

Pre-service Teachers' Engagement With Generative AI for Culturally Grounded Visual Teaching Resources

Macharious Nabang

`machariousnabang@gmail.com`

Bagabaga College of Education

Harry Barton Essel

Kwame Nkrumah University of Science and Technology

Akosua Tachie-Menson

Kwame Nkrumah University of Science and Technology

Prosper Setsoafia

Kwame Nkrumah University of Science and Technology

Charles Atta Kodua

Kwame Nkrumah University of Science and Technology

Juliet Appiah-Kubi

Kwame Nkrumah University of Science and Technology

Edmund Boamah Acheampong

Kwame Nkrumah University of Science and Technology

George Attah Aboagye

Kwame Nkrumah University of Science and Technology

Matthew Nyaaba

University of Georgia

Research Article

Keywords: Generative artificial intelligence, Pre-service teachers, Technology acceptance, system usability, culturally responsive teaching, Visual resource quality, Lesson delivery effectiveness, Teacher education in Ghana

Posted Date: June 4th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9684535/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Generative artificial intelligence (GenAI) is increasingly being adopted to support the creation of instructional visuals in teacher education, yet there is little evidence of how preservice teachers can potentially use GenAI to design culturally responsive visual teaching resources, especially in low-resources areas like Ghana. In this study, we adopted a cross-sectional explanatory design involving 201 pre-service teachers in Ghana to explore their self-efficacy in using AI-generated visual resources during their teaching practicum. Data was collected using a survey and analyzed quantitatively. Composite-based PLS-SEM was used to evaluate the measurement and structural models, complemented with Necessary Condition Analysis (NCA) to identify bottlenecks for high outcomes. Group differences by programme and gender were examined using ANOVA with Tukey HSD and effect sizes. All hypothesized structural paths were supported. The model explained 39.9% of variance in TA, 39.3% in VRQ and 61.6% in LDE. The strongest structural effects were $SU \rightarrow TA$ ($\beta = .330$) and $VRQ \rightarrow LDE$ ($\beta = .309$). NCA showed that VRQ and TA were necessary conditions for high LDE, with bottleneck evidence indicating that achieving 90% LDE requires at least $SU = 3.2$ and $VRQ = 3.0$ (on a five-point scale). Programme differences were gradually small but were most notable for LDE ($\eta^2 \approx .066$, medium), with Primary Education consistently reporting lower mean levels than JHS. Gender differences were mostly small, except for TA (Cohen's $d = .45$, moderate), favouring males. Generally, the findings suggest that effective integration of AI-generated visuals in teacher education depends on strengthening usability and culturally responsive capability alongside acceptance, with targeted supports for lower-performing programme cohorts and equity-focused onboarding to reduce acceptance gaps.

1. Introduction

Generative artificial intelligence at an increasing rate is influencing how teachers develop instructional materials, assess learning and support pedagogical decision-making across the higher education landscape (Crompton & Burke, 2023). The arrival of multimodal AI systems including GPT-based text and image generators has birthed new opportunities for creating customized visuals, explanations and examples that can enhance the preparation of lessons. These tools are now integrated into learning management systems, feedback platforms and digital content creation workflows, encouraging renewed interest in how teachers conceptualise authorship, creativity and instructional design in an AI-mediated environment (Zhang et al., 2023; Nikou, 2024).

Even though AI adoption is increasing globally, its acceptance in African teacher education institutions remains uneven due to limitations in infrastructure, policy uncertainty and varied levels of digital readiness among educators (Essel et al., 2022; Akanzire et al., 2025). Yet the potential benefits are significant. In contexts such as Bagabaga College of Education in Tamale, Ghana, where teaching resources are not always readily available, generative AI can aid pre-service teachers to create contextually grounded visual aids that support meaningful learning. AI tools can generate culturally appropriate diagrams, images and examples which can reduce the time for preparation and expand access to high-quality teaching resources (Alejandro et al., 2024; Spasopoulos et al., 2025).

Also, the integration of AI into teacher education raises essential concerns. Scholars caution that generative models may produce cultural inaccuracies, misrepresent indigenous and local artefacts, or embed Western-centric assumptions if users are not trained to recognize and correct these biases (Conner, 2024; Nyaaba et al., 2025). In Ghana, effective teaching often depends on the ability draw on learners' cultural backgrounds, community knowledge and linguistic differences. Therefore, it is essential that AI-generated teaching and learning resources reflect these cultural dimensions for equity, authenticity and relevance.

A study on AI adoption among pre-service teachers shows that psychological and technological factors such as perceived usefulness, perceived ease of use, system usability and self-efficacy strongly influence willingness to engage with AI tools (Zhang et al., 2023; Alejandro et al., 2024; Runge et al., 2025). System usability which is very often evaluated as using the System Usability Scale, has consistently emerged as a major determinant of acceptance in digital learning environments (Hidayat et al., 2024; Roncal Belzunce et al., 2025). However, studies have focused largely on behavioural intention rather than actual pedagogical use and few have explored how teachers' cultural orientations influence the integration of AI into teaching and learning.

Teachers' culturally responsive teaching self-efficacy (CRTSE) has been shown to predict the ability to select adapt and critique instructional materials in ways that affirm students' cultural identities and resist deficit narratives (Ialuna et al., 2024; Chalar Mahali et al., 2022; Mensah, 2025). As generated AI becomes a common component of curriculum design, teachers need both cultural and technological confidence to evaluate AI generated outputs, detect cultural inaccuracies and adapt visuals for Ghanaian classrooms (Li et al., 2025). Despite this, empirical studies linking CRTSE to AI-supported instructional design remain limited, especially in the Global South.

Research also shows that positive perceptions of AI do not always translate into actual classroom use when institutional, infrastructural or contextual barriers persist (Holmes et al., 2022; Akanzire et al., 2025). Understanding how acceptance interacts with cultural competence, usability, and institutional quality is therefore important for designing effective AI-supported teacher education programmes.

Against this background, this current study investigates how technological self-efficacy, culturally responsive teaching self-efficacy, and system usability influence pre-service teachers' acceptance and use of AI tools for creating culturally responsive visual teaching and learning resources. The study focuses on pre-service teachers at Bagabaga College of Education in Tamale, Ghana and forms part of a broader initiative to explore the contribution of AI to instructional quality in Global South contexts. The study examines six latent constructs: Self-Efficacy (SE), Culturally Responsive Teaching Self-Efficacy (CRTSE), System Usability (SU), Technology Acceptance (TA), Visual Resource Quality (VRQ) and Lesson Delivery Effectiveness (LDE). It tests twelve hypotheses that reflect both direct and mediated relationships among these constructs.

Research Questions

1. How do culturally responsive teaching self-efficacy shape pre-service teachers' acceptance of GenAI for visual instructional design?
2. How does technology acceptance influence the quality of GenAI-generated visual teaching resources?
3. What conditions are required for AI-generated visual resources to support effective lesson delivery among pre-service teachers in Ghana?

2. Literature Review and hypotheses development

2.1 AI supported instruction and teacher education in Ghana

The rise of generative artificial intelligence has transformed how teachers develop instructional materials, especially in environments where access to high quality teaching resources is limited (Crompton & Burke, 2023). AI systems that produce diagrams, texts and culturally contextualized visuals offer new possibilities for pre-service teachers who often struggle with inadequate resource availability in the Global South settings including Ghana. When adopted effectively, generative AI can enhance lesson preparation efficiency, diversify instructional strategies and support innovation in teacher education (Alejandro et al., 2024; Spasopoulos et al., 2025).

Despite these opportunities, researchers highlight significant variation in AI integration across African institutions due to limited digital literacy gaps, infrastructure, and inconsistent institutional readiness (Essel et al., 2022; Akanzire et al., 2025). This makes it very important to understand the psychological and technological factors that shape how pre-service teachers begin to adopt AI for institutional purposes.

2.2 Self Efficacy and Technology Use in Teacher Preparation

Self-efficacy is central to teachers' decision to adopt and use instructional technologies. Grounded in Bandura's social cognitive theory, self-efficacy refers to a person's belief in their ability to perform a specific task successfully. Teachers with higher technological self-efficacy are more likely to explore, evaluate and integrate emerging tools in their teaching (Tondeur et al., 2020). In AI-mediated learning, self-efficacy influences not only intention but also the quality of outputs produced, including the ability to refine prompts, assess accuracy and generate useful learning materials (Zhang et al., 2023).

With these findings, technological self-efficacy is expected to shape pre-service teachers' acceptance of AI tools, their ability to produce high quality visual aids and effective delivery of lessons.

Hypotheses

H1: SE has a positive effect on TA

H2: SE has a positive effect on VRQ

H3: SE has a positive effect on LDE

2.3 Culturally Responsive Teaching Self Efficacy and AI Generated Learning Materials

Culturally Responsive Teaching Self Efficacy (CRTSE) describes teachers' confidence in designing instruction that affirms learners' cultural identities, community values and lived experiences (Gay, 2018; Chahar Mahali et al., 2022). Teachers with high CRTSE are better prepared to evaluate whether AI images and examples are culturally appropriate, especially as Generative AI models may misrepresent African Artefacts, traditions or contexts (Green, 2020; Nyaaba et al., 2025)

In Ghana, where teaching and learning are strongly influenced by cultural narratives and local languages, the ability to assess and adapt AI generated is particularly important. CRTSE therefor becomes a key determinant of whether pre-service teachers are able to use AI tools meaningfully and responsibly (Ialuna et al., 2024; Mensah, 2025). Studies show that culturally confident teachers are more likely to refine AI prompts, detect bias in representation and select visuals that add cultural and pedagogical value to lessons (Li et al., 2025). Based on this evidence, CRTSE is hypothesized to influence TA, VRQ and LDE.

Hypotheses

H4: CRTSE has a positive effect on TA.

H5: CRTSE has a positive effect on VRQ.

H6: CRTSE has a positive effect on LDE.

2.4 System Usability and User Interaction with AI Tools

System usability represents the extent to which an educational technology is easy to learn, efficient to use and satisfactory for the user. In higher education, usability is a major predictor of the adoption of digital tools including chatbots, learning management systems and AI supported instructional technologies (Hidayat et al., 2024; Roncal Belzunce et al., 2025). When systems are intuitive and require minimal cognitive efforts, educators report higher acceptance and demonstrate stronger behavioural engagement (Alejandro et al., 2024; Zhang et al., 2023).

For pre-service teachers in resource constrained contexts, usability is even more crucial because users may have limited access to devices and inconsistent internet connectivity (Essel et al., 2022). Therefore,

if a generative AI tool for producing visual aids feels overly complex or difficult to control, teachers may not adopt it regardless of potential benefits. System Usability is therefore expected to influence acceptance, visual resource quality and lesson delivery.

Hypotheses

H7: SU has a positive effect on TA

H8: SU has a positive effect on VRQ

H9: SU has a positive effect on LDE

2.5 Technology Acceptance and Instructional Quality

The Technology Acceptance Model postulates that attitude toward a technology influence its adoption and actual usage (Davis, 1989). As generative AI makes its way into teacher education, acceptance plays an important role in determining how pre-service teachers integrate AI into planning of lessons, assessment and instructional preparation (Nikou, 2024; Runge et al., 2025). Accepted technologies are more likely to be used consistently, which in turn contributes to higher quality teaching resources and more effective lesson delivery.

Technology Acceptance (TA) is therefore conceptualized as a mediator that connects cultural responsiveness, usability and self-efficacy to instructional outcomes. It is also modelled as a direct predictor of visual resource quality and effective teaching.

Hypotheses

H10: TA has a positive effect on VRQ.

H11: TA has a positive effect on LDE.

2.6 Visual Resource Quality and Lesson Delivery Effectiveness

Visual Resource Quality (VRQ) captures clarity, cultural appropriateness, relevance to lesson objectives and aesthetic design. High quality visuals enhance comprehension, engagement and coherence of lesson delivery (Mayer, 2014). When visuals align with learners' cultural background and cognitive level, they support instructional flow, build conceptual understanding and improve classroom interaction (Li et al., 2025).

In AI assisted instructional design, quality depends not only on the capability of the system but also on the user's ability to evaluate and refine generated outputs. Enhanced VRQ is therefore expected to

contribute directly to Lesson Delivery Effectiveness (LDE).

Hypothesis

H12: VRQ has a positive effect on LDE

Summary

The reviewed literature demonstrates that self-efficacy, cultural responsiveness, and system usability are critical to understanding how pre-service teachers adopt and use AI tools for producing culturally grounded visual instructional materials. Together with technology acceptance, visual resource quality and lesson delivery effectiveness, these variables form the basis for structural model that tests the twelve hypotheses across six latent constructs.

3. Methods

3.1 Research Design and Procedure

A cross-sectional explanatory research design was adopted for the study to examine the relationship among technological self-efficacy, culturally responsive teaching self-efficacy, system usability, technology acceptance, visual resource quality and lesson delivery effectiveness in the context of AI-supported instructional design. This design was appropriate because the aim was to identify the predictive and mediating effects among multiple latent constructs at a single point in time.

The customized ChatGPT-based generative AI model

The Culturally Responsive Visual Aid Generator (CReVAG) is a creative assistant designed to support teachers in the development of culturally responsive visuals for teaching and learning. It assists educators in creating high-definition classroom images, posters, diagrams, story scenes, flashcards, and other visual learning materials that are inclusive, contextually relevant, and sensitive to learners' cultural backgrounds.

The assistant is particularly useful in helping teachers translate lesson objectives into visual concepts or image-generation prompts while avoiding stereotypical or generalized representations of culture. Instead, it encourages specificity, authenticity, and respectful representation. Its work is guided by key contextual information, including the country, region and district of the school, class or level, subject, and topic.

CReVAG is especially aligned with Ghanaian educational contexts and can support visual design in relation to Ghana Education Service curriculum expectations. Its approach reflects the curriculum's emphasis on inclusive and gender-responsive pedagogy, learner-centred teaching, creativity, collaboration, communication, and respect for national and global identity.

The knowledge section of the tool as shown in Fig. 3 is filled with the GES (NaCCA) curriculum materials which are the official teaching and learning materials of all subjects for the basic schools where the pre-service teachers undertake their practicum.

In the capabilities section of the tool as shown in Fig. 3, four functions are available: Web Search, Canvas, Image Generation, and Code Interpreter & Data Analysis. Of these four functions, only Canvas and Image Generation were activated. The Image Generation function was used to produce the visuals, while the Canvas function provided the option to edit the visuals where necessary.

3.2 Participants and Context

Level 400 pre-service teachers (360) from Bagabaga College of Education in the Northern Region of Ghana were considered for the study. The sample size was determined using Krejcie and Morgan’s (1970) sample size determination table, which recommends a minimum sample size of approximately 186 respondents for a population of 360 at a 95% confidence level and a 5% margin of error. Since 5% of the population represents 18 respondents, the minimum sample size was increased by 18 to account for possible non-response and incomplete questionnaires. This resulted in a total projected sample size of 204 respondents. Of the 204 questionnaires administered, 201 were completed and returned, representing a response rate of 98.5% (Holtom et al., 2022). Participants were enrolled in three programme specialisations, namely Junior High School (JHS) Education, Primary Education and Early Grade Education. Table 1 presents the demographic characteristics of the sample. The sample was predominantly male (58.7%) and heavily concentrated in the JHS programme (77.6%), with smaller representation from Primary Education (12.9%) and Early Grade Education (9.5%). The majority of participants (80.6%) reported prior experience with AI tools, primarily conversational agents such as ChatGPT.

Table 1
Demographic Characteristics of Participants

Characteristic	Category	n	%
Gender	Male	118	58.7
	Female	83	41.3
Programme	JHS	156	77.6
	Primary	26	12.9
	Early Grade	19	9.5
Prior AI experience	Yes	162	80.6
	No	39	19.4

3.3 Instruments

The survey instrument was adapted from the Teachers' Acceptance of Artificial Intelligence (TAAI) scale developed and validated by Guo et al. (2025). The TAAI is a 27-item instrument measuring five dimensions of teachers' AI acceptance, namely Perceived Usefulness, Perceived Ease of Use (PEU), Behavioural Intention (BI), Self-Efficacy (SE) and Anxiety (ANX), with robust psychometric properties ($\alpha = 0.92$; RMSEA = 0.061; CFI = 0.981; all factor loadings > 0.60). Three TAAI dimensions were adapted for this study. SE items were recontextualised around confidence in generating AI-based visual teaching aids. PEU items were narrowed to the usability of specific AI tools for visual resource creation, and the term "AI-based tools" was used rather than "AI technologies" to improve respondent comprehension, consistent with wording refinements reported by Guo et al. (2025). Perceived Usefulness and BI items were merged into a single Technology Acceptance construct reflecting willingness to integrate AI-generated visuals into teaching. Three additional constructs were developed for this study. Culturally Responsive Teaching Self-Efficacy items were adapted from Siwatu's (2007) scale on preservice teacher culturally responsive self-efficacy and contextualised using the cultural tenets of indigenous language, indigenous knowledge, ethnicity/race and religious beliefs as identified in the GenAI-CRSciA framework (Nyaaba et al., 2024). Visual Resource Quality and Lesson Delivery Effectiveness items were newly created to capture perceptions of AI-generated outputs and their impact on instruction, respectively. In total, the instrument comprised 30 items across six constructs, each measured on a five-point Likert scale.

The semi-structured interview protocol was developed using the GenAI Culturally Responsive Science Assessment framework (Nyaaba et al., 2024) and Siwatu's (2007). The protocol comprised 12 open-ended questions designed to address each research question. Questions addressing the first research question explored how pre-service teachers evaluated the cultural appropriateness of AI-generated visuals, drawing on the indigenous knowledge and ethnicity/race tenets to probe whether outputs reflected local artefacts, community practices and students' cultural identities. Questions addressing the second research question examined participants' prompting strategies and use of local languages when interacting with AI tools, drawing on the indigenous language tenet and on the finding by Nyaaba et al. (2024) that culturally tailored outputs depend on the specificity of cultural background information provided by users. Questions addressing the third research question investigated how participants navigated religious and community considerations when integrating AI visuals into lesson delivery, operationalising the religious beliefs and family/community tenets of the framework.

Both instruments underwent content validation before deployment. For the survey, a panel of four experts reviewed all items for construct alignment, clarity and contextual appropriateness for the Ghanaian teacher education setting, followed by cognitive pre-testing with six pre-service teachers using the think-aloud procedure described by Guo et al. (2025). Three items were reworded and one was replaced based on this process. A pilot administration ($n = 25$) confirmed adequate item distributions and a completion time of 15 to 20 minutes. For the interview protocol, piloting with three pre-service teachers led to the simplification of two questions and the addition of one probe to elicit specific examples of cultural mismatch in AI outputs. Final interviews lasted 25 to 40 minutes.

Data Collection

Data were collected during the mandatory twelve-week practicum undertaken by final-year pre-service teachers in partner basic schools. During the practicum, participants engaged in regular school-based professional activities, including classroom observation, lesson preparation and delivery, assessment of learners, and participation in extracurricular activities within their respective host schools. The intervention phase was introduced during the practicum period, during which participants used a customized ChatGPT-based generative AI model to create culturally grounded visual resources for selected lessons. Ethical approvals were sought from the Humanities and Social Sciences Research Ethics Committee, KNUST (Re: HuSSREC/AP/346/Vol.4) and the Bagabaga College of Education Research Committee before collection was carried out. Pre-service teachers were briefed in their classrooms about the aims of the study, the voluntary nature of participation and confidentiality procedures. Students were informed that they could decline or withdraw from the study without personal academic consequences. After consenting, participants completed the questionnaire either through a secure Google Form or a paper-based version, depending on class logistics and individual preference. This is after participants have been given a weeklong orientation and hands-on training on how to use the customized ChatGPT tool to create culturally-grounded visuals for teaching and learning. Actual data collection commenced after the sixth week of the practicum, allowing participants sufficient time to settle into their host schools and gain practical experience using the AI tool in authentic classroom contexts. The data collection exercise lasted for two weeks. Questionnaire data were collected using both online and paper-based formats. The online version was administered through Google Forms, while printed questionnaires were provided to participants who could not conveniently access or complete the online form. Both versions contained identical items to ensure consistency across responses. To reduce the likelihood of duplicate submissions, the Google Form was restricted to one response per participant. Completed paper questionnaires were collected immediately after completion and reviewed for completeness before data entry. Students completed the questionnaire independently in supervised environments to avoid peer influence and to address potential issues arising.

Data Analysis

The data were analyzed quantitatively to examine the relationships among the study variables. Descriptive statistics were first computed to summarize participants' demographic characteristics and the central tendencies of the main constructs. Thereafter, composite-based Partial Least Squares Structural Equation Modeling (PLS-SEM) was used to assess both the measurement model and the structural model. The measurement model was evaluated using indicators of reliability, convergent validity, discriminant validity, and collinearity. The structural model was then examined to test the hypothesized relationships among self-efficacy, culturally responsive teaching self-efficacy, system usability, technology acceptance, visual resource quality, and lesson delivery effectiveness. In addition, Necessary Condition Analysis (NCA) was conducted to identify the minimum levels of specific conditions required to achieve high outcomes. Group differences by programme and gender were further

examined using analysis of variance (ANOVA), followed by Tukey HSD post hoc tests and effect size estimates.

4. Results

4.2 Descriptive Statistics

Composite scores for the six constructs were computed as the mean of their respective items. Table 2 shows that mean scores for all constructs were above the scale midpoint, indicating generally positive perceptions and experiences of AI-aided visual resource generation and lesson delivery.

Table 2
Descriptive Statistics for Latent Constructs (N = 201)

Construct	Mean	SD	Min	Max
SE	3.49	0.84	1.00	5.00
CRTSE	3.46	0.85	1.00	5.00
SU	3.54	0.66	1.00	5.00
TA	3.89	0.72	1.00	5.00
VRQ	3.75	0.71	1.00	5.00
LDE	3.80	0.71	1.00	5.00

4.3 Measurement Model Assessment

4.3.1 Construct Reliability and Convergent Validity

Reliability and Convergent Validity were assessed using Cronbach’s Alpha, Composite Reliability and Average Variance Extracted (AVE). As shown in Table 3, Cronbach alpha values range from 0.738 (SU) to 0.859 (CRTSE), which acceptable to excellent internal consistency. Composite reliability values (rho_c) range from 0.836 to 0.898, exceeding recommended threshold of 0.70. AVE values range from 0.563 to 0.648, all above the 0.50 cutoff, which support convergent validity.

Table 3
Construct reliability and Convergent validity

Construct	Cronbach's α	ρ_A (rho_a)	ρ_c (Composite Reliability)	AVE
CRTSE	0.859	0.860	0.898	0.639
LDE	0.849	0.851	0.893	0.625
SE	0.817	0.828	0.871	0.576
SU	0.738	0.767	0.836	0.563
TA	0.819	0.827	0.880	0.648
VRQ	0.827	0.828	0.879	0.592

All items loaded strongly on their intended constructs. Outer loadings range from 0.652 to 0.839 for the SU items and from 0.709 to 0.821 for the other constructs, which are well above the 0.70 guideline for established scales. Outer weights for the indicators were all positives and of reasonable magnitude, confirming that all items contributed meaning fully to their respective latent variables.

4.3.2 Discriminant Validity

Discriminant Validity was first assessed using the Fornell-Larcker criterion. Table 4 presents the square roots of the AVE values on the diagonal, with the correlations between constructs in the off-diagonal cells. In every case, the square root of AVE for each construct exceeded its correlation with other constructs supporting discriminant validities.

Table 4
Fornell-Larcker criterion

CRTSE	CRTSE	LDE	SE	SU	TA	VRQ
	0.799					
LDE	0.605	0.790				
SE	0.588	0.607	0.759			
SU	0.581	0.619	0.521	0.751		
TA	0.516	0.595	0.512	0.559	0.805	
VRQ	0.528	0.654	0.487	0.529	0.486	0.769

Discriminated validity was further evaluated using Heterotrait-Monotrait (HTMT) ratio. All HTMT values were below 0.85, as shown in Table 5, which provides additional evidence that the constructs are empirically distinct.

Table 5
HTMT ratios among latent construct

CRTSE	CRTSE	LDE	SE	SU	TA	VRQ
LDE	0.705					
SE	0.698	0.714				
SU	0.717	0.777	0.658			
TA	0.608	0.707	0.603	0.701		
VRQ	0.619	0.776	0.576	0.660	0.585	

Table 6 presents the Inter-constructs correlation matrix derived from the PLS latent variable scores. All correlations are positive and moderate to strong, which is consistent with the theoretical expectation that these constructs are related but not redundant.

Table 6
Inter-construct correlations

CRTSE	CRTSE	LDE	SE	SU	TA	VRQ
	1.000					
LDE	0.605	1.000				
SE	0.588	0.607	1.000			
SU	0.581	0.619	0.521	1.000		
TA	0.516	0.595	0.512	0.559	1.000	
VRQ	0.528	0.654	0.487	0.529	0.486	1.000

4.4 Collinearity Assessment

Collinearity diagnostics were examined at both the indicator and structural levels. At the indicator level, all outer VIF values range between 1.30 and 2.37, and at the structural level, predictor VIF values for the endogenous constructs range between 1.63 and 1.96 as indicated in Table 7. These values fall below the commonly used threshold of 3.3 and well below 5, indicating that multicollinearity is not a concern in this model.

Table 7
Structural VIF Values for Predictor Paths

CRTSE	CRTSE	LDE	SE	SU	TA	VRQ
		1.957			1.818	1.877
LDE						
SE		1.777			1.652	1.738
SU		1.899			1.631	1.813
TA		1.710				1.664
VRQ		1.647				

4.5 Structural Model Results

4.5.1 Explained Variance and Effect sizes

The model explained a substantial proportion for variance in the key endogenous constructs as shown in Table 8, the predictors accounted for 39.9% of the variance in Technology Acceptance (TA), 39.3% of the variance in Visual Resource Quality (VRQ) and 61.6% of the variance in Lesson Delivery Effectiveness (LDE). Adjusted R^2 values were only slightly lower, indicating stable explanatory power.

Table 8
 R^2 and Adjusted R^2 for endogenous constructs

Construct	R^2	Adjusted R^2
TA	0.399	0.390
VRQ	0.393	0.380
LDE	0.616	0.606

Effect sizes (f^2) were computed for each predictor. System Usability (SU) had the largest effect on Technology Acceptance ($f^2 = 0.146$), while Self Efficacy (SE) and CRTSE made smaller but non trivial contributions to both TA and LDE. Overall, effect sizes ranged from small to moderate, suggesting that the predictors jointly play meaningful roles in shaping acceptance, source quality and lesson delivery.

4.5.2 Direct effects

Table 9 presents the standardized path coefficients for the twelve hypothesized relationships. All predictor had positive coefficients and many paths were statistically significant in the bootstrapped model.

Table 9
Direct Path Coefficients

CRTSE	CRTSE	LDE	SE	SU	TA	VRQ
		0.133			0.189	0.219
LDE						
SE		0.195			0.228	0.153
SU		0.178			0.330	0.228
TA		0.177				0.167
VRQ		0.309				

System Usability was the strongest predictor of Technology Acceptance ($\beta = 0.330$), with both SE and CRTSE also contributing positively. CRTSE exhibited notable direct effects on both VRQ and LDE, indicating that culturally confident pre-service teachers tend to produce better quality AI-generated visuals and deliver more effective lessons. TA had positive effects on VRQ and LDE, although these effects were smaller in magnitude than those of CRTSE and SU.

4.5.3 Total and indirect effects

To capture both direct and indirect influences, total effects were computed (Table 10). For LDE, the strongest total effects were observed for SU (0.324), SE (0.295) and CRTSE (0.244), indicating that lesson delivery effectiveness is influenced by a combination of usability, self-efficacy and cultural responsiveness. For VRQ, SU again had the largest total effect (0.284), followed by CRTSE (0.251) and SE (0.191). TA's total effect on LDE through VRQ was modest but positive.

Table 10
Total effects

Path	Total effect
CRTSE → LDE	0.244
CRTSE → VRQ	0.251
SE → LDE	0.295
SE → VRQ	0.191
SU → LDE	0.324
SU → VRQ	0.284
TA → LDE	0.228
TA → VRQ	0.167

Table 11 summarises total indirect effects. Indirect paths from CRTSE, SE and SU to LDE and VRQ through TA and VRQ were small in magnitude. For example, SU had an indirect effect of 0.146 on LDE and 0.055 on VRQ, while Se had indirect effects of 0.099 on LDE and 0.038 on VRQ. The specific indirect effects show that the chains CRTSE → TA → VRQ and SU → TA → VRQ → LDE are present but modest. Taken together, these results suggest that mediation through TA exists but is relatively weak compared to the direct paths.

Table 11
Total and specific indirect effects

Path	Total effect
CRTSE → LDE (indirect)	0.111
CRTSE → VRQ (indirect)	0.032
SE → LDE (indirect)	0.099
SE → VRQ (indirect)	0.038
SU → LDE (indirect)	0.146
SU → VRQ (indirect)	0.055
TA → LDE via VRQ	0.052

4.5.4 summary of measurement and structural model quality

The combined evidence from reliability, convergent validity and discriminant validity confirms that the six constructs are measured robustly. VIF values show that collinearity is not problematic. The model explains a substantial proportion of variance in Technology Acceptance, Visual Resource Quality and Lesson Delivery Effectiveness. Direct paths indicate that Self Efficacy, Culturally Responsive Teaching Self Efficacy and System Usability are all important drivers of both acceptance and pedagogical outcomes, while mediation through TA is present but comparatively modest.

4.6 Hypotheses Testing Results

The proposed twelve hypotheses that reflected the direct and indirect relationships among Self Efficacy (SE), Culturally Responsive Teaching Self Efficacy (CRTSE), System Usability (SU), Technology Acceptance (TA), Visual Resource Quality (VRQ) and Lesson Delivery Effectiveness (LDE). Hypothesis testing was conducted using a 5000-sample bootstrapping in Smart PLS 4, with significance determined by 95 percent confidence intervals. Table 12 summarises the results of the hypothesis testing indicating statistically significant positive associations among the proposed constructs (Hair et al., 2019; Khan et al., 2019).

Table 12
Summary of the hypothesis testing

Hypothesis	Structural Path	β	Supported?
H1	SE $\rightarrow$ TA	0.228	Supported
H2	SE $\rightarrow$ VRQ	0.153	Supported
H3	SE $\rightarrow$ LDE	0.195	Supported
H4	CRTSE $\rightarrow$ TA	0.189	Supported
H5	CRTSE $\rightarrow$ VRQ	0.219	Supported
H6	CRTSE $\rightarrow$ LDE	0.133	Supported
H7	SU $\rightarrow$ TA	0.330	Supported
H8	SU $\rightarrow$ VRQ	0.228	Supported
H9	SU $\rightarrow$ LDE	0.178	Supported
H10	TA $\rightarrow$ VRQ	0.167	Supported
H11	TA $\rightarrow$ LDE	0.177	Supported
H12	VRQ $\rightarrow$ LDE	0.309	Supported

4.7 Interpretation of hypothesis testing

4.7.1 Effects of SE, CRTSE and SU on TA

H1, H4 and H7 were all supported. System Usability (SU) had the strongest effect on Technology Acceptance (TA), followed by Self Efficacy (SE) and Culturally Responsive Teaching Self Efficacy (CRTSE). This indicates that pre-service teachers are more likely to accept AI visual generation tools when they feel confident using digital technologies, when they possess strong culturally responsive instructional beliefs and when the system is perceived as intuitive and easy to explore.

Predictors of Visual Resource Quality

H2, H5, H8 and H10 were supported. SU, CRTSE, SE and TA each contributed positively to Visual Resource Quality (VRQ). Culturally responsive teaching beliefs and system usability played critical roles, suggesting that both cultural competence and familiarity with the tool influence the quality of visuals produced.

Predictors of Lesson Delivery Efficiency

H3, H6, H9, H11 and H12 were supported. Direct effects indicated that SE, CRTSE, SU and TA improved the effectiveness of lesson delivery. The strongest predictor, however, was Visual Resource Quality

(VRQ), confirming H12. This finding indicates that when AI-generated visuals are of high quality, they substantially improve the clarity, communicative power and cultural significance of lesson delivery.

Mediation observations

Although TA mediated several relationships, the indirect effects were modest. TA strengthened the influence of SE, SU and CRTSE on VRQ and LDE, but most of the explanatory power came from direct effects. This shows that pre-service teachers' beliefs and usability perceptions exert an immediate influence on their instructional performance, even before acceptance fully develops.

4.8 Necessary Condition Analysis

Necessary Condition Analysis (NCA) results together with effect sizes and permutation p-values for potential necessary conditions are reported in Table 13. Bottleneck tables (Tables 14–16) specify the minimum levels of each condition required to reach selected outcome levels.

Table 13
NCA effect sizes (CR-FDH) and permutation p-values

X	Y	NCA effect size d (CR-FDH)	P (perm)	Magnitude
VRQ	LDE	0.19	0.002	Medium
TA	LDE	0.158	0.0015	Medium
SU	LDE	0.137	0.0005	Medium
CRTSE	LDE	0.13	0.0005	Medium
SE	LDE	0.095	0.004	Small
CRTSE	TA	0.12	0.0005	Medium
SU	TA	0.038	0.4983	Small
SE	TA	0.0	1.0	Small
TA	VRQ	0.135	0.0295	Medium
SU	VRQ	0.128	0.002	Medium
CRTSE	VRQ	0.113	0.001	Medium
SE	VRQ	0.11	0.002	Medium
<i>Note: Effect size (d) magnitude follow established NCA conventions (Dul, 2016).</i>				

For Lesson Delivery Effectiveness (LDE), Table 13 shows several statistically significant necessary conditions with most exhibiting medium necessity effects. The strongest constraint is VRQ → LDE ($d = .190$, $p = .002$), followed by TA → LDE ($d = .158$, $p = .0015$), SU → LDE ($d = .137$, $p = .0005$) and CRTSE → LDE ($d = .130$, $p = .0005$). Practically, this means high lesson delivery is impossible if the quality of visual

resources is not adequate, and it is also constrained if acceptance, usability or culturally responsive teaching self-efficacy are too low. SE → LDE is necessary but weaker ($d = .095$, small), suggesting self-efficacy helps, but it is not the tightest “gate” on lesson delivery in this model.

For Technology Acceptance (TA) as an outcome, CRTSE is the only meaningful necessary condition (CRTSE → TA: $d = .120$, $p = .0005$). In contrast, SU → TA is not necessary ($p = .4983$) and SE → TA shows no necessity effect ($d = .000$, $p = 1.000$). In simple terms, high TA can still occur without high SE and without SU being above a minimum but it cannot occur when culturally responsive teaching self-efficacy is too low, a grounded finding that fits the Ghanaian teacher-education context.

For Visual Resource Quality (VRQ), multiple conditions are necessary at medium magnitude: TA → VRQ ($d = .135$, $p = .0295$), SU → VRQ ($d = .128$, $p = .002$), CRTSE → VRQ ($d = .113$, $p = .001$), and SE → VRQ ($d = .110$, $p = .002$). So, high-quality AI-generated visuals are constrained unless teachers have at least minimum levels of acceptance, usability experience, cultural responsiveness and confidence.

4.8.1 Bottleneck analyses

Across the bottleneck analyses (Tables 14–16), the results specify the minimum “threshold” levels of key conditions required to make high performance possible (Dul, 2016, 2020). For Lesson Delivery Effectiveness (LDE), Table 14 shows that the requirements tighten evidently as the targeted LDE level increases, with the strongest constraints at the high end (e.g., 90% LDE requiring at least SU = 3.2 and VRQ = 3.0, alongside CRTSE = 2.8, SE = 2.2 and TA = 1.8), indicating that excellent delivery is bottlenecked mainly by tool usability and visual quality rather than acceptance alone. For Visual Resource Quality (VRQ), Table 15 shows a similar pattern in which the usability threshold remains highest (e.g., 90% VRQ requiring SU = 3.2), while CRTSE = 2.4 and SE = 2.2 also become binding and constrained mainly by usability plus culturally responsive capacity and confidence. Finally, for Technology Acceptance (TA), Table 16 identifies CRTSE as the dominant bottleneck at higher outcome levels (e.g., 90% TA requiring CRTSE = 2.2), whereas SU = 1.8 and SE = 1.0 remain low and stable across TA levels, implying that very high acceptance is predominantly constrained by cultural responsiveness rather than general self-efficacy or usability minima.

Table 14
Bottleneck table for LDE

X	Min X for 50% LDE	Min X for 70% LDE	Min X for 80% LDE	Min X for 90% LDE
SE	1.0	1.4	1.4	2.2
CRTSE	1.2	1.2	1.4	2.8
SU	1.8	2.6	2.8	3.2
TA	1.0	1.8	1.8	1.8
VRQ	1.2	1.8	2.2	3.0

Table 15
Bottleneck table for VRQ

X	Min X for 50% VRQ	Min X for 70% VRQ	Min X for 80% VRQ	Min X for 90% VRQ
SE	1.0	1.8	1.8	2.2
CRTSE	1.0	1.0	1.2	2.4
SU	1.8	1.8	3.2	3.2
TA	1.0	1.8	1.8	1.8

Table 16
Bottleneck table for TA

X	Min X for 50% TA	Min X for 70% TA	Min X for 80% TA	Min X for 90% TA
SE	1.0	1.0	1.0	1.0
CRTSE	1.2	1.2	1.2	2.2
SU	1.8	1.8	1.8	1.8

4.9 Post-hoc Tukey HSD Comparisons by Programme of study

The Tukey HSD post-hoc comparisons by programme (Tables 17) show a very consistent pattern: the only statistically reliable pairwise differences occur between JHS and Primary Education, while Early Grade does not differ significantly from either group across the tested constructs. Specifically, Primary Education scores lower than JHS on System Usability with mean difference = -0.3577, $p\text{-adj} = .0266$, Technology Acceptance recording 18: -0.4218, $p\text{-adj} = .0152$, with Visual Resource Quality showing 19: -0.4154, $p\text{-adj} = .0148$ and most strongly on Lesson Delivery Effectiveness with 20: -0.5423, $p\text{-adj} = .0007$. This post-hoc pattern directly mirrors the descriptive means in Table 21, where JHS is the highest and Primary Education is the lowest for these constructs, indicating that the general programme effect is largely driven by the JHS-Primary gap, especially for the instructional endpoint (LDE)

Table 17
Tukey HSD post-hoc comparisons of programme differences across study constructs

Construct	Group 1	Group 2	Mean difference	p-adj	95% CI lower	95% CI upper	Significant?
System Usability	Early Grade	JHS	0.2038	.4001	-0.1681	0.5758	No
System Usability	Early Grade	Primary	-0.1538	.7118	-0.6158	0.3081	No
System Usability	JHS	Primary	-0.3577	.0266	-0.6819	-0.0335	Yes
Technology Acceptance	Early Grade	JHS	0.1392	.6989	-0.2679	0.5463	No
Technology Acceptance	Early Grade	Primary	-0.2826	.7386	-0.7883	0.2231	No
Technology Acceptance	JHS	Primary	-0.4218	.0152	-0.7767	-0.0669	Yes
Technology Acceptance	Early Grade	JHS	0.1417	.6803	-0.2580	0.5414	No
Technology Acceptance	Early Grade	Primary	-0.2737	.3958	-0.7701	0.2228	No
Technology Acceptance	JHS	Primary	-0.4154	.0148	-0.7638	-0.0670	Yes
Lesson Delivery Effectiveness	Early Grade	JHS	0.0597	.9318	-0.3340	0.4535	No
Lesson Delivery Effectiveness	Early Grade	Primary	-0.4826	.0540	-0.9717	0.0065	No
Lesson Delivery Effectiveness	JHS	Primary	-0.5423	.0007	-0.8856	-0.1990	Yes

Tables 18–21 show meaningful but generally modest subgroup variation by programme and gender. By programme (Table 18), JHS reports the highest mean scores on the downstream outcomes TA (3.96), VRQ (3.82), and LDE (3.88) whereas Primary Education is consistently lowest across all constructs, especially LDE (3.34) and VRQ (3.40); Early Grade tends to sit between the two, with comparatively strong SE (3.67) and similar LDE (3.82) to JHS. This description pattern is supported by the programme effect sizes (Table 21): programme differences are small for SE and CRTSE ($\eta^2 \approx .01 - .014$), but become more salient for SU, TA and VRQ (still small, $\eta^2 \approx .038 - .040$) and are most practically meaningful for the instructional endpoint LDE ($\eta^2 \approx .066$; Cohen's $f = .265$, medium), indicating that programme context matters most for how well AI-generated visuals translate into effective lesson delivery. By Gender (Table 19), males report higher mean scores than females on every construct, with the most significant gaps on TA (4.06 vs. 3.74) and VRQ (3.87 vs. 3.64) and more minor differences for SE/CRTSE/SU. The

effect-size estimates (Table 23) confirm that most gender differences are small ($d = .08-.19$ for SE, CRTSE, SU). At the same time, TA shows a moderate difference ($d = .45$) and VRQ ($d = .34$) and LDE ($d = .27$) fall in the small-to-moderate range, suggesting that gender disparities are most pronounced at Technology Acceptance stage and may carry through to perceived visual quality and lesson delivery. Taken together, these results indicate that equity-sensitive implementation should prioritise (i) programme-tailored supports, especially for Primary Education Cohorts where Acceptance, VRQ and LDE are lowest, and (ii) gender-responsive onboarding that strengthens acceptance and sustained engagement, since TA is where the most significant gender gap appears.

Table 18
Mean scores of constructs by programme of study

Programme of study	SE	CRTSE	SU	TA	VRQ	LDE
Early Grade	3.67	3.45	3.40	3.82	3.67	3.82
Junior High School	3.51	3.50	3.60	3.96	3.82	3.88
Primary Education	3.27	3.24	3.25	3.54	3.40	3.34
Note. Values represent mean composite scores on a five-point Likert scale						

Table 19
Mean scores of constructs by gender

Gender	SE	CRTSE	SU	TA	VRQ	LDE
Female	3.46	3.41	3.48	3.74	3.64	3.72
Male	3.53	3.53	3.61	4.06	3.87	3.91
Note. Values represent mean composite scores on a five-point Likert scale						

Table 20
Mean differences and effect sizes by gender

Construct	Female Mean	Male Mean	Cohen's d
SE	3.46	3.53	0.08
CRTSE	3.41	3.53	0.14
SU	3.48	3.61	0.19
TA	3.74	4.06	0.45
VRQ	3.64	3.87	0.34
LDE	3.72	3.91	0.27
Note. Cohen's d values of 0.20, 0.50 and 0.80 represent small, medium and large effects respectively			

Table 21
Programme-level effect sizes (η^2 and Cohen's f)

Construct	F	p	η^2	Cohen's f	Interpretation
SE	1.400	.2491	.014	.119	Small
CRTSE	1.096	.3364	.011	.105	Small
SU	3.871	.0224	.038	.198	Small
TA	4.044	.0190	.039	.202	Small
VRQ	4.083	.0183	.040	.203	Small
LDE	6.965	.0012	.066	.265	Medium

5. Discussions

This study examined how pre-service teacher's perceived system usability (SU), culturally responsive teaching self-efficacy (CRTSE) and Self-Efficacy (SE) relate to technology acceptance (TA), uptake of an AI-based visual tool and downstream perceptions of visual resource quality (VRQ) and lesson delivery effectiveness (LDE) in a Ghanaian college-of-education context. In all, the finding support a coherent explanatory story: usability and culturally responsive confidence matter, but much of their influence operates through acceptance process (TA), while high-quality lesson delivery depends not only on "what helps" (sufficiency logic) but also on "what must be present" (necessity logic). This double lens is increasingly recommended in educational technology and IS-style explanatory studies because sufficiency results alone can hide binding constraints that impedes high performance in real classrooms (Dul, 2016; Dul et al., 2020; Richter et al., 2020).

5.1 Interpretating the structural model: acceptance as a central mechanism

The structural model indicates that both perceived usability and culturally responsive teaching self-efficacy are meaningful antecedence of technology acceptance for the AI-generated visual tool (Table 9). These aligns with the long standing view in educational technology research that adoption is really driven by tool availability alone; instead, adoption is shaped by whether users perceive the tool as workable, supportive and align with their professional goals and values (Crompton & Burke, 2023; Kasneci et al., 2023)

Importantly, the model suggests that TA functions as a key explanatory "bridge" between upstream perceptions (Usability and Culturally responsive confidence) and actual tool use outcome (Table 9 and Table 11). Methodologically, the study 's use of PLS-SEM is consistent with recommendations for estimating predictive, composite-based models in educational technology and information systems context, especially when the focus is on explaining variance in key outcomes and testing mediated

relationships (Hair et al., 2019; Khan et al., 2019). Substantively, this means that improving the interface experience and strengthening culturally responsive teaching confidence are likely to translate into stronger tool adoption partly because they raise teacher overall acceptance and readiness to integrate the tool into their practice not merely because they make the tool “easier”

A notable pattern in the model is that usability may show limited direct influence on usage outcomes once TA is included (Table 9), while its indirect influence through TA remains meaningful (Table 11). This is a familiar empirical pattern in acceptance research: Ease of Use and Usability and often matter most because they shape perceived usefulness, attitudes and broader acceptance process, which then drives sustained use (Teo, 2015). In practical terms, this points to an implementation lesson for teacher education: usability improvements should not be taken lightly. They are foundational inputs into acceptance and acceptance is what turns to move day-to-day classroom behaviour.

5.2 Moving from “use” to “quality” and “effective delivery”

Beyond adoption, the study extends the discussion from “whether teachers use AI visual” to “whether the visuals support high quality resource and effective lesson delivery”. This is essential because generative AI tools can increase production of materials without guaranteeing pedagogical or cultural quality. Recent higher-education research on AI tools repeatedly cautions that educational value depends on alignment, evaluation and responsible integration, not mere availability (Kasneci et al., 2023; Wang et al., 2024).

In this study, VRQ and LDE show meaningful explained variance (Table 9), suggesting that the proposed pathway of adoption and instructional outcomes is not trivial. Conceptually, this supports a “quality-through-use” argument that is particularly relevant to visual pedagogy in Ghana: AI-generated visuals can enhance lesson delivery when teachers are not only willing to use them (acceptance and uptake), but also able to judge quality and adapt outputs to local culture and curriculum expectations. That teacher judgement component is consistent with work showing that visual aids support learning most effectively when educators can select, adapt and contextualise materials rather than use them uncritical (Al-Barakat, 2023)

5.3 Necessary condition analysis: identifying what must be in place for high outcomes

A major contribution of the manuscript is the integration of Necessary Condition Analysis (NCA) with PLS-SEM (section 4.8). Whereas, SEM tests sufficiency logic (whether higher X tends to produce higher Y reaches a minimum threshold). This combine approach is increasingly encouraged because it reveals bottlenecks that remain invisible when researchers only interpret path coefficients (Dul, 2016; Dul et al., 2020; Richter et al., 2020)

In the NCA results (Table 13), key predictors functioned as necessary conditions for achieving high levels of the focal outcomes, reinforcing a practical: some factors are not merely helpful, they are non-negotiable minimums. The bottleneck tables (Tables 14–16) make this particularly actionable by

translating necessity into “minimum required levels”. For instance, the bottleneck patterns show that very high LDE requires minimum thresholds across multiple conditions, including VRQ and usability-related factors, meaning that even motivated users cannot achieve top-tier delivery outcomes if quality and baseline usability perceptions are too low. Likewise, high VRQ depends on meeting minimum levels across acceptance-related inputs. These findings align with the basic NCA interpretation: achieving excellence is constrained by the weakest binding condition, so interventions should prioritize lifting the limiting factor first (Dul, 2020).

Crucially, the necessity results strengthen the manuscript’s relevance. In teacher education, it is common to focus on raising mean scores through workshops and trainings. The NCA results suggest a more disciplined strategy: identify which conditions are currently below the “must-have” threshold for high outcomes and invest there. That is a more traditional and practical approach to professional preparation because it mirrors how craft knowledge develops: you secure the basis before expecting mastery.

5.3 Programme and Gender differences: Targeted support rather than one-size-fits-all

Group comparisons (section 4.9) indicate that programme context matters. Primary Education students recorded consistently lower mean scores across SE, CRTSE, SU, TA, VRQ and LDE relative to Early Grade and JHS cohorts (Table 21). Although, most programme-level effects are small (Table 24), LDE reaches a medium effect magnitude, suggesting a practically meaningful difference in perceived lesson delivery effectiveness by programme (Table 24). This pattern implies that programme structures, teaching cultures and practicum expectations may shape how confidently students adopt and translate AI-generated visuals into effective lesson delivery. Post-hoc comparisons (Section 4.9 tables) further imply that the largest gaps often involve Primary Education relative to the other cohorts, which is consistent with the descriptive pattern in Table 21. Gender patterns show males scoring higher on all constructs (Table 22), with the most noticeable practical gap on TA (Table 23), where the effect size is in the small to moderate range, while other differences are small (Table 23). These resonate with earlier work in teacher technology acceptance reporting that gender differences can appear in acceptance-related beliefs, although effects are context-dependent and should not be treated as fixed traits (Teo, 2015). Interpreting this within a culturally responsive teacher education frame, the implication is not “male students are better with technology”, but that access, encouragement, prior exposure and confidence-building opportunities may not be evenly distributed. A fair response is to build structured, scaffolded practice and a peer-supported studio routines that normalize experimentation for all students especially in cohorts showing lower acceptance or usability perceptions.

5.5 Theoretical and Practical implications for teacher education in Ghana

Theoretically, the results support extending acceptance-oriented explanations by explicitly incorporating culturally responsive confidence as a meaningful driver of acceptance and use in context where culturally-grounded representation is pedagogically central. These matters because AI tools are often on

global datasets that can under-represent local Ghanaian cultural forms, so teacher confidence in culturally responsive adaptation becomes part of the adoption story, not an add-on. Recent teacher education scholarship also emphasises that AI literacy should include critical evaluation, ethics and culturally grounded judgements, not just tool operation (Ng et al., 2021; Dilek, 2025)

Practically, pre-service teachers should be guided in the use of prompt refinement, critical evaluation of AI-generated content and adaptation of visuals to reflect Ghanaian cultural contexts. System usability emerged as a powerful driver of acceptance and instructional outcomes, which means that institutions adopting AI tools prioritize platforms that are intuitive, accessible on low cost devices and supportive of intermittent internet connectivity. The strong effect of visual resource quality on lesson delivery suggests that generative AI can play an important role in overcoming resource scarcity in Ghanaian colleges of education provided that teachers receive training in the evaluating and improving AI outputs. Overall, AI-integration strategies in the Global South should balance technological capability with cultural authenticity, ease of use and ongoing pedagogical support.

5.6 Limitations and Future research

The findings should be interpreted with appropriate caution. First, the design is cross-sectional, so the structural paths support explanatory associations rather than strong causal claims. Longitudinal or Quasi-experimental designs could test whether improvements in usability or culturally responsive confidence produce sustained increases in acceptance, use and lesson delivery outcomes over time. Second, self-report data can inflate relationships through common method variance; while procedural and statistical checks can reduce this concern, future work could triangulate with classroom observations, artefact scoring of generated visuals, and practicum-based performance measures. (Kock, 2015; Tehseen et al., 2017; Podsakoff et al., 2024).

Finally, future research should deepen the Ghana-specific contribution by examining how culturally localized prompting supports (for example, curated Ghanaian cultural motifs and curriculum-aligned prompt libraries) change both the quality of AI visuals and the confidence with which teachers use them. This would respond directly to the broader higher education AI literature's call to move from novelty adoption toward disciplined, context-sensitive pedagogical integration (Crompton & Burke, 2023; Zawacki-Richter et al., 2019; Wang et al., 2024).

Declarations

Author Contribution

M.N. conceptualized the study and wrote the main manuscript. H.B.E., A.T-M, P. S., C.A.K., J.A-K., E.B.A. and G.A.A. contributed to writing, editing, and manuscript refinement. M.Ny. contributed to writing, review, and revision of the manuscript. All authors reviewed and approved the final version of the manuscript.

Acknowledgement

The authors thank the participating pre-service teachers and Bagabaga College of Education for their support during the study.

References

1. Akanzire, B. N., Nyaaba, M., & Nabang, M. (2025). Generative AI in teacher education: Teacher educators' perception and preparedness. *Journal of Digital Educational Technology*, 5(1), ep2508. <https://doi.org/10.30935/jdet/15887>
2. AlAli, A. H., & Al-Barakat, A. A. (2023). Visual aids in education: Effectiveness and teacher perspectives. *Journal of International Students*.
3. Alejandro, I. M. V., Sanchez, J. M. P., Sumalinog, G. G., Mananay, J. A., Goles, C. E., & Fernandez, C. B. (2024). Pre-service teachers' technology acceptance of artificial intelligence (AI) applications in education. *STEM Education*, 4(4), 445–465. <https://doi.org/10.3934/steme.2024024>
4. Chahar Mahali, S., & Sevigny, P. R. (2022). Multicultural classrooms: Culturally responsive teaching self-efficacy among a sample of Canadian preservice teachers. *Education and Urban Society*, 54(8), 946–968. <https://doi.org/10.1177/00131245211062526>
5. Conner, C. M. (2024). *Exploring Culturally Responsive Teaching Self-Efficacy: A review of the psychometric properties of the Culturally Responsive Teaching Self-Efficacy Scale - ProQuest*. <https://www.proquest.com/openview/3957419fe5820ca484a0c439dfd7d61f/1?pq-origsite=gscholar&cbl=18750&diss=y>
6. Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(22). <https://doi.org/10.1186/s41239-023-00392-8>
7. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
8. Dilek, M., Baran, E., & Aleman, E. (2025). AI Literacy in Teacher Education: Empowering Educators Through Critical Co-Discovery. *Journal of Teacher Education*, 76(3), 294–311. <https://doi.org/10.1177/00224871251325083>
9. Dul, J. (2016). Necessary condition analysis (NCA): Logic and methodology of necessary but not sufficient causality. *Organizational Research Methods*, 19(1), 10–52. <https://doi.org/10.1177/1094428115584005>
10. Dul, J., & Kuik, R. (2020). A Statistical Significance Test for Necessary Condition Analysis. *Organizational Research Methods*. <https://doi.org/10.1177/1094428118795272>
11. Dul, J., Hauff, S., & Bouncken, R. B. (2023). Necessary condition analysis (NCA): review of research topics and guidelines for good practice. *Rev Manag Sci*, 17, 683–714. <https://doi.org/10.1007/s11846-023-00628-x>

12. Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., et al. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *Int J Educ Technol High Educ*, 19, 57. <https://doi.org/10.1186/s41239-022-00362-6>
13. Eyal, L. (2025). Rethinking artificial-intelligence literacy through the lens of teacher educators: The adaptive AI model. *Computers and Education Open*, 9, 100291. <https://doi.org/10.1016/j.caeo.2025.100291>
14. Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice* (2nd ed.). Teachers College.
15. Green, B., & Viljoen, S. (2020, January). Algorithmic realism: expanding the boundaries of algorithmic thought. In Proceedings of the 2020 conference on fairness, accountability, and transparency (pp. 19–31). <https://doi.org/10.1145/3351095.3372840>
16. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
17. Hidayat-ur-Rehman, I., & Ibrahim, Y. (2024). Exploring factors influencing educators' adoption of ChatGPT: a mixed method approach. *Interactive Technology and Smart Education*, 21 No, 499–534. <https://doi.org/10.1108/ITSE-07-2023-0127>. 4.
18. Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European journal of education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
19. Holtom, B., Baruch, Y., Aguinis, H., & Ballinger, A., G (2022). Survey response rates: Trends and a validity assessment framework. *Human Relations*, 75(8), 1560–1584. <https://doi.org/10.1177/00187267211070769>
20. Ialuna, F., Civitillo, S., Schachner, M. K., & Jugert, P. (2024). Culturally responsive teaching self-efficacy and cultural diversity climate are positively associated with the academic and psychological adjustment of immigrant and nonimmigrant students. *Cultural Diversity & Ethnic Minority Psychology*, 31(4), 683–698. <https://doi.org/10.1037/cdp0000697>
21. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
22. Khan, G. F., Sarstedt, M., Shiau, W. L., Hair, J. F., Ringle, C. M., & Fritze, M. P. (2019). Methodological research on partial least squares structural equation modeling (PLS-SEM). *Internet Research*, 29(3), 407–429. <https://doi.org/10.1108/IntR-12-2017-0509>
23. Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, 11(4), 1–10. <https://doi.org/10.4018/ijec.2015100101>
24. Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610. <https://doi.org/10.1177/001316447003000308>
25. Lewis, J. R. (2018). The System Usability Scale: Past, present, and future. *International Journal of Human–Computer Interaction*, 34(7), 577–590. <https://doi.org/10.1080/10447318.2018.1455307>

26. Li, M., Enkhtur, A., Yamamoto, B. A., Cheng, F., & Chen, L. (2025). Potential societal biases of ChatGPT in higher education: A scoping review. *Open Praxis*, 17(1), 79–94.
10.55982/openpraxis.17.1.750
27. Mahali, S. C., & Sevigny, P. R. (2022). Multicultural Classrooms: Culturally Responsive Teaching Self-Efficacy among A Sample of Canadian Preservice Teachers. *Education and Urban Society*.
<https://doi.org/10.1177/00131245211062526>
28. Mayer, R. E. (2014). Multimedia Instruction. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology*. Springer.
https://doi.org/10.1007/978-1-4614-3185-5_31
29. Mensah, F. S. (2025). Reclaiming agency in instructional technology integration: a praxeological analysis of pre-service mathematics teachers' roles in a GeoGebra-based model lesson. *Front Educ*, 10, 1670170. 10.3389/feduc.2025.1670170
30. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
<https://doi.org/10.1016/j.caeai.2021.100041>
31. Nikou, S. A. (2024). Factors influencing student teachers' intention to use mobile augmented reality in primary science teaching. *Educ Inf Technol*, 29, 15353–15374. <https://doi.org/10.1007/s10639-024-12481-w>
32. Nyaaba, M., & Zhai, X. (2025). Developing a theory-grounded AI tool for the generation of culturally responsive lesson plans. *Computers and Education Artificial Intelligence*, 9, 100474.
<https://doi.org/10.1016/j.caeai.2025.100474>
33. Podsakoff, P. M., Podsakoff, N. P., Larry, J., Williams, C., & Huang, J. Yang (2024). Common Method Bias: It's Bad, It's Complex, It's Widespread, and It's Not Easy to Fix. *Annual Review Organizational Psychology and Organizational Behavior*, 11, 17–61. <https://doi.org/10.1146/annurev-orgpsych-110721-040030>
34. Richter, N. F., Schubring, S., Hauff, S., Ringle, C. M., & Sarstedt, M. (2020). When predictors of outcomes are necessary: Guidelines for the combined use of PLS-SEM and NCA. *Industrial management & data systems*, 120(12), 2243–2267.
35. Roncal-Belzunce, V., Gutiérrez-Valencia, M., Martínez-Velilla, N., & Ramírez-Vélez, R. (2025). System Usability Scale for Gamified E-Learning Courses: Cross-Cultural Adaptation and Measurement Properties of the Spanish Version. *International Journal of Human–Computer Interaction*, 41(24), 15750–15760. <https://doi.org/10.1080/10447318.2025.2499947>
36. Runge, I., Hebibi, F., & Lazarides, R. (2025). Acceptance of pre-service teachers towards artificial intelligence (AI): The role of AI-related teacher training courses and AI-TPACK within the technology acceptance model. *Education Sciences*, 15(2), 167.
37. Spasopoulos, T., Sotiropoulos, D., & Kalogiannakis, M. (2025). Generative AI in Pre-Service Science Teacher Education: A Systematic Review. *Advances in Mobile Learning Educational Research*, 5(2), 1501–1523. <https://doi.org/10.25082/AMLER.2025.02.007>

38. Tehseen, S., Ramayah, T., & Sajilan, S. (2017). Testing and controlling for common method variance: A review of available methods. *Journal of Management Sciences*, 4(2), 146–175.
39. Teo, T. (2015). Technology acceptance among pre-service teachers: Does gender matter? *Australasian Journal of Educational Technology*, 31(3). <https://doi.org/10.14742/ajet.1672>
40. Tondeur, J., Scherer, R., Siddiq, F., et al. (2020). Enhancing pre-service teachers' technological pedagogical content knowledge (TPACK): a mixed-method study. *Educational Technology Research And Development*, 68, 319–343. <https://doi.org/10.1007/s11423-019-09692-1>
41. Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems With Applications*, 252, 124167. <https://doi.org/10.1016/j.eswa.2024.124167>
42. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16, 39. <https://doi.org/10.1186/s41239-019-0171-0>
43. Zhang, Z., Maeda, Y., Newby, T., Cheng, Z., & Xu, Q. (2023). The effect of preservice teachers' ICT integration self-efficacy beliefs on their ICT competencies: The mediating role of online self-regulated learning strategies. *Computers & Education*, 193, 104673. <https://doi.org/10.1016/j.compedu.2022.104673>

Figures

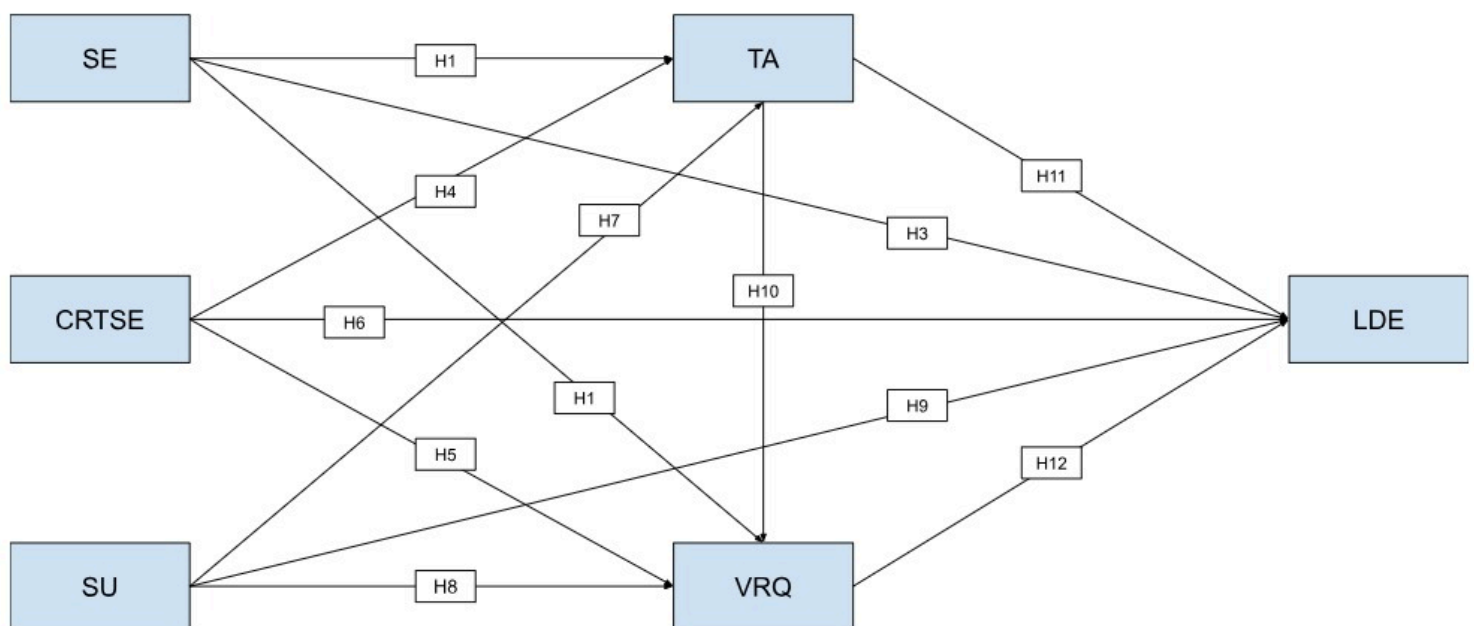


Figure 1

Conceptual/Structural Model

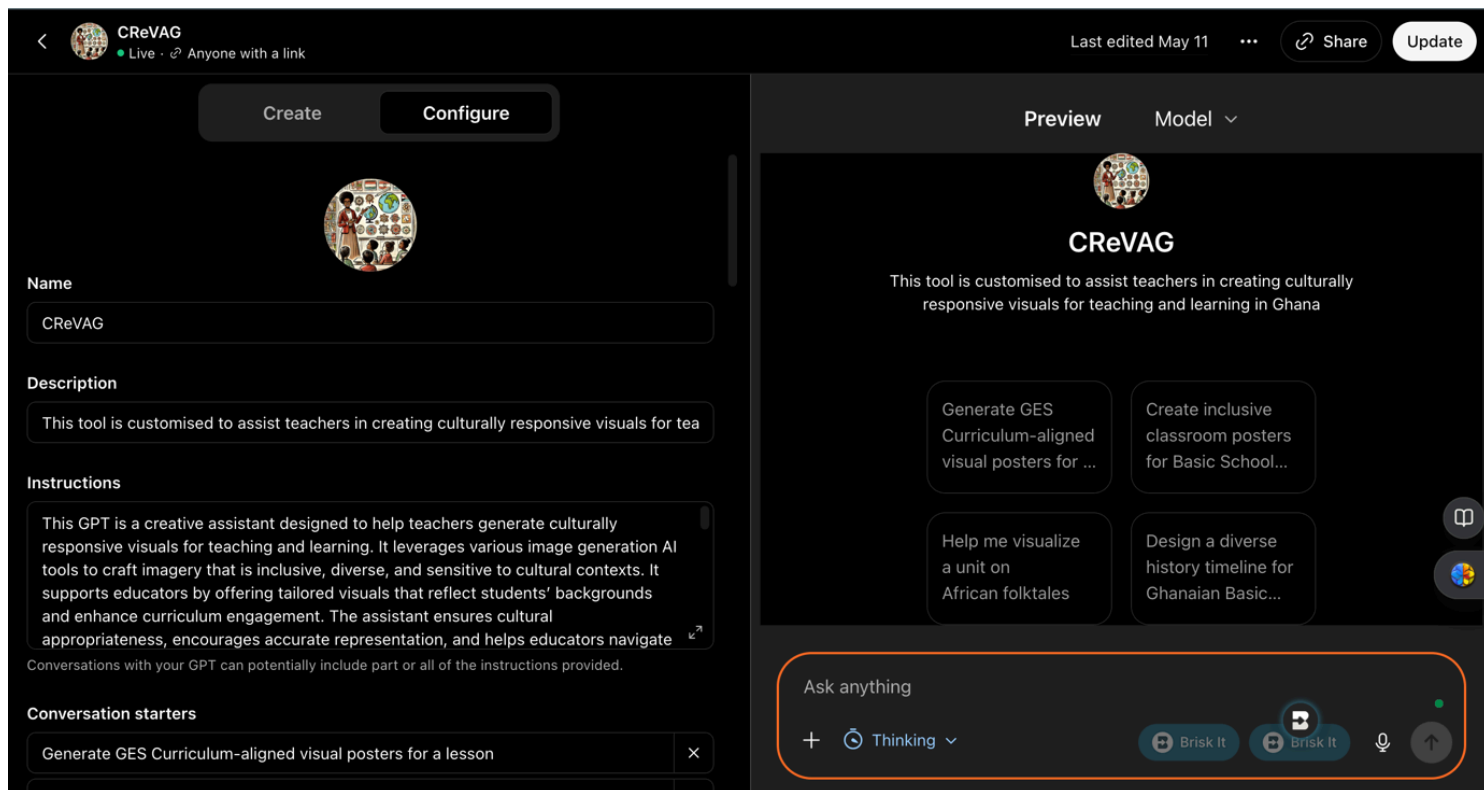


Figure 2

The Configuration panel and the user interface

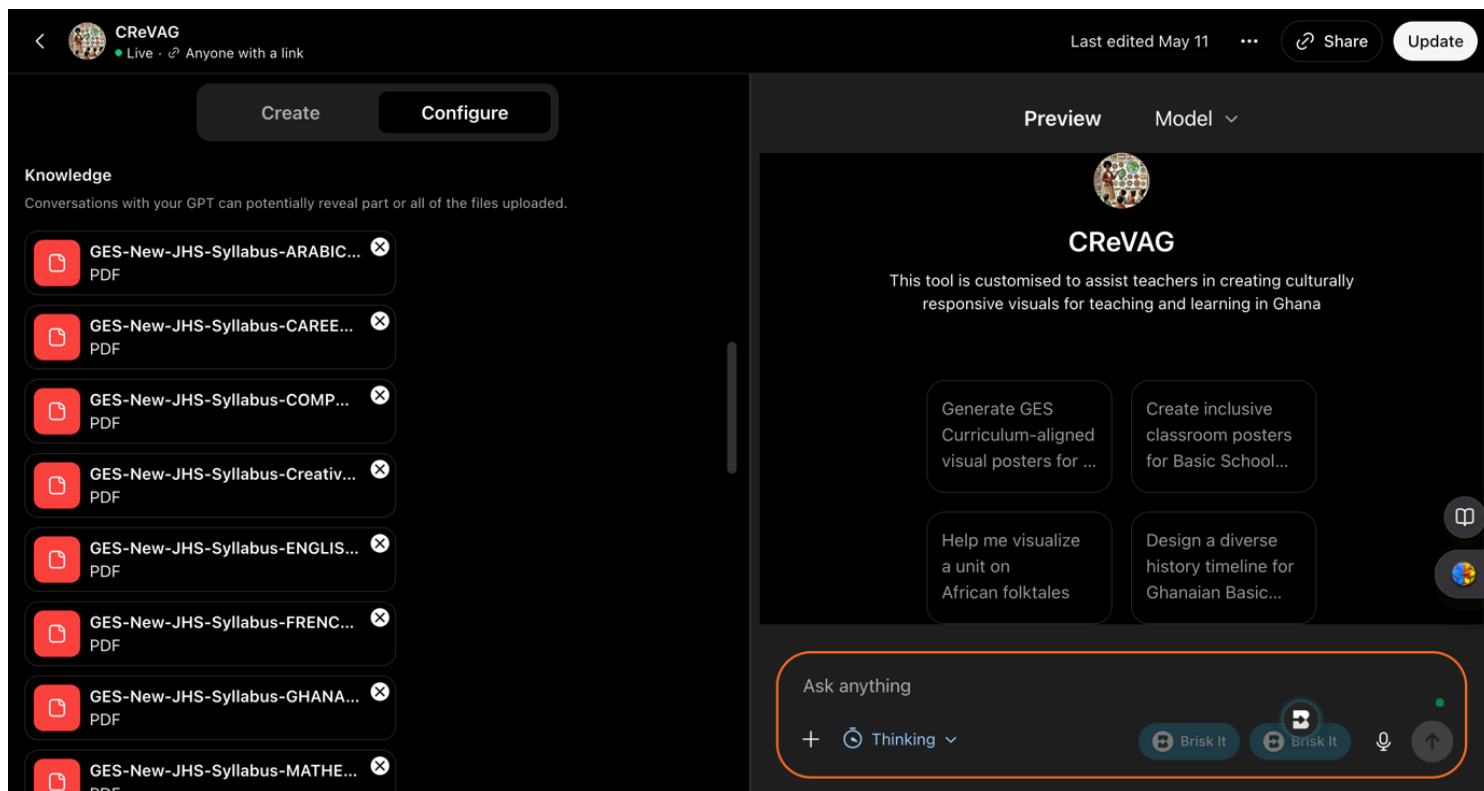


Figure 3

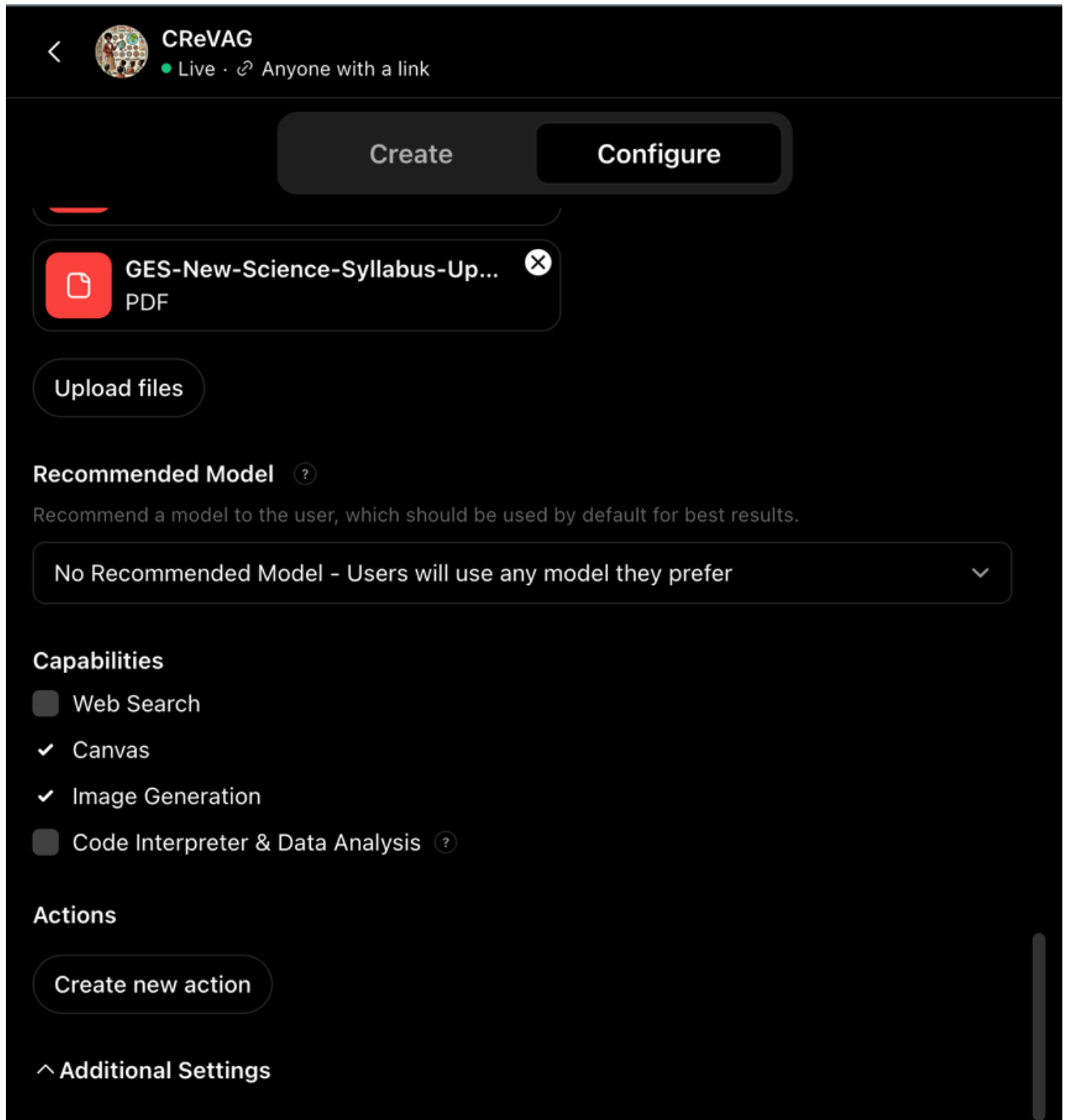


Figure 4

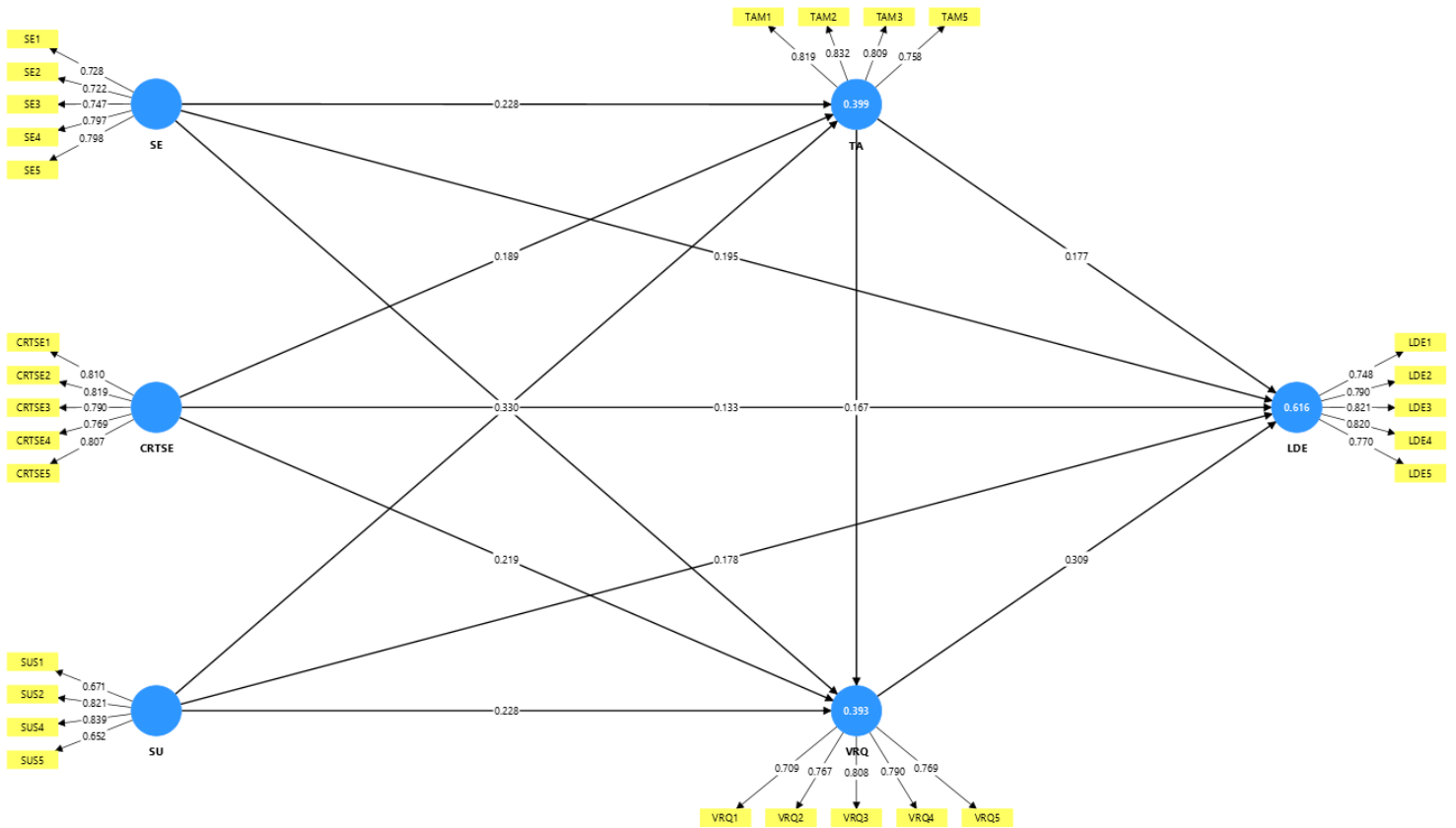


Figure 5

Assessment Model