27TH INTERNATIONAL CONFERENCE ON ENGINEERING AND PRODUCT DESIGN EDUCATION
11-12 SEPTEMBER 2025, UNIVERSITY OF MALTA, MALTA

ACADEMICALLY INFORMED AI VS. HUMAN-DESIGNED MILK PACKAGING: A COMPARATIVE EVALUATION

Henry P. LEE[1], Asa R. JACKSON[2], Blake GIBBONS[3], Laura K. JEFFERIES[4] and Bryan F. HOWELL[4]
[1]Parsons School of Design, United States of America
[2]Kolding School of Design, Denmark
[3]RunDiffusion, United States of America
[4]Brigham Young University, United States of America

ABSTRACT
As advancements in language model-powered text-to-image AI platforms accelerate, individuals can increasingly generate high-fidelity visual content more efficiently, regardless of background. These platforms are powerful tools for rapidly iterating and visualising packaging design concepts. This study assesses whether AI-generated milk package designs, steered by academic packaging research, will produce packaging outcomes that perform equal to or better than human-designed outcomes with minimal designer input. For this study, researchers curated, summarised, and combined leading academic articles on packaging design into textual AI prompts. The textual prompts were input into the platform RunDiffusion to generate seven milk packaging designs. The images created by the platform were reviewed by participants and compared to seven human-made milk packaging designs and a control across seven perceptual dimensions: purchase likelihood, nutrition, ordinary, sustainable, colour, imagery, and typography. Results indicate that AI-generated and human-made designs did not differ significantly, suggesting that AI platforms can efficiently produce milk packaging design outcomes that compete with human designs. This study helps identify emerging opportunities for best practices and knowledge transfer between design, technology, industry and education.

Keywords: Generative packaging design, artificial intelligence collaborations, design student academic research, human-computer interaction, design education

1 INTRODUCTION
Artificial Intelligence (AI) technology has advanced rapidly over the past few years, developing new tools for designers to apply in their processes and outputs. Advancements include text-to-image software that can generate realistic imagery quickly and with minimal input [1]. These tools can increasingly be used by industry design teams and will alter traditional workflows by providing efficient design outcomes. These tools are subsequently relevant to design educators and their students. In general, design students need to be taught and equipped with the latest industry tools to support their education and transfer to industry, including the rapidly evolving technology that is AI.
One area of design that is liable to be transformed by AI is packaging design. It is now possible to use AI tools to develop packaging designs for food, beverages, hard and soft products, and other goods using text-to-image tools. There are ample academic studies regarding packaging design that designers can use to inform their design decisions. However, little research has been published on the viability of using AI tools in packaging design. Because scholarly outputs are typically textual, generating textual prompts for text-to-image software as a reference is possible and efficient. For example, academic studies determine that a specific packaging colour or material elicits a particular reaction from a consumer. In that case, anyone can quickly apply that knowledge to an AI prompt and the technology will produce a visual output representing those insights.
This study assesses whether AI-generated visuals generated by the text-to-image platform RunDiffusion with minimal designer input, steered by academic packaging research, will produce packaging outcomes that perform equal to or better than human-designed outcomes. In an online survey, seven design

student-generated milk carton designs and one control design were compared to seven AI-generated designs. We expect this will uncover new, efficient, and impactful packaging design methods. If successful, industry designers and academic design educators should consider how scholarly outputs can impact design outcomes.

2 METHOD

2.1 Participants

A convenience sample of 48 adult participants completed the survey. Participants included University students and professionals, between the ages of 20 and 32, located in California, Utah and Denmark. We acknowledge this is a limited sample and accompanies a cultural bias that may not accurately represent broader consumer preferences.

2.2 Academic Packaging Research

For the study, six articles were selected from approximately 70 packaging-related academic manuscripts based on two constraints: their relevance to the field (citation count), and their focused outcomes on design elements including colour, typography, and the presence of imagery [2, 3, 4, 5, 6, 7].

2.3 AI Framework

The articles were manually summarised into a bulleted list, condensed into a descriptive textual prompt, and input into the AI platform RunDiffusion (which supported the research), running the text-to-image model Flux 1.1 [Pro] (Blueberry) by BlackForestLabs. This program was chosen because it reliably produces accurate typography and has a simple user experience. Researchers determined that developing a prompt without the help of AI would allow for a more specific study, such as incorporating imagery outcomes as discussed in the manuscripts. In the future, it is possible to study how a large language model, LLM, might generate a prompt from academic articles independently, further streamlining the design process. The prompt developed for this study reads:

> *Design the front panel of a milk carton with a predominantly [colour] colour scheme. The front panel should prominently feature bold, large, Roman, upper-case letters that read "Milkhaus" horizontally across the front. Beneath the main label, include the text "Grade A Milk" in smaller lettering. At the bottom of the front panel, include the words "Half Gallon" in small lettering. Include a simple illustrated cow on the front of the milk carton. The overall design should use horizontal lines. The milk carton should only be displayed from the front straight on. There should be a white minimalist background.*

2.4 Milk Carton Designs

Based on the scholarship outcomes, different colours were alternated in the prompt, including blue, black-and-white, red, purple, and green [3]. In addition, the prompt for all the images generated gave instructions to include the image of a cow on the front of the carton because of research stating that "For the majority of heavy (milk) users, images such as a cow or the countryside reinforce the origin and the naturalness of the product in the mind of the participants" [2]. Researchers for this study had basic AI prompt experience. The AI prompts and 44 images were generated in less than two hours. Seven were chosen for their colour, imagery, and typography quality. The seven human-made designs were selected based on their colour and performance in a previous study on milk packaging [8]. Within each set of seven, three designs included blue, purple and green colour schemes, two included red schemes, and two were black-and-white. A fifteenth image was used as a control image and selected from the same previous milk packaging study [8]. All the images used in the study are shown in Figure 3.

2.5 Survey Questions

The fifteen designs were presented to participants in a Qualtrics survey, who assessed them against seven perceptual dimensions. Four questions were utilitarian: "How likely are you to purchase this milk"? And "To what degree is this milk nutritious, ordinary, and sustainable"? Three questions were hedonic: "How attractive is this milk based on colour, imagery, and typography"? These questions were chosen based on a previous study on milk packaging design [8] and food-related labelling [9]. Additionally, the survey asked participants to agree to participate in the study and asked, "How often do

you consume milk?" with the options of 'Daily," "2-5 Times a Week", "1-2 Times a Week," "Few Times a Month," and "Rarely or Never."

3 RESULTS & DISCUSSIONS

A repeated-measures experiment was conducted to compare participants' ratings of milk-carton designs across three different design sources: a control design, AI-generated designs, and human-designed milk cartons, across seven perceptual rating dimensions: purchase likelihood, nutritious, ordinary, sustainable, colour, imagery, and typography. Initially, the study recruited any willing participant and did not filter them by milk consumption. Based on a bias identified in the rarely/never participants, shown in section 3.1, this was a weakness in the study, and this group was removed from the analysis in sections 3.2 and 3.3.

3.1 Averaged Design Scores by Milk Consumption Group

Although not the study's initial focus, one unanticipated outcome revealed that participant milk consumption patterns influenced ratings. Participant ratings shown in Fig. 1 report the average design scores of the seven perceptual dimensions by consumption group. Positive scores reflect the human-made designs, and negative scores reflect the AI-generated designs. Surprisingly, participants who rarely/never drank milk were heavily biased to human-designed packaging across all dimensions.

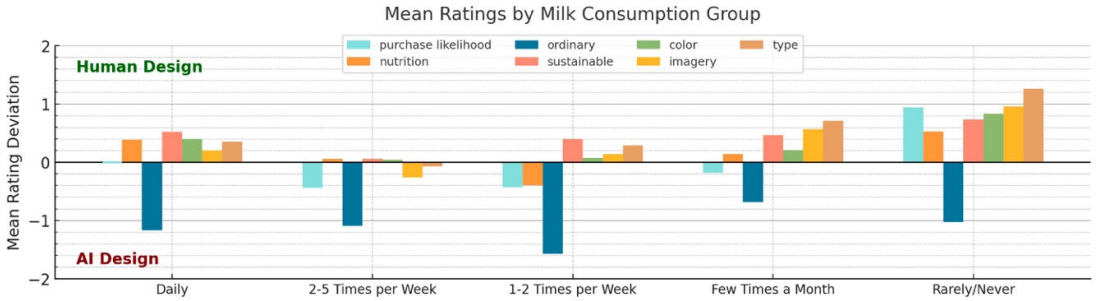

Figure 1. Average design scores by milk consumption group

Similarly, daily consumption participants rated all perceptual dimensions, except ordinariness, positively between 0.0 and 0.5. Arguably, they are the only group favouring human design outcomes in all dimensions except purchase likelihood, which was rated at 0.0. Exploring the purchasing context may explain these outcomes. If participants rarely or frequently purchase milk, they may be more inclined to review the designs solely as a graphical label or ignore the packaging content, as it is a well-practised decision. The sustainable and colour dimensions were consistently rated higher for human-made designs across all groups. Only the ordinary dimension was rated consistently as AI-generated across all groups. Future studies exploring how participant consumption impacts preferences would add depth and understanding to milk and other food packaging.

3.2 AI and Human-made Mean Scores Across all Dimensions by Source

A two-way ANOVA with Source and Dimension as within-subjects factors revealed significant main effects of Source ($F(14, 518) = 8.48$, $p < 0.0001$) and Dimension ($F(6, 222) = 3.31$, $p = 0.0039$), as well as a strong Source × Dimension interaction ($F(84, 3108) = 7.57$, $p < 0.0001$), indicating that differences among control, AI, and human designs varied by dimension. The mean scores for the control, AI, and human-made designs across the seven dimensions are shown in Figure 2. Statistical significance between the three categories was only found in: ordinary and imagery. Human-made designs received higher ratings in sustainable, colour, imagery and typography. The control received higher ratings in ordinary and nutrition. An AI-generated design received the highest rating for purchase likelihood. For the ordinary rating, AI and human designs were seen as significantly less ordinary than the control design ($p < .0001$). For the imagery rating, the human designs perform significantly better than the control and AI designs ($p < .0001$). The AI imagery may perform slightly better than the control but did not meet the Bonferroni cutoff. All other dimensions showed no significant pairwise differences, indicating that the three sources perform similarly in purchase likelihood, nutrition, colour, and typography.

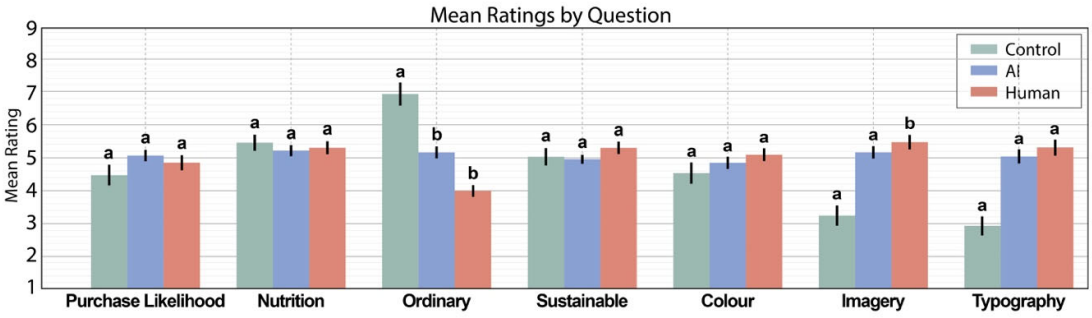

Figure 2. Control, AI and Human Mean Scores by Question

The initial hypothesis was that AI-generated designs could perform equal to or better than human-generated designs. The results indicate that although the AI did not significantly outperform the human designs, it performed the same as the human-designed outcomes, except in imagery. The hypothesis also outlined that existing academic packaging outputs would successfully steer the AI prompt to generate positive packaging results, which proved correct. Students and pedagogues can leverage these results to validate exploring alternative AI-influenced steps in the design process to expand designers' design possibilities and efficiency.

3.3 Package Design Mean Scores Across Perceptual Dimensions by Source

The mean score for all three sources, control, AI-generated, and human-made designs, of individual designs across the seven dimensions is shown in Figure 3. The mean for all AI-generated designs is 5.07, and human-made designs are 5.05, indicating no performance difference between the two sources, but both scored higher than the control at 4.66. However, individual designs in both groups showed measurable differences. The five highest ranking scores are not significantly different from each other. The top two were human-made, and the next three were AI-generated designs. All five were significantly different from the control. The control shared significance with three AI-generated designs and four human-made designs. Both AI and human designs had one design that performed significantly worse than the control.

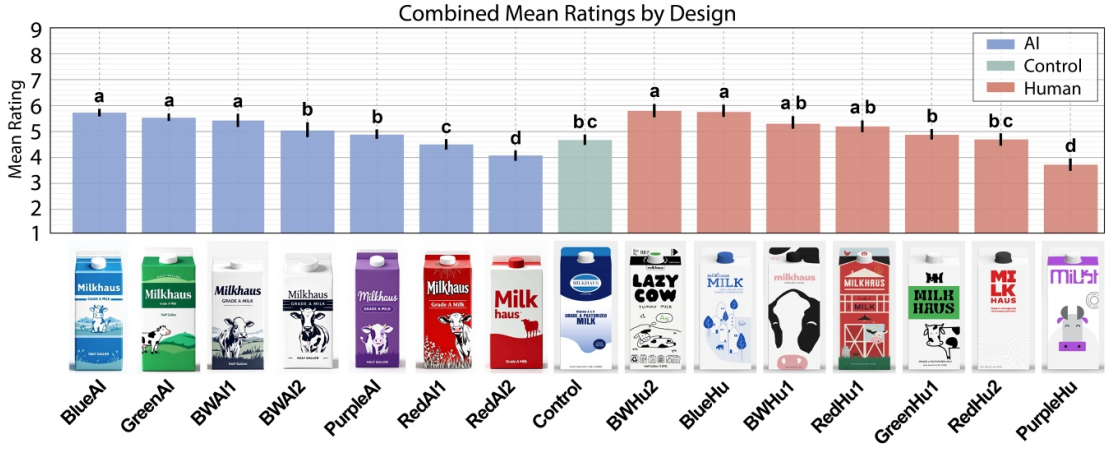

Figure 3. Participants' ratings for the control, AI, and Human-made packaging designs

The human-generated designs from the previous study at E&PDE'23 took four to six weeks to complete. In contrast, AI-generated designs only required a few hours to research appropriate articles, synthesise the content, and develop a prompt for image generation. It's possible that one designer, student, or even a non-designer could brand a milk product quickly and efficiently using AI based on academic research. This design process would combine academia and industry outcomes in meaningful ways.

AI packaging design informed by academic outputs could be perceived to threaten graphic designer careers while being a viable service to other disciplines. Positive use cases could include companies quickly and efficiently personalising their brands through varied packaging based on location or specific demographic values. The study's results demonstrate the importance of teaching design students AI tools

to fill the gap between the evolving industry possibilities and academia. Design students should be prepared to enter the industry with a complete understanding of these emerging powerful tools.

3.4 Positive Dimensional Insights by Design

The AI-generated design, BlueAI, the front-facing blue cow, received the highest rating in two dimensions: purchase likelihood (6.08) and the most attractive imagery (6.63). The human-generated design, BlueHu, the blue farm, received the highest rating in three dimensions: nutritional (6.16), sustainability (6.18), colour (6.08), and the second-highest rating in typography (5.95). The Lazy Cow received the highest rating in typography (6.33) and scored second highest in five dimensions: purchase likelihood (5.84), nutrition (5.95), sustainability (6.24), colour (5.82), and imagery (6.58). The three designs in Figure 4 received the highest ratings across all dimensions except for ordinary.

Figure 4. These three designs received the highest ratings across all dimensions except for ordinary

3.5 Negative Dimensional Insights by Design

The control received the highest ordinary rating (6.92), and the highest of all scores recorded. It also received the lowest scores for imagery (3.26) and typography (2.95). The Human-generated design, PurpleHu, the purple cow, received the lowest scores in four dimensions: purchase likelihood (3.45), nutritious (4.11), sustainable (4.22), and colour (3.29). The human-generated design, RedHu2, with the square red logo, received the lowest score as the least ordinary design (3.39). The AI-generated design, RedAI2, red cow silhouette scored the second least sustainable (4.24) and tied with RedAI1 as the poorest colours (4.16). The five designs in Figure 5 were the weakest performers in the study.

3.6 Designer Insights

Insights from these outcomes for future designers of milk packages may be to focus on blue or black-and-white colour schemes over red or purple ones. The positive designs also included stylised images of cows in a farming context. The human-made designs scored highest in all but two dimensions, purchase likelihood and imagery, indicating that there are specific design elements that humans could focus on, such as typography. In contrast, AI can adequately design other elements, such as imagery. Furthermore, this suggests the possibility of a collaborative approach to packaging design, with AI designing certain elements and humans designing other elements. Further study will provide insights into this collaborative approach.

Figure 5. These five designs received the lowest scores across all dimensions except for ordinary, where the control design on the far left scored the highest

4 CONCLUSIONS

When averaged across all designs and dimensions, AI-generated and human-designed cartons did not differ significantly, indicating that AI tools can match human designers in consumer perceptions of milk packaging. Including AI in the packaging design process is a viable method to reduce design development time. Incorporating academic research outputs to inform AI prompts is effective. Human designs scored higher in all dimensions except purchase likelihood and imagery, suggesting human curation is important in refining and determining final design details. A collaborative approach to packaging design, with AI designing certain aspects and humans designing others, appears reasonable. Outcomes also demonstrate that design educators might consider the impacts of AI programs on packaging design education and prepare students to incorporate them appropriately in their design process. Students and teachers should understand how to ideate, create, and communicate with AI, while also learning to distinguish their impact and influence as designers in this evolving creative ecosystem. Additionally, this knowledge contributes to an understanding of AI and human strengths, respectively, allowing students and designers to focus their skills on where they excel within the design discipline, like typography and sustainable messaging.

REFERENCES

[1] Crowson K. et al. VQGAN-CLIP: Open Domain Image Generation and Editing with Natural Language Guidance. *Computer Vision – ECCV 2022*. Lecture Notes in Computer Science, 2022, vol 13697. Springer, Cham. https://doi.org/10.1007/978-3-031-19836-6_6.

[2] Hollywood L., Wells L., Armstrong G. and Farley H. Thinking outside the carton: Attitudes towards milk packaging. *British Food Journal*, 2013, 115. 10.1108/BFJ-Jul-2010-0127.

[3] Ampuero O. and Vila N. Consumer perceptions of product packaging, *Journal of Consumer Marketing*, 2006, Vol. 23 No. 2, pp. 100-112. https://doi.org/10.1108/07363760610655032.

[4] Mushili B. and Mwanza B. G. The Influence of Packaging Attributes on Consumer Buying Behaviour: A Case Study of Milk Packages. *Management Journal for Advanced Research*, 2024, 4(1), 54–59. https://doi.org/10.5281/zenodo.10652880.

[5] Su J. and Wang S. Influence of food packaging colour and foods type on consumer purchase intention: the mediating role of perceived fluency. *Front. Nutr.* 2024, 10:1344237. doi: 10.3389/fnut.2023.1344237.

[6] Labrecque L. I. and Milne G. R. Exciting red and competent blue: the importance of colour in marketing. *J Acad Mark Sci*. 2012, 40:711–27. doi: 10.1007/s11747-010-0245-y.

[7] Wang H., Ab Gani M. A. A. and Liu C. Impact of Snack Food Packaging Design Characteristics on Consumer Purchase Decisions. *Sage Open*, 2023, 13(2). https://doi.org/10.1177/21582440231167109.

[8] Howel, B. F., Lee K. S., Holbrook R. N., Houghton S. J., Reynolds L. and Jefferies L. K. Multidisciplinary Dairy Milk Packaging Collaboration, *Proceedings of the International Conference on Engineering and Product Design Education (E&PDE 2023*), Barcelona, Spain., September 2023. DOI number: 10.35199/EPDE.2023.38.

[9] Howell B. F., Newcomb E., Loh D. W., Jackson A. R., Dunn M. L. and Jefferies L. K. Shaping Consumer Perceptions of Genetically Modified Foods: The Influence of Engineering, Science, and Design Signifiers in Packaging Disclosure Statements, *Foods*, 2025.