are not interpreted as substantive.

Table 2. Measurement diagnostics for the primary learner constructs

Construct	Mean	SD	Cronbach α	McDonald ω	CR	AVE
Vigor	3.01	0.79	.891	.893	.893	.583
Dedication	3.45	0.79	.905	.906	.904	.656
Absorption	3.22	0.81	.904	.905	.906	.616
AI Learning Support	3.47	0.83	.830	.835	.835	.561
Intrinsic Motivation	3.82	0.71	.886	.889	.890	.575
Learning Engagement (17 items)	3.23	0.73	.954	.955	—	—

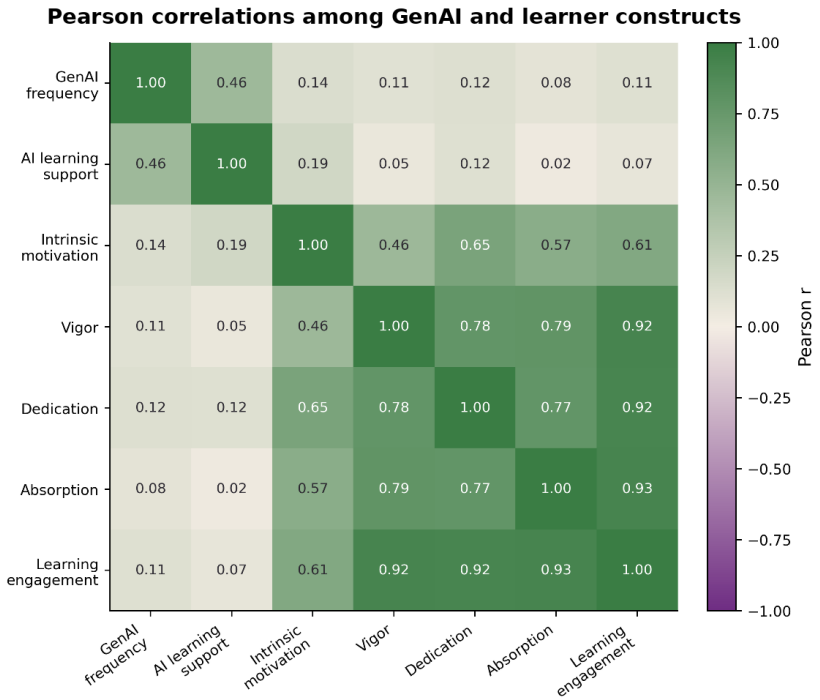

Figure 3. Pearson correlations among observed GenAI and learner constructs. Values describe co-variation only and are not causal effects.

Table 2 and **Figure 3** summarize the correlation structure. The frequency and AI Learning Support were moderately correlated (r = .461), while Learning Engagement and Intrinsic Motivation were weakly correlated with GenAI (r = .111, r = .136, respectively). Intrinsic Motivation was significantly correlated with overall engagement (r = .607), Absorption (r = .567) and Dedication (r = .650). A complete correlation matrix is presented in **Figure 3**. These are bivariate relationships which are not directional or mechanistic. **Table 2** gives an overview of the measurement diagnostics for the main constructs of the learner.

3.2. Adjusted associations and indirect associations

GenAI frequency, AI Learning Support, Intrinsic Motivation, gender, age, grade and broad field of study were the variables included in this adjusted model, which explained 39.8% of the variance in Learning Engagement (adjusted R^2 = .383). The clearest adjusted association was for Intrinsic Motivation (β = .617, 95% robust CI .538 to .696, p < .001). GenAI frequency had β = .051 (95% CI −.029 to .131, p = .208), and AI Learning Support had β = −.064 (95% CI −.153 to .025, p = .160). Neither GenAI frequency nor AI Learning Support was found to have an independent association at the .05 level after adjustment.

This trend was repeated for each measure of engagement. The results show an association between Intrinsic Motivation and Vigor (β = .470, p < .05, FDR), Dedication (β = .653, p < .05, FDR), and Absorption (β = .587, p < .05, FDR) which remained significant after FDR correction. There was no dimension that was related to GenAI frequency. A negative coefficient (−.100) was observed for AI Learning Support, but was interpreted as exploratory and was not robust. The standardized coefficients and robust 95% confidence intervals are shown in **Figure 4**.



Figure 4. Robust standardized coefficients and 95% confidence intervals for overall Learning Engagement and its three dimensions. The vertical dashed line marks zero.

A more detailed picture was obtained from the bootstrap decomposition. There was an indirect path from GenAI to Learning Engagement through Intrinsic Motivation that was .086 (95% Bootstrap interval: .030, .145; Bootstrap p = .003). The estimate via AI Learning Support was −.029 (−.072 to .010, p = .152). The direct association was .051 (−.029 to .130), the total association was .109 (.009 to .202), and the total indirect association was .057 (.000 to .094). This is because some of the variance between GenAI frequency and engagement overlaps with that of Intrinsic Motivation; this does not mean that it is causal or in the temporal order that is presented, and other orders are explored in Section 4.

These estimates were subject to three sensitivity checks. First, categorical coding was a better fit than the linear term ($F(3, 503) = 5.66$, $p < .001$): very frequent users (n = 80) had higher engagement scores than the other four groups, which did not differ from one another (all $p > .13$). This pattern is a post hoc finding and the null linear estimate might mask a non-monotonic association, so it must be confirmed. Second, one of the items from the intrinsic-motivation (IM2, '…make me concentrate and forget everything') scales was semantically related to an Absorption item, and this item was dropped, but the association remained similar (β = .597, 95% CI .516 to .678). Third, in Harman's single-factor test the first unrotated factor accounted for only 43.5% of item variance, suggesting that common method variance is unlikely to explain the results fully, although it cannot be ruled out with single-source self-report data. **Table 3** summarizes the main adjusted and bootstrap estimates.

Table 3. Main adjusted and bootstrap association results

Analysis	Predictor or path	Estimate	95% interval	p
Adjusted overall model	GenAI frequency	β = .051	−.029 to .131	p = .208
Adjusted overall model	AI Learning Support	β = −.064	−.153 to .025	p = .160
Adjusted overall model	Intrinsic Motivation	β = .617	.538 to .696	p < .001
Bootstrap indirect	Via AI Learning Support	−.029	−.072 to .010	p = .152
Bootstrap indirect	Via Intrinsic Motivation	.086	.030 to .145	p = .003
Bootstrap direct	GenAI frequency – engagement	.051	−.029 to .130	p = .220

3.3. Predictive robustness and explainable modeling

Nested repeated cross-validation was used to check whether the explanatory pattern depended on the linear specification. Elastic Net reached a mean R^2 of .353 (RMSE = .584, MAE = .457). Random Forest reached .416 (.553, .425), XGBoost .423 (.550, .423), and LightGBM .427 (.548, .422). The tree-based models performed similarly, and their fold-to-fold SDs (≈ .08) were far larger than the differences between them, so no model is considered superior; all modestly exceeded Elastic Net.

Two baselines clarify the source of the predictive signal. A linear model with Intrinsic Motivation alone reached a cross-validated R^2 of .355, whereas a model with all features except Intrinsic Motivation did not predict held-out data ($R^2 = -.026$), and GenAI frequency alone gave $R^2 = -.008$. In the SHAP summary, Intrinsic Motivation had the largest mean absolute SHAP value (.302), followed distantly by GenAI

frequency (.056); the sophomore indicator (.041), AI Learning Support (.032), and the remaining demographic indicators (≤ .024) contributed little. **Figure 5** presents the nested cross-validated performance and **Figure 6** the SHAP distribution. The SHAP ranking agrees with the regression in identifying Intrinsic Motivation as the leading correlate.

No stable evidence following FDR correction for exploratory interaction tests. There was no significant interaction between GenAI frequency and the HSS indicator ($p = .098$, $q = .197$) and there was no interaction between gender and GenAI frequency ($p = .465$). Prior to correction, the age omnibus interaction was significant at $p = .020$, but after the correction, it was $q = .079$, and no grade interactions were found. Given the small sizes of some age groups, age differences in the GenAI–engagement association cannot be considered reliable.

Nested cross-validated predictive performance is summarized in **Table 4**.

Table 4. Nested repeated five-fold cross-validated predictive performance

Model	R^2 mean ± SD	RMSE mean ± SD	MAE mean ± SD
Elastic Net	.353 ± .060	.584 ± .048	.457 ± .036
Random Forest	.416 ± .086	.553 ± .040	.425 ± .028
XGBoost	.423 ± .081	.550 ± .042	.423 ± .028
LightGBM	.427 ± .083	.548 ± .045	.422 ± .030

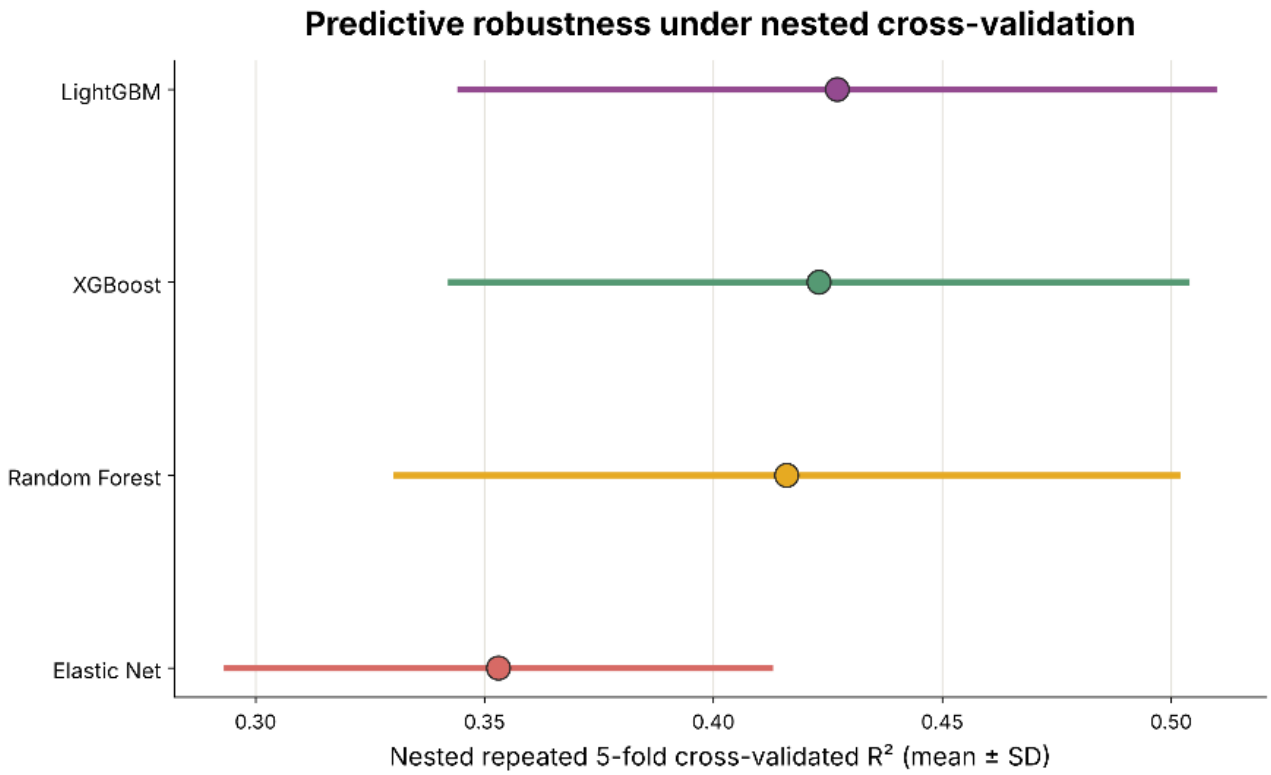

Figure 5. Nested repeated five-fold cross-validated R^2 values (mean ± SD across 25 outer folds) for the predictive models. The models are robustness checks, not causal estimators.

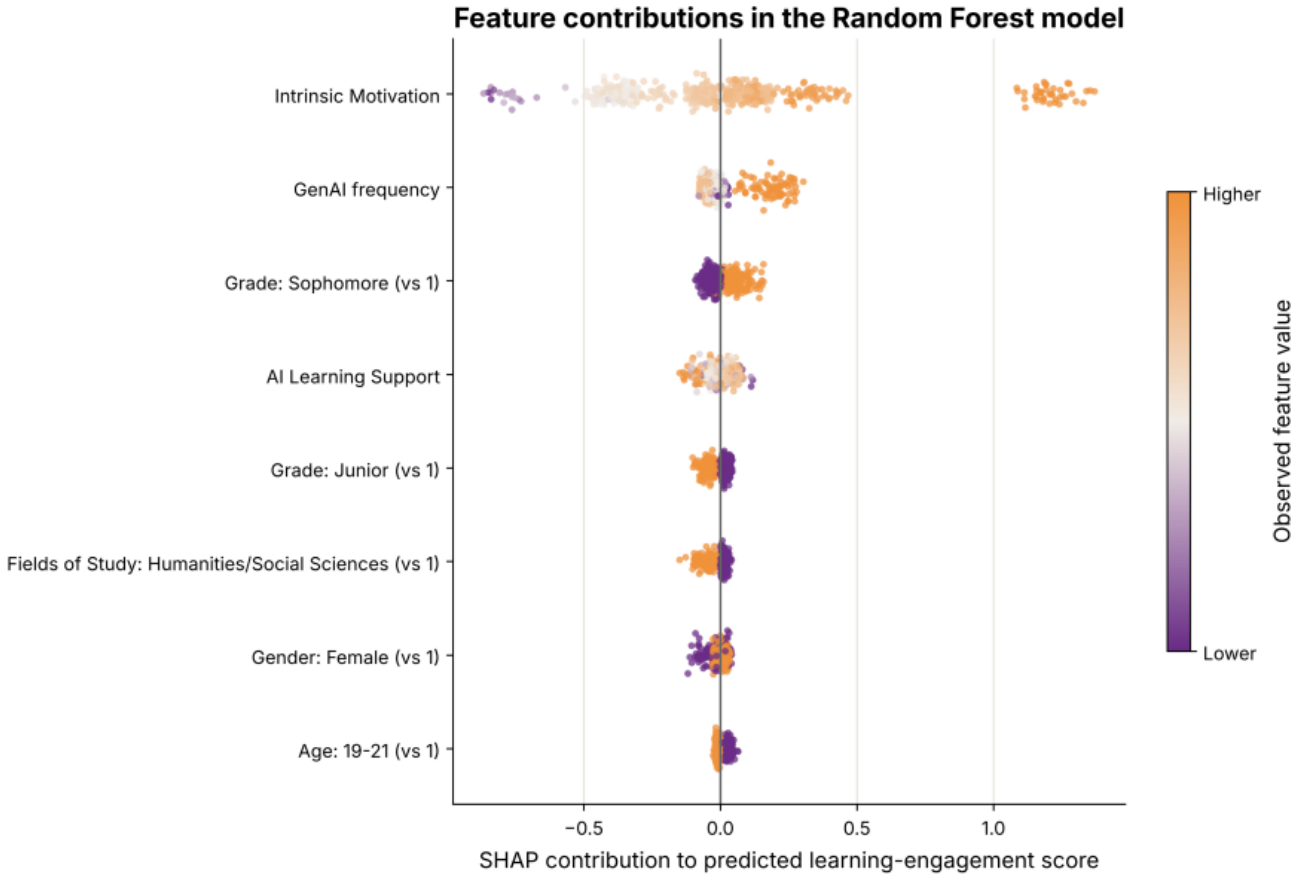

Figure 6. SHAP contributions in the full-sample Random Forest model. Horizontal position shows contribution to the predicted Learning Engagement score; point color represents lower-to-higher observed feature values.

3.4. Art and design VR evidence, 3D infrastructure, and cross-source integration

The art and design VR source provided the domain-specific layer that the GenAI dataset could not. Wen et al. [27] reported that VR equipment-use frequency had the lowest mean of the ten items (1.72, SD = .993). Interest in VR technology was 3.73, acceptance of VR in art and design teaching 3.96, perceived applicability of VR to teaching 3.89, perceived importance of experiencing VR 3.87, willingness to share VR experience 4.13, and perceived impact on art and design majors 4.03. Attitudes toward VR were thus relatively positive, whereas reported equipment use was low. Because only means and standard deviations were reported, item distributions and inter-item correlations could not be examined in this source. The ten item-level means and standard deviations are summarized in **Table 5** and visualized in **Figure 7**.

Table 5. Reported VR item statistics in the Chinese art and design sample (n = 226)

Item	Mean	SD
VR equipment use frequency	1.72	.993
Interest in VR technology	3.73	.931
Accept VR in art and design teaching	3.96	.903

Continuation Table:

Item	Mean	SD
VR applicable to teaching	3.89	.828
VR experience is important	3.87	.869
Willing to share VR experience	4.13	.752
Like VR after learning about it	3.94	.817
VR may become a university teaching mode	3.91	.812
VR may change life/social/work	3.91	.778
VR may impact art and design majors	4.03	.771

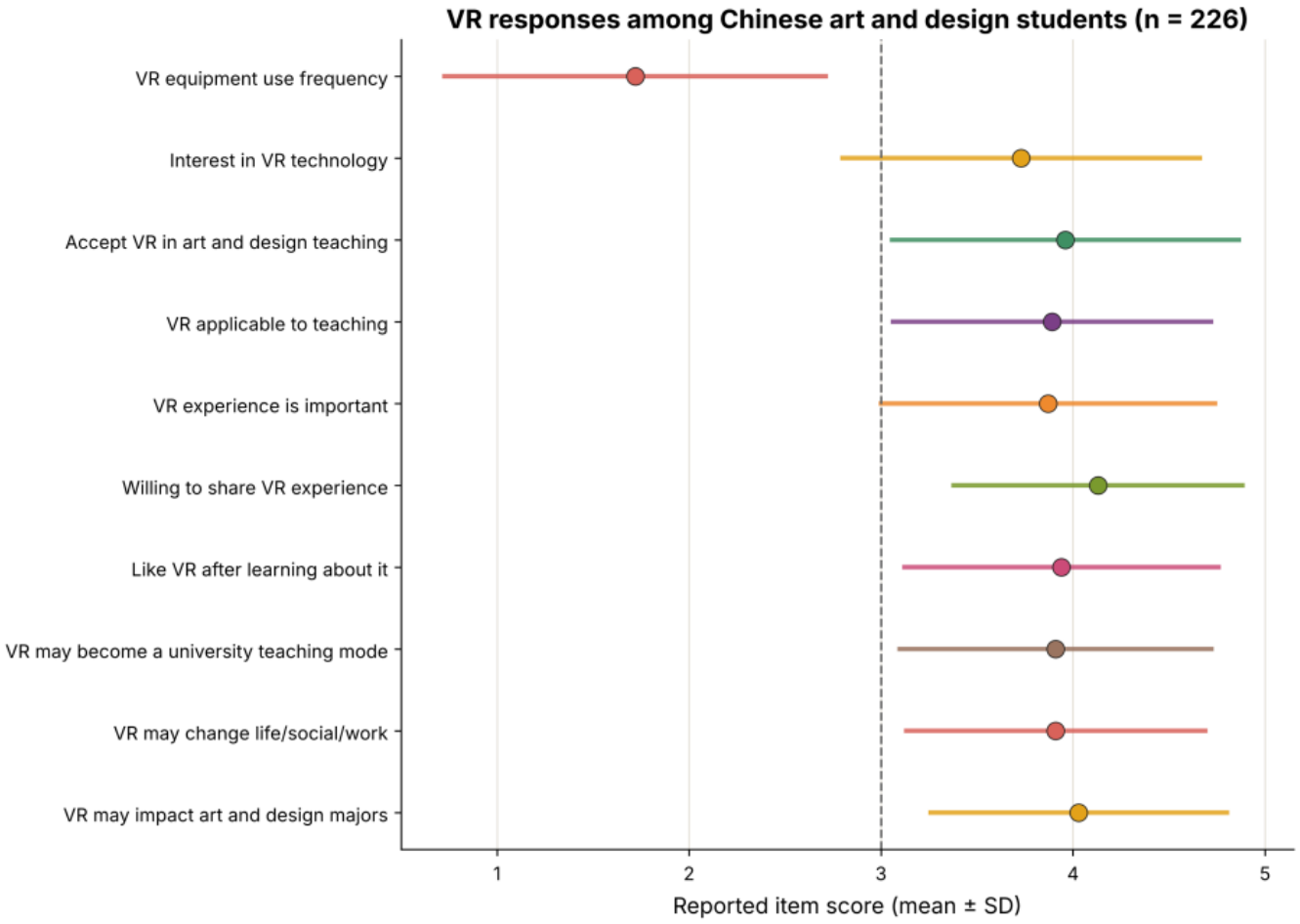

Figure 7. Reported means ± SD for ten VR items in the Chinese art and design study (n = 226). The dashed line at 3 is a visual reference only.

When we examined the 3D metadata (**Table 6**), we saw that there were 9,990 records in the model-information file. There were 16,562 rows in the category-mapping file, and 49 detailed categories with labels, one of which was blank, and 8 super-categories, one of which was 'Others'. A join returned 9,989 records containing bounding-box details, 34 detailed categories and 7 super-categories. The median x/y/z dimensions were 110.9/67.5/81.4 in the metadata's native units (which are not described in the repository but are consistent with centimeters). The x and z

dimensions of the mapped assets were generally quite variable, as is illustrated in **Figure 8**, which is a log-scale plot of the groups of cabinet/shelf/desk, sofa, lighting, chair, bed, table, and pier/stool. The assets therefore span many interior furniture types and vary widely in size. They show no evidence of learner presence, engagement, creativity or attainment.

The quantified 3D metadata inventory is summarized in **Table 6**, while spatial variation is illustrated in **Figure 8**.

Table 6. Quantified inventory from the publicly available 3D toolbox metadata

Metric	Value	Interpretation
model_info records	9,990	3D model metadata with category ID and bounding-box dimensions
Category-mapping rows	16,562	Semantic model-to-category mappings
Unique detailed categories	49 (34 joined)	Labeled detailed categories; 1,751 rows unlabeled
Unique super-categories	8 (7 joined)	Higher-level semantic groups, including 'Others'
Joined bounding-box records	9,989	Records available for geometric analysis
Median x/y/z dimensions	110.9 / 67.5 / 81.4	Native metadata units (consistent with centimeters)

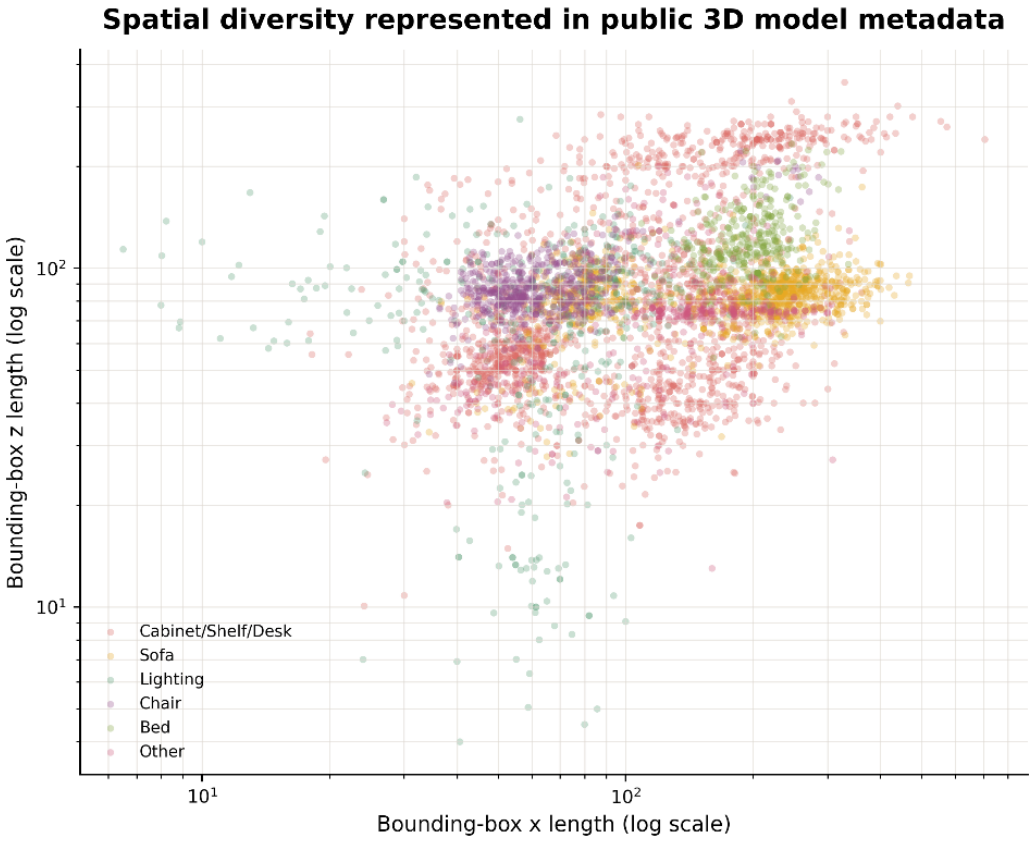

Figure 8. Bounding-box x-length versus z-length for a readability sample of joined 3D model records on logarithmic axes. The plot describes spatial diversity in the technical asset metadata.

According to the toolbox documentation, the pipeline is intended to extract the initial camera parameters (position, target, field of view) for the camera from the scene JSON (layouts, floors, furniture) and output them to a file – these are coarse initial solutions and may need to be adjusted. The lighting category had 968 of the joined records. The toolbox by itself offers no immersive or VR interaction.

Table 7 is a combination of the three main layers and one contextual source. Direct measures of GenAI-use frequency, intended AI-assisted learning support, Intrinsic Motivation and engagement are included in the GenAI learner layer. For the art and design VR layer direct measures of willingness and perceived teaching relevance are incorporated. The 3D layer provides direct technical information of asset categories, geometry and camera processing. The layers are complementary; they are not combined into one dataset. Convergence is examined only where constructs partly overlap: the strongest learner correlate was Intrinsic Motivation; willingness to share VR experience was the highest-rated VR item, but the two are not treated as a pooled estimate.

Table 7. Construct-level triangulation matrix

Ecosystem dimension	GenAI learner layer (own analysis)	Art and design VR layer (published statistics [27])	3D technical layer (own audit of public metadata)	Defensible interpretation
GenAI-use frequency	Direct: GenAI frequency in 519 learners	Not measured	Not measured	General learner-level GenAI evidence (recency-anchored)
Intended AI-assisted use	Direct: LS1–LS4	Not measured	Not measured	Intended use, not support received
Learner engagement	Direct: Vigor, Dedication, Absorption	Acceptance/interest only	Not measured	Do not treat VR acceptance as an engagement outcome
Intrinsic Motivation	Direct: IM1–IM6	Interest/willingness only	Not measured	Strongest learner-level correlate of engagement
Art and design specificity	Not isolated in GenAI sample	Direct: Chinese art and design students	Indoor furniture assets for interior and spatial design	Art and design claims anchored to VR source
Immersive/VR readiness	Not measured	Direct descriptive VR items	Technical spatial support	Supports readiness/ affordance, not causal effectiveness
Spatial digital environment	Not measured	VR experience context	Direct metadata and toolbox	Technical enabling layer only

4. Discussion

This study has been concerned with one such typical issue that is prevalent in the field of educational technology research: the presence of relevant evidence, but its distribution in a variety of sources. The use of GenAI is typically reported at the student level and VR evidence in art and design is typically reported as acceptance or readiness. 3D resources can explain the contents of what can be created and rendered, but not the contents of what students learn. It would be easier to tell the

story if it were assumed that these sources were interchangeable, but that would be indefensible. The primary contribution of this paper is a systematic way of combining complementary evidence while staying within each evidence source and not pooling effect estimates.

Intrinsic Motivation was the most consistent correlate of engagement for the GenAI learner layer in both adjusted and dimension-level analyses, and in the SHAP analysis. It is in line with the need to consider learner agency and critical competence in addition to access to AI [8,11–18] and also in line with research reporting that the students' use of AI was more beneficial when they were active evaluators of the tool, rather than passive recipients [1,3,9]. There is no evidence in the data that GenAI raises motivation. There are a number of different ways these data might be explained equally well, such as: more motivated students might use GenAI more often; engagement might bolster motivation; there could be a reciprocal influence between engagement and motivation; and there could be an influence of a third variable (workload) on both engagement and motivation. These can only be separated by longitudinal designs (or by experimental designs). Additionally, part of the association could be due to conceptual overlap between two self-reported constructs, but not because of the similar wording of items in the two scales (as suggested by the IM2 sensitivity analysis).

When all the other factors are accounted for, there is not much correlation between the frequency of student use of GenAI tools and engagement. The correlation between the two variables ($r = .111$) indicates that about 1.2% of the variance is common between the two variables, that each increment on the frequency scale was associated with 0.09 engagement points before adjusting for each other, and 0.04 engagement points after adjusting for each other, and that frequency alone did not predict held-out data (cross-validated $R^2 \approx 0$). While less frequent users may be using GenAI for reflection or feedback, frequent users may be using GenAI for convenience, availability or an academic need or habit, such as writing routine text or verifying facts. Task design, policy, and critical evaluation are factors suggested by previous studies as creating educational value [4–10]. It should therefore be emphasized that adoption or usage counts is not a measure of the quality of the education offered by a particular institution. This is by no means inconsistent with the exploratory discovery that the heaviest users reported higher engagement, which was post hoc and cannot be interpreted in terms of direction.

Caution is advised with the AI Learning Support scale as well. The items include items that you would use GenAI to answer questions, complete assignments, get information, or evaluation and feedback, rather than what kind of support you received. The scale had good internal consistency, and it was related to the frequency with which it was used, but not specifically to overall engagement. The predictors are overlapping and the goal can vary significantly from student to student, as well as being too general for a four item scale. There is nothing in these data that suggests that AI support has a negative impact; it only means that in this model, which controlled for all other factors, no statistically significant association between AI support and engagement was observed (the adjusted coefficient for AI Learning Support was negative but statistically nonsignificant).

Evidence in the form of art and design is another type of evidence. Published art and design sample [27] received relatively high ratings for willingness and the perceived relevance, while relatively low rates of equipment use were reported. This discrepancy is referred to as a readiness–access gap. The word is descriptive,

not explanatory—high reported willingness, low reported equipment use, there are other things that might explain that, but they were not measured in the sources. Some barriers could be accessibility, curriculum integration, staffing, and quality of VR tasks, and adoption research indicates that intention to use immersive environments is dependent on conditions of the implementation, ease of use and usefulness of VR [19–26]. Other issues in art and design may include the nature of the image, aspects of space control, the process of critique and transition between the virtual and real making. No scores of achievement or design-quality ratings were reported in the source and this manuscript does not claim that VR is beneficial for art and design learning. Likewise, the results of the pre/post engagement with the material in [28] are not applied to students in art and design. The term ‘metaverse-enabling' is defined in this paper, and clarified by the 3D technical layer. The toolbox includes different semi-structured representations of layouts, floors, furniture layout and geometry, as well as camera parameters to which a spatial scene can be created. The properties are important to the studio activities in four ways. The exercises where students select, arrange and combine objects are supported by semantic categories to help with the composition. Second, size of bounding boxes is a means for practicing scale and proportion. Thirdly, viewpoint and perspective studies (which are related to the still-life sketching in [27]) can be assisted by camera generation, and perhaps the improvement or generation of the initial cameras can be a learning activity. Fourth, for the 968 lighting assets, they could be used for lighting exercises post-render.

These are "affordances" and not "learning effects". A technical asset library can't be a learning environment, it requires task design, feedback, interaction design and assessment. This is in line with the debate on the metaverse and its application in education, where the importance of a clear definition and empirical facts to support the positive impact of the immersive element in education is highlighted [19–26]. It also separates the framework from the general models of educational technology, which rarely include information about the evidence and integration of spatial content, immersive access and AI support prior to evaluation of a metaverse learning space.

The results show there are three considerations that need to be taken into account in the design, but these are not designed to be effective, merely suggestions. First, the highest correlation with engagement was Intrinsic Motivation, meaning that the use of GenAI in studio courses might be related to personally meaningful activities, such as exploring design options to solve the students' design problem; rather than engagement in the use of GenAI. Second, it is more instructive to have evidence of the process (prompt logs, feedback on AI-generated content, revision histories) and expert evaluation of the work than it is to have usage counts. Third, students may need to evaluate, reject and/or improve ideas generated by AI tools that are suggested by the critical AI literacy [8,9]. Theoretically, the predictive models served as robustness checks, since they were consistent across the analytic approaches used, the top model was a motivational construct akin to engagement, and they were similar to regression in that they did not represent a process, but rather described the data.

These conclusions have a number of limitations. First, the data are cross sectional and the learners' data are self-reports, which means that it is not possible to find the order of the events. Second, the broad field-of-study variable does not differentiate between art and design students, and the structure of the engagement with the motivations is not clear for these students. The GenAI data set is also collected from three universities in two Chinese provinces, but lacks information on

time period of data collection, or sampling frame, thus generalizability and selection bias cannot be evaluated. Third, the GenAI items didn't reference image generation tools as being used in the art and design practice, and there was no evidence of visual AIGC in the studio to anchor the frequency item; it was anchored by the learners' evidence that was general GenAI evidence. Fourth, the art and design VR source gives only item-level summary statistics, and offers only descriptive triangulation. Fifth, the study in Ganzhou is small and only focuses on mathematics. Sixth, diagnostic CFA should be based on 5-point, both maximum likelihood on Pearson correlations, and on ordinal estimation (e.g., WLSMV) should be used in replication. Last, the technical analysis was carried out for the toolbox metadata, and not the 3D-FRONT scene corpus.

The framework must now be piloted with a single population of art and design studio students, identifying the amount of GenAI usage, the motivation for using GenAI, a 3D task, repeated measures, expert ratings of composition, spatial reasoning, originality and task completion, as well as process measures (prompts, edits, camera movement, time on task, and outputs). At this stage it is not an evidence-based teaching system, rather a framework of evidence.

We propose an innovative application of the framework – a 4-week second year studio unit called composing and rendering an interior still-life scene. The frequency scale and AI Learning Support, motivation, and engagement scales are filled before and after Phase 1 when students create and reflect on their thoughts for the composition using GenAI and reflect on their prompts. In Phase 2 students watch a staged scene from multiple viewpoints in VR, and draw perspective drawings as they view it and collect readiness items [27] and equipment-access logs. Students will use the 3D-FRONT assets to place a room during Phase 3, and will use the category metadata and the bounding-box dimensions to ensure scale and proportion, and optimize the toolbox generated cameras. The use of a conventional technique on a comparison section would enable comparisons to be made based on blind expert ratings. In the design of this study there are three propositions that are tested: P1, the strength of the relationship between Intrinsic Motivation and engagement will be higher compared to the strength of the relationship between the frequency of GenAI and engagement; P2, the gap between willingness and actual use will be narrowed when access to VR is provided; and P3, the use of 3D-asset support will influence process measures (viewpoint exploration), where the influence will be assessed on design quality. The institutions are not required to have all of GenAI, VR and 3D assets on a single platform before they can take part in such a “bounded and auditable” assessment.”

5. Conclusion

The primary methodological contribution of this study is a source-bounded framework integrating learner-level GenAI evidence, art and design VR readiness evidence, and 3D technical resources at the construct level. The framework was tested on public data, revealing that the frequency of using GenAI was not significantly linked to the students' perceived Learning Engagement, whereas Intrinsic Motivation was strongly linked to it; the published art and design source revealed positive attitudes towards VR, but with low use of the equipment reported. There are well-defined boundaries to the framework. The learner data are self-reported, cross-sectional, and are not specific to art and design; the art and design evidence is summary data only; the 3D layer is technical affordances, not learning outcomes; and

the indirect associations do not establish causal mediation. Importantly, the next step is a single population study measuring GenAI usage, motivations, VR access, 3D task processes and their relationships to expert-rated design outcomes, using longitudinal or comparative designs, to test the propositions advanced here.

Author contributions: Conceptualization, KY and LW; methodology, KY and AV; software, KY; validation, KY, AV and LW; formal analysis, KY; investigation, KY; resources, AV and LW; data curation, KY; writing—original draft preparation, KY; writing—review and editing, KY, AV and LW; visualization, KY; supervision, AV and LW; project administration, LW; funding acquisition, LW. All authors have read and agreed to the published version of the manuscript.

Data availability: All data analyzed in this study are publicly available. GenAI learner dataset and questionnaire: Wang W. (2026), A Study on the Influence of GenAI Use on Learning Engagement, Mendeley Data, V1, https://doi.org/10.17632/8cjd7chtcv.1 (CC BY 4.0) [33]. Art and design VR evidence: summary statistics published in Wen et al. [27], https://doi.org/10.6007/IJARBSS/v12-i4/12990. Contextual VR engagement evidence: Guo [28], https://doi.org/10.47191/ijsshr/v7-i06-99. 3D technical resources: 3D-FRONT-ToolBox, https://github.com/3D-FRONT-FUTURE/3D-FRONT-ToolBox, and the 3D-FRONT-FUTURE website repository, https://github.com/3D-FRONT-FUTURE/3D-FRONT-FUTURE.github.io (consulted for documentation only).

Code availability: The complete analysis scripts (data preparation, descriptive statistics, reliability and CFA diagnostics, regression models with HC3 robust standard errors and Benjamini–Hochberg correction, nested cross-validation with hyperparameter settings, and SHAP analysis) are provided as Supplementary Material S1.

http://kbweb.pubreviewsync.com/www/webSiteFile/downloadFile?siteId=0&token=59729629cc55484fa94fbd5f6279f0bf

Ethical scope: This study used publicly available, de-identified secondary data and published aggregate statistics; no new participants were recruited and no new data were collected, so additional ethical approval was not required.

Funding: This research received no external funding.

Conflicts of interest: The authors declare no conflicts of interest.

References

1. Kasneci E, Seßler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences. 2023. https://doi.org/10.1016/j.lindif.2023.102274
2. Dwivedi YK, Kshetri N, Hughes L, et al. Opinion Paper: 'So what if ChatGPT wrote it?' Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management. 2023. https://doi.org/10.1016/j.ijinfomgt.2023.102642
3. Chan CKY, Hu W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education. 2023. https://doi.org/10.1186/s41239-023-00411-8
4. Lo CK. What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. Education Sciences. 2023. https://doi.org/10.3390/educsci13040410
5. Tlili A, Shehata B, Adarkwah MA, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments. 2023. https://doi.org/10.1186/s40561-023-00237-x
6. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations

in Education and Teaching International. 2023. https://doi.org/10.1080/14703297.2023.2190148
7. Chan CKY. A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education. 2023. https://doi.org/10.1186/s41239-023-00408-3
8. Walter Y. Embracing the future of Artificial Intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education. 2024. https://doi.org/10.1186/s41239-024-00448-3
9. Chiu TKF. The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. Interactive Learning Environments. 2023. https://doi.org/10.1080/10494820.2023.2253861
10. Bond M, Khosravi H, de Laat M, et al. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. International Journal of Educational Technology in Higher Education. 2024. https://doi.org/10.1186/s41239-023-00436-z
11. Holmes W, Tuomi I. State of the art and practice in AI in education. European Journal of Education. 2022. https://doi.org/10.1111/ejed.12533
12. Ng DTK, Leung JKL, Su J, Ng CW, Chu SKW. Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. Educational Technology Research and Development. 2023. https://doi.org/10.1007/s11423-023-10203-6
13. Çelik İ, Dindar M, Muukkonen H, Järvelä S. The promises and challenges of artificial intelligence for teachers: a systematic review of research. TechTrends. 2022. https://doi.org/10.1007/s11528-022-00715-y
14. Kamalov F, Santandreu Calonge D, Gurrib I. New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. Sustainability. 2023. https://doi.org/10.3390/su151612451
15. Gligorea I, Cioca M, Oancea R, Gorski AT, Gorski H, Tudorache P. Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. Education Sciences. 2023. https://doi.org/10.3390/educsci13121216
16. Wu R, Yu Z. Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. British Journal of Educational Technology. 2024. https://doi.org/10.1111/bjet.13334
17. Song C, Song Y. Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. Frontiers in Psychology. 2023. https://doi.org/10.3389/fpsyg.2023.1260843
18. Yusuf A, Pervin N, Román-González M. Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. International Journal of Educational Technology in Higher Education. 2024. https://doi.org/10.1186/s41239-024-00453-6
19. Tlili A, Huang R, Shehata B, et al. Is Metaverse in education a blessing or a curse: a combined content and bibliometric analysis. Smart Learning Environments. 2022. https://doi.org/10.1186/s40561-022-00205-x
20. Zhang X, Chen Y, Hu L, Wang Y. The metaverse in education: Definition, framework, features, potential applications, challenges, and future research topics. Frontiers in Psychology. 2022. https://doi.org/10.3389/fpsyg.2022.1016300
21. Al-Adwan AS, Li N, Al-Adwan A, Abbasi GA, Albelbisi NA, Habibi A. Extending the Technology Acceptance Model (TAM) to Predict University Students' Intentions to Use Metaverse-Based Learning Platforms. Education and Information Technologies. 2023. https://doi.org/10.1007/s10639-023-11816-3
22. Xi N, Chen J, Gama F, Riar M, Hamari J. The challenges of entering the metaverse: An experiment on the effect of extended reality on workload. Information Systems Frontiers. 2022. https://doi.org/10.1007/s10796-022-10244-x
23. Makransky G, Mayer RE. Benefits of Taking a Virtual Field Trip in Immersive Virtual Reality: Evidence for the Immersion Principle in Multimedia Learning. Educational Psychology Review. 2022. https://doi.org/10.1007/s10648-022-09675-4
24. Lee H, Hwang Y. Technology-Enhanced Education through VR-Making and Metaverse-Linking to Foster Teacher Readiness and Sustainable Learning. Sustainability. 2022. https://doi.org/10.3390/su14084786
25. AlGerafi MAM, Zhou Y, Oubibi M, Wijaya TT. Unlocking the Potential: A Comprehensive Evaluation of Augmented Reality and Virtual Reality in Education. Electronics. 2023. https://doi.org/10.3390/electronics12183953
26. Dimitriadou E, Lanitis A. A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. Smart Learning Environments. 2023. https://doi.org/10.1186/s40561-023-00231-3
27. Wen Q, Zalay@Zali AA, Wong WL, Chuah KP. The Willingness of China University Students towards Virtual Reality Technology in Still Life Sketching. International Journal of Academic Research in Business and Social Sciences. 2022;12(4):557–565. https://doi.org/10.6007/IJARBSS/v12-i4/12990

28. Guo S. Implementation of Immersive Virtual Reality Games and Learning Engagement for Selected Mathematical Concepts among Middle School Students in Ganzhou City, China. International Journal of Social Science and Human Research. 2024;7(6):4395–4399. https://doi.org/10.47191/ijsshr/v7-i06-99
29. Hassija V, Chamola V, Mahapatra A, et al. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. Cognitive Computation. 2024. https://doi.org/10.1007/s12559-023-10179-8
30. Molnar C, König G, Herbinger J, et al. General pitfalls of model-agnostic interpretation methods for machine learning models. In: xxAI – Beyond Explainable AI. Lecture Notes in Computer Science, vol 13200. Springer; 2022:39–68. https://doi.org/10.1007/978-3-031-04083-2_4
31. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems. 2017;30:4765–4774.
32. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence. 2019;1(5):206–215. https://doi.org/10.1038/s42256-019-0048-x
33. Wang W. A Study on the Influence of GenAI use on Learning Engagement [dataset]. Mendeley Data, V1. 2026. https://doi.org/10.17632/8cjd7chtcv.1
34. Fu H, Cai B, Gao L, et al. 3D-FRONT: 3D Furnished Rooms with layOuts and semaNTics. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021. https://doi.org/10.1109/ICCV48922.2021.01075
35. Fu H, Jia R, Gao L, et al. 3D-FUTURE: 3D Furniture Shape with TextURE. International Journal of Computer Vision. 2021;129:3313–3337. https://doi.org/10.1007/s11263-021-01534-z