Machine Learning Applications in Marketing
Individual Research Project
Richard Koch



Can Generative AI create better Visual Marketing Content than Humans?

Machine Learning Applications in Marketing
Individual Research Project
Can Generative AI Create Better Visual Marketing Content than Humans?

Richard Koch

Cover image: Conceptual representation of human and AI-generated creativity in marketing imagery. Image generated using Midjourney (2026).

Table of contents

1. Introduction ……..……..……. 3
2. Data Collection ……..……..……... 4
3. AI Image Generation ……..…..…… 4
4. Survey Design ……. 7
5. Data Analysis ……...……... 9
6. Findings ……....……. 10
7. Discussion ……..…… 16
8. Limitations ……..... 17
9. Reflection ……..……....……. 18
10. Conclusion ……... 19

References …….……..... 20

1 Introduction

Artificial intelligence has already reshaped how marketers analyse and optimise decisions. What is changing now is production. This shift is reflected at an organisational level, with AI adoption reaching close to 90% of firms by 2025 (McKinsey & Company, 2025). Over the past decade, AI has primarily been used as a predictive tool to analyse consumer behaviour, optimise pricing, and personalise communication (Agrawal, Gans and Goldfarb, 2018). More recently, however, the focus has shifted towards generative AI (GenAI), which enables the creation of new content, including text, images, and video.

This transition is driven by advances in transformer architectures and diffusion models, which allow systems such as DALL·E to generate highly realistic visual outputs based on text prompts and learned visual patterns (Bommasani et al., 2021). In marketing, where visual content plays a crucial role in shaping brand perception, these developments are particularly significant. Firms have traditionally invested substantial resources in professional photography, design, and campaign production to create high-quality visual assets. Generative AI now offers the possibility to produce such assets at scale, with greater speed and at lower cost (Huang and Rust, 2018).

However, an important question remains: how do consumers perceive AI-generated marketing imagery compared with human-created visuals? While generative systems are capable of producing impressive images, it is not yet clear whether audiences perceive them as realistic, authentic, or of comparable quality.

This study addresses this question by investigating whether AI-generated images can compete with or outperform human-created images in terms of perceived quality, realism, and aesthetic appeal. To do so, a small-scale empirical study was conducted. A dataset of

marketing images was collected, AI-generated counterparts were created, and a perception survey was conducted to compare evaluations. The results were then analysed using descriptive statistics to compare perceived performance across both image types.

The objective of this study is to provide a more grounded assessment of generative AI's role in marketing content creation and to evaluate whether it represents a complementary creative tool or a potential substitute for traditional processes.

2 Data Collection

Approximately 60–70 marketing images were collected from Instagram because it reflects a real-world environment in which professional and user-generated marketing content coexist under similar conditions. The selected images represent typical content, including product, lifestyle, and promotional photography and imagery. The dataset included a range of image qualities and formats, including both high-resolution images and lower-quality visuals derived from video content. This reflects the visual inconsistencies of real-world social media environments. From this broader dataset, a subset of 28 images was selected based on their suitability for comparison with AI-generated counterparts.

3 AI Image Generation

Generative AI images were created using a combination of tools, including MidJourney, ChatGPT-based image generation, and other freely available platforms. Tool selection

depended on the specific visual requirements of each image, particularly with regard to composition, lighting, and the presence of text or branded elements.

While the assignment emphasises text-to-image generation, this study does not interpret generative AI as a purely text-driven process. Instead, image-guided prompting was incorporated where necessary to ensure comparability between AI-generated and human-created images. This reflects realistic usage conditions, where generative AI is typically embedded in an iterative workflow combining prompts, visual references, and selection rather than being used as a one-step generation tool.

From the initial pool of approximately 60–70 images, AI-generated counterparts were produced and iteratively refined. Initial exploratory tests revealed systematic limitations in generating comparable images, especially in cases involving text, branded products, or complex technical structures. For example, attempts to recreate campaigns featuring recognisable products (e.g. Apple or Dyson) or text-heavy visuals resulted in artefacts, distorted elements, or nonsensical text outputs. These observations informed the selection process by excluding images where structural comparability could not be achieved.

Only images that achieved a high degree of visual comparability were selected for the perception study. This reflects a common professional workflow in marketing, where multiple variations are generated and curated rather than used in their raw form.

This approach is consistent with the researcher's professional experience in two distinct marketing contexts. In a consumer-focused environment at heo, generative AI was used to create visual content for a campaign under significant time constraints.

The absence of strict visual constraints allowed the generative process to prioritise speed, variation, and stylistic flexibility.

In contrast, experience at Bucher Hydraulics highlights the limits of generative AI in constrained environments. Here, AI-generated imagery is explored for representing hydraulic systems and machinery targeted at engineering audiences. While AI-generated images may appear convincing at first glance, structural inconsistencies quickly reduce credibility.

A key challenge in this context is consistency. Marketing visuals often need to be reused across multiple formats, perspectives, and media, including video and animation. Current generative AI systems struggle to produce stable, reusable visual assets, as each generated image represents a new interpretation rather than a controlled variation of an underlying model.

This distinction is also reflected in the design of the present study. While free-form generation may be sufficient in unconstrained creative contexts, the experimental comparison of AI-generated and human-created images requires a high degree of visual alignment. As a result, image-guided prompting was necessary to ensure that AI-generated outputs closely matched the underlying visual concepts.

In future applications, including subsequent marketing campaigns, similar constraints would likely require the use of reference-based workflows, particularly when visual consistency, brand identity, or recognisable assets must be maintained.

4 Survey Design

The perception study was implemented using Google Forms. Participants were recruited through a convenience sample, including Imperial classmates, professional networks (excluding current employer), friends, and family.

Three survey versions were conducted:

- Survey 1 (n = 14): paired comparison
- Survey 2 (n = 12): independent image evaluation
- Survey 3 (n = 25): counterbalanced version of Survey 2

The main analysis is based on Survey 2 and Survey 3, resulting in a combined sample of 37 respondents.

In Survey 2 and Survey 3, images were presented individually in a fixed sequence. Participants evaluated each image using a 7-point Likert scale for:

- perceived quality
- perceived realism
- aesthetic appeal

Importantly, the allocation of AI-generated and human-created images was alternated between Survey 2 and Survey 3, while keeping the image sequence constant (see Figure 2). This approach reduces stimulus-specific bias and ensures that results are not driven by individual image characteristics.

Survey 1 used paired images to test direct comparison effects under more demanding conditions (see Figure 1)

Data were exported from Google Forms and initially explored using Microsoft Excel to calculate average ratings and identify patterns. The dataset was then structured for further analysis, enabling comparison across image types and survey versions.

Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7	Pair 8
img 1: AI	img 1: AI	img 1: Real	img 1: AI	img 1: Real	img 1: AI	img 1: AI	img 1: Real
Img 2 : Real	Img 2 : Real	Img 2 : AI	Img 2 : Real	Img 2 : AI	Img 2 : Real	Img 2 : Real	Img 2 : AI

Figure 1: Paired comparison survey design.
Each pair consisted of two visually matched images (one AI-generated, one human-created), presented side by side. The position of AI and human images was alternated across pairs to reduce position bias.

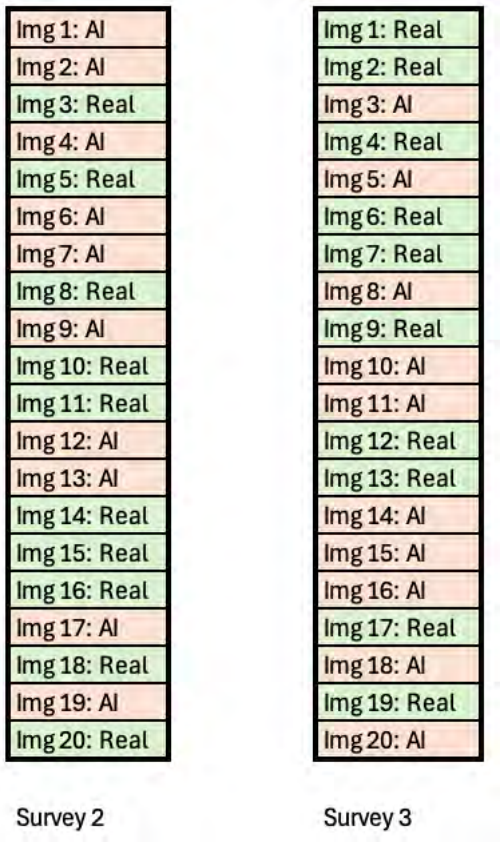

Figure 2: Counterbalanced survey structure.
The figure shows the two survey versions with identical image sequences but reversed AI/human assignment, reducing stimulus-specific bias.

5 Data Analysis

The collected survey responses were analysed using descriptive statistics. The analysis focused on comparing average ratings for human-created and AI-generated images across the three perception dimensions: quality, realism, and aesthetics. In addition, images were categorised based on key characteristics, including presence of text or logos, human presence, and technical complexity, to enable more detailed subgroup comparisons.

For each variable, the mean score was calculated separately for the two image categories. The results were then visualised, providing an intuitive overview of perceptual differences between the two image types. The comparison tables for all 3 surveys are shown below (Figures 3-5):

Survey 1	AI + Real	AI + Real avg.		AI total	AI avg.		Real total	Real avg.
I find the visual appearance of this image appealing	67.7143	4.2321		37.5000	4.6875		30.2143	3.7768
This image appears professionally produced	80.8571	5.0536		43.0714	5.3839		37.7857	4.7232
This image appears visually believable	74.3571	4.6473		39.0000	4.8750		35.3571	4.4196

Figure 3: Mean ratings for AI-generated vs human-created images (Survey 1 – paired comparison, n = 14). The figure presents average scores for aesthetic appeal, perceived professional quality, and realism. Images were evaluated side by side, with AI-generated visuals consistently achieving higher ratings across all dimensions.

Survey 2	AI + Real	AI + Real avg.		AI total	AI avg.		Real total	Real avg.
I find the visual appearance of this image appealing	86.2500	4.3125		47.6667	4.7667		38.5833	3.8583
This image appears professionally produced	92.6667	4.6333		49.0000	4.9000		43.6667	4.3667
This image appears visually believable	89.1667	4.4583		44.6667	4.4667		44.5000	4.4500

Figure 4: Mean ratings for AI-generated vs human-created images (Survey 2 – independent evaluation, n = 12). The figure shows average scores across the three evaluation dimensions. AI-generated images outperform human-created images in aesthetic appeal and perceived professional quality, while realism remains comparable.

Survey 3	AI + Real	AI + Real avg.		AI total	AI avg.		Real total	Real avg.
I find the visual appearance of this image appealing	91.7200	4.586		47.0000	4.7000		44.7200	4.4720
This image appears professionally produced	96.1200	4.8060		49.4400	4.9440		46.6800	4.6680
This image appears visually believable	96.0000	4.8000		49.3600	4.9360		46.6400	4.6640

Figure 5: Mean ratings for AI-generated vs human-created images (Survey 3 – counterbalanced evaluation, n = 25). The figure presents average scores across all dimensions with reversed image assignment. AI-generated images maintain higher ratings in aesthetic appeal and realism, while human-created images show a slight advantage in perceived professional quality.

This analysis allows for a direct comparison of how audiences perceive AI-generated visuals relative to traditional human-created marketing imagery. No inferential statistical tests were conducted, which limits the ability to assess whether observed differences are statistically significant. While inferential and machine learning approaches could provide deeper insights, the present analysis remains exploratory in nature (James et al., 2021).

6 Findings

The results of the perception study suggest a consistent pattern across all survey designs, with AI-generated images performing at least as well as, and in many cases better than, human-created marketing images.

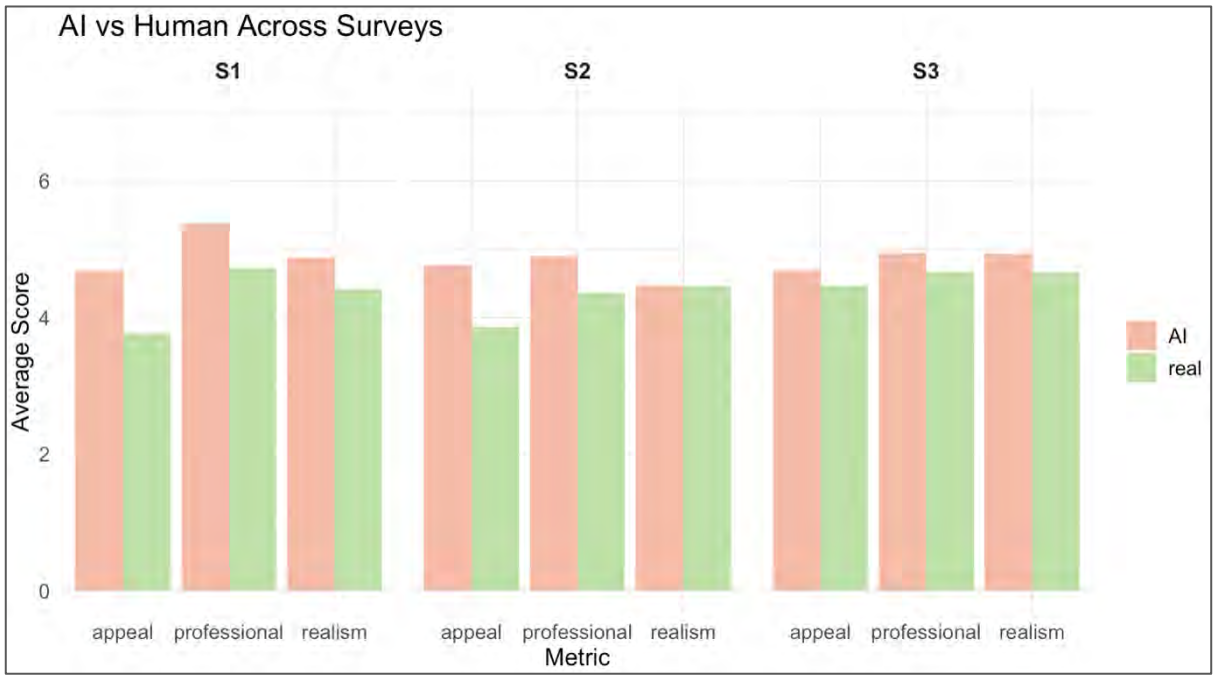

Figure 6: Comparison of AI-generated and human-created images across surveys and evaluation dimensions. Mean scores for aesthetic appeal, professional quality, and realism are shown for all three survey designs.

Based on the combined dataset of Survey 2 (n = 12) and Survey 3 (n = 25), resulting in a total of n = 37 respondents, AI-generated images tended to achieve equal or higher ratings across most evaluation dimensions (see Figure 6). In Survey 2, AI-generated images outperformed human-created images in aesthetic appeal (M = 4.77 vs 3.86) and perceived professional quality (M = 4.90 vs 4.37), while perceived realism was nearly identical (M = 4.47 vs 4.45).

In Survey 3, this pattern remained largely stable despite the counterbalanced design. AI-generated images achieved higher ratings in aesthetic appeal (M = 4.70 vs 4.47) and perceived realism (M = 4.94 vs 4.66), while human-created images showed a slight advantage in perceived professional quality (M = 5.23 vs 4.94). Overall, differences between the two image types were moderate but consistent, with AI-generated images performing strongly across all dimensions (see Figure 7). This indicates that the observed differences are not driven by specific images, but reflect broader perceptual tendencies.

Results from the paired comparison Survey 1 (n = 14) further reinforce these findings. When images were evaluated side by side, AI-generated images achieved higher ratings across all three dimensions, including aesthetic appeal (M = 4.12 vs 3.78), perceived professional quality (M = 5.38 vs 4.72), and realism (M = 4.88 vs 4.42), suggesting that AI-generated visuals are not penalised under direct comparison.

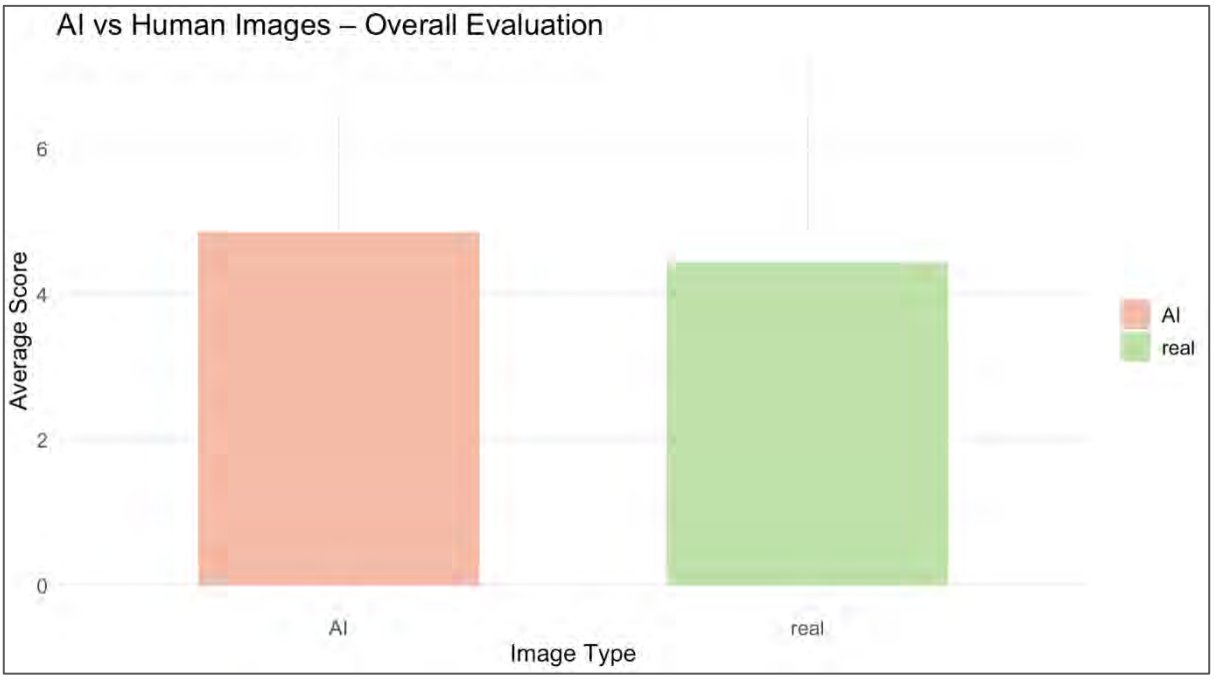

Figure 7: Mean ratings for AI-generated vs human-created images (Survey 1 – paired comparison, n = 14). Images were evaluated side by side across all three dimensions.

An important contextual factor is the viewing environment. Most participants completed the survey on mobile devices, where images are consumed quickly and at relatively small scale. Under these conditions, minor artefacts in AI-generated images are less likely to be detected, while lower-resolution human-created images, particularly those derived from video content, may be rated more negatively. As shown in Figure 8, AI-generated images tend to achieve slightly higher and more consistent evaluation scores.

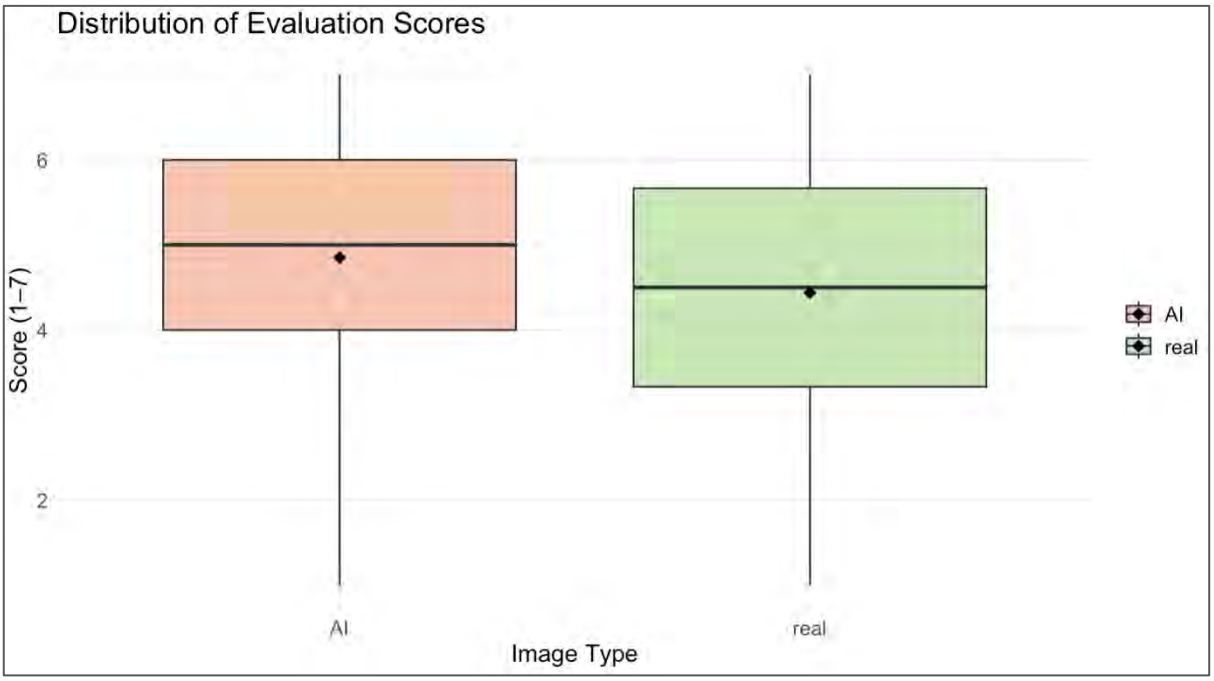

Figure 8: Distribution of evaluation scores for AI-generated and human-created images. Boxplots illustrate central tendency and variance across all observations.

Participant feedback supports this interpretation. Several respondents reported difficulty distinguishing between AI-generated and human-created images.

A central finding of this study is that the effectiveness of AI-generated imagery is not absolute, but dependent on context and audience expertise. In mobile-first environments, perception is driven primarily by clarity, composition, and aesthetic coherence, allowing AI-generated images to perform on par with or exceed human-created visuals. In contrast, in expert-driven contexts such as industrial marketing, evaluation criteria shift towards structural accuracy and technical correctness, making the limitations of generative AI more apparent. Figure 9 demonstrates that increasing technical complexity reduces evaluation scores, supporting the argument that structural demands influence perceived quality.

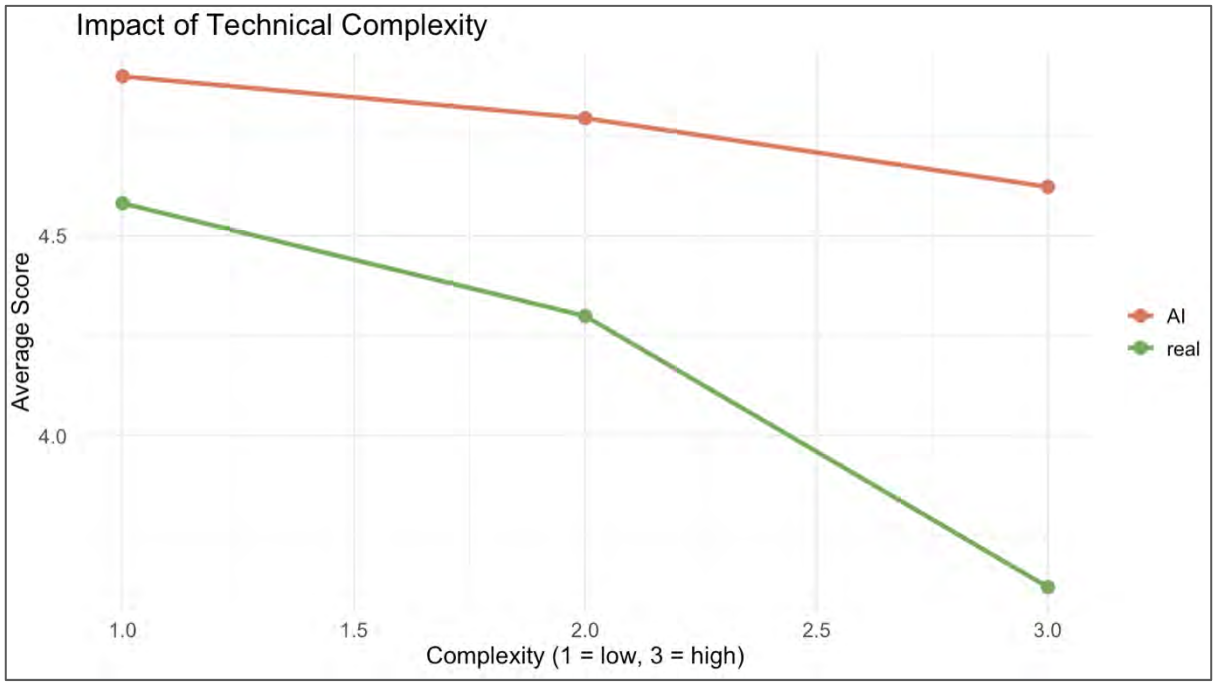

Figure 9: Impact of technical complexity on image evaluation.
Average scores are shown across increasing levels of structural complexity.

The influence of image-level characteristics is further illustrated in Figures 10 and 11, where text/logos and human presence appear to influence evaluation scores.

Overall, the findings suggest that AI-generated marketing imagery is not only competitive with human-created content, but can outperform it in key perceptual dimensions, particularly in aesthetic and professional evaluation under typical digital consumption conditions, consistent with recent evidence on the performance of generative AI in visual marketing contexts (Hartmann, Exner and Domdey, 2024).

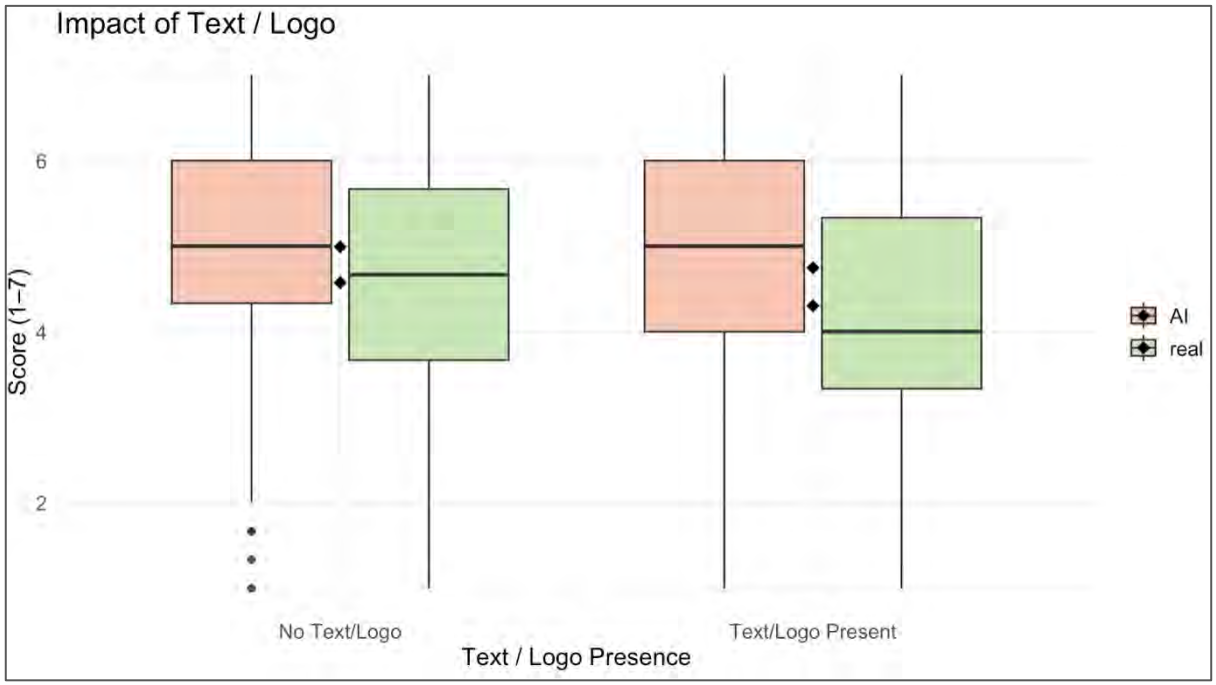

Figure 10: Impact of text and branding on image evaluation
Distribution of scores by image type (AI-generated vs human-created) and presence of text/logos, showing lower and more variable ratings for images containing text or branded elements.

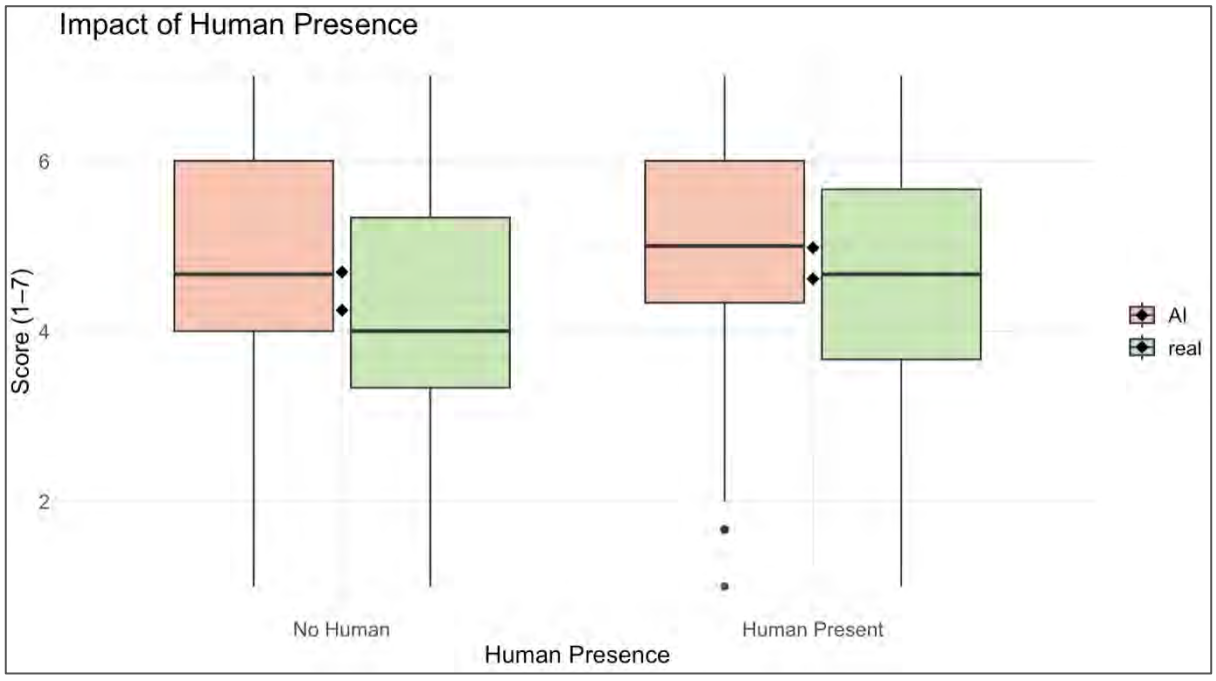

Figure 11: Impact of human presence on image evaluation
Distribution of scores by image type (AI-generated vs human-created) and presence of human subjects, indicating differences in perceived quality and realism across conditions.

7 Discussion

A central tension in the findings is that AI-generated images appear to outperform human-created visuals, while the underlying production processes differ fundamentally. The AI images represent curated outputs from an iterative generation process, whereas human-created images reflect real-world variability, including lower-quality executions. As a result, the comparison contrasts optimised outputs with naturally occurring variation, likely leading to an overestimation of AI performance.

The findings also suggest a shift in how visual content is evaluated. In mobile-first environments, where images are consumed quickly and at small scale, perception is driven less by technical precision and more by immediate visual coherence. Under these conditions, generative AI performs well, as it tends to optimise composition, colour balance, and aesthetic clarity. Minor artefacts are less likely to be detected, particularly when compared to lower-resolution or less controlled human-generated content.

This advantage does not generalise across contexts. In more constrained environments, such as industrial marketing or technically complex product communication, evaluation criteria shift towards accuracy, consistency, and structural reliability. In these settings, generative AI remains limited, as even small inaccuracies reduce credibility for expert audiences.

A further limitation concerns consistency over time. While generative AI can produce high-quality individual images, it struggles to generate stable, reusable visual assets across campaigns, formats, or perspectives. Each output represents a new interpretation rather than a controlled variation, complicating its use in structured brand communication.

Taken together, the findings do not suggest a simple replacement of human creativity, but a redistribution of roles within the creative process. Generative AI shifts value from execution towards selection, direction, and validation. This aligns with the concept of "co-intelligence" (Mollick, 2024), where AI expands the space of possible outputs while human input remains essential in shaping and aligning them with strategic intent.

8. Limitations

The study is based on a convenience sample and should therefore be interpreted as exploratory rather than generalisable. The selection of AI-generated images followed a curated process, reflecting realistic marketing workflows, but potentially leading to an overestimation of average AI performance. The use of image-guided prompting deviates from a strictly text-to-image approach, although this was necessary to ensure comparability between images and aligns with practical usage.

Ethical and legal considerations around copyright and training data remain unresolved and may constrain the broader adoption of generative AI in marketing (Bommasani et al., 2021; McKinsey & Company, 2025).

The results indicate that image-level characteristics may influence evaluation independently of whether an image is AI-generated or human-created. High- and low-performing images often cluster across both categories, suggesting that differences may reflect preferences for specific visual concepts rather than the generation method itself (see Figure 12)



Figure 12: Influence of image-level characteristics on evaluation scores
Evaluation scores across individual image pairs, showing that top-performing AI-generated and human-created images achieve very similar ratings. Several matched pairs cluster closely in score, suggesting that performance is driven more by the underlying visual concept than by the generation method.

9 Reflection

The findings of this study suggest that generative AI is unlikely to fully replace human creatives in the near term, but the boundary between augmentation and substitution is already beginning to shift. In high-volume, low-risk environments such as social media content, AI can effectively operate as a partial substitute. The ability to generate large volumes of visually coherent, aesthetically optimised images at low cost reduces the need for traditional production processes in these contexts. The role of human creatives shifts towards selection and curation.

However, this substitution effect weakens as the requirements for precision, consistency, and contextual understanding increase. In areas such as brand building, storytelling, and

technically complex communication, human input remains essential. A more realistic interpretation is therefore that generative AI restructures creative work rather than eliminating it. It introduces a workflow in which ideation and execution are increasingly decoupled. The ability to generate content becomes less valuable than the ability to frame problems, guide systems, and critically assess outputs. In this sense, generative AI does not reduce the importance of human contribution, but raises the level at which it operates.

10 Conclusion

This study examined whether AI-generated marketing imagery can outperform human-created visuals in terms of perceived quality, realism, and aesthetics. A small-scale perception study compared human-created marketing images with AI-generated counterparts. The findings indicate that generative AI can produce images that are highly competitive with human-created marketing visuals, particularly in terms of aesthetic appeal and perceived quality. However, minor differences in perceived realism suggest that audiences may still detect subtle cues that differentiate AI imagery from traditional photography.

Overall, the results highlight the growing potential of generative AI as a powerful tool in marketing content production. While AI imagery may not fully replace human creativity, it offers significant opportunities to enhance creative workflows, increase production efficiency, and support rapid marketing experimentation.

As generative technologies continue to evolve, the collaboration between human designers and AI systems will likely become a central component of future marketing practice.

References

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence.* Harvard Business Review Press.

Bommasani, R. et al. (2021). *On the opportunities and risks of foundation models*. Stanford Center for Research on Foundation Models. DOI:10.48550/arXiv.2108.07258 Available at: https://www.researchgate.net/publication/353941945_On_the_Opportunities_and_Risks_of_Foundation_Models (Accessed: 4 April, 2026)

Hartmann, J., Exner, Y., & Domdey, S. (2024). *The power of generative marketing: Can generative AI create superhuman visual marketing content?* International Journal of Research in Marketing. DOI:10.2139/ssrn.4597899 Available at: https://www.researchgate.net/publication/375480180_The_power_of_generative_marketing_Can_generative_AI_create_superhuman_visual_marketing_content (Accessed: 4 April, 2026)

Huang, M.-H. and Rust, R.T. (2018). *Artificial intelligence in service*. Journal of Service Research, 24(1), pp. 3–19. DOI:10.1177/1094670517752459 Available at: https://www.researchgate.net/publication/322958062_Artificial_Intelligence_in_Service (Accessed: 4 April, 2026)

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021). *An Introduction to Statistical Learning: with Applications in R*. 2nd ed. New York: Springer.

McKinsey & Company (2025) *The state of AI in 2025: Agents, innovation, and transformation*. Survey, 5 November. Available at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai (Accessed: 4 April, 2026)

Mollick, E. (2024) *Co-Intelligence: Living and Working with AI*. London: WH Allen.

AI-References

ChatGPT
Used for: Image Generation, R-Code Debugging
OpenAI (2026). *ChatGPT*. Available at: https://chat.openai.com/ (Accessed: 4 April, 2026).

Midjourney
Used for: Image Generation
Midjourney Inc. (2026. *Midjourney)*. Available at: https://www.midjourney.com/ (Accessed: 4 April, 2026).