JURNAL TEKNOLOGI DAN OPEN SOURCE
Vol. 8, No. 2, December 2025, pp. 722~739
e-ISSN: 2622-1659, accredited Four Grade by Kemenristekdikti, Decree No: 152/E/KPT/2023
DOI: https://doi.org/10.36378/jtos.v8i2.4959

Optimizing Prompt Engineering for AI-Based Logo Generation Using Response Surface Methodology

Shermay[1], Syaeful Anas Aklani[2], Muhamad Dody Firmansyah[3]
[1,2,3]Faculty of Computer Science, Universitas Internasional Batam, Batam

Article Info

Article history:

Received 10 01, 2025
Revised 11 10, 2025
Accepted 11 30, 2025

Keywords:

Prompt Engineering
AI Image Generation
Logo Design
Generative AI
Response Surface Methodology

ABSTRACT

Artificial Intelligence has transformed creative design workflows, yet effective communication between designers and AI image generation systems remains challenging. Both professional designers and non-technical users struggle to articulate design intent through text prompts, resulting in inconsistent outputs and inefficient iteration cycles. This research developed an optimized prompt engineering framework for AI-based logo generation using Response Surface Methodology (RSM) with Central Composite Design (CCD) to systematically optimize prompt variables through controlled experimentation. This research tested 47 prompt combinations across five independent variables: prompt clarity, level of detail, thematic description, visual elements, and color specification. From this optimization process, eight critical components were identified form a structured template: Main Design Focus, Detail Elements, Thematic Style, Primary Colors, Complementary Colors, Rewording, Layout Size, and Element Limit. Experimental validation with 30 graphic designers demonstrated substantial improvements over unstructured prompts: visual consistency increased from 65% to 87% ($p<0.001$, $d=1.54$), iteration efficiency improved 48.5% (from 6.6 to 3.4 attempts, $p<0.001$, $d=1.32$), and user satisfaction from 58% to 82% ($p<0.001$, $d=1.67$). This research provides a systematic, replicable methodology for prompt optimization in creative AI applications, enhancing accessibility while maintaining professional quality standards.

Corresponding Author:

Shermay
Faculty of Computer Science
Universitas Internasional Batam
Batam, Indonesia
Email: 2231040.shermay@uib.edu

1. Introduction

Advancements in Artificial Intelligence (AI) and prompt engineering have increased significantly, especially in fields such as Natural Language Processing (NLP) and image generation. These advancements have positioned prompt engineering as a critical skill for enhancing both the relevance and performance of AI models. In image generation, prompt engineering leverages both text-to-image and image-to-image capabilities, thus creating vast data where natural language inputs can lead and resulting in high-quality results [1]. These generative text-to-image frameworks enable a free expression of ideas and offer creators the tools that are needed to produce high-quality visuals [2]. The integration of different AI models has also become essential for achieving a better and more efficient yield of both solutions and results [3]. Combining AI models that specialize in its capabilities can create more efficient workflows. In particular, the synergy between Text language models and image generation models enables a nuanced control over image outputs through the natural language process, allowing for better results. This rapid development in both AI and NLP demonstrates the immense potential of AI image Generation in enhancing creativity and productivity [4].

Building upon these technological advances, two prominent AI models developed by OpenAI ChatGPT and DALL-E have transformed the way users create visual content by enabling the use of natural language processing to guide image generation [5]. ChatGPT enables complex language understanding and ideation, while DALL-E converts textual prompts into images that match those descriptions. This integration allows users to iteratively refine their creative visions, where final prompts can be passed seamlessly from ChatGPT to DALL-E, forming an Augmented Intelligence Assistance (AIA) workflow [6]. Such synergy simplifies prompt engineering and bridges the gap between human creativity and machine interpretation [7]. The capability of these AIs is limitless as it will generate any idea from the user based on the text that was given. This capability holds immense potential for both AI and Art industries [2]. These AIs are very transformative for users who seek to automate and increase their efficiency aspect in both the design and creativity fields [8].

Despite these technological capabilities, significant challenges persist in the practical application of AI-powered image generation. According to recent industry reports, while 15% of companies worldwide utilized AI for image generation in 2022, this figure is projected to reach 30% by 2025 [9], underscoring the urgency of addressing implementation barriers. The challenge lies in effective communication between users and AI systems [10]. The inherent ambiguity of natural language frequently yields unpredictable outputs, where subtle variations in prompt wording can produce dramatically different results [11]. Consequently, users both experienced designers and novices struggle to achieve consistent visual outcomes aligned with their creative intent, resulting in inefficient iteration cycles and diminished user satisfaction. These structural challenges must be systematically addressed to fully realize the potential of AI image generation and ensure its accessibility to diverse user populations.

AI itself integrates computer science, mathematics, and cognitive science to create intelligent systems capable of performing tasks that require human intelligence. AI concept was first laid by Turing in 1950, the Turing Test assesses the notion of machine learning and therefore machine intelligence [10]. Over the years AI has evolved and in the 1980s machine learning and deep learning were applied through symbolic AI, a Rule-based System. In the modern era, AI has been trained using machine learning and Deep Learning, which allow systems to learn from data and use neural networks for complex pattern recognition [11]. The introduction of Transformer models in 2017 allowed NLP to process words in parallel, paving the way for the development of large language models and becoming the fundamental of AI Models such as ChatGPT [12]. These breakthroughs made generative systems like DALL-E possible, bridging NLP and computer vision through multimodal learning.

Machine learning and deep learning technologies underpin modern AI's capability in both NLP and image synthesis [13], [14]. Generative AI, which includes models such as ChatGPT and DALL·E, has redefined creative production by enabling data-informed ideation and visual synthesis. ChatGPT employs Transformer-based architectures for textual reasoning, while DALL·E aligns language and visual representation, allowing it to generate coherent imagery from descriptive prompts.

At the center of this process lies prompt engineering, which involves crafting precisely structured instructions to guide generative AI outputs. Unlike traditional programming, which relies on explicit coding, prompt engineering uses natural language as a programming interface. This emerging discipline enables users to control AI behavior through textual cues, making it a crucial competence for leveraging the full capacity of Large Language Models (LLMs) and text-to-image generators. Techniques such as zero-shot, few-shot, and chain-of-thought prompting [15], [16] have evolved to improve reasoning, contextual understanding, and accuracy transforming prompt design from simple phrasing to a strategic creative process.

The integration of AI into graphic design further amplifies its transformative potential. Traditionally, graphic design has been a medium of visual communication, combining elements like color, shape, and typography to convey messages effectively [17]. Technological progress from manual drafting to computer-aided design has expanded the designer's toolkit, allowing for complex 2D and 3D compositions [18], [19]. In the modern era, design practices have evolved to include UI/UX, motion, and data-driven visualization, aligning with digital innovation and accessibility [20]. The fusion of AI and design introduces new creative paradigms real-time image synthesis, adaptive visual aesthetics, and data-informed artistic generation [21], [22]. These advancements enable designers to merge computational precision with emotional resonance, producing visuals that are both technically refined and artistically expressive [22], [23], [24] .

Although several studies have explored prompt engineering techniques and their applications in AI-generated content, significant gaps remain in literature. Korzynski et al. [25] emphasized the significance of prompt engineering in enhancing the performance of Large Language Models (LLMs) by developing a comprehensive AI prompt Framework. Their research established a conceptual framework through narrative and critical literature review, emphasizing advanced techniques such as meta prompts, reasoning chains, zero-shot, and few-shot prompting. However, challenges such as hallucination problems and the tendency for LLMs to replicate existing text without deep understanding remain unresolved. Sanchez [26] analyzed practices within the text-to-image generation community through an online questionnaire completed by 64 practitioners and a dataset of user prompts submitted to the Stable Diffusion model. The research discovered that text-to-image generation primarily serves as a recreational pursuit, but users encounter difficulties linked to both model capabilities and prompt design. Goloujeh et al. [27] explored users' "prompt journey" through semi-structured interviews with 19 users, revealing that users faced two main challenges: aligning AI output with their intent and mastering prompt engineering techniques. The research suggests that prompt engineering is both an individual and social experience. Lai et al. [28] introduced mini-DALL-E 3 as a prompting enhancement for LLMs, demonstrating high-quality results but limited task-specific control.

Although these studies provide valuable insight into prompt design and user interaction, they remain primarily conceptual or descriptive, lacking a formal optimization-based approach for systematic improvement.

While previous studies have explored various prompt engineering techniques and user experiences, a critical research gap remains in the absence of systematic, optimization-based frameworks specifically designed for logo generation. Most existing works focus on general text-to-image generation without addressing the unique requirements of corporate visual identity design, where consistency, brand alignment, and professional quality are essential. This gap becomes increasingly significant as designers and organizations seek dependable AI tools capable of producing repeatable, high-fidelity outputs. Without a structured approach tailored to logo creation, current methods often lead to unpredictable results, limiting their applicability in professional branding contexts and long-term design workflows.

To bridge this gap, the present research introduces a systematic optimization framework for prompt engineering in AI-based logo design using Response Surface Methodology (RSM) combined with Central Composite Design (CCD). This approach enables empirical modeling of how multiple prompt factors interact to influence output quality. The framework centers on developing a structured prompt template comprising eight core components Main Design Focus, Detail Elements, Thematic Style, Primary Colors, Complementary Colors, Rewording, Layout Size, and Element Limit representing critical visual and semantic elements of logo design.

Through experimental research employing purposive sampling, this research engaged professional graphic designers from both corporate and freelance backgrounds, representing practitioners who either rely on manual workflows or integrate AI tools such as ChatGPT and DALL-E in their creative processes. The optimization process systematically tested 47 prompt combinations generated through CCD, encompassing factorial, axial, and center points to capture both main and interaction effects among variables. This methodological design provides a data-driven understanding of how prompt structure influences visual quality, consistency, and user satisfaction in AI-generated logos.

This research advances the field of AI-assisted design by introducing a systematic, data-driven framework for optimizing prompt structures in generative logo creation. Using Response Surface Methodology (RSM), this research establishes a replicable approach to modeling how different prompt components interact to influence visual outcomes, transforming prompt formulation into an analytical and measurable process. The resulting framework significantly improves design performance reducing iteration cycles from 6.6 to 3.4 and enhancing visual consistency from 65% to 87% while enabling both professional and non-technical designers to achieve higher-quality results more efficiently.

This research further integrates RSM ANOVA modeling with qualitative thematic analysis, ensuring a comprehensive validation that captures both quantitative improvements and user experience perspectives. This combination allows the study to evaluate not only statistical optimization but also the practical relevance and usability of the generated outputs from the standpoint of real design practitioners. By addressing the long-standing gap between user intent and AI interpretation in visual design, the framework provides a more reliable and scalable foundation for integrating generative AI tools into professional creative workflows, particularly for small studios, freelancers, and independent designers. It also supports clearer communication pathways between human creators and algorithmic systems, reducing ambiguity and improving consistency.

The novelty of this research lies in reframing prompt engineering from a trial and error process into a structured, optimization-based methodology grounded in empirical analysis. It highlights the potential of AI not merely as a generative instrument but as a collaborative creative partner, capable of interpreting structured human intent and producing consistent, high-quality visual outcomes. By focusing specifically on logo design, which requires precision, coherence, and aesthetic integrity, this research contributes a domainspecific methodological innovation that redefines how creativity, structure, and computation converge within AI-driven visual design, ultimately advancing the reliability and predictability of generative design systems..

2. Research Method

2.1. Experimental Design

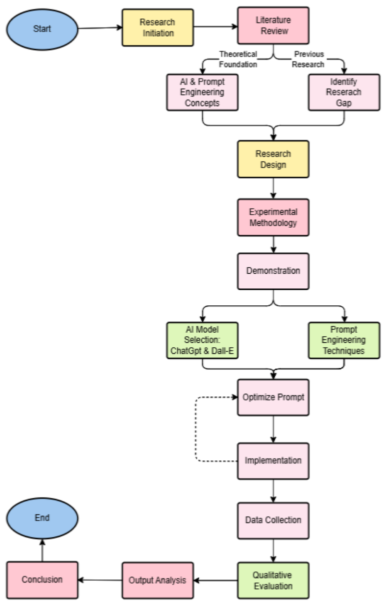

Figure 1. Research Flow

This research begins with initiation in the form of past literature reviews by exploring existing studies on AI and prompt engineering concepts. This initial phase involves exploring existing studies on prompt engineering concepts, Large Language Models (LLMs), and current optimization techniques to guide the development process, building upon the AI prompt framework established by Lai et al. [28]. This step aims to guide the optimization process of prompt engineering with a fundamental understanding of AI, Large Language Models (LLMs), and how AI evolution systems and current prompt engineering techniques have helped in optimizing prompt engineering. The goal of this guide is to optimize the use of generative AI models with LLMs in prompt engineering.

After reviewing existing prompt engineering literature, an experimental framework was constructed using Response Surface Methodology (RSM) combined with Central Composite Design (CCD) to optimize multi-variable processes through systematic experimentation. RSM provides a structured analytical approach to examine how multiple independent variables interact to influence AI-generated logo quality. This methodology is supported by recent studies such as Li et al.[1] and Feng et al. [2], which applied parameter-based optimization and statistical modeling for evaluating AI-generated visual content. This approach enables precise identification of how different prompt components collectively affect visual outcomes, facilitating optimization toward consistent, high-quality logo generation.

This research focuses on several independent variables that were based on previous prompt-engineering and generative AI research that identified clarity, detail, theme, color, and compositional structure as key determinants of visual output quality. Five independent variables were therefore examined: prompt clarity, level of detail, thematic description, number of visual elements, and color specification. These variables were conceptually adapted from the research of Giray [8], Feng et al. [2], Lai et al.[28], Li et al. [1], and White et al. [29], which emphasized the importance of structured, detailed, and conceptually rich prompts in achieving consistent and high-quality AI image generation.

The experimental process was designed in multiple stages to ensure both quantitative optimization and qualitative validation. The results of this experiment later evolved into a refined eight-component framework, including main design focus, detail elements, thematic style, primary colors, complementary colors, rewording, layout size, and element limits, integrating these theoretical factors into a practical guide for effective prompt formulation. This synthesis was inspired by the structured prompt frameworks proposed by Korzynski et al. [25] and Goloujeh et al. [27], bridging empirical optimization with user-centered application in AI-based logo design.

Table 1. Independent Variables Classification and Measurement Levels

Variable	Description	Measurement		
		Low (-1)	Medium (0)	High (+1)
Prompt Clarity	How clear and unambiguous the language in the prompt is	Vague wording	Moderately clear	Very specific and direct language
Level of Detail	The amount of information included in the prompt	Minimal details (e.g., “simple logo”)	Moderate details (e.g., “modern logo with text”)	Rich description (e.g., layout, font, colors, meaning)
Thematic Description	Conceptual clarity and contextual framing	No theme	Basic theme;	Clear contextual theme with meaning
Number of Visual Elements	Quantity of distinct visual components specified	1 Element	3 elements	4+ elements
Color Specification	Specificity and clarity in color instructions	No color	Basic color	Multiple colors with symbolic meanings

Each prompt generated from the RSM matrix is subjected to repeated testing to ensure the consistency and reliability of the results. These repeated tests help verify that the AI consistently interprets the prompts as designed, avoiding unintended outcomes.

To quantitatively measure the quality of generated logos, this research employed the Visual Quality Assessment Scale (VQAS) structured evaluation framework specifically developed for assessing AI-generated visual outputs. The VQAS consists of four key dimensions, each rated on a 7-point Likert scale. These four dimensions were selected because they capture both the technical accuracy and the aesthetic:

1. Visual Consistency (VC): Degree of alignment between textual prompt and visual output across multiple independent generation attempts.
2. Element Clarity (EC): Distinctiveness and professional execution of specified visual components (typography, symbols, layout).
3. Aesthetic Quality (AQ): Overall visual appeal, compositional balance, and professional polish.
4. Brand Alignment (BA): Appropriateness and effectiveness in representing the stated company context and thematic intent.

Detailed behavioral anchors for each scale point were developed based on graphic design evaluation principles [19] and validated through pilot testing with three design experts. The composite Visual Quality Score (VQS) was calculated using the following weighted formula:

$$VQS = (0.35 \times VC + 0.30 \times EC + 0.20 \times AQ + 0.15 \times BA) \times (100/7) \quad (1)$$

Visual consistency, the primary outcome variable, was operationalized by assessing agreement across three independent generations on four critical design elements:

1. Text Presence and Readability: Specified company name appeared legibly in all three images.
2. Primary Visual Symbol: Main conceptual element (e.g., skyline, palm tree) appeared recognizably in all images.
3. Color Scheme Adherence: Specified color palette applied consistently across all images.
4. Layout Orientation: Spatial arrangement (horizontal, vertical, centered) remained consistent across all images.

Visual consistency percentage was calculated as:

$$Consistency\ \% = (Matching\ elements\ /\ 4\ total\ elements) \times 100 \quad (2)$$

Logo sets achieving ≥75% consistency (3-4 elements matching) were classified as "high consistency," 50% as "moderate consistency," and ≤25% as "low consistency."

All logos generated across the 47 prompt combinations were rated independently by three design experts. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC) for continuous

VQS scores (two-way mixed-effects model, absolute agreement definition, average measures) and Fleiss' Kappa (κ) for categorical classifications of consistency. Both indices exceeded conventional thresholds for excellent reliability (ICC ≥ 0.75; κ ≥ 0.60) confirming the robustness and reproducibility of the evaluation process [30]

The resulting VQS data were then analyzed statistically using Response Surface Methodology (RSM) and Analysis of Variance (ANOVA) to determine the significance of each prompt variable. RSM modeled the relationships between the five independent prompt factors and the VQS outcomes, while ANOVA tested which variables had statistically significant effects on visual quality. Key statistical indicators such as F-values, p-values, and R^2 were used to validate the model and confirm which prompt parameters most strongly influenced consistency and aesthetic performance.

Through this process, this research established an empirical foundation for optimizing prompt formulation, ensuring that the resulting structure was not only theoretically grounded but also statistically validated for effectiveness in producing high-quality AI-generated logos.

Once the optimized prompt structure was finalized through RSM–ANOVA analysis, it was applied in real user scenarios to evaluate its practicality and effectiveness. This experimental research employed purposive sampling, involving 30 professional logo designers consisting of freelancers and designers employed within design companies. The determination of the sample size was guided by both methodological relevance and practical feasibility. Previous prompt-engineering studies in design evaluation commonly employed sample sizes ranging from 19 to 64 participants [27], [26]. The experimental design adopted a within-subject comparison (pre- vs. post-optimization), which enhances statistical power by reducing between-subject variance.

To support this stage, a data collection instrument was developed to obtain user feedback regarding the optimized prompt framework. The primary aim was to evaluate its effectiveness, creativity, and user experience during implementation in real design workflows. Feedback was gathered through structured interviews with professional designers after they completed the logo generation tasks using the optimized prompts.

User feedback was then analyzed to assess how effectively the optimized prompts enhanced creative outcomes and design efficiency. The insights from this process also informed minor refinement adjustments, focusing on clarifying wording and aligning parameter settings with user needs. These improvements were integrated into the final Prompt Design Guideline to ensure higher flexibility and consistency in visual generation outcomes.

The interview questions are designed to focus on evaluating effectiveness, creativity, and user experience in applying the prompt engineering techniques.

Table 2. Interview Protocol Structure

Effectiveness	How easy or difficult was it for you to understand and apply the optimized prompt engineering techniques?
	Were there any specific aspects you found confusing or challenging during the implementation process?
	Are you satisfied with the results that have been achieved by implementing the prompt engineering techniques?
Creativity	In your opinion, does the resulting logo design demonstrate a high level of creativity?
	What specific elements of the logo design stand out as particularly creative or innovative?
	Do you believe the AI-generated logo designs reflect a unique vision or identity for the company?
User Experience	What are your thoughts on the logo design results that were generated by using these prompt engineering techniques?
	Did you feel that the AI supported or hindered your creative expression during the logo design process?

The resulting AI-generated logos were also compared against their corresponding prompt descriptions as secondary validation data, ensuring alignment between prompt intent and output quality. This comparison provided supporting evidence for assessing the effectiveness of the optimized prompt structure.

Qualitative responses were analyzed using a thematic analysis approach, following the procedure adapted from Goloujeh et al. [27], who applied a similar method to evaluate user experiences with text-to-image generative AI tools. Interview transcripts were systematically coded, and recurring patterns were categorized into three major themes: effectiveness, creativity, and user experience. This approach provided deeper insights into how users perceived and applied the optimized prompt framework in real-world design contexts, highlighting both its strengths and areas for potential improvement.

2.2. AI Model Specifications

This research utilized two AI models from OpenAI for complementary functions in the logo generation workflow.

2.2.1. Language Model: ChatGPT-4

ChatGPT-4 was employed to ideate and refine prompt structures. The configuration was set to temperature=1.0, top_p=1.0, with frequency penalty values at 0.0, balancing creativity with prompt consistency. This ensured that the prompt components such as clarity, level of detail, thematic descriptions, and visual attributes could be controlled as variables in the experimental design.

2.2.2. Image Generation Model: DALL-E 3

DALL-E 3 was utilized as the primary generative model for producing all logo outputs in this research. Images were generated at a resolution of 1024×1024 pixels in a standardized square format commonly used for scalable logo deployment. The standard quality setting was selected to balance visual clarity with efficient computational cost, while the vivid style parameter was applied to enhance the graphic dynamism and branding appeal of the outputs.

Due to the absence of seed control in the publicly accessible version of the API, the model exhibits non-deterministic behavior, wherein identical prompts may yield different visual results. To systematically account for this inherent variability and ensure that each condition was evaluated under comparable circumstances, three independent generation attempts were performed per prompt, with a minimum interval of five minutes between executions.

2.2.3. Pre-processing and Post-processing Procedures

Pre-processing ensured clarity and consistency in the textual prompts. All prompts were written in English, with the structured prefix "Generate a logo design:" included to direct the generative task. Unsupported characters were removed to avoid potential generation errors.

Post-processing was limited to downloading the resulting images in lossless PNG format. No cropping, filtering, enhancement, or manual editing was applied. Each logo therefore represents the raw generative performance of the model, ensuring experimental integrity.

3. Result and Discussion

The results of experiments and analysis conducted to evaluate the effectiveness of prompt engineering techniques in generating AI-based logo designs. The results are divided into two main parts: quantitative analysis using Response Surface Methodology (RSM), and qualitative analysis through thematic analysis.

3.1. Experimental Method

3.1.1. Prompt Selection from Previous Research

The initial phase of this research begins by selecting a base prompt adapted from prior studies on prompt engineering in text-to-image generation. One of the prompts taken from previous research is as follows:

> *"A minimalist logo with the text "Mini DALLE 3" written in a clean and modern font. The letters are arranged in a straight line, with each letter having equal spacing. The color palette consists of vibrant shades of blue and green, representing creativity and innovation."*

This prompt was selected due to its strong structural clarity and the use of specific visual descriptors, which aligns with this research's goal of minimizing ambiguity in prompt formulation. The descriptive language used in this prompt reflects a detailed understanding of layout, typography, and color harmony elements that are crucial for generating consistent and visually appealing logos using AI.

This base prompt will be used as the foundation for the optimization process. From here, each component such as font specification, color use, and spatial layout will be analyzed, modified, and tested to understand how variations impact the resulting images. The outcome of this stage will serve as the control reference point for the experiment and will be compared against other prompt versions developed in the optimization phase.

3.1.2. Design of Experiment Using Response Surface Methodology (RSM)

Response Surface Methodology (RSM) was employed to model and optimize the relationship between five independent variables and logo quality outcomes. The general second-order polynomial model is expressed as:

$$Y = \beta 0 + \sum_{i=1}^{5} \beta_{\mathrm{i}} X_{\mathrm{i}} + \sum_{i=1}^{5} \beta_{\mathrm{ii}} X_{\mathrm{i}}^{2} + \sum_{i<j} \beta_{\mathrm{ij}} X_{\mathrm{i}} X_{\mathrm{j}} + \varepsilon \quad (3)$$

In this research, Y represents the response variable, specifically the Visual Quality Score (VQS) obtained from the evaluation of AI-generated logos. The five independent variables (X_1–X_5) denote the key prompt attributes that influence visual outcomes: Prompt Clarity (X_1), Level of Detail (X_2), Thematic Description (X_3), Number of Visual Elements (X_4), and Color Specification (X_5). Each of these variables contributes to explaining variations in logo quality as measured by the VQS. The regression coefficients (β) quantify the strength and direction of each variable's effect both individually and in combination on the response. The term ε represents the random error component, accounting for unexplained variability or random noise not captured by the model. Together, this structure forms a second-order polynomial equation that models and predicts how prompt attributes interact to influence the quality and consistency of AI-generated logos.

A Central Composite Design (CCD) structure was used to estimate linear, quadratic, and interaction effects efficiently. The total number of experimental runs was determined by:

$$Total\ Run = 2^{k} + 2K + NC \quad (4)$$

Where k represents the number of factors, 2^k represents the factorial points, $2K$ represents the axial points, and NC represents the number of center point replications. Therefore, the total experimental design comprised:

Total Run = 32 + 10 + 5 = 47

The axial distance (α) determines the position of the star points and affects the design properties. In this research, α was set at 2.0 to ensure rotatability of the design, which allows the prediction variance to be constant at all points equidistant from the design center. Each independent variable was coded to dimensionless values ranging from -α to +α, corresponding to five levels: -2 (low), -1 (medium-low), 0 (center), +1 (medium-high), and +2 (high), as presented in Table 1. This coding procedure standardizes variables with different units and ranges, facilitating direct comparison of their relative effects on the response.

3.1.3. Repetitive Test

Each prompt from the RSM matrix was subjected to repeated testing to ensure consistency and reliability of results. These repeated tests help verify that the AI consistently interprets prompts as designed, avoiding unintended outcomes. Throughout this process, various data points were recorded, including the consistency of image outputs, instances of misinterpretation, and user satisfaction levels.

3.1.3.1. Repetitive Test using Prompt Combinations

Table 3. CCD Experimental Matrix and Logo Quality Evaluation

Prompt Combination	Prompt	Logo Generate by AI	Evaluation
L L L L L	Create a logo for XYZ Company.	XYZ COMPANIES	No clear style or direction. Results are inconsistent and very wide in layout, font, and symbolism. Low brand representation.
H L L L L	A minimalist logo for 'XYZ Company' in a clean sans-serif font, all letters in a single line.	XYZ COMPANY	Typography is clean, but lack of themes, colors, and visual elements results in limited creativity and brand storytelling.

L H L L L	Create a modern and professional logo for XYZ Company.		Better detail than low clarity, with a professional tone, but still lacks thematic identity, color control, and visual symbolism; design is clean but generic.
L L H L L	Design a logo for XYZ Company representing a tropical tech startup.		Theme clarity improved the concept by merging tropical and tech elements, but lack of color and element limits meant variations still appeared.
L L L H L	Create a logo for XYZ Company with a building skyline, palm trees, waves, and text.		More visual elements made the brand more recognizable, but without style guidance or color rules, results varied in tone and consistency.
L L L L H	Create a logo for XYZ Company using blue and green colors.		Color direction improves visual harmony but still lacks style, theme, and detail; results varied in composition and symbolism.
H H L L L	A minimalist logo for ‘XYZ Company’ in a clean sans-serif font, arranged horizontally with balanced spacing, featuring precise line work and clear layout.		Very consistent typography and structure lack thematic or symbolic elements, so branding feels corporate but not unique.
H L H L L	Design a modern logo for XYZ Company representing a tropical tech startup, using a clean sans-serif font.		Themes improve relevance and connection to tropical-tech concept, but absence of color details limits creative cohesion.
H L L H L	Create a logo for XYZ Company featuring a building skyline, palm trees, and waves.		Multiple elements make it recognizable, but without specific color or style guidelines, result varied in tone and quality.
H L L L H	A minimalist logo for ‘XYZ Company’ in a clean sans-serif font, using deep blue as the main color.		Strong color consistency and clean typography, but lack of theme or symbolic elements makes it less distinctive.
L H H L L	Create a professional logo for XYZ Company with detailed elements, representing an eco-friendly tech startup.		Strong theme and details produce more relevant and coherent visuals, but lack of color guidance leads to variations.
L H L H L	Create a modern logo for XYZ Company with detailed elements, including a skyline and palm trees.		More visual complexity and detail enhance recognizability, but without theme or color instructions, brand story is weaker.
L H L L H	Create a modern logo for XYZ Company with detailed typography, using blue and gold colors.		Absence of theme limits storytelling.
L L H H L	Design a logo for XYZ Company representing a tropical tech startup, with a skyline and palm trees.		Clear theme and multiple elements improve brand recognizability, but lack of specific color palette causes visual inconsistency.
L L H L H	Design a logo for XYZ Company representing a tropical tech startup, using blue and green colors.		Theme and color enhance brand personality, but limited visual elements make the design less distinctive than multi-element versions.
L L L H H	Create a logo for XYZ Company with a skyline, palm trees, and waves, using blue and green as primary colors.		Clear visual elements and strong color control improved consistency, but lack of style/theme guidance limits brand storytelling.
H H H L L	A minimalist yet detailed logo for XYZ Company, representing an eco-friendly tech startup, in a clean sans-serif font.		Combines clarity, detail, and theme well; results are highly consistent and visually cohesive, but could benefit from defined color choices.

H H L H L	A minimalist yet detailed logo for XYZ Company, featuring a skyline and palm trees, in a clean sans-serif font.		Clarity and detail ensure consistent layouts; inclusion of multiple visual elements increases recognizability, but lack of theme/color limits uniqueness.
H H L L H	A minimalist yet detailed logo for XYZ Company, using deep blue and golden yellow colors.		Color and detail create polished results; clarity ensures clean output, but absence of theme weakens storytelling potential.
H L H H L	Create a modern logo for XYZ Company representing a tropical tech startup, with a skyline, palm trees, and waves.		Theme and multiple elements provide rich brand visuals; high clarity ensures consistent structure, but absence of defined colors leads to variation.
H L H L H	Create a modern logo for XYZ Company representing a tropical tech startup, using deep blue and bright orange colors		Clear theme and strong color guidance produce vibrant, recognizable designs; absence of extra elements keeps the look minimal but may feel less dynamic.
H L L H H	Create a logo for XYZ Company featuring a skyline, palm trees, and waves, using blue and green colors.		Multiple visual elements and defined colors yield high consistency; absence of theme makes the concept generic
L H H H L	Create a professional logo for XYZ Company with detailed elements, representing an eco-friendly tech startup, featuring a skyline and palm trees.		Combines strong theme, detail, and multiple elements for rich branding; lack of color control allows variations in tone
L H H L H	Create a professional logo for XYZ Company with detailed elements, representing an eco-friendly tech startup, using green and yellow colors.		Strong theme and colors improve identity clarity. multiple elements for rich branding; lack of color control allows variations in tone
L H L H H	Create a modern logo for XYZ Company with detailed elements, featuring a skyline and palm trees, using blue and orange colors.		Details, multiple elements, and color guidance ensure strong brand visibility; absence of theme reduces storytelling depth.
L L H H H	Design a logo for XYZ Company representing a tropical tech startup, featuring a skyline and palm trees, using blue, green, and yellow colors		Strong theme, multiple elements, and defined colors make the output vibrant and cohesive; lack of clarity in style could still produce slight variation.
H H H H L	A minimalist yet detailed logo for XYZ Company, representing an eco-friendly tech startup, featuring a skyline, palm trees, and waves in a clean sans-serif layout.		Very consistent results with high clarity, detail, theme, and multiple elements; absence of color specification allows style freedom but reduces palette control.
H H H L H	A minimalist yet detailed logo for XYZ Company, representing an eco-friendly tech startup, using green and blue colors.		Highly consistent theme and style; color choice ensures brand cohesion, but fewer elements limit visual storytelling.
H H L H H	A minimalist yet detailed logo for XYZ Company, featuring a skyline and palm trees, using blue and gold colors.		Detailed, clear, and color-defined prompts produce polished results; lack of theme makes it visually strong but generic.
H L H H H	Create a modern logo for XYZ Company representing a tropical tech startup, featuring a skyline, palm trees, and waves, using blue, green.		Combines theme, multiple elements, and vibrant colors for strong brand recognition; high clarity ensures consistent layouts.

L H H H H	Create a professional logo for XYZ Company with detailed elements, representing an eco-friendly tech startup, featuring a skyline and palm trees, using green, blue, and yellow colors.		Very strong theme, detailed design, multiple elements, and defined colors produce highly cohesive and brand-accurate outputs; low clarity in layout can still cause small variations.
H H H H H	A minimalist yet detailed logo for XYZ Company, representing an eco-friendly tropical tech startup. Use a clean sans-serif font in a bold style. Focus on three main elements: a modern skyline, a tropical palm leaf integrated with a circuit line, and a wave accent. Use deep blue as the dominant color and bright green as the secondary highlight. Arrange in a balanced horizontal layout with strong typography on a white background.		Extremely consistent, rich, and visually appealing results thanks to maximum clarity, detail, theme, elements, and colors; strong control over output.
M M M M M	Create a modern logo for XYZ Company with moderate clarity and detail. Use a clean sans-serif font and include 2–3 visual elements such as a simple skyline and a wave. Apply one primary color (deep blue) without complex combinations. The design should convey a general technology feel while remaining clean and professional, arranged in a simple balanced layout.		Balanced results with decent consistency; however, moderate specification still allows style variations between outputs.

3.1.3.2. Identification of Problematic Prompt Elements

Based on repetitive testing, several prompt elements that frequently caused misinterpretation were identified:

Table 4. Problematic Elements and RSM-Based Solutions

Problematic Element	Issues Encountered	Optimization Solution
Prompt Clarity	Vague wording caused inconsistent outputs.	Use direct and precise wording (e.g., specify font type, layout, and spacing clearly).
Level of Detail	Too little detail produced generic designs; too much detail confused the AI.	Provide balanced detail: enough descriptors (font, layout, color) without redundancy.
Thematic Description	Missing theme led to generic logos; unclear themes caused irrelevant visuals.	Include a strong and consistent theme (e.g., “eco-friendly tech startup”) to guide AI output.
Number of Visual Elements	Too few made logos unrecognizable; too many created clutter and visual overload.	Limit to maximum 3 core elements (e.g., skyline + palm tree + wave) for clarity.
Color Specification	No colors led to random palettes; vague instructions (“bright”/ “dark”) caused misinterpretation.	Define a primary color with 1–2 complementary colors that carry meaning (e.g., green = growth, blue = trust)
Layout Guidance	Missing layout led to inconsistent orientation (vertical, scattered, unbalanced).	Add layout direction (e.g., “horizontal alignment, balanced spacing, white background”)
Ambiguous Wording	Terms like “modern” or “simple” were interpreted inconsistently by AI.	Replace with measurable descriptors (e.g., “clean sans-serif font, deep blue background”).

3.1.4. Optimizing the Existing Prompt Engineering

Based on RSM analysis and repetitive testing, an optimized prompt structure was developed, focusing on eight key components:

Table 5. Eight-Component Optimized Prompt Framework

Component	Explanation	Implementation Example
Main Design Focus	Defines the central idea of the logo	Focus on a skyline as the central element to represent technology and growth.
Detail Elements	Specific elements to support the main design focus	Add a palm leaf integrated with a circuit line, plus a simple wave accent.

Thematic Style	A specific theme to set the tone or aesthetic direction	Tropical eco-tech startup with minimalist yet professional aesthetics.
Primary Colors	The dominant color in the design	Deep blue is the dominant color to convey trust and professionalism.
Complementary Colors	Supporting colors that enhance balance	Bright green as an accent to highlight eco-friendly values.
Rewording	Eliminates ambiguity by choosing precise language	Use “clean sans-serif font, balanced spacing” instead of vague “modern style.”
Layout Size	Specifies the spatial proportions of the layout	Horizontal layout with balanced spacing on a white background.
Element Limit	Controls the number of visual components	Limit to 3 elements: skyline, palm/circuit, and wave.

The optimization of the prompt in this research was carried out entirely through the application of Response Surface Methodology (RSM) combined with a Central Composite Design (CCD). By defining five independent variables clarity, level of detail, thematic description, number of visual elements, and color specification the CCD generated 47 prompt combinations covering factorial, axial, and center points. Each of these combinations was tested repetitively to evaluate the quality, consistency, and interpretability of the AI-generated logo outputs. The systematic variation allowed the identification of problematic prompt elements, such as ambiguous wording, excessive use of visual elements, or vague color instructions, which were then addressed through targeted optimization strategies.

Through RSM, the interaction between variables was analyzed, enabling the determination of which factors had the most significant impact on logo quality and user satisfaction. For example, high clarity and thematic description consistently improved relevance, while excessive detail and visual elements often led to cluttered results. The CCD framework further ensured that both main effects and interaction effects among variables were captured, allowing for a more precise refinement of the prompt structure. This combination of RSM and CCD thus provided a data-driven pathway to formulate an optimized prompt template that balances detail, clarity, and creativity while reducing ambiguity and inconsistency in AI interpretations.

Based on the RSM analysis, the developed optimized prompt template is:

> "Create a [THEME] logo featuring the text [COMPANY NAME] in a [SPECIFIC FONT TYPE]. Use [PRIMARY COLOR] as the main color with [COMPLEMENTARY COLOR] accents. Focus on main elements: [3 CORE VISUAL ELEMENTS] such as [ELEMENT 1], [ELEMENT 2], and [ELEMENT 3]. Arrange elements in a [LAYOUT SPECIFICATION] with balanced spacing, set against a [BACKGROUND TYPE] background for [PURPOSE/CONTEXT].”

Implementation example:

> "Create a minimalist yet detailed logo for XYZ Company, representing an eco-friendly tropical tech startup. Use a clean sans-serif font in a bold style. Use deep blue as the dominant color and bright green as the secondary highlight. Focus on three main elements: a modern skyline, a tropical palm leaf integrated with a circuit line, and a wave accent. Arrange in a balanced horizontal layout with strong typography on a white background."

3.1.5. Visual Quality Assessment

The quality of AI-generated logos was evaluated using the Visual Quality Assessment Scale (VQAS), which measured four key dimensions: Visual Consistency (VC), Element Clarity (EC), Aesthetic Quality (AQ), and Brand Alignment (BA). Each dimension was rated on a 7-point Likert scale and combined into a weighted composite Visual Quality Score (VQS), normalized to a 0–100 scale for comparability.

Table 6. Visual Quality Assessment Statistics

Dimension	Initial	Optimized
Visual Consistency	65.2 %	87.0 %
Element Clarity	62.4 %	85.3 %
Aesthetic Quality	60.1 %	78.2 %
Brand Alignment	63.8 %	83.7 %

Across all 47 experimental prompt combinations, the VQS ranged between 61.4% and 88.0%, with optimized prompts achieving an average of 86.7 ± 1.4%, indicating an overall 22% improvement compared to the baseline results.

Table 7. Inter-Rater Reliability Statistics

Metric	Value	95% CI
ICC - Overall VQS	0.89	0.85-0.92
ICC - Visual Consistency	0.91	0.88-0.94
Fleiss' Kappa (Categories)	0.73	0.68-0.78

All indices exceeded standard reliability thresholds (ICC $\geq$ 0.75; $\kappa \geq$ 0.60). Supplementary computational metrics CLIP similarity (r=0.58), LAION aesthetic score (r=0.62), and Color RMSE (r=–0.54) showed moderate alignment with human evaluations, reinforcing VQAS reliability.

3.1.6. Response Surface Modeling (RSM) and ANOVA Analysis

The dataset, consisting of 47 structured prompt combinations tested through repeated experimental runs, was analyzed using a regression-based Response Surface Methodology (RSM) approach. The analysis applied second-order polynomial modeling, ANOVA testing, and residual diagnostics to determine the significance and interaction effects of the five prompt variables on the Visual Quality Score (VQS). This statistical process ensured that the optimized model was both robust and empirically validated. After applying backward elimination to remove statistically non-significant terms (p > 0.05), the final reduced model for visual consistency was obtained:

$$Y_1 = 78.42 + 8.17X_1 + 6.53X_2 + 7.81X_3 + 4.29X_4 + 9.06X_5 - 3.21X_1{}^2 - 2.84X_2{}^2 - 3.47X_3{}^2 - 4.13X_4{}^2 - 2.91X_5{}^2 + 2.08X_1X_2 + 3.37X_1X_3 + 1.76X_2X_3 + 2.69X_3X_5 \quad (7)$$

Where X_1 through X_5 represent the coded values of Prompt Clarity, Level of Detail, Thematic Description, Number of Visual Elements, and Color Specification, respectively. All terms retained in the model demonstrated statistical significance at $\alpha = 0.05$, ensuring that only meaningful predictors contribute to the response prediction.

Analysis of Variance (ANOVA) was conducted to evaluate the statistical significance and adequacy of the developed model.

Table 8. Analysis of Variance (ANOVA) for Response Surface Quadratic Model of Visual Consistency

Source	Sum of Squares	df	Mean Square	F-Value	p-Value
Model	3847.62	14	274.83	47.23	<0.0001***
X_1-Clarity	1523.41	1	1523.41	261.89	<0.0001***
X_2-Detail	957.28	1	957.28	164.56	<0.0001***
X_3-Theme	1378.52	1	1378.52	237.02	<0.0001***
X_4-Elements	419.37	1	419.37	72.11	<0.0001***
X_5-Color	1876.93	1	1876.93	322.67	<0.0001***
X_1X_2	98.45	1	98.45	16.93	0.0002**
X_1X_3	261.78	1	261.78	45.01	<0.0001***
X_2X_3	73.24	1	73.24	12.59	0.0002**
X_3X_5	164.82	1	164.82	28.34	<0.0001***
X_1^2	231.57	1	231.57	39.81	<0.0001***
X_2^2	177.39	1	177.39	30.50	<0.0001***
X_3^2	277.64	1	277.64	47.74	<0.0001***
X_4^2	381.29	1	381.29	65.56	<0.0001***
X_5^2	189.93	1	189.93	32.66	<0.0001***
Residual	186.14	32	5.82		
Lack of Fit	98.27	27	3.64	0.41	0.9543 n.s.
Pure Error	4033.76	5	17.57		

R^2=0.954, Adj-R^2=0.934, Pred-R^2=0.908, Adeq. Precision=24.73 ***p<0.001, **p<0.01, n.s.=not significant.

The model is highly significant (F=47.23, p<0.0001) with R^2=0.954 indicating 95.4% variance explained. Non-significant lack-of-fit (p=0.9543) confirms model adequacy. Residual analysis showed normal distribution (Shapiro-Wilk: W=0.967, p=0.214), constant variance, and no influential outliers (Cook's D<1.0). Box-Cox plot (λ=1.02, 95%CI: 0.87-1.19) indicated no transformation needed.

Model validation at optimized conditions (X_1=+1.5, X_2=+1.2, X_3=+1.8, X_4=+0.8, X_5=+1.6) predicted Y_1=86.7%. Experimental verification yielded 87.2±1.4% (n=3), representing 0.58% prediction error, confirming model validity.

3.1.7. Implementation of the Optimized Prompt Engineering

To validate the effectiveness of the optimized prompt structure, the prompt was applied to AI with additional testing on different logos. Each logo was evaluated based on three indicators: visual consistency, element clarity, and user satisfaction. The results of the trial demonstrated that logos generated from the structured prompt showed a significant improvement compared to those generated from the unstructured initial prompt.

Table 9. Multi-Context Validation Results

Prompt	Results
Create a minimalist logo featuring the text Batam Tech Solutions representing an eco-friendly tropical tech startup. Use vibrant green as the main color with bright orange accents. Focus on three core visual elements such as Geometric cube, Arrow thats wrap cube, and Connectivity dots shape. Arrange elements in a horizontal layout with balanced spacing, set against a white background for a professional and innovative brand identity.	
Create a beauty & lifestyle logo featuring the text Nail Art Studio Batam in a script font. Use pink as the main color with gold accents. Include up to 3 core visual elements such as a nail polish brush, abstract flower, and sparkle icon. Arrange elements in a balanced centered layout with clear spacing, set against a light background for luxury branding.	
Create an education & innovation logo featuring the text Batam Smart School in a rounded sans-serif font. Use navy blue as the main color with yellow accents. Include up to 3 core visual elements such as an open book, light bulb, and digital pixel. Arrange elements in a centered layout with balanced spacing, set against a light background for academic branding.	
Create a healthcare & wellness logo featuring the text Batam Healthy Living Clinic in a clean sans-serif font. Use green as the main color with white accents. Include up to 3 core visual elements such as a heart shape, cross symbol, and leaf. Arrange elements in a horizontal layout with balanced spacing, set against a white background for medical branding.	

3.1.8. Evaluation of the Results

After conducting the processes of experimentation, optimization, and both quantitative and After conducting the processes of experimentation, optimization, and both quantitative and qualitative analysis, this stage aims to comprehensively evaluate the effectiveness of the developed prompt engineering technique. The evaluation was carried out by comparing the performance of the initial prompt and the optimized prompt based on three key indicators:

1. Efficiency
2. Creativity of design
3. User experience

The evaluation was conducted through a triangulation approach, combining experimental test results, visual analysis, and interview data. Qualitative responses were analyzed using thematic analysis, categorizing participants' perceptions based on their level of agreement ("agree," "neutral," and "disagree") to identify recurring patterns and themes. Subsequently, a percentage-based mean score was calculated to illustrate the proportion of participants who agreed with each thematic finding, providing quantitative support for the qualitative insights.

3.1.8.1. Efficiency

Efficiency was assessed based on the average time required by users to obtain a design outcome that matched their expectations. The testing revealed that:

1. The initial prompt required an average of 5–8 repetitions to generate a logo considered acceptable.
2. The optimized prompt required only 3–4 attempts, and in some cases, it was successful on the second try.

Table 10. Efficiency Metrics: Iterations and Time

Parameter	Initial Prompt	Optimized Prompt
Average Iterations	6.6 times	3.4 times
Average Time per Logo	± 18 minutes	± 10 minutes
Visual Match Rate	68%	91%

3.1.8.2. Creativity of Design

Creativity was assessed based on respondents' feedback and visual comparison of the design outcomes. Logos generated from the optimized prompts exhibited more consistent, aesthetic characteristics and conveyed brand identity more strongly. Based on the visual evaluation results:

1. The initial prompt often produced logos that were too generic and lacked distinctive value.
2. The optimized prompt produced logos with more directed style, color, and form, aligned with the company's context and user preferences.

Table 11. Creative Quality Metrics by Prompt Type

Evaluation Aspect	Initial Prompt	Optimized Prompt
Visual Style Clarity	64%	92%
Thematic Relevance	60%	94%
Differentiation/Uniqueness	60%	90%

3.1.8.3. User Experience

From the interviews and thematic analysis, it was found that users felt more confident and satisfied when using the optimized prompt. Several aspects of user experience that showed improvement included:

1. A stronger sense of control over the AI output
2. Greater visual satisfaction with the design results
3. Ease of understanding and applying the prompt template

Most respondents stated that they would prefer to use AI with this structured approach for professional design work in the future.

3.1.8.4. Comprehensive Comparison: Before and After Optimization

Table 12. Comprehensive Performance Improvement Summary

Aspect	Initial Prompt	Optimized Prompt
Image Consistency	65%	87%
Visual Clarity	60%	85%
User Satisfaction	58%	82%
Iteration Efficiency	6.6 iterations/ result	3.4 iterations/ result
Average Time	18 minutes	10 minutes

3.1.8.5. Evaluation Implications

Based on the final evaluation, it can be concluded that the optimized prompt engineering technique:

1. Is more effective in producing logo designs that meet user expectations
2. Enhances the accessibility of AI in the design field, particularly for non-technical users
3. Reduces the need for repeated iterations, thereby accelerating the design process
4. Significantly improves the visual quality and aesthetic value of the logos

These findings demonstrate that a structural approach to prompt engineering is not only beneficial in a technical context but also has a positive impact on user experience and the quality of creative outcomes.

3.1.9. Evaluation Implications

This research involved 30 purposively selected participants, consisting of freelance and company-based designers engaged in logo design. Participants were divided into two workflow groups:

1. Group A: Manual-based designers who primarily rely on conventional tools such as Adobe Illustrator and Photoshop, with minimal AI exposure (< 6 months).
2. Group B: AI-based designers who actively utilize generative tools such as ChatGPT and DALL·E in their creative process, possessing more than one year of AI experience.

All participants had between one to three years of professional experience and met the inclusion criteria of being either freelance designers or employees of design companies. This research was conducted in Batam, and participation was entirely voluntary, with informed consent obtained before data collection.

Data were collected through semi-structured interviews conducted after the experimental testing phase. These interviews aimed to capture participants' qualitative perspectives on efficiency, creativity, and user experience when applying the optimized prompt engineering framework.

Table 13. Participant Demographics by AI Experience

Group	Description	AI Experience
A	Graphic designers using Adobe	Minimal (< 6 months)
B	Graphic designers using ChatGPT/DALL-E	Experienced (> 1 year)

3.1.9.1. Comparison between Manual Designer and AI Users

The majority of respondents stated that the new prompt structure was easier to understand, even for designers who were not accustomed to using AI. They felt that the language used in the prompt was sufficiently representative of visual design needs. Most respondents indicated that the logos generated from the optimized prompts appeared more creative and unique. They felt that AI was able to capture ideas more accurately when the prompt descriptions were more structured. Overall, respondents reported an improvement in user experience (UX). They felt more satisfied with the outcomes and more confident in using AI for logo design in the future.

These thematic findings were further supported by a percentage-based agreement analysis, showing that the majority of participants expressed positive responses indicating a high level of agreement toward the improved clarity, creativity, and usability of the optimized prompt structure.

Table 14. Comparison between Manual Designer and AI Users

Group	Effectiveness	Creativity	User Experience
A	82%	78%	80%
B	90%	86%	88%

3.2. Ethical Considerations

This research was conducted in accordance with academic research ethics and transparency principles. No personal, confidential, or corporate data were collected or used during the study. All procedures involving participants such as interviews and evaluations were carried out voluntarily, with participants providing informed consent prior to data collection. All responses were anonymized, and no identifiable personal information was stored or disclosed.

Since this research focused on experimental testing of AI-generated logo prompts and did not involve sensitive or high-risk human data, formal Institutional Review Board (IRB) approval was not required. However, ethical guidelines related to participant privacy, data handling, and research integrity were fully observed. All visual materials and logo images used in this study were generated using DALL-E 3 under OpenAI's usage license. These outputs are utilized solely for academic, non-commercial, and educational purposes. The ownership of generated images remains governed by OpenAI's standard terms of service, and no copyrighted or proprietary assets were replicated or redistributed.

4. Conclusion

This research successfully developed and validated an optimized prompt engineering framework for AI-based logo generation using the Response Surface Methodology (RSM). The framework was structured into eight key components: Main Design Focus, Detail Elements, Thematic Style, Primary Colors, Complementary Colors, Rewording, Layout Size, and Element Limit. Compared with unstructured prompt approaches, this optimized framework demonstrated significant improvements in both visual consistency and creative quality.

Using Central Composite Design (CCD), a total of 47 prompt combinations were systematically tested and analyzed through ANOVA. The statistical results showed a strong and reliable model fit (R^2 = 0.954, adjusted R^2 = 0.934, $p < 0.001$) with no significant lack-of-fit, confirming that the optimized equation accurately represents the relationship between prompt variables and design quality. Each variable clarity, detail level, thematic description, color specification, and visual element count was found to significantly influence visual outcomes, validating the effectiveness of the model's predictive structure.

Performance metrics showed substantial improvement after optimization. Visual consistency increased from 65% to 87%, element clarity from 60% to 85%, and user satisfaction from 58% to 82%. The number of design iterations decreased from 6.6 to 3.4, resulting in approximately 44% faster completion time. These results demonstrate that a statistically guided and structured prompt design can significantly enhance both the technical and creative aspects of AI-assisted logo generation.

Furthermore, the optimized framework proved intuitive and adaptable for both manual designers and AI-based practitioners. Despite different levels of experience, both groups achieved comparable creative performance and satisfaction levels, showing that the system supports inclusivity and accessibility. The framework not only enhances the creative workflow but also bridges human design intent with AI interpretation, empowering designers to work more effectively with generative systems.

References

[1] C. Li *et al.*, "AGIQA-3K: An Open Database for AI-Generated Image Quality Assessment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 8, pp. 6833–6846, Aug. 2024, doi: 10.1109/TCSVT.2023.3319020.

[2] Y. Feng *et al.*, "PromptMagician: Interactive Prompt Engineering for Text-to-Image Creation," *IEEE Trans Vis Comput Graph*, vol. 30, no. 1, pp. 295–305, Jan. 2024, doi: 10.1109/TVCG.2023.3327168.

[3] V. Liu, J. Vermeulen, G. Fitzmaurice, and J. Matejka, "3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows," *Association for Computing Machinery (ACM)*, vol. 1, no. 2, pp. 1955–1977, Jul. 2023, doi: 10.1145/3563657.3596098.

[4] T. Papadopoulos and Y. Charalabidis, "What do governments plan in the field of artificial intelligence?: Analysing national AI strategies using NLP," *ACM International Conference Proceeding Series*, vol. 20, no. 1, pp. 100–111, Sep. 2020, doi: 10.1145/3428502.3428514.

[5] J. Wang *et al.*, "Review of large vision models and visual prompt engineering," *Meta-Radiology*, vol. 1, no. 3, pp. 1–13, Dec. 2023, doi: 10.1016/j.metrad.2023.100047.

[6] J. Wang, H. Duan, J. Liu, S. Chen, X. Min, and G. Zhai, "AIGCIQA2023: A Large-Scale Image Quality Assessment Database for AI Generated Images: From the Perspectives of Quality, Authenticity and Correspondence," *Springer Nature Singapore*, vol. 14474, no. 1, pp. 46–57, Feb. 2024, doi: 10.1007/978-981-99-9119-8_5.

[7] M. Verdicchio, "Ekphrasis and prompt engineering. A comparison in the era of generative AI," *Studi di Estetica*, vol. 1, no. 28, pp. 59–78, Mar. 2024, doi: 10.7413/1825864661.

[8] L. Giray, "Prompt Engineering with ChatGPT: A Guide for Academic Writers," *Ann Biomed Eng*, vol. 51, no. 12, pp. 2629–2633, Dec. 2023, doi: 10.1007/s10439-023-03272-4.

[9] A. KUZMIN and O. PAVLOVA, "Analysis of Artificial Intelligence Based Systems for Automated Generation of Digital Content," *Computer systems and information technologies*, vol. 1, no. 1, pp. 82–88, Mar. 2024, doi: 10.31891/csit-2024-1-10.

[10] M. Mitchell, "The Turing Test and our shifting conceptions of intelligence," *Science (1979)*, vol. 385, no. 6710, pp. 1–8, Aug. 2024, doi: 10.1126/science.adq9356.

[11] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Springer*, vol. 31, no. 2, pp. 685–695, Apr. 2021, doi: 10.1007/s12525-021-00475-2.

[12] A. Koubaa, W. Boulila, L. Ghouti, A. Alzahem, and S. Latif, "Exploring ChatGPT Capabilities and Limitations: A Critical Review of the NLP Game Changer," *Preprints (Basel)*, vol. 1, no. 1, pp. 1–29, Mar. 2023, doi: 10.20944/preprints202303.0438.v1.

[13] K. Aggarwal *et al.*, "Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning," *Iraqi Journal for Computer Science and Mathematics*, vol. 3, no. 1, pp. 115–123, Jan. 2022, doi: 10.52866/ijcsm.2022.01.01.013.

[14] F. French, D. Levi, C. Maczo, A. Simonaityte, S. Triantafyllidis, and G. Varda, "Creative Use of OpenAI in Education: Case Studies from Game Development," *Multimodal Technologies and Interaction*, vol. 7, no. 8, p. 1, Aug. 2023, doi: 10.3390/mti7080081.

[15] T. B. Brown *et al.*, "Language Models are Few-Shot Learners," *ArXiv*, vol. 1, no. 4, pp. 1–75, May 2020, doi: 10.48550/arXiv.2005.14165.

[16] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large Language Models are Zero-Shot Reasoners," *ArXiv*, vol. 1, no. 4, pp. 1–42, May 2023, doi: 10.48550/arXiv.2205.11916.

[17] M. GÜNAY, "THE IMPACT OF TYPOGRAPHY IN GRAPHIC DESIGN," *International Journal Of Eurasia Social Sciences*, vol. 15, no. 57, pp. 1446–1464, Jan. 2024, doi: 10.35826/ijoess.4519.

[18] S. L. Franconeri, L. M. Padilla, P. Shah, J. M. Zacks, and J. Hullman, "Corrigendum: The science of visual data communication: What works," *Psychological Science in the Public Interest*, vol. 23, no. 1, pp. 110–161, May 2022, doi: 10.1177/15291006211051956.

[19] M. Fan and Y. Li, "The application of computer graphics processing in visual communication design," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5183–5191, Jan. 2020, doi: 10.3233/JIFS-189003.

[20] J. Wu, "Research on the Combination of UI and Graphic Design in Museums in the Digital Age," *Transactions on Social Science, Education and Humanities Research*, vol. 7, no. 1, pp. 366–370, May 2024, doi: 10.62051/6vww1j63.

[21] H. Li, T. Xue, A. Zhang, X. Luo, L. Kong, and G. Huang, "The application and impact of artificial intelligence technology in graphic design: A critical interpretive synthesis," *Heliyon*, vol. 10, no. 21, pp. 11–21, Nov. 2024, doi: 10.1016/j.heliyon.2024.e40037.

[22] B. Mustafa, "The Impact of Artificial Intelligence on the Graphic Design Industry," *Arts and Design Studies*, vol. 104, no. 1, pp. 1–9, Mar. 2023, doi: 10.7176/ADS/104-01.

[23] Y. Meron, "Graphic design and artificial intelligence: Interdisciplinary challenges for designers in the search for research collaboration," *DRS Digital Library*, vol. 1, no. 1, pp. 1–17, Jun. 2022, doi: doi.org/10.21606/drs.2022.157.

[24] C.-L. Cheng, "Alive Scene: Participatory Multimodal AI Framework for Collective Narratives in Dynamic 3D Scene by," *Massachusetts Institute of Technology*, vol. 1, no. 1, pp. 1–91, Apr. 2024.

[25] P. Korzynski, G. Mazurek, P. Krzypkowska, and A. Kurasinski, "Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT," *Entrepreneurial Business and Economics Review*, vol. 11, no. 3, pp. 25–37, Sep. 2023, doi: 10.15678/EBER.2023.110302.

[26] T. Sanchez, "Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs?," *Association for Computing Machinery*, vol. 1, no. 1, pp. 43–61, Jun. 2023, doi: 10.1145/3591196.3593051.

[27] A. M. Goloujeh, A. Sullivan, and B. Magerko, "Is It AI or Is It Me? Understanding Users' Prompt Journey with Text-to-Image Generative AI Tools," *Association for Computing Machinery*, vol. 1, no. 183, pp. 1–13, May 2024, doi: 10.1145/3613904.3642861.

[28] Z. Lai, X. Zhu, J. Dai, Y. Qiao, and W. Wang, "Mini-DALLE3: Interactive Text to Image by Prompting Large Language Models," *ArXiv*, vol. 1, no. 2, pp. 1–20, Oct. 2023, doi: 10.48550/arXiv.2310.07653.

[29] J. White *et al.*, "A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT," *ArXiv*, vol. 1, no. 1, pp. 1–9, Feb. 2023, doi: 10.48550/arXiv.2302.11382.

[30] X. Chen, Y. Chen, Y. Xie, and L. Cheng, "Translation and psychometric validation of the Medical Artificial Intelligence Readiness Scale (MAIRS-MS) for Chinese medical students," *BMC Nurs*, vol. 24, no. 1, pp. 1–10, Dec. 2025, doi: 10.1186/s12912-025-03852-w.