Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1337–1347
https://doi.org/10.54941/ahfe1008362

A Human–AI Collaborative Mechanism for Commercial Packaging Design With Generative AI

Zibin Liu

China Academy of Art, Hangzhou 310002, China

ABSTRACT

This study the rapid adoption of generative artificial intelligence (GenAI) has significantly expanded its role in commercial packaging design, enabling designers to produce visual concepts with unprecedented speed. However, existing studies have primarily emphasized generation efficiency and visual quality, while paying limited attention to the stability of generated results, their commercial applicability, and the evolving relationship between designers and AI throughout the design process. This study investigates the aesthetic generation characteristics and practical limitations of GenAI in commercial packaging design and proposes a designer-centered human–AI collaborative creation framework. A mixed-method approach combining literature review, controlled experiments, and case analysis was employed. Under identical generation conditions—including the same model, prompts, and generation parameters—commercial packaging images were repeatedly generated and comparatively analyzed in terms of visual style, packaging structure, brand consistency, cultural expression, and commercial feasibility. The results indicate that GenAI can efficiently produce visually coherent design concepts while maintaining a consistent overall stylistic direction. Nevertheless, substantial variations remain in packaging structure, typographic legibility, brand identity, material organization, and cultural interpretation due to probabilistic sampling, training data distribution, and model-specific aesthetic preferences. These findings suggest that current GenAI systems are effective for early-stage ideation but remain insufficient for directly producing commercially deployable packaging solutions. Based on these findings, this study proposes a designer-led human–AI collaborative workflow in which designers retain responsibility for brand strategy, cultural interpretation, aesthetic judgment, and final decision-making, while GenAI supports concept exploration, design expansion, and iterative refinement. The proposed framework provides both theoretical insights and practical guidance for integrating generative AI into commercial packaging design while preserving design quality, brand coherence, and commercial viability.

Keywords: Generative artificial intelligence (GenAI), Commercial packaging design, Human–AI collaboration, Aesthetic generation, Collaborative design

INTRODUCTION

Text-to-image generative artificial intelligence (GenAI), often described as AIGC, now supports packaging visualization, brand promotion, and cultural-product design. Rapid generation shifts design work from manual

Received August 26, 2026; Revised September 23, 2026; Accepted October 5, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

production toward prompt construction, selection, aesthetic judgment, and correction (Tang et al., 2025; Pu and Yang, 2023).

However, model defaults can weaken shelf differentiation and misread brand tone, cultural symbols, mandatory information, and relations among packaging components. An attractive rendering may therefore remain incomplete as a brand and packaging system, particularly for premium, cultural, and series-based products (Xu et al., 2023).

This study asks how GenAI's visual exploration can be translated into stable, commercially applicable packaging through designer-led collaboration. It proceeds from problem identification to framework construction and operational validation. Commercial cases and repeated outputs inform a three-part framework, and an 18-stimulus consumer study tests structured prompting as one operational proposition.

AESTHETIC GENERATION AND PROBLEM FRAMING

Aesthetic Generation and Packaging Requirements

Generative art uses rules and stochastic processes to produce outcomes partly independent of direct manual control. Contemporary models learn statistical relations from image-text datasets and recombine color, composition, material, and style under textual or visual constraints (Gatys et al., 2015; Ramesh et al., 2021; Rombach et al., 2022).

Model aesthetics reflect training distributions, platform optimization, and probabilistic sampling rather than lived experience or independent judgment. A coherent image may therefore lack understanding of a symbol's product, brand, or cultural significance. Designers must define the problem, judge alternatives, correct errors, and integrate selected material into a deployable system (Liang et al., 2025; Spence and Van Doorn, 2022; Tang et al., 2025; Wang et al., 2025).

Accordingly, image coherence is not treated as design completion. The analysis examines product meaning, packaging-system consistency, and brand differentiation as requirements connecting aesthetic generation to commercial packaging.

LIMITATIONS IN COMMERCIAL PACKAGING DESIGN

Limits of Cultural Interpretation

Generated novelty mainly reflects recombination rather than lived experience or cultural judgment. Commercial packaging must connect product position, brand memory, use context, and visual symbols; broad labels such as premium or traditional Chinese style may produce recognizable surfaces without product-specific meaning (Wang, 2024; Zhao, 2025).

The Wanglaoji and Yili cases illustrate this boundary. Landscape imagery gives the Wanglaoji can immediate cultural recognition, while ink scenery and osmanthus link Yili packaging to Jiangnan and the Hangzhou Asian Games. However, these references do not by themselves explain their relation to herbal tea, milk nutrition, or daily use. Cultural interpretation thus links recognizable imagery to a product-specific narrative and remains a designer responsibility (Liang et al., 2025).



(a) (b)

Figure 1: Cultural translation in commercial packaging: (a) Wanglaoji AI Thousand Li of Rivers and Mountains can; (b) Yili AI Memories of Jiangnan Asian Games milk package. Sources: public promotional materials released by the respective brands.

Stochastic Output and Packaging-System Instability

Probabilistic sampling causes outputs to vary even under the same model, prompt, and parameters, including composition, product count, box structure, logo placement, supporting materials, and scene relations. Such variation aids exploration, but commercial series require repeatable hierarchy, structural logic, and identity across components (Chong et al., 2025).

Figure 2 retains a blue-green palette, landscape motifs, and a premium gift tone across repeated tea-packaging outputs. However, opening methods, divisions, containers, cards, text, and visual centers change across the batch. Broad style consistency supports direction finding but not a coherent, producible packaging system.



Figure 2: Commercial tea-packaging outputs generated by the same model under the same prompt and parameter conditions. Source: generated and compiled by the author.

AI expands alternatives, while designers stabilize packaging-system relations through selection, reconstruction, and production calibration. Generated sketches become useful when these decisions are documented and tested against explicit structural and series requirements (Tang et al., 2025; Yang and Li, 2025).

Table 1: Qualitative comparison of repeated generation under fixed conditions.

Dimension	Stable Across Batch	Observed Variation	Design Implication
Visual style	Blue-green palette; landscape; premium gift tone	Color depth, light, motif scale, emphasis	Useful for direction, not a stable brand system
Package structure	Gift box, bag, tea pack, bottle, card	Opening, divisions, hierarchy, display order	Not ready for structural development
Brand and type	Names, marks, ornaments attempted	Position, scale, clarity, legibility	Manual identity and typesetting required
Material system	Multi-component series scenes	Quantity, proportion, function, priority	Weak hierarchy and series continuity
Cultural meaning	Landscape, bamboo, seals, lattice, utensils	Decorative mixing; unstable semantics	Human cultural and brand integration required

Style Homogenization and Brand Differentiation

Style homogenization reflects overlapping training data, similar pipelines, recurring prompt conventions, and platform preferences for saturated color, centered products, dramatic light, dense texture, and decorative scenery. These features create immediate impact but can narrow outcomes and reinforce generic motifs (Hintze et al., 2026).

Across the Nestle series in Figure 3, different products repeatedly appear as centered packages within idealized landscapes, bright color, atmospheric light, and decorative motifs. Distinct use settings and emotional positions thus converge into the same polished natural world, sometimes allowing the background to overshadow the product message. Visual completeness can therefore conceal weak category and brand differentiation.



Figure 3: Nestle commercial poster series using AI-generated imagery. Source: public brand promotional materials.

A DESIGNER-LED COLLABORATIVE FRAMEWORK

Integrated Collaboration Architecture

The framework responds to cultural ambiguity, structural variation, and stylistic convergence through a shared design brief. The brief defines product position, users, benefits, cultural evidence, structure, required information,

production limits, channel, and use setting, thereby narrowing the search space and establishing comparison criteria.

Prompts translate the brief into structured design conditions: cultural packaging specifies motif sources and meanings; food packaging defines information hierarchy and consumption context; gift packaging states opening methods, component relations, and series requirements. AI explores visual alternatives, while the designer evaluates them against the brief, revises constraints, and reconstructs the selected direction. Figure 4 connects these constraints, AI exploration, and designer reconstruction within one collaboration architecture.

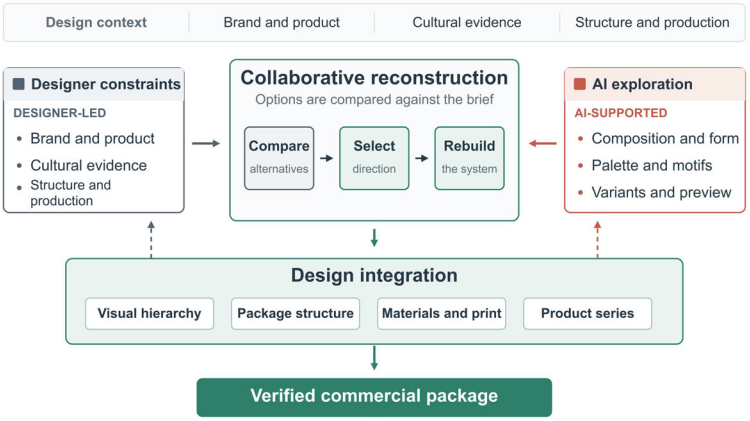


Figure 4: Overall architecture for designer-led human-AI collaboration in commercial packaging. Source: developed by the author from the study findings.

Operational Mechanism and Responsibility

Roles are assigned by task and deliverable. Designers own brand strategy, cultural interpretation, aesthetic criteria, compliance, and release decisions; AI supports sketch generation, asset expansion, style testing, and local refinement. Generated images remain provisional concept material, not production files (Wang and Shao, 2025).

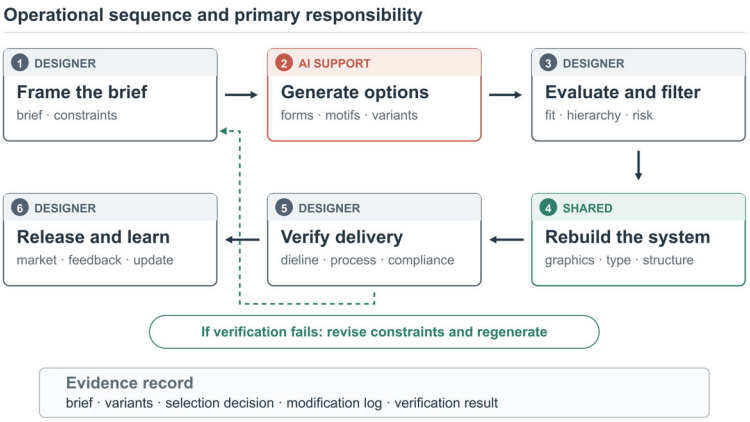


Figure 5: Operational mechanism for packaging-design implementation under designer-led human-AI collaboration. Source: developed by the author from the study findings.

After selection, the designer rebuilds typography and approved brand assets, organizes mandatory information, and adapts graphics to dielines, materials, finishes, and printing. Each iteration records inputs, selection reasons, rejections, and modifications, while the project lead verifies cultural sources, copy, structure, regulation, and intellectual property. Figure 5 connects this implementation loop to explicit responsibility at each decision point.

Collaborative Value Creation and Evaluation

Collaborative value can be evaluated across exploration, development, and production through usable-option rate, direction-selection time, identity and information continuity, dieline fit, production compatibility, correction effort, and prototype defects. These measures connect output quality to design decisions rather than generation volume.

Cultural packaging adds verification of motif sources, product-narrative relations, and possible simplification or misuse. Figure 6 organizes value at efficiency, brand, aesthetic-cultural, and application levels so consequential decisions remain verifiable and attributable (Liang et al., 2025; Ishihara et al., 2023; Yang and Li, 2025).

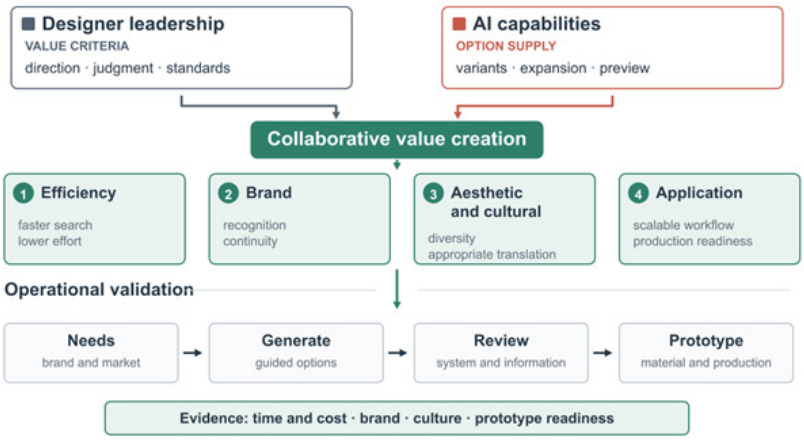


Figure 6: Human-AI collaborative value-creation system for commercial packaging design. Source: developed by the author from the study findings.

CONSUMER VALIDATION: DESIGN AND RESULTS

Validation Design and Analytical Scope

The consumer study examines one operational proposition of the framework: whether designer-led structured prompts improve cultural relevance and information organization while preserving commercial appeal. An 18-stimulus comparison and criterion-based review assess this proposition.

The 18 stimuli comprise six real Chinese tea packages (Condition A), six ordinary-prompt images (B), and six designer-led structured-prompt images (C). Conditions B and C use the fictional HUAXU Wuyi rock-tea brand, the same rigid-box and tin format, and six matched cultural concepts. Condition C adds rules for brand and information hierarchy, cultural evidence, palette, materials, printing, and physical plausibility.

Images were generated once with OpenAI's image-generation tool on 11 August 2026 (1254 × 1254 pixels), using a common viewpoint and studio setting without retouching.

Nine balanced questionnaire blocks each present six images, two per condition. Every stimulus appears in three blocks and sequence positions. The 287 adults report recent purchases of tea-related products, rate 12 bipolar items on a seven-point scale, select visible defects, and identify one preferred package. The items form five constructs: visual quality, brand, information, cultural performance, and commercial response (2, 2, 2, 3, and 3 items, respectively).

Averaging construct items and the two images per condition within participants yields 287 paired B-C observations. Paired t tests compare means; Wilcoxon signed-rank tests assess ordinal-data sensitivity; Holm correction addresses five comparisons. A is descriptive because brands and photography differ. Reliability uses Spearman-Brown coefficients for two-item constructs and Cronbach's alpha for three-item constructs.

All 287 respondents meet consent, adult-age, recent-purchase, attention-check, and complete-response criteria.

The author reviews all 18 stimuli for brand logic, typography, cultural relevance, information hierarchy, structure, material cues, series consistency, and apparent production readiness, connecting consumer responses to visible design features.

Validation Stimuli and Comparison Conditions

Figure 7 compares the three conditions. B and C are matched by cultural concept; commercial benchmarks in A provide external market references.



Figure 7: Complete set of 18 stimuli used in consumer validation. Rows show commercial benchmarks (A), ordinary prompts (B), and designer-led structured prompts (C); columns pair B and C by cultural concept.

Reliability and Construct Comparisons

The five constructs show good internal consistency, with Spearman-Brown coefficients of 0.843–0.862 and Cronbach's alpha of 0.897–0.902.

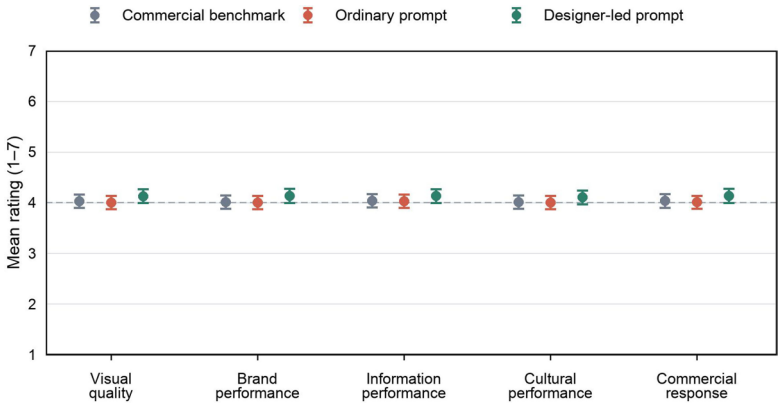


Figure 8: Consumer construct means and 95% confidence intervals by condition. The dashed line marks the neutral midpoint of 4. Source: B01-B09 consumer survey data.

Table 2: Mean comparison between ordinary and designer-led prompt conditions.

Construct	B mean	C mean	Difference (C-B)
Visual quality	3.999	4.125	0.125
Brand performance	4.003	4.128	0.125
Information performance	4.023	4.129	0.106
Cultural performance	4.000	4.105	0.105
Commercial response	4.007	4.134	0.127

C shows slightly higher mean ratings than B across all five constructs (Table 2; Figure 8), with differences of 0.105–0.127 points. The comparisons remain nonsignificant after Holm correction (all $p = 0.214$).

Defect Diagnosis and Preference

Figure 9 shows fewer cultural-use, information-hierarchy, material-quality, and text-legibility problems in C, but more unclear brand positioning. The defect pattern is selective rather than uniform.

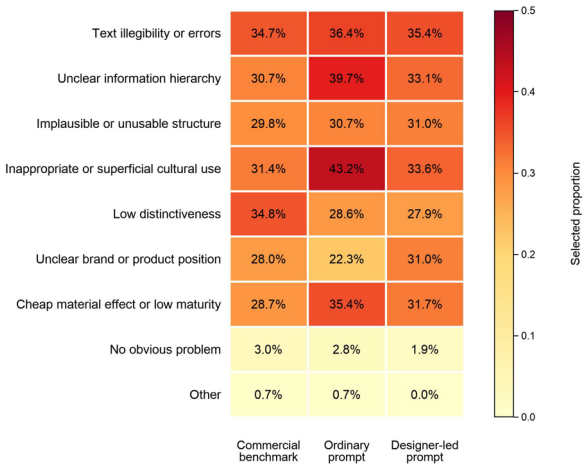


Figure 9: Proportion of stimulus evaluations selecting each visible defect by condition. Multiple defects may be selected for one stimulus. Source: B01-B09 consumer survey data.

Preferred-package shares are 38.7% for C, 26.8% for B, and 34.5% for A, providing complementary descriptive evidence.

Design Review Synthesis

The author's review connects this consumer pattern to design decisions: structured prompts clarify motif-product relations and information hierarchy in several concepts, while typography, brand assets, and container relationships remain priorities for designer reconstruction.

DISCUSSION

The three-part framework connects the preceding diagnosis to design action: cultural ambiguity is addressed through sourced narratives, stochastic variation through component and hierarchy rules, and homogenization through brand-specific selection criteria. Its architecture defines constraints, its operational mechanism assigns responsibility, and its evaluation criteria connect collaboration to packaging outcomes. The consumer comparison illustrates prompt structuring as one practical entry point into this framework.

The practical contribution is to make designer intervention explicit and traceable across generation, selection, reconstruction, and verification. Consumer feedback informs these decisions, while typography, dielines, materials, compliance information, and approved brand assets remain part of technical development.

The single tea category, six stimulus pairs, and short median completion time (176 seconds) call for replication. Future studies should recruit across channels and combine digital ratings with physical inspection, correction-time records, and dieline, material, and print-prototype checks.

CONCLUSION

This study addresses how GenAI's rapid visual exploration can be translated into commercially applicable packaging. The proposed designer-led mechanism connects cultural interpretation, system stability, and brand differentiation to a shared brief, explicit responsibilities, and evaluation criteria. AI expands visual alternatives; designers establish product-specific meaning, coordinate components and information, and reconstruct selected concepts for technical development. Linking problem diagnosis, collaboration architecture, and implementation provides a practical route from generated images to packaging decisions.

The 18-stimulus study with 287 consumers illustrates the prompt-level application of this mechanism: structured prompts show slightly higher mean ratings and preference, with selective reductions in reported defects; the construct comparisons are not statistically significant. These findings point to cultural and information requirements as priorities for designer intervention. The contribution is a traceable framework connecting generative exploration, design judgment, and packaging development, with consumer feedback informing refinement. Physical-package and production testing can further assess its commercial application.

ACKNOWLEDGMENT

The author sincerely thanks all participants who contributed to the consumer evaluation. Their time and feedback support the empirical analysis and the refinement of the designer-led human–AI collaborative framework presented in this study.

REFERENCES

Chong, L., Lo, I.P., Rayan, J., et al. (2025). Prompting for products: Investigating design space exploration strategies for text-to-image generative models. Design Science, 11, e2. https://doi.org/10.1017/dsj.2024.51

Gatys, L.A., Ecker, A.S. and Bethge, M. (2015). A neural algorithm of artistic style. arXiv:1508.06576.

Hintze, A., Proschinger Åström, F. and Schossau, J. (2026). Autonomous language-image generation loops converge to generic visual motifs. Patterns, 7, 101451. https://doi.org/10.1016/j.patter.2025.101451

Ishihara, S., Kuo, R. and Ishihara, K. (2023). AI image generation boosts Kansei engineering design process. Kansei Engineering, 101, 23–33. https://doi.org/10.54941/ahfe1002988

Liang, X.Y., Xie, X.F. and Dou, H. (2025). Innovative design of Guangxi ethnic minority IP and derivative products based on generative artificial intelligence technology. Packaging Engineering, 46(8), 364–376. https://doi.org/10.19554/j.cnki.1001–3563.2025.08.034

Pu, Z.H. and Yang, B. (2023). Application of artificial intelligence-assisted technology in packaging design. Packaging Engineering, 44(12), 273–281. https://doi.org/10.19554/j.cnki.1001-3563.2023.12.030

Ramesh, A., Pavlov, M., Goh, G., et al. (2021). Zero-shot text-to-image generation. arXiv:2102.12092.

Rombach, R., Blattmann, A., Lorenz, D., et al. (2022). High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 10684–10695. https://doi.org/10.1109/CVPR52688.2022.01042

Spence, C. and Van Doorn, G. (2022). Visual communication via the design of food and beverage packaging. Cognitive Research: Principles and Implications, 7, 42. https://doi.org/10.1186/s41235-022-00391-9

Tang, X., Windham, J. and Thomas, J. (2025). Human-AI collaboration within industrial design: The Argyle Design Framework. Interdisciplinary Practice in Industrial Design, 183, 40–50. https://doi.org/10.54941/ahfe1006425

Wang, N., Kim, H., Peng, J., et al. (2025). Exploring creativity in human-AI co-creation: A comparative study across design experience. Frontiers in Computer Science, 7, 1672735. https://doi.org/10.3389/fcomp.2025.1672735

Wang, W. and Shao, J.F. (2025). Human-intelligence interactive collaborative design. Packaging Engineering, 46(2), 1–12. https://doi.org/10.19554/j.cnki.1001–3563.2025.02.001

Wang, Y. (2024). Research on application strategies of AIGC technology in packaging design. China Packaging, 44(6), 60–63.

Xu, C., Du, X.Z. and Yu, K.D. (2023). Challenges and strategies of AIGC application in the design industry. AI-View, 4, 51–60. https://doi.org/10.16453/j.2096-5036.2023.04.006

Yang, Y. and Li, H. (2025). Design evaluation system of AI-generated content in the industrial design of construction machinery. Human-Computer Interaction and Emerging Technologies, 195, 61–70. https://doi.org/10.54941/ahfe1006224

Zhao, Y. (2025). Aesthetic construction of generative artificial intelligence based on the theory of image and artistic conception. Modern Communication (Journal of Communication University of China), 47(3), 123–130. https://doi.org/10.19997/j.cnki.xdcb.2025.03.003