AI Augmented Mascot Design Workflow for Digital Learning Media with Collaborative Intelligence

Manissaward Jintapitak*

College of Arts, Media and Technology, Chiang Mai University, Chiang Mai, 50200, Thailand

* Corresponding author. E-mail address: manissaward.j@cmu.ac.th

Received: 16 December 2025; Revised: 18 March 2026; Accepted: 25 March 2026; Available online: 26 March 2026

Abstract

Mascot characters are increasingly central to digital communication and learning environments, yet their creation remains dominated by labor-intensive manual workflows. This study investigates whether generative AI can enhance mascot design while preserving coherent character identity. A traditional manual workflow was compared with an AI-assisted collaborative workflow that employed Gemini AI, a structured prompting protocol. This model utilized clear role separation, where AI supported ideation while humans retained identity control and final decision-making. Ten evaluators rated outputs from both workflows on identity coherence, emotional clarity, visual appeal, and variation richness. Results showed that AI assistance substantially increased exploratory breadth and stylistic diversity, yielding significantly higher scores for variation richness and near-significant gains in visual appeal, while identity coherence and emotional clarity remained comparable to the manual condition. Correlation analyses further indicated that greater variation was positively associated with visual appeal. However, it was only weakly related to identity stability, suggesting that AI-generated diversity did not fragment character meaning under human oversight. Overall, the findings support a human-centered collaborative-intelligence framework in which generative AI functions as an exploratory partner rather than a replacement for designers. The proposed workflow offers practical guidance for integrating AI into character and mascot development, with promising implications for branding and educational media.

Keywords: Artificial Intelligence, Generative Image Synthesis, Collaborative Intelligence, Mascot Design, Digital Learning Media

Introduction

Mascots and character-based visual agents play an increasingly significant role in digital communication, branding, and educational media. In contemporary digital environments, mascots function not merely as decorative illustrations but as communicative interfaces that mediate identity, emotion, and meaning between content creators and audiences. Within digital learning platforms in particular, mascots are frequently deployed as pedagogical agents to sustain learner engagement, personalize content delivery, and support affective connection with learning material—functions that make their design quality directly relevant to user experience and learning effectiveness (Lundy, 2008; Mayer, 2009; Moreno & Mayer, 2007). As digital learning environments grow in scale and complexity, the demand for diverse, adaptive, and contextually responsive character assets has increased accordingly. However, despite their growing pedagogical and communicative importance, mascot design workflows remain largely dependent on traditional manual processes that are time-intensive, skill-dependent, and difficult to scale within the production constraints of rapidly evolving educational media contexts.

From a design perspective, mascot and character creation are fundamentally concerned with identity construction, emotional expressiveness, and narrative coherence. Prior studies emphasize that effective character design relies on clear identity cues, silhouette readability, gesture clarity, and emotional resonance to ensure recognizability and communicative power across platforms (Bancroft, 2006; White, 2006). Research in

branding and marketing further positions mascots as symbolic figures that embody organizational values and shape audience perception through visual and affective cues (Brown & Ponsonby-McCabe, 2014). Cognitive and perceptual studies also demonstrate that viewers rapidly extract emotional and structural information from simplified visual forms, highlighting the importance of expressive clarity and proportion in character interpretation (Smith & Rossit, 2018). In digital media environments, mascots additionally operate as narrative agents that support storytelling and self-representation, enabling users to engage with abstract ideas and identities through embodied visual narratives (Lundy, 2008). Taken together, these characteristics underscore the conceptual depth required in mascot design while simultaneously revealing the limitations of manual workflows in producing diverse and adaptive character variations within constrained production timelines (Jintapitak, 2023).

To manage this complexity, Design Thinking has been widely adopted as a creative framework for character and mascot development. Rooted in Design Thinking, the current workflow utilizes iterative prototyping to harmonize functional goals with expressive identity (Cross, 2011; Curedale, 2017), ensuring a human-centric approach to AI-assisted design. Within mascot design practice, this approach facilitates the translation of persona attributes, narrative intent, and emotional goals into coherent visual forms through successive refinement cycles. Previous work on the IRIS personal character project demonstrates how Design Thinking-based workflows can produce identity-consistent and narratively grounded mascot designs, while also revealing the time and labor demands associated with manual iteration (Jintapitak, 2023). These constraints become particularly pronounced in digital learning and communication contexts that require rapid content production and stylistic flexibility.

Recent advances in generative artificial intelligence have introduced new possibilities for visual design workflows. The emergence of diffusion-based architectures has shifted the design paradigm from manual execution to prompt-based iteration. This transition allows rapid synthesis of complex visual concepts, significantly lowering the barrier to iterative conceptualization and stylistic experimentation (Ramesh et al., 2021; Rombach et al., 2022). The emergence of multimodal systems further enhances this capability by integrating linguistic, visual, and contextual cues within a unified generative process, supporting greater coherence across iterative outputs (Google DeepMind, 2023). Beyond technical performance, studies on AI-generated imagery suggest that such systems can significantly influence aesthetic perception and visual appeal, particularly when outputs maintain internal consistency and compositional quality (Cha & Wang, 2025; Bianchi et al., 2025). These developments indicate that generative AI has the potential to expand creative exploration and accelerate early-stage prototyping in mascot design. Nevertheless, the practical value of AI-generated outputs in specialized design contexts—particularly those requiring cultural specificity, narrative consistency, and pedagogical appropriateness—remains an open empirical question that warrants systematic investigation.

Nevertheless, the integration of generative AI into creative practice raises important challenges. Human-centered AI research emphasizes that generative systems should augment rather than replace human judgment, particularly in tasks involving cultural interpretation, emotional nuance, and ethical responsibility (Shneiderman, 2020). Studies on AI co-creative tools further show that while AI excels at generating variation and supporting divergent thinking, human designers remain essential for maintaining conceptual coherence, narrative integrity, and contextual appropriateness (Ning et al., 2023; Kadenhe et al., 2025). Additional concerns related to bias, representational fairness, and authorship highlight the need for structured frameworks that clearly define the roles of human and artificial agents in collaborative design processes (Samala & Rawas, 2025).

Despite growing interest in AI-assisted creativity, existing research has not yet articulated a unified methodology that integrates mascot design theory, Design Thinking principles, multimodal generative AI, and systematic evaluation within a single workflow. In particular, there is limited empirical work examining how AI can be embedded into established mascot design processes in ways that preserve identity coherence and narrative meaning while simultaneously addressing the scalability and diversity demands characteristic of digital learning media production. This gap is practically significant: instructional designers and educators who develop character-based learning content frequently face resource constraints that limit the number of viable design variations that can be produced manually, yet the quality and expressiveness of mascot assets directly influence learner engagement and the effectiveness of character-based pedagogical communication (Mayer, 2009; Veletsianos, 2010).

To address this gap, this study proposes an AI-enhanced mascot design framework grounded in collaborative intelligence. Using the IRIS personal character as a baseline, the research compares a traditional manual design workflow with an AI-assisted pipeline employing Gemini AI. The study pursues three explicit objectives: (1) to examine how generative AI influences stylistic diversity and perceived design quality in mascot development; (2) to clarify the complementary roles of human designers and AI systems within a structured collaborative workflow; and (3) to evaluate the extent to which AI assistance can expand creative exploration without compromising core character identity and emotional clarity. By formalizing a collaborative intelligence approach to mascot design, this research contributes a scalable and human-centered framework applicable to digital learning media, branding, and contemporary visual communication.

To establish a reference point for subsequent comparison, the traditional mascot design process used in the IRIS character project is summarized in Fig. 1. The workflow illustrates the sequential and labor-intensive nature of manual character development, beginning with briefing and visual research, followed by hand-drawn sketching, refinement, and final rendering. The progression from early two-dimensional (2D) sketches to fully rendered three-dimensional (3D) models highlights both the creative depth and the substantial time demands of traditional mascot workflows, underscoring the need for more scalable and flexible design approaches.

The Fig. 1 illustrates the sequential stages of the human-only mascot design process, including briefing, reference research, sketching, refinement, and final rendering. This workflow serves as the baseline condition for comparison with the proposed AI-enhanced collaborative intelligence framework.

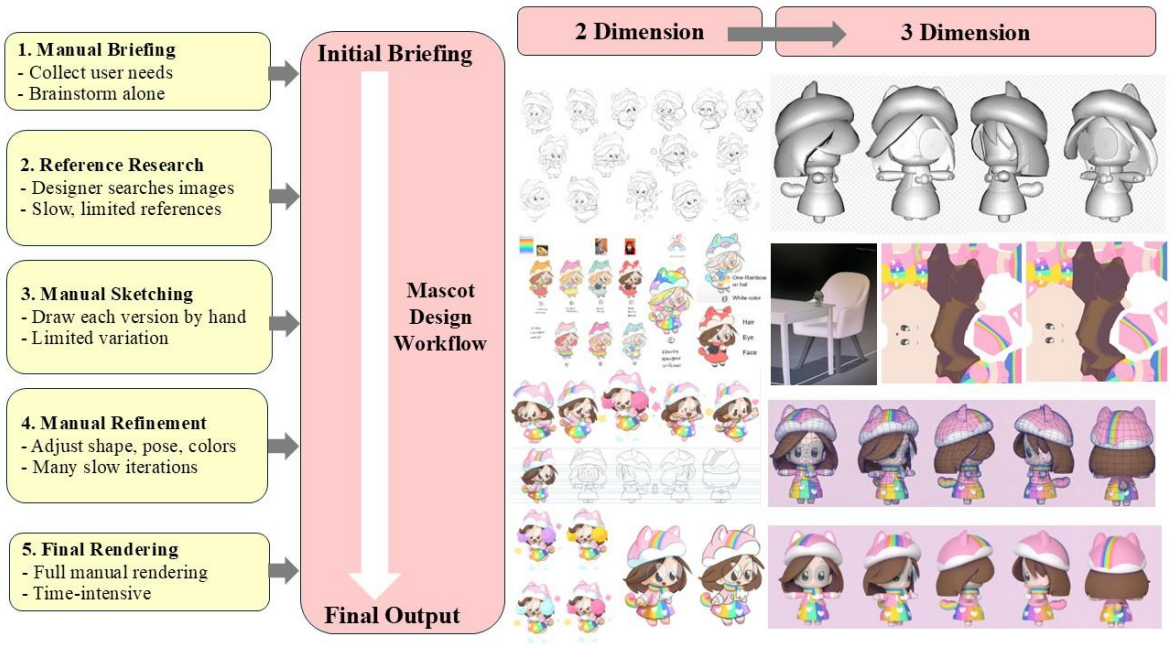

Figure 1 Traditional Manual Mascot Design Workflow for the IRIS Character (2D-3D Pipeline)

Materials and Methods

This study used a qualitative-experimental hybrid design to develop and evaluate an AI-enhanced mascot design framework grounded in collaborative intelligence. A comparative exploratory approach was adopted by examining two design conditions: (1) a traditional manual mascot design workflow and (2) an AI-assisted workflow integrating Gemini AI. The IRIS personal character project served as a common design brief to ensure comparability across conditions. Data sources comprised (1) design artifacts generated in both workflows, (2) workflow process records, including iteration time and the number of viable design variations, and (3) human evaluator ratings assessing identity coherence, emotional clarity, visual appeal, and variation richness.

A. Study Design

The study was designed as a comparative exploratory investigation aimed at examining how generative artificial intelligence can augment key stages of mascot development. Rather than testing causal hypotheses, the research focused on identifying differences in workflow efficiency, stylistic and conceptual variation, and perceived design quality between manual and AI-assisted design processes.

Two parallel design conditions were established. The baseline condition followed a conventional human-only mascot design workflow, encompassing briefing, visual research, manual sketching, refinement, and final rendering. The experimental condition restructured this workflow into an AI-assisted pipeline in which generative AI supported ideation and visual exploration while human designers retained conceptual control and decision-making authority.

To ensure methodological consistency, both conditions were aligned to the same character persona, narrative intent, and visual constraints defined by the IRIS design brief. Comparative analysis was conducted using design artifacts, workflow records, and evaluator assessments collected under identical evaluation conditions.

B. Materials and Data Sources

Three categories of materials and data sources were employed in this study.

First, manual design artifacts were collected from the baseline mascot design workflow. These artifacts included hand-drawn sketches, refined line drawings, colored turnarounds, texture maps, and finalized 3D character models produced through conventional manual processes. The artifacts represent sequential stages of mascot development and reflect the iterative nature of human-only design practice.

Second, AI-assisted design artifacts were generated during the experimental workflow using generative artificial intelligence. The study analyzed a multistage dataset of AI-generated visual outputs produced over iterative design cycles. These materials encompassed initial conceptual variations, refined character propositions, and high-fidelity prototype-level representations, documenting the evolution of character identity through the generative process. To maintain methodological parity, all AI-generated artifacts were synthesized under the identical design briefs and identity constraints utilized in the manual workflow. This alignment was strictly enforced to ensure comparability between conditions and isolate the impact of the generative process on the final output.

Third, evaluation data were collected from human evaluators using structured rating instruments. These data comprised quantitative ratings across predefined evaluation dimensions, including identity coherence, emotional clarity, visual appeal, and variation richness, for both manual and AI-assisted design outputs.

All design artifacts and evaluation data were generated specifically for this study and were not derived from external datasets. The materials were prepared and presented under consistent conditions to support reliable comparative analysis.

C. AI System and Settings

The AI-augmented workflow utilized Gemini AI, a multimodal generative system developed by Google DeepMind (2023). This architecture was selected for its sophisticated ability to synthesize textual descriptors and visual cues into high-fidelity outputs, facilitating a more nuanced and responsive character design process than unimodal systems. Unlike conventional diffusion-based text-to-image models such as Stable Diffusion (Rombach et al., 2022) or DALL·E (Ramesh et al., 2021), which operate primarily through latent diffusion processes optimized for photorealistic or stylized image synthesis, Gemini's multimodal architecture supports tighter integration of linguistic context and visual generation within a unified interactive environment. This characteristic facilitates an iterative design architecture that prompts refinement, contextual continuity, and responsive visual adjustment across successive exchanges. Such a conversational framework allows for a recursive loop between the designer's intent and the generative output. Additionally, Gemini's conversational interface allows the human designer to supply corrective and consistency-control prompts directly within the same session, supporting the structured human-AI interaction principles central to the collaborative intelligence framework adopted in this study.

To ensure results reflect accessible design workflows, Gemini AI was utilized in its standard production configuration. The study relied exclusively on direct inference without model fine-tuning or external dataset integration. All generative outputs were elicited through structured natural-language prompting executed by the human researcher. The system operated as a generative support tool rather than an autonomous decision-making agent.

Within the experimental workflow, Gemini AI was applied primarily during the ideation and refinement stages to generate multiple visual variations corresponding to the same character brief. The AI system did not automatically select or rank outputs. All generated images were reviewed, filtered, and curated manually by the human designer to ensure consistency with predefined character identity, narrative intent, and visual constraints.

No post-generation image editing, automated enhancement, or algorithmic modification was applied before evaluation. The final set of AI-assisted design artifacts used for analysis consisted solely of outputs selected through human judgment following iterative prompt refinement.

To illustrate the integration of generative AI into the mascot design process, an overview of the proposed AI-enhanced workflow is presented. Fig. 2 summarizes the collaborative intelligence framework adopted in this study, highlighting the distribution of roles between Gemini AI and human designers across key stages of mascot development.

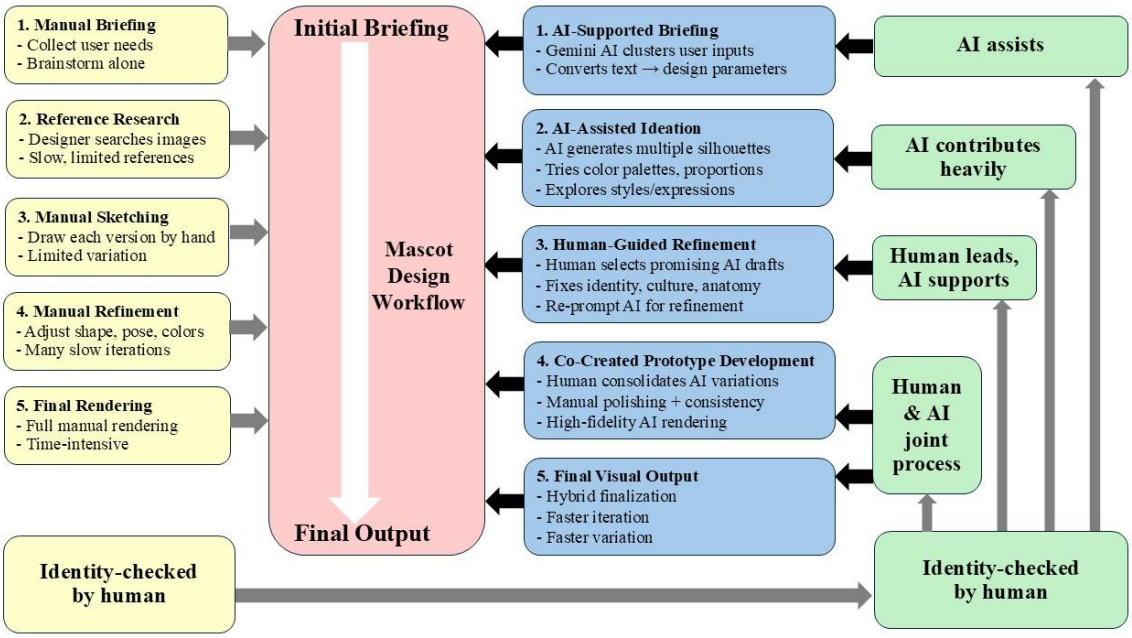

Figure 2 Proposed AI-enhanced mascot design workflow integrating Gemini AI and human expertise

The Fig. 2 illustrates a collaborative intelligence pipeline in which generative AI supports ideation, variation generation, and rendering, while human designers retain control over identity definition, refinement, and final decision-making. The detailed prompting strategy and step-by-step workflow procedures associated with this framework are presented below.

D. Prompting Protocol and Workflow Procedure

Building upon the previously defined system configuration, this section details the prompting protocol and procedural workflow utilized to operationalize the AI-assisted design process. This documentation ensures the reproducibility of the human-AI collaborative framework. A structured prompting protocol was developed to ensure that AI-generated outputs aligned with the predefined identity, narrative intent, and visual constraints of the IRIS character. The protocol was designed to support reproducibility, iterative refinement, and clear role separation between generative AI and human designers within the collaborative intelligence framework.

1. Prompt Structure (Table 1)

Each prompt followed a consistent four-component structure:

a. Character persona and narrative attributes; specifying core identity traits, personality, emotional tone, and narrative role of the character.

b. Visual and stylistic constraints, including references to shape language, color palette, proportion, pose characteristics, and overall stylistic direction.

c. Functional and emotional design goals; indicating the intended communicative purpose of the character and the emotional response to be conveyed.

d. Refinement instructions: providing corrective or selective guidance based on evaluation of previously generated outputs.

Table 1 Representative prompts used in the AI-assisted mascot design workflow.

Prompt type	Purpose in workflow	Example prompt
Ideation prompt (broad exploration)	Generate multiple stylistic and conceptual alternatives during early exploration.	"Create a friendly female mascot inspired by the IRIS concept, with soft rounded shapes, warm colors, and expressive eye gestures. The character should communicate empathy and curiosity, suitable for educational environments."
Refinement prompt (style narrowing)	Adjust and guide outputs toward improved fit with identity and design constraints.	"Keep the same character identity, face structure, and costume elements, but refine the proportions to appear slightly more youthful. Maintain consistency with the previous version while improving clarity and pose readability.
Consistency-control prompt	Maintain visual continuity across iterations and prevent identity drift.	"Regenerate the mascot while preserving the same hairstyle, costume details, and color palette. Avoid major structural changes. Keep proportions stable across outputs."

Note. Examples are illustrative and were adapted iteratively during the design process rather than used as fixed instructions.

These prompts were not treated as static instructions but as iterative tools that evolved alongside designer judgment. Together, they formed the foundation for the subsequent workflow stages and operationalized the collaborative intelligence framework adopted in this study.

2. Workflow Stages

The AI-assisted mascot design workflow consisted of five sequential stages, corresponding to the framework illustrated in Figure 2.

a. Stage 1: AI-Supported Briefing The initial design brief was translated into structured prompts reflecting the IRIS character's persona, narrative context, and visual identity. This stage established the conceptual boundaries within which AI-generated exploration was conducted.

b. Stage 2: AI-Driven Ideation and Variation Generation Using the structured prompts, Gemini AI generated multiple visual variations exploring alternative poses, silhouettes, facial expressions, color schemes, and stylistic interpretations. This stage emphasized breadth of exploration and rapid ideation, producing a diverse set of candidate designs within a short timeframe.

c. Stage 3: Human-Guided Refinement Human designers critically reviewed the AI-generated outputs to assess identity coherence, emotional clarity, and visual appropriateness. Selected outputs were retained, while others were discarded. Based on this evaluation, prompts were iteratively refined to correct deviations, emphasize desirable attributes, and progressively narrow the design space.

d. Stage 4: Co-Created Prototype Development and Consolidation Refined outputs were consolidated into a set of identity-consistent mascot proposals. This stage focused on stabilizing visual characteristics and preparing representative prototypes suitable for evaluation. Final selection decisions were made exclusively by human designers.

e. Stage 5: Final Visual Output Following prototype consolidation, selected mascot designs were prepared as final visual outputs representing the completed AI-assisted workflow. These outputs underwent a final human review to confirm alignment with the original design brief, character identity, and quality standards before being submitted for evaluator assessment.

3. Human-AI Interaction Principles

Throughout the workflow, generative AI functioned as a supportive ideation and visualization tool rather than an autonomous design agent. Human designers maintained control over conceptual framing, evaluative judgment, and final decision-making at all stages. No AI-generated output was accepted without human review and approval. Particular attention was paid to consistency across iterations: when AI outputs drifted from the established identity profile, prompts were adjusted or discarded rather than redefining the character concept.

4. Output Selection and Documentation

For each iteration cycle, workflow records were maintained, including the number of generated variations, selected candidates, and iteration duration. Only outputs that met predefined identity and quality criteria were included in the final evaluation dataset. These records supported subsequent comparative analysis between the manual and AI-assisted workflows.

E. Baseline (Manual) Condition

The baseline condition followed a traditional human-only mascot design workflow, as illustrated in Fig. 1. This workflow comprised sequential stages of briefing, visual research, manual sketching, refinement, and final rendering, reflecting established mascot design practice before the integration of generative AI.

Design development under this condition relied entirely on manual processes. Initial concepts were generated through individual brainstorming and user-need interpretation by the designer, followed by visual research to establish stylistic references. Design variations were produced through hand-drawn sketches and incremental refinements, requiring repeated adjustments to pose, silhouette, color, and proportion.

Refinement and finalization were conducted through successive manual iterations, progressing from two-dimensional line drawings to colored turnarounds, texture mapping, and three-dimensional modeling. Each iteration required direct human intervention, limiting the number of viable variations that could be developed within practical time constraints.

The baseline workflow served as the reference condition for comparative analysis, enabling examination of differences in workflow efficiency, stylistic diversity, and perceived design quality relative to the AI-assisted workflow.

F. Participants (Human Evaluators)

Ten human evaluators participated in the assessment of mascot design outputs generated under both the manual and AI-assisted conditions. All evaluators had prior experience or academic background in design-related fields, including visual communication, animation, and character or mascot design.

Participants were recruited using convenience sampling based on their familiarity with visual design evaluation and their ability to interpret character identity and expressive qualities. To ensure consistency, all evaluators assessed the same set of design outputs under identical viewing conditions.

Evaluations were conducted individually to avoid group influence or discussion effects. No personal identifying information was collected, and all responses were anonymized prior to analysis.

It should be acknowledged that a sample of ten evaluators represents a limited basis for inferential statistical analysis. This sample size was considered appropriate for the exploratory aims of the study; however, it constrains statistical power and the generalizability of any significance testing conducted. All quantitative findings derived from evaluator ratings should therefore be interpreted as preliminary observations rather than definitive conclusions, and replication with larger and more diverse evaluator samples is recommended in future research.

G. Measures and Instruments

A structured evaluation instrument was developed to assess and compare mascot design outputs produced under the manual and AI-assisted workflows. The instrument employed a five-point Likert scale (1 = very low, 5 = very high) and focused on four evaluation dimensions relevant to character and mascot design quality.

The first dimension, identity coherence, assessed the clarity and consistency of the character's core identity across visual representations, including recognizability of persona traits and stability of defining features. The second dimension, emotional clarity, evaluated the effectiveness of visual elements—such as facial expression, posture, and gesture—in conveying intended emotional cues. The third dimension, visual appeal, measured the perceived aesthetic quality of the design, including overall attractiveness, balance, and stylistic harmony. The fourth dimension, variation richness, assessed the perceived diversity and breadth of design alternatives generated within each workflow.

All evaluators independently rated each design output across the four dimensions using the same assessment instrument. Scores were recorded for subsequent quantitative comparison between the manual and AI-assisted conditions.

To assess the internal consistency of the evaluation instrument, Cronbach′s alpha was calculated across the four rating dimensions for both the manual and AI-assisted conditions. This measure was included to provide an indication of inter-rater reliability and to support transparent reporting of the measurement properties of the instrument, consistent with recommendations for small-sample evaluative research.

H. Data Analysis

Data analysis was conducted using non-parametric statistical methods appropriate for ordinal data and small sample sizes. Because the same ten evaluators assessed outputs under both conditions, the data constitute a paired design. Before comparative analysis, Shapiro–Wilk tests were conducted to assess the normality of rating distributions for each evaluation dimension. Results confirmed violations of normality across all four attributes (all $p < .001$), supporting the use of non-parametric procedures throughout.

For descriptive purposes, ratings are reported as Median (Mdn) and Interquartile Range (IQR), as these measures are more appropriate than means and standard deviations for ordinal Likert-scale data. Group comparisons between the manual and AI-assisted conditions were conducted using the Wilcoxon signed-rank test, which is suited to paired ordinal data without the assumption of normality. Effect sizes are reported as rank-biserial correlation (r), with values of .10, .30, and .50 interpreted as small, medium, and large, respectively (Cohen, 1992).

To examine associations among evaluation dimensions within the AI-assisted condition, Spearman's rank-order correlation (ρ) was computed, as this method is appropriate for ordinal data and does not assume a normal distribution. Internal consistency of the evaluation instrument was assessed using Cronbach's alpha across both conditions. All inferential results should be interpreted with caution, given the limited sample size (N = 10), which yields low statistical power; findings are presented as exploratory observations intended to identify patterns and trends rather than to establish definitive conclusions.

Results

This section presents the results of the comparative evaluation between the traditional manual mascot design workflow and the AI-assisted workflow. Results are reported across four evaluation dimensions assessed by human evaluators. Given the exploratory nature of this study and the small sample size (N = 10), results should be interpreted with appropriate caution regarding generalizability and statistical power.

A. Evaluation Results of Manual and AI-Assisted Workflows

Human evaluators rated design outputs generated under both workflows using the standardized assessment instrument described in the Materials and Methods section. Because the same 10 evaluators assessed outputs from both conditions, the data constitute a paired design. Given the ordinal nature of Likert-scale data and the results of Shapiro-Wilk tests, which indicated significant violations of normality across all four attributes ($p < .001$), non-parametric statistical methods were utilized for all subsequent analyses. Descriptive statistics are reported as Median (Mdn) and Interquartile Range (IQR), and group comparisons were conducted using the Wilcoxon signed-rank test. Effect sizes are expressed as rank-biserial correlation (r), with values of .10, .30, and .50 interpreted as small, medium, and large, respectively (Cohen, 1992).

The descriptive data (Table 2) indicate that the AI-assisted workflow achieved notably higher ratings in Variation Richness (Mdn = 5.0 vs. 3.0) and Visual Appeal (Mdn = 5.0 vs. 4.0). Scores for Identity Coherence and Emotional Clarity were comparable across conditions, suggesting that while AI excels in creative expansion, the human-led manual workflow remains highly effective for maintaining character consistency.

Statistical testing supported these observations (Table 2). A statistically significant difference was found for Variation Richness (W = 0.0, $p = .002$, $r = .886$), indicating a large effect and confirming that AI assistance substantially increased stylistic diversity. Visual Appeal showed a positive trend approaching significance (W = 0.0, $p = .063$, $r = .640$), with a large effect size suggesting practical relevance despite the absence of statistical significance at the conventional threshold—a pattern likely reflecting limited statistical power rather than the absence of a true effect. No significant differences were found for Identity Coherence (W = 2.5, $p = .625$, $r = .289$) or Emotional Clarity (W = 7.0, $p = .688$, $r = .232$), indicating that the AI-assisted workflow remained comparable to manual design in preserving core character meaning.

Table 2 Comparative ratings (Mdn [IQR]) between manual and AI-assisted workflows

Attribute	Manual Workflow	AI-Assisted Workflow	W	p	r
Identity Coherence	5.0[0.0]	5.0[1.0]	2.5	.625	.289
Emotional Clarity	5.0[0.75]	4.5[1.0]	7.0	.688	.232
Visual Appeal	4.0[1.0]	5.0[0.0]	0.0	.063	.640
Variation Richness	3.0[0.0]	5.0[0.0]	0.0	**.002***	.886

*Note. W = Wilcoxon signed-rank statistic. Effect size r = rank-biserial correlation. *p < .05. Results should be interpreted with caution due to the limited sample size (N = 10).*

B. Correlation Analysis of the AI-Assisted Workflow

To examine relationships among evaluation dimensions within the AI-assisted condition, Spearman's rank-order correlation (ρ) was computed, as this method is appropriate for ordinal data and does not assume normality. Given the small sample size (N = 10), all correlation values should be treated as preliminary and interpreted with caution; they describe associations within this specific sample only and do not imply causation.

A statistically significant positive correlation (Table 3) was observed between Variation Richness and Visual Appeal (ρ = .67, p = .035), indicating that greater stylistic diversity within the AI-assisted condition was associated with higher perceived attractiveness. The moderate positive correlation between Emotional Clarity and Visual Appeal (ρ = .50, p = .141) and a weak negative correlation between Variation Richness and Identity Coherence (ρ = −.27, p = .447) were not statistically significant at this sample size. The latter suggests that increased stylistic diversity did not substantially compromise character consistency within this collaborative framework, though this observation warrants confirmation in larger samples.

Table 3 Spearman rank-order correlations among evaluation dimensions for the AI-assisted workflow (N = 10)

Attribute	Identity Coherence	Emotional Clarity	Visual Appeal	Variation Richness
Identity Coherence	-	0.00	.10	-.27
Emotional Clarity	0.00	-	.50	.33
Visual Appeal	0.10	0.50	-	**.67***
Variation Richness	-0.27	0.33	**.67***	-

*Note. Values represent Spearman's ρ. *p < .05 (two-tailed). Given N = 10, all non-significant values are preliminary and may lack sufficient power to detect true effects.*

C. Workflow Output Characteristics

In addition to evaluator ratings, notable differences were observed in the characteristics of design outputs produced under each workflow. The manual workflow yielded a significantly constrained volume of fully refined design variations. This limitation is attributed to the sequential, labor-intensive nature of hand-drawn iteration, which creates an inherent trade-off between exploratory breadth and production time. In contrast, the AI-assisted workflow produced a substantially larger set of viable design alternatives within a shorter timeframe, reflecting increased exploration capacity.

These differences were most evident during the early ideation and refinement stages. AI-generated outputs enabled rapid exploration of alternative poses, silhouettes, and stylistic interpretations, allowing multiple design directions to be evaluated simultaneously. Manual outputs, while more time-intensive, demonstrated careful incremental refinement across successive stages, resulting in highly controlled and consistently polished designs.

To illustrate these differences, representative design outputs from both workflows are presented in Fig. 3.

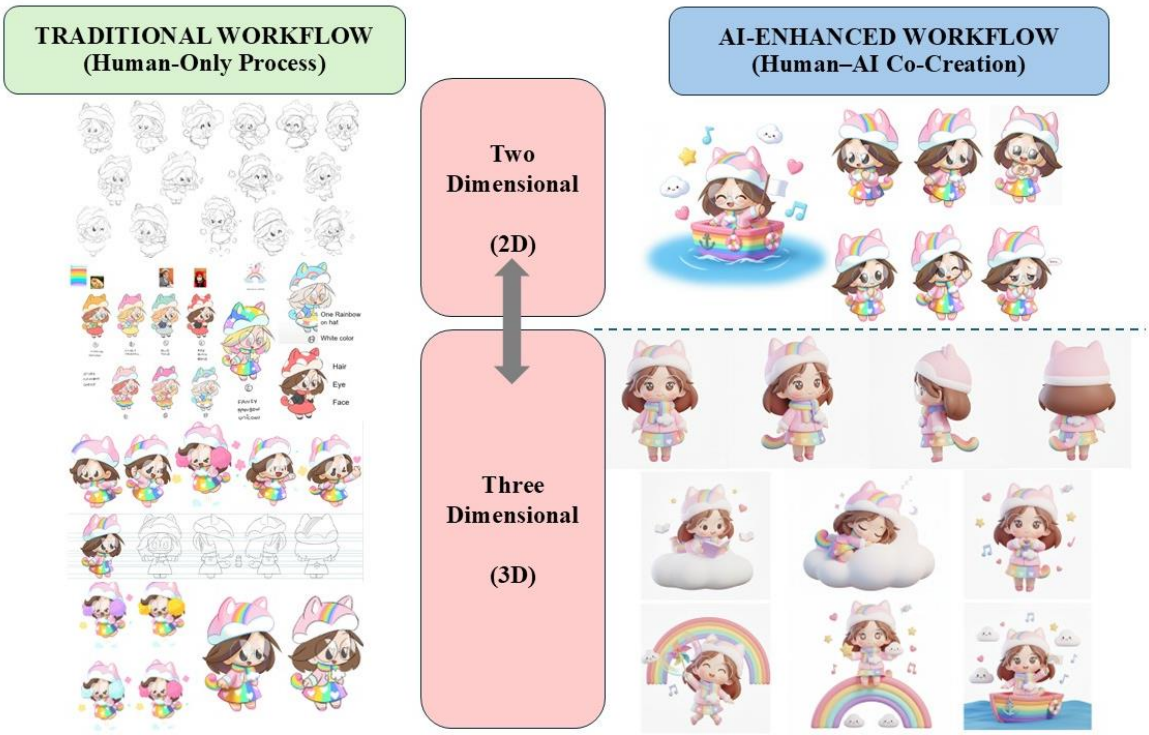

Figure 3 Representative design outputs from the traditional manual workflow and the AI-enhanced collaborative workflow

The Fig. 3 compares visual outputs across the two workflows. The traditional human-only workflow (left) progresses from manual sketching and sequential refinement to final 2D and 3D renderings. The AI-enhanced workflow (right) shows a broader range of stylistically consistent alternatives produced more rapidly through human–AI co-creation. Together, these examples visually illustrate the increased variation and exploratory breadth enabled by AI assistance, while maintaining recognizable character identity.

D. Results Summary

Overall, these preliminary findings suggest that the AI-assisted mascot design workflow enhanced exploratory breadth and visual variation while maintaining levels of identity coherence and emotional clarity comparable to those achieved through traditional manual processes. The statistically significant improvement in variation richness and the large effect size observed for visual appeal are particularly noteworthy, though both should be interpreted cautiously given the small sample size (N = 10) and the exploratory nature of this study. Taken together, these results offer initial empirical support for the potential of collaborative intelligence approaches to augment mascot design workflows without displacing the core role of human designers, although replication with larger and more diverse samples is needed before broader conclusions can be drawn.

Discussion

The findings of this study indicate that integrating generative AI into a Design Thinking–based mascot design workflow can enhance exploratory breadth and visual variation without compromising core identity coherence or emotional clarity. Compared with the traditional manual workflow, the AI-assisted workflow generated a greater number of viable design alternatives within shorter iteration cycles, while maintaining comparable evaluator ratings for identity-related dimensions. These results suggest that a collaborative intelligence approach extends

creative capacity rather than substituting manual expertise. Statistical analyses further elucidate this pattern: while visual appeal and variation richness were significantly augmented in the AI-assisted condition, identity coherence and emotional clarity remained remarkably stable. These results suggest that AI-enhanced exploration does not inherently dilute the conceptual anchor of the character when governed by a structured human-AI protocol.

These outcomes align with broader research on text-to-image generation, which highlights aesthetic quality, perceptual coherence, and output diversity as key performance indicators (Tian et al., 2025). The ability of generative models to rapidly propose multiple stylistically consistent directions from a single brief helps explain the elevated ratings for visual appeal and variation richness. At the same time, comparable identity and emotional scores demonstrate that increased variation did not fragment the character concept, supporting the view that structured prompting and human-guided filtering can stabilize meaning within AI-assisted workflows. These tendencies are visible in Fig. 3, where AI-assisted outputs show a broader stylistic and expressive range while retaining recognizable identity cues.

The correlation analysis offers additional insight. The significant positive relationship between variation richness and visual appeal ($\rho = .67$, $p = .035$) suggests that diversity did not simply multiply alternatives but actively contributed to perceived attractiveness. Conversely, the weak and slightly negative relationship between variation richness and identity coherence ($\rho = -.27$, $p = .447$) indicates that stylistic expansion can be pursued without substantially destabilizing character identity when guided by human oversight. A moderate association between emotional clarity and visual appeal ($\rho = .50$, $p = .141$) further implies that characters conveying clearer emotional cues may also be perceived as more aesthetically appealing. It should be noted, however, that all correlation values are preliminary and must be interpreted with caution due to the limited sample size ($N = 10$), which yields low statistical power; these associations describe patterns within this specific sample only and do not warrant causal interpretation or broad generalization.

From a human-AI collaboration perspective, these findings reinforce arguments that generative systems are most effective when positioned as partners that extend creative exploration rather than autonomous designers. Human-centered AI frameworks emphasize that reliability and meaning depend on human interpretation, oversight, and value-sensitive decision-making (Shneiderman, 2020). Prior work similarly shows that AI is particularly well suited to divergent ideation, while humans remain responsible for framing problems, curating outputs, and making final judgments (Ning et al., 2023; Kadenhe et al., 2025). The current study operationalizes this principle through a strict role-delineation framework: the AI functions as a generative ideation engine, while the human designer maintains sovereign control over identity parameters and final evaluative selection. This ensures that creative agency remains human-centric despite the high volume of machine-assisted output.

The collaborative workflow also extends conventional mascot and character design practice, which traditionally relies on iterative manual refinement grounded in expertise and narrative insight. Design Thinking approaches encourage user-centered exploration and reflective prototyping to translate persona attributes into coherent visual forms (Cross, 2011; Curedale, 2017). Within this context, mascots must balance recognizability, expressiveness, and narrative resonance to function as symbolic communicators (Bancroft, 2006; Brown & Ponsonby-McCabe, 2014; Lundy, 2008; Jintapitak, 2023). Embedding generative AI within this framework enables designers to test broader alternatives early, then converge intentionally through human

interpretation. Practically, the results suggest that AI tools may be most effective during early ideation and exploration, after which manual refinement consolidates identity and meaning.

The results are particularly significant for digital learning media, the foundational motivation for this research. They demonstrate a viable path for integrating AI to enhance educational engagement through consistent and appealing character-driven narratives. In e-learning environments, mascots serve as pedagogical agents that sustain learner engagement, personalize content delivery, and support affective connection with learning material (Lundy, 2008; Mayer, 2009). Research on multimedia learning further demonstrates that the quality and diversity of visual elements in instructional media directly influence cognitive engagement and learning outcomes (Moreno & Mayer, 2007). The AI-assisted workflow rapidly generates multiple character poses, expressions, and stylistic variants, addressing a practical constraint frequently encountered by instructional designers and educators: producing diverse visual assets within limited time and production resources. For example, a single mascot brief could yield a range of expressive configurations suitable for different instructional scenarios—such as encouraging, questioning, or celebrating—without requiring repeated manual iterations. This scalability suggests that the proposed collaborative intelligence framework may be particularly well suited to asynchronous and interactive learning environments that demand rich, varied character-based content (Veletsianos, 2010). Nevertheless, systematic evaluation of the framework's pedagogical effectiveness, including its impact on learner motivation, comprehension, and long-term retention, remains an important direction for future research.

At the same time, the results should not be interpreted as evidence that generative AI can reliably replace human expertise in all design contexts. Studies in high-stakes domains show that advanced multimodal systems can exhibit inconsistency, context-insensitive errors, and unpredictable failure modes, underscoring the need for sustained human oversight (Morishita et al., 2024). Broader analyses also highlight ethical concerns related to bias, representation, and authorship when generative systems are deployed without safeguards (Samala & Rawas, 2025). In this study, these risks were mitigated by constraining AI use to ideation and exploratory visualization while reserving identity definition, cultural interpretation, and final decision-making for human designers.

Visual consistency presents an additional practical consideration. Although AI amplified stylistic diversity, some outputs displayed subtle drift in proportions, costume details, or facial features across iterations—a phenomenon attributable in part to the inherent difficulty of maintaining variable continuity across successive generations in the model's latent space, where small distributional shifts between prompt iterations can accumulate into perceptible changes in visual identity (Rombach et al., 2022). For mascots—where recognizability and continuity are essential—such drift may erode brand stability if unmanaged. Here, consistency was preserved through structured prompting, iterative human filtering, and manual refinement at key checkpoints, reinforcing that AI functions best as an exploratory partner while identity control remains human-led.

Several limitations warrant acknowledgment. The study focused on a single mascot project (IRIS) and one AI system, which may limit generalizability across character types, cultural settings, or model architectures. The evaluator sample was relatively small and design-oriented, potentially differing from broader user perceptions. Moreover, the evaluation emphasized perceptual qualities rather than longitudinal outcomes such as memorability, audience engagement, or brand impact. Future research should replicate this framework across multiple characters and domains, employ more diverse evaluator groups, and examine real-world performance in communication, branding, and learning contexts.

Despite these limitations, the study provides preliminary empirical evidence that a carefully structured collaborative-intelligence workflow can integrate generative AI into mascot design in ways that expand exploration capacity while preserving human oversight of meaning, identity, and cultural fit. These insights offer practical guidance for professional design practice and suggest pedagogical models for incorporating AI tools thoughtfully into character and mascot design education.

Conclusion and Suggestions

This study examined whether generative AI can enhance mascot design workflows while preserving coherent character identity. By comparing a traditional manual process with an AI-assisted collaborative workflow, the preliminary results suggest that AI meaningfully expands exploratory breadth and visual variation without substantially reducing emotional clarity or identity consistency, though these findings should be interpreted cautiously given the small sample size (N = 10) and the exploratory nature of the study.

The main contribution lies in proposing a human-centered collaborative-intelligence framework that clarifies the appropriate role of AI: supporting rapid ideation and variation generation, while human designers retain control over identity definition, narrative intent, and final decision-making. In this sense, AI functions as an exploratory partner rather than a replacement for professional expertise. The statistically significant improvement in variation richness and the large effect size observed for visual appeal provide initial empirical support for this framework, while comparable ratings for identity coherence and emotional clarity confirm that increased stylistic diversity did not compromise core character meaning under human oversight.

Practically, the framework guides designers and educators seeking to integrate AI tools into character and mascot development. The findings suggest that AI is most effective in early ideation stages, with human-led selection and refinement used to stabilize meaning, style, and continuity. Given the increasing demand for diverse and adaptive visual assets in digital learning environments, the framework offers particular promise for instructional designers and educators who require scalable character-based content within constrained production timelines. Nevertheless, its pedagogical value—including effects on learner engagement, motivation, and comprehension—still requires systematic evaluation before broader adoption can be recommended.

Future research should replicate this framework across broader cultural contexts, character types, and AI systems, and assess longer-term outcomes such as brand recognition, audience engagement, and learning effects in authentic educational settings. Overall, thoughtfully integrating AI within a structured, human-centered workflow can expand creative capacity while preserving the human meaning-making central to successful mascot and character design.

Acknowledgements

The author gratefully acknowledge that this study was conducted as part of the development process for producing visual materials to support asynchronous learning and online instructional media for the course Mascot Design. This project was funded by the Teaching and Learning Innovation Center (TLIC), Chiang Mai University, whose support, facilities, and resources enabled the integration of generative AI technologies into

educational media production and contributed to advancing research on collaborative intelligence in mascot and character design.

Author Contributions

Author: Conceptualization; methodology; preparation of research materials and subject inputs; investigation, data collection, and data analysis; interpretation of findings; writing—original draft preparation; and writing—review and editing.

Conflict of Interests

The authors declare that there are no conflicts of interest associated with this study.

Funding

This study was supported by the Teaching and Learning Innovation Center (TLIC), Chiang Mai University. Funding for developing visual and digital learning materials for the Mascot Design course enabled the creation of asynchronous instructional media and the integration of generative AI into the mascot design workflow.

Declaration of Generative AI and AI-assisted Technologies

During the preparation of this manuscript, the author used Google Gemini and ChatGPT to assist with English grammar checking. Following its use, the author thoroughly reviewed and edited the content and accepted full responsibility for the final version of the publication.

References

Bancroft, T. (2006). *Creating characters with personality*. Watson-Guptill.

Bianchi, I., Branchini, E., Uricchio, T., & Bongelli, R. (2025). Creativity and aesthetic evaluation of AI-generated artworks: Bridging problems and methods from psychology to AI. *Frontiers in Psychology, 16*, Article 1648480.

Brown, S., & Ponsonby-McCabe, S. (2014). *Brand mascots and other marketing animals*. Routledge.

Cha, E., & Wang, D. (2025). A study of the relationship between artificial intelligence generated image advertising and consumer brand awareness. *Creative Business and Sustainability Journal, 47*(1), 1-19. https://doi.org/10.58837/CHULA.CBSJ.47.1.1

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Cross, N. (2011). *Design thinking: Understanding how designers think and work*. Berg Publishers.

Curedale, R. (2017). *Design thinking: Process & methods* (4th ed.). Design Community College Inc.

Google DeepMind. (2023). *Gemini: A family of highly capable multimodal models* (Technical Report). https://arxiv.org/abs/2312.11805

Jintapitak, M. (2023). Equivalent of character design for design thinking in personal character design process. In *Proceedings of the Joint International Conference on Digital Arts, Media and Technology (ECTI DAMT & NCON)* (pp. 16-20). https://doi.org/10.1109/ECTIDAMTNCON57770.2023.10139327

Kadenhe, N., Al Musleh, M., & Lompot, A. (2025). Human-AI co-design and co-creation: A review of emerging approaches, challenges, and future directions. *Proceedings of the AAAI Symposium Series, 6*(1), 265-270.

Lundy, P. (2008). *Digital storytelling and self-representation in new media*. Peter Lang.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309-326. https://doi.org/10.1007/s10648-007-9047-2

Morishita, M., Fukuda, H., Muraoka, K., Nakamura, T., Hayashi, M., Yoshioka, I., Ono, K., & Awano, S. (2024). Evaluating GPT-4V's performance in the Japanese national dental examination: A challenge explored. *Journal of Dental Sciences, 19*(3), 1595-1600. https://doi.org/10.1016/j.jds.2023.12.007

Ning, B., Liu, F., & Liu, Z. (2023). Creativity support in AI co-creative tools: Current research, challenges and opportunities. In *Proceedings of the International Conference on Computer Supported Cooperative Work in Design (CSCWD)* (pp. 5-10). https://doi.org/10.1109/CSCWD57460.2023.10152832

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. (2021). Zero-shot text-to-image generation. *In Proceedings of the International Conference on Machine Learning (ICML)*. https://arxiv.org/abs/2102.12092

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (pp. 10674-10685). https://doi.org/10.1109/CVPR52688.2022.01042

Samala, A., & Rawas, S. (2025). Bias in artificial intelligence: Smart solutions for detection, mitigation, and ethical strategies in real-world applications. *IAES International Journal of Artificial Intelligence, 14*(1), 32-43. https://doi.org/10.11591/ijai.v14.i1.pp32-43

Shneiderman, B. (2020). Human-centered AI: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction, 36*, 495-504. https://doi.org/10.1080/10447318.2020.1741118

Smith, F. W., & Rossit, S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLOS ONE, 13*(5), e0197160. https://doi.org/10.1371/journal.pone.0197160

Tian, Y., Liu, Y., Wang, S., & Kwong, S. (2025). Quality assessment for text-to-image generation: A survey. *IEEE MultiMedia, 32*(2), 44-52. https://doi.org/10.1109/MMUL.2025.3538862

Veletsianos, G. (2010). Emerging technologies in distance education. AU Press.

White, T. (2006). *Animation from pencils to pixels: Classical techniques for the digital animator*. Routledge.