Designing Hyper-Personalized Retail Product Packaging through Generative AI and Augmented Reality

Muhammad Minhajuddin
Faculty of Computer Science & Technology
Algoma University
Brampton, Canada
mminhajuddin@algomau.ca

Joel Varghese
Faculty of Computer Science & Technology
Algoma University
Brampton, Canada
jovarghese@algomau.ca

Sathaporn “Hubert” Hu
Faculty of Computer Science & Technology
Algoma University
Sault Ste. Marie, Canada
hus@algomau.ca

Somang Nam
Faculty of Computer Science & Technology
Algoma University
Brampton, Canada
somang.nam@algomau.ca

***Abstract*—Personalized product packaging can enhance consumer engagement, but conventional design processes cannot easily tailor packaging to individual preferences in real time. We present HYPAR (Hyper-Personalized Packaging in AR), a cross-platform WebAR prototype that integrates a generative diffusion model with marker-based augmented reality (AR) to allow consumers to design and preview personalized packaging on a physical product in real time. The system is positioned as a pre-purchase preview tool for in-store, made-to-order retail, where a shopper selects emotive descriptors to generate a unique packaging texture to overlay on the physical product via the device's camera. We evaluated HYPAR through a user study with 15 participants using a perfume box as the target product, combining the System Usability Scale (SUS) with semi-structured interviews. The system achieved a high SUS score, indicating strong perceived usability, while a complementary qualitative analysis of the interviews indicated increased perceived ownership, perceived luxury, and expressed purchase intent relative to a generic alternative. These findings suggest that integrating real-time generative AI with AR can shift retail packaging from static, mass-produced designs towards personalized, emotionally resonant consumer experiences.**

***Index Terms*—Generative AI, Augmented Reality, Personalized Packaging, Real-time Interaction, Retail Product Customization**

I. Introduction

In contemporary retail, physical packaging is a critical touchpoint in consumer decision-making, shaping first impressions at the point of sale. Yet, despite extensive personalization in digital marketing, packaging itself remains largely static and mass-produced, yielding a one-size-fits-all experience that often fails to resonate with individual consumer identities. The emerging metaverse, understood not as a single virtual destination but as a continuum that interweaves physical and digital worlds [1], offers a different basis for thinking about this gap. Retail scholarship has begun to treat this continuum as a concrete commercial setting. Hyper-personalized experiences have been identified as a purchase-stage differentiator for metaverse retailing [2], phygital customer experiences that layer digital content onto physical goods have been shown to depend primarily on sight and touch [3], and the coupling of physical products with digital counterparts is now studied as a strategic retail decision in its own right [4]. This framing motivates treating a single physical shelf product as an anchor for many possible digital identities, each adapted to the consumer at the point of purchase. It is this specific, bounded, phygital interpretation of the metaverse, not a fully virtual storefront, that we adopt throughout. Realizing this vision in everyday commerce depends on two recently matured, co-evolving technologies. Artificial intelligence (AI) techniques such as Latent Diffusion Models now allow high-fidelity, context-aware visuals to be synthesized from natural language in seconds [5], providing the generative substrate for personalized content, while AR has evolved from a novelty into a strategic communication channel capable of anchoring digital content to physical goods at the point of purchase [6]. Their convergence makes it possible, for the first time, to operationalize a physical retail experience in which a generic shelf item becomes a canvas for an individualized digital twin.

We present HYPAR (Hyper-Personalized Packaging in AR), a cross-platform WebAR prototype positioned within this physical-digital continuum. HYPAR allows consumers to design and preview personalized packaging on a physical product in real time, without requiring a dedicated mobile application or headset. The intended deployment is in-store, where shoppers interact with a tester or a display product on the retail floor. Instead of replacing the packaging already on the shelf, HYPAR functions as a pre-purchase preview tool for made-to-order fulfilment: a shopper selects emotive

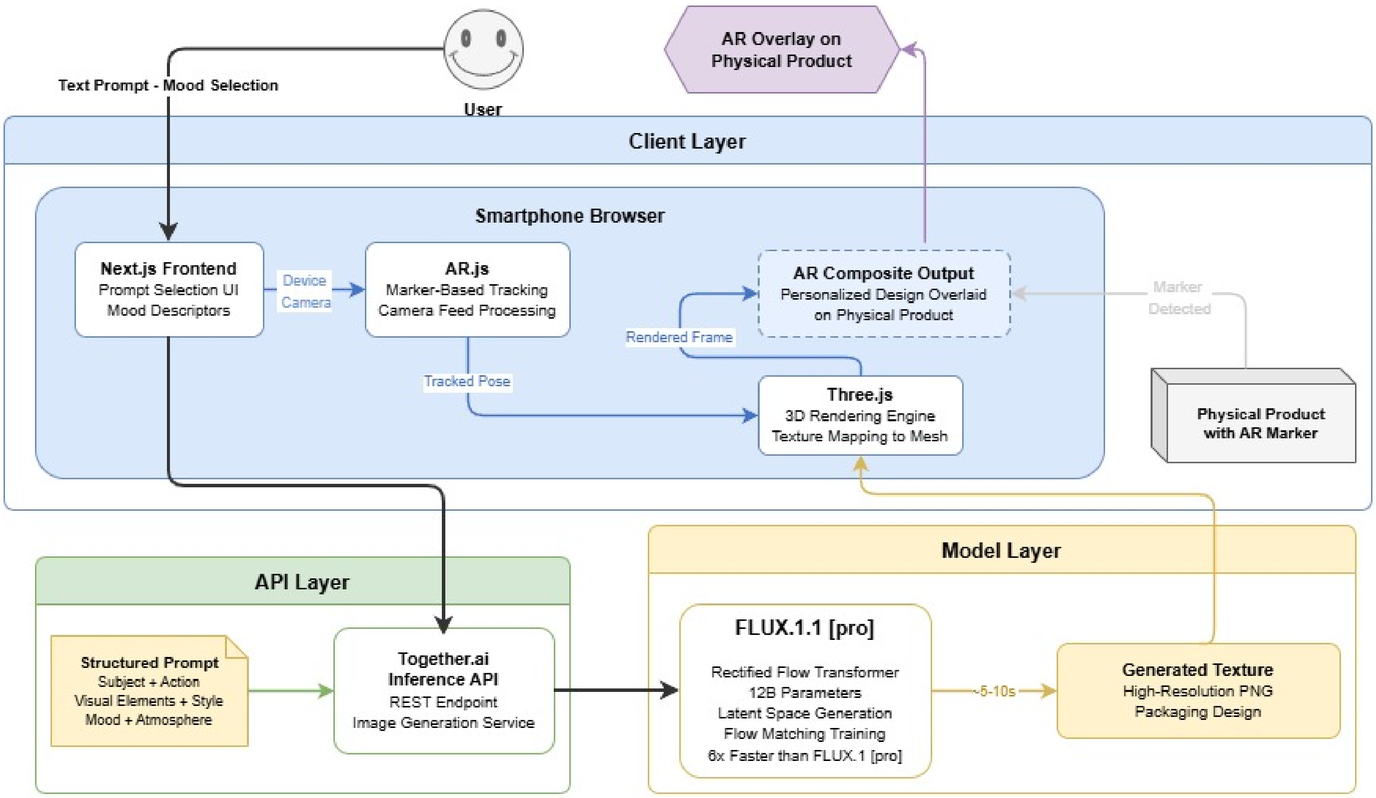

Fig. 1. System architecture of HYPAR showing the client-side WebAR rendering pipeline, the Together.ai inference API, and the FLUX.1.1 [pro] generative model.

descriptors (e.g., "elegant" or "nostalgic"), and the system generates a unique packaging texture and overlays it on the physical box through the device's camera. Once the prototype is finalized, the personalized design is then printed and fulfilled post-purchase. Through this loop, AR enables an impromptu spatial-commerce interaction that can augment the physical shelf. Through the psychological mechanism of the endowment effect, by which a higher value is assigned to objects that one helps to create [7], the system is designed to establish a deeper emotional connection between consumers and products, illustrating how a low-friction, browser-based slice of the metaverse can deliver tangible consumer value in everyday retail.

This paper addresses the following research question: *How does an augmented reality system that dynamically generates hyper-personalized packaging influence consumers' first impressions, perceived attractiveness, and engagement with a physical product?*

To investigate this, we conducted a user study and collected basic usability data of the system using the System Usability Scale and added qualitative feedback through semi-structured interviews.

This work does not propose a new generative model, but instead shows how existing, off-the-shelf components can be composed into a novel end-to-end retail experience and evaluated with users. Specifically, the contributions would be the following:

1) The design and implementation of an end-to-end WebAR pipeline that integrates a hosted generative diffusion model with a marker-based AR overlay on a physical product, demonstrating a hardware-free entry point into physical spatial commerce;
2) Qualitative evidence that hyper-personalized AR packaging is associated with increased perceived ownership, perceived luxury, and expressed purchase intent relative to a generic alternative, obtained from a technology-probe study; and
3) A discussion of the conditions under which the convergence of AI and AR adds value in physical retail, including the boundary case of in-store contexts where the physical product is already accessible.

II. Related Works

A. Generative AI for Product and Packaging Design

Denoising Diffusion Probabilistic Models [8] and their optimization through Latent Diffusion Models [5] established the technical basis for high-fidelity image synthesis from natural language, enabling complex textures, artistic styles, and structural concepts that were previously labour-intensive to produce manually. In retail packaging, this capability is shifting workflows from static templates towards dynamic creation. On the design side, AI-generated content has been operationalized for category-specific packaging methods, such as a tea-product workflow, demonstrating that AI can compress design cycles while increasing creative diversity [9]. On the consumer side, satisfaction with AI-generated packaging has been shown to depend on how well the generated visuals align with product expectations [10], with related work integrating multivariate interfaces and AI-generated content to personalize e-commerce experiences [11]. Across both threads, virtual iteration before physical production is identified as the central value proposition [12], but existing systems stop short of applying generated outputs to a specific physical product in real time during the consumer's decision moment, which is the gap to be targeted.

B. Augmented Reality (AR) for Interactive Packaging

In parallel, AR has transitioned from a novelty to a strategic communication channel that overcomes the physical limitations of traditional labels by providing virtually unlimited digital real estate for information [13]. Recent work has applied AR specifically to enhance packaging design [6] and to improve carton packaging structure [14], demonstrating feasibility but typically operating with pre-authored static content. A consumer-facing distinction has emerged between AR that provides utility (e.g., nutritional facts) and AR that provides immersion (e.g., storytelling or aesthetic enhancement) [15], with immersive approaches more strongly associated with engagement and recall.

C. Personalization, Emotion, and Storytelling

The most compelling opportunity lies in the convergence of AI and AR to facilitate emotional resonance and narrative storytelling. Interactive packaging has been explored as a medium for storytelling through the concept of "unfolding the self," allowing users to uncover layers of narrative embedded in the product [16], transforming a generic product into a unique artifact. Psychologically, this heightened sense of ownership can be linked to the endowment effect, in which individuals value items more highly when they feel a personal connection to them [7]. Several recent systems have begun to combine generative AI with AR directly: GenAIR investigates design factors for employing generative AI in AR [17]; CARING-AI explores context-aware AR instruction generation [18]; and PILAR delivers personalized LLM-based explanations within AR [19]. These works confirm the feasibility of generative AI as a content engine for AR displays, but they remain largely application-specific or research-facing. A gap persists in end-to-end, hardware-free solutions that allow retail consumers to generate, visualize, and apply custom designs to a specific physical product in real time. This gap can be addressed by combining a hosted diffusion model with marker-based WebAR in a browser-only deployment, operationalizing the theoretical benefits of personalization-as-ownership for retail consumer use.

III. System Overview

Hyper-Personalized Packaging in AR (HYPAR) is designed to enable real-time, user-driven customization of retail product packaging through a pipeline that connects generative AI with AR visualization. The system was implemented as a web application to provide cross-platform access without requiring participants to install dedicated software, allowing consumers to engage with the experience directly through any modern mobile browser. Fig. 1 presents the full system architecture, organized into three layers: a client-side WebAR rendering layer, an API inference layer, and a generative model layer. Marker tracking and rendering update at the device display frame rate, producing a stable, continuously registered overlay, whereas the generative step incurs a one-off interactive latency of 5–10 seconds before a new texture becomes available. The assumptions required for this interactivity are a modern mobile browser with WebGL support, a functioning camera, adequate lighting for reliable marker detection, and network connectivity to the hosted inference endpoint.

A. Client Layer

The client is a Next.js single-page application that runs entirely in the user's mobile browser. It handles the prompt selection interface, dispatches generation requests, and is responsible for rendering the AR composite output. AR.js handles marker-based tracking from the device camera feed, computing the pose of an AR marker affixed to the physical perfume box, while Three.js renders a 3D mesh of the box at that tracked pose and applies the generated texture to its material. Concretely, AR.js processes each camera frame to detect a custom pattern marker and estimates the marker pose as a 4×4 model-view matrix. Three.js consumes this matrix to position a box-shaped mesh within a WebGL scene aligned to the physical product, so that the overlay tracks the box as the device moves. The generated 1024×1024 PNG is applied as a UV-mapped texture across the mesh faces, and the scene is re-rendered at the browser animation-frame rate, decoupled from the one-off generation step. Marker detection, pose estimation, and rendering all run client-side in the browser; only the texture-generation request is sent off-device.

B. Generative Pipeline

When the user selects a preset descriptor, the client constructs a structured prompt and submits it to the Together.ai inference API, which serves the Black Forest Labs FLUX.1.1 [pro] model [20] as a hosted endpoint. We selected FLUX.1.1 [pro] because it operates as a rectified flow transformer in latent space and offers approximately $6\times$ faster generation than its predecessor while maintaining strong prompt adherence and stable latent representations [20], both of which are important for an interactive retail scenario where wait time directly affects the experience. The model is built on a hybrid architecture of multimodal and parallel diffusion transformer blocks scaled to 12 billion parameters, trained using flow matching, a general method for training generative models that includes diffusion as a special case [21]. Hosting inference on Together.ai avoided the GPU infrastructure burden from the deployment and produced end-to-end generation latencies in the 5–10 second range under the conditions of our user study.

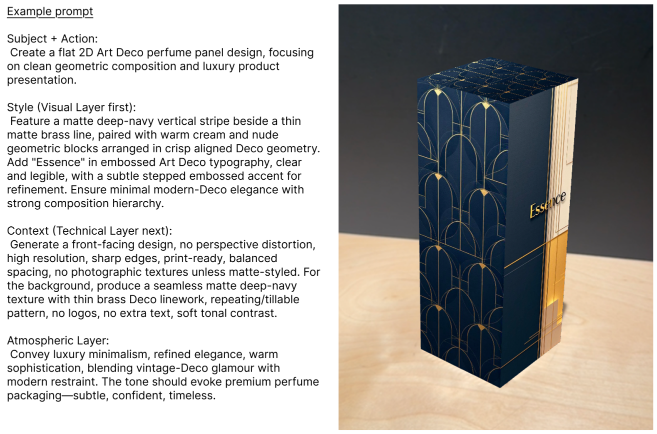

Fig. 2. Example of the internal prompt for the *Art Deco Navy Stripe* preset and the corresponding AI-generated packaging visual rendered in HYPAR.

C. Prompt Framework and Presets

Prompts are constructed using the structured framework recommended by Black Forest Labs [22]: *Subject + Action + Style + Context*. The subject specifies the product class and design intent, the action describes what the design should do, the style specifies the visual language, and the context covers technical and atmospheric details such as lighting, mood, and rendering constraints. This framework was used because word ordering and structural consistency materially affect FLUX output quality, and a fixed structure reduces variance across repeated generations.

To make the system accessible to retail consumers without prompt engineering expertise, the user-facing interface presents five preset descriptors spanning Art Deco, mid-century, 80s maximalist, and 60s–70s pop-art aesthetics: Art Deco Navy Stripe, Art Deco Deep Plum, Sage & Cream Deco, Ultramarine 80s Maximalist, and Vibrant Funky Retro Pop.[1] Each preset is bound to a fully specified internal prompt that follows the Subject + Action + Style + Context framework; the user selects the descriptor, but the prompt sent to FLUX.1.1 [pro] is substantially longer and explicitly controls subject, design action, style elements (typography, colour, motif), and atmospheric layer (mood, finish, lighting). Fig. 2 shows a representative example for the Art Deco Navy Stripe preset and the corresponding texture rendered on the physical perfume box. This decoupling between the user-facing label and the underlying prompt is a deliberate design choice because it ensures consistent, high-quality output for an exploratory user study while preserving a clear path to free-text input in future deployments.

D. End-to-End Flow

The complete user interaction proceeds as follows. The user opens the WebAR application via a QR code and grants camera permission. They select one of the five preset descriptors, which triggers a request to the Together.ai endpoint with the corresponding internal prompt. After 5-10 seconds, the response returns a high-resolution PNG that is uploaded as a Three.js material texture, MeshBasicMaterial, and bound to the 3D box mesh. The user then aims their smartphone camera at the physical perfume box; AR.js detects the marker, recovers the pose, and Three.js composites the textured mesh in real time, producing an AR overlay in which the personalized design appears applied to the actual product. The user can iterate by selecting a different descriptor and repeating the cycle.

By delivering this pipeline through a standard web browser, HYPAR provides a lightweight integration path for retailers that requires no changes to existing physical inventory or supply chains because the same generic perfume box used as a shelf tester can anchor any number of personalized previews.

[1]The preset visual designs and their prompts were developed by Mahnoor Hasan (Digital Futures, OCAD University).

IV. Methodology

The research followed a structured approach to evaluate the HYPAR technology probe in the case of a perfume package. The study design was reviewed by the Institutional Research Ethics Board (Reference #044-202526). While 16 individuals expressed interest in participating, we excluded one participant because they did not meet the study criteria. The technical apparatus consisted of a hybrid setup: each participant's own smartphone served as the primary interface for the WebAR application, while a physical perfume box modified with AR markers served as the tangible anchor for the digital overlays. Preparation involved generating digital access points (QR codes for consent, application launch, and post-study feedback), preparing the marker-augmented physical assets, and provisioning the Together.ai inference endpoint.

A. Participants

The final cohort of 15 participants was composed of 13 male and 2 female participants. The majority (14 of 15) were aged 18–24, with one participant aged 35–44. All participants held at least a college or university degree. Reported frequency of purchases of luxury or specialized fragrance and cosmetic products in the prior 12 months was distributed as follows: 8 of 15 participants reported occasional purchases (1–2 times per year), 6 of 15 reported seasonal purchases (3–5 times per year), and one participant reported 6–10 purchases per year. Before engaging with the prototype, participants completed a pre-study questionnaire to establish baseline attitudes towards packaging aesthetics and personalization. All Likert items used a 5-point scale (1 = strongly disagree, 5 = strongly agree). Participants reported a relatively low self-rated frequency of luxury fragrance and cosmetic purchases (M=2.93, SD=1.39). On packaging attentiveness, most participants agreed or strongly agreed that they notice and appreciate small design details in product packaging (M=3.73, SD=1.10). Approximately 73% indicated that they are more likely to purchase a product if the packaging matches their personal style (M=3.93, SD=0.88), and 53% found current retail packaging options to be too

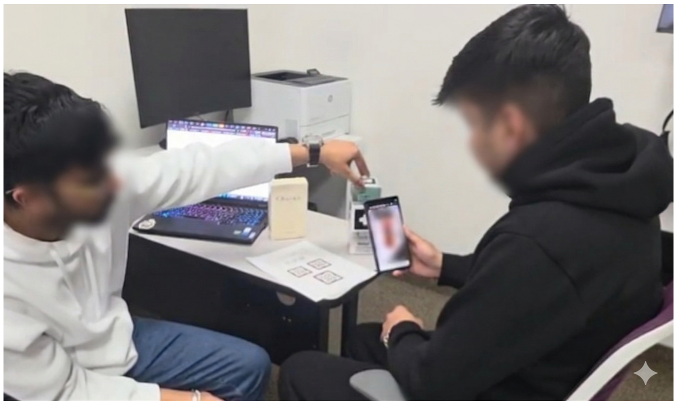

Fig. 3. A participant interacting with the HYPAR prototype on a mobile device during an individual user study session.

generic or repetitive (M=3.53, SD=0.74). When asked about willingness to pay a premium for personalized packaging, six agreed or strongly agreed, five were neutral, and four disagreed (M=3.33, SD=1.11). These baseline data indicate moderate-to-strong receptiveness to packaging personalization in principle, which the interaction phase of the study was designed to test in a concrete artifact-based setting.

B. Procedure

Each participant attended an individual session conducted on the university campus, lasting approximately 15 minutes (Fig. 3). Sessions followed a standardized protocol divided into four stages: (1) onboarding, (2) generative design & AR interaction, (3) comparative reflection, and (4) evaluation.

1) Onboarding and Initialization: Upon arrival, participants used their mobile devices to scan a QR code leading to a digital consent form. After consent was documented, a second QR code launched the HYPAR web application in the mobile browser. Participants then provided basic demographic information, including age, gender, and prior experience with cosmetic shopping.

2) Generative Design and AR Interaction: In the core interaction phase, participants navigated to the prompt selection interface and chose one of the five preset descriptors to trigger the generative pipeline. The system contacted the Together.ai endpoint to produce a high-resolution texture, a process requiring approximately 5–10 seconds. Once the texture was generated, participants pointed their smartphone camera at the physical perfume box, and the application composited the AI-designed texture onto the 3D model in real time. To ensure participants experienced the system's interactivity, each was encouraged to perform at least one design iteration, generating a new texture from a different preset and comparing it to their initial result.

3) Comparative Reflection: Following the interactive design phase, a standard, non-personalized perfume box was placed alongside the AR-enhanced version. Participants were asked to observe both items simultaneously and to compare their emotional responses and perception of the product's luxury value between the generic box and their custom design.

4) Evaluation: The session concluded with a two-part evaluation. First, participants scanned a final QR code to complete a post-study questionnaire on system usability. Second, a semi-structured interview was conducted using a qualitative interview guide. The interview covered participants' likes and dislikes, the effectiveness of seeing the design in AR versus on a 2D screen, and how such a tool might influence their purchasing decisions in a real retail context.

C. Data Analysis

Quantitative data combined the pre-study Likert items and the post-study System Usability Scale (SUS) [23]. Descriptive statistics were used throughout. SUS was selected as the primary post-interaction quantitative measure because the central technical question addressed by this study is whether an end-to-end pipeline can be made usable for a non-expert retail consumer. Usability is the necessary precondition for any subsequent claim about engagement or purchase intent. As an exploratory study, the SUS is acknowledged to measure the preliminary perceived usability only and does not measure perceived ownership, luxury, or purchase intent. Those constructs are examined qualitatively through the interview analysis and are treated as exploratory observations. Qualitative data from the semi-structured interviews were analyzed using the Semantic Thematic Analysis [24]. The qualitative analysis was used to contextualize the participants' experiences of the system, particularly their emotional responses.

V. Results

A. System Usability

The mean System Usability Scale (SUS) score for HYPAR was 89.50 (SD=5.11), which can be interpreted as an *Excellent* range [25], or to a Grade A [23]. This suggests a high *perceived* usability for a first-time interaction with a novel interface combining generative AI with marker-based AR. The narrow standard deviation further suggests that this perception was relatively uniform across participants.

B. Thematic Analysis

The qualitative interviews produced 109 coded segments of transcripts, organized into three themes: *Engagement*, *Retail Impact*, and *Usability*. The entity of analysis was determined to be a sentence. Table I presents the full thematic structure with sub-theme descriptions. We use the qualitative material here primarily to substantiate and qualify the SUS result, with a secondary focus on engagement-related observations relevant to the research question.

1) Engagement (T1): The Engagement theme accounted for 46 coded segments across five sub-themes. *Ownership* (14/46) was the most frequently coded: participants reported a stronger personal connection to the product after contributing to its design, often framing the personalized box as theirs to keep and the generic box as disposable. P14 captured this directly: *"the*

TABLE I
THEMATIC ANALYSIS RESULTS SHOWING THREE OVERARCHING THEMES, THEIR SUB-THEMES, AND DESCRIPTIONS.

Theme	Sub-theme	Description
T1: Engagement	Ownership	Participants felt a stronger personal attachment to the product when they had a role in designing its packaging.
	Luxury	The personalized design was perceived as more premium or luxurious than the standard generic box.
	Creative Control	Participants valued the ability to choose preset descriptors and see their creative intent materialized on a product.
	Customization	Multiple participants wanted more granular control, such as colour pickers, font choices, or the ability to upload their own images.
	Iterative Design	The ability to generate multiple designs and compare them was valued as part of the creative process.
T2: Retail Impact	Purchase Intent	Participants indicated they would be more likely to purchase a product if this personalization tool were available in-store.
	Retail Experience	Participants envisioned how this tool could change their interactions with products in a physical retail environment.
T3: Usability	Ease of use	The browser-based approach was appreciated for requiring no app download or special hardware.
	Wait Time	The 5-10 second wait for texture generation was noted by some as a minor source of friction, though it was generally acceptable.
	Tracker Reliability	Occasional "jitter" or tracking loss momentarily broke the immersion for some users.
	Prompt Clarity	Some participants expressed a desire to write their own prompts rather than select from presets.

generic one makes me feel part of the crowd . . . the personalized one makes me feel special." *Creative Control* (10/46) and *Luxury* (7/46) supported this pattern: participants described the AI-generated designs as more premium-looking than the standard box and valued seeing their aesthetic preferences materialize on a real product, though one participant qualified that perceived luxury is ultimately driven by brand reputation (P7). *Customization* (12/46) was distinct from these in that it surfaced as participant requests for capabilities beyond the current presets, including vector or image upload, multi-face design control, and bottle-side customization, signalling an appetite for engagement that exceeds what the current prototype offers. *Iterative Design* (3/46) was the smallest sub-theme but consistent: the ability to compare multiple generated outputs was treated as a meaningful part of the creative process.

2) Retail Impact (T2): The Retail Impact theme comprised 40 coded segments across two sub-themes. *Retail Experience* (23/40) was the single most frequently coded sub-theme in the dataset. Participants repeatedly contrasted the spatial AR overlay with conventional 2D web views, describing the AR experience as easier to evaluate and closer to handling the actual product. P14: *"It's always difficult, because it's a 2D shape . . . This was far easier, because I had the object in my hand."* The same sub-theme also surfaced a boundary case: P13 noted that *"since you're already there in the store, you can, 90% of the time, pick up the actual product in your hand and feel and know,"* suggesting the AR overlay may add less value when the physical product is already directly accessible. *Purchase Intent* (17/40) captured a related but distinct effect: participants reported that previewing the design in AR reduced the uncertainty typically associated with purchasing a personalized product, as they would use the feature in-store to own a customized box.

3) Usability (T3): The Usability theme accounted for 23 coded segments and substantiates the high SUS result while surfacing two specific frictions. *Ease of Use* (8/23) aligns directly with the SUS score result: participants described the browser-based interface as immediately learnable and contrasted it favourably with applications that require dedicated downloads or onboarding. The two friction points were *Tracker Reliability* (8/23), where several participants reported jitter, tracking loss under finger occlusion, and texture cropping at oblique angles, with one participant (P9) describing a specific failure mode in which placing a thumb on the marker disrupted detection; and *Prompt Clarity* (6/23), where participants expressed a preference for free-text input over selecting from the five preset descriptors. *Wait Time* (1/23) was raised once and was characterized as minor. The qualitative pattern here is consistent with the SUS score: aggregate usability is high, but the specific frictions cluster around the AR tracking layer and the bounded expressivity of the preset interface, both of which we address in Section VII.

VI. DISCUSSION

This section interprets the findings reported in Section V to address the research question: *How does an augmented reality system that dynamically generates hyper-personalized packaging influence consumers' first impressions, perceived attractiveness, and engagement with a physical product?*

A. System Usability

The SUS score of 89.50 indicates that HYPAR could be perceived as accessible and immediately learnable. The Ease of Use sub-theme corroborated this, with participants emphasizing the absence of an installation step and the conventional smartphone interaction model. This is a non-trivial result for AI-AR convergence systems, which often surface their underlying complexity at the interface layer; the present finding is consistent with prior HCI work showing that careful design factor selection can keep generative AI integration tractable for end users in AR contexts [17], [18].

However, the Tracker Reliability sub-theme made clear that aggregate usability is only part of the picture. Several participants reported jitter, tracking loss under finger occlusion, and texture cropping at oblique angles, issues that momentarily disrupted the AR illusion without being severe enough to lower the SUS score substantially. Combined with comments on landing-page design, this suggests that targeted improvements to tracking robustness and onboarding interface design are necessary for deployment readiness, even though the prototype is demonstrably usable.

B. First Impressions and Perceived Attractiveness

The pre-study data established a baseline in which 73% of participants reported being more likely to purchase a product if its packaging matched their personal style, and 53% described current retail packaging as too generic. The thematic findings indicate that HYPAR addressed both attitudes within a single interaction. The comments in the Luxury sub-theme can be interpreted as a demonstration of participants associating the AI-generated designs with higher quality and exclusivity, with the visual subtlety and colour coherence of the generated textures contributing to this perception. This connects to Chen et al.'s finding that consumer satisfaction with AI-generated packaging is closely tied to expectation alignment [10]: the preset-driven outputs were perceived by participants as expectation-aligned for a luxury product, though the present study does not isolate the contribution of the structured BFL prompt framework [22] from other system factors.

However, this effect was not universal. We argue that the concept of luxury is culturally and socioeconomically situated; what participants in this sample associated with premium aesthetics may not generalize across different cultural contexts, income levels, or consumer backgrounds. One participant noted that perceived luxury is ultimately driven by brand reputation rather than visual customization, suggesting that the system's influence on perceived attractiveness may be moderated by prior brand awareness. For an unknown or emerging brand, personalization can have greater weight in shaping first impressions and for established luxury brands, it can function as a complement.

The AR visualization layer further amplified first impressions by allowing participants to assess the design in a spatial context. Participants described 2D representations as *dull and static* (P1) and noted that the AR overlay made the design feel present and real, a level of assessment that static product images cannot provide. This aligns with the immersive AR results reported by Accardi et al. [15], who found that immersive AR is more strongly associated with engagement and recall than informational AR. Side-by-side comparison with the generic box reinforced such a finding. The same physical product was characterized as common and disposable in its generic form and as unique and worth preserving in its personalized form, indicating that the visual personalization itself significantly enhances the product's perceived attractiveness and shapes a highly positive first impression.

C. Engagement

The qualitative data indicate that active participation in the design process fostered emotional attachment. The Ownership sub-theme, the most frequently coded in Theme 1, captured this directly: participants who selected a preset and observed the resulting output felt a personal investment in the outcome and explicitly contrasted the personalized box with the generic alternative as personally meaningful versus interchangeable. This is consistent with the endowment effect [7], in which involvement in the creation of an object increases its perceived value, and with prior framings of interactive packaging as a medium for narrative and identity expression [16].

Engagement extended beyond ownership to the creative process itself. Participants described the experience as an act of self-expression in which the system served as a medium for translating internal aesthetic preferences into a visible, tangible form. Iterative comparison across multiple generations, though less frequently coded, suggested a secondary engagement mechanism in which participants converged on a personally satisfying output. In aggregate, generative AI provided creative agency while AR provided spatial confirmation by anchoring the generated design to the physical product, producing an experience that was simultaneously participatory and informative.

The data also revealed boundaries to this effect. P13's observation that in-store consumers can typically handle the actual product directly suggests that the AR overlay's added value may be context-dependent, with the strongest effects likely in scenarios where the final personalized product is not yet physically accessible, such as online shopping, gifting, or in-store made-to-order previews of unprinted packaging. Several participants also felt the five presets constrained their creative expression. Their requests for free-text prompting, image upload, multi-face design control, and bottle customization suggest that expanding the input design space could further increase the depth and duration of engagement, while raising new design challenges around prompt scaffolding for non-expert users.

> *"If having this in the retail store actually happens, I would probably do it for every product, because having something custom-made for every product is really cool." – P15*

VII. Limitations and Future Work

Several limitations of this study should be acknowledged. The small sample size of the user study, while appropriate for an exploratory technology probe study, limits the generalizability of the quantitative findings, in addition to a possible novelty effect. The demographic composition was skewed, which may not reflect the broader retail consumer population. The sample consisted of general consumers, which is an imperfect match for a study framed around perceived luxury; habitual luxury-fragrance consumers, being more familiar with the category, might weigh brand and material cues differently from personalization. Consistent with a technology-probe design, the descriptive statistics do not compute inferential hypothesis tests, so the quantitative results should be read as characterizing this sample, not as generalizable effects. Future studies should recruit a larger and more demographically diverse sample to validate these preliminary findings.

A more substantive methodological limitation is the absence of a comparison condition that isolates the contribution of the AR visualization layer from the contribution of personalization itself. Because the study compared HYPAR's AI-generated, AR-rendered output against a standard generic box, the observed effects on perceived ownership, luxury, and purchase intent are most defensibly attributed to the combination of personalization and AR visualization instead of testing either factor in isolation. Prior work indicates that personalization alone can shift consumer satisfaction with packaging [10], and the present study cannot distinguish how much of the observed effect is attributable to AR specifically versus to participant involvement in the creative process. A follow-up study with a within-subjects ablation design, comparing HYPAR's full AR pipeline against a 2D web-based personalization condition and a non-personalized baseline, would directly address this confound and would be a next step.

The marker-based AR tracking introduced technical friction that was reported by multiple participants. Jitter, tracking loss under finger occlusion, texture cropping at oblique angles, and difficulty with simultaneous multi-marker detection were all noted. Future iterations could explore markerless tracking approaches, such as object detection models (e.g., YOLO [26]) or plane detection via WebXR's hit-test API, to improve tracking robustness directly on the product surface without requiring physical fiducials.

The current system offers five preset descriptors. While this design decision ensured consistent, high-quality generative output for the user study, participant feedback indicated a clear demand for greater input flexibility. Recommended features included custom text prompting, the ability to upload personal images or vector graphics, the option to apply different designs to different faces of the box, and the extension of customization to the product container itself. Incorporating these features would require careful prompt scaffolding to maintain output quality and avoid overwhelming non-expert users while expanding the design space.

The study was conducted in a controlled environment. Factors such as ambient lighting, time pressure, social context, and competing products may influence both usability perceptions and engagement in ways the current study could not capture. As noted in Section VI, P13's observation that in-store consumers can typically handle the actual product suggests the AR overlay may add less value in physical retail than in online or remote contexts. Future work should explore deployment in both physical retail environments and online shopping contexts where the AR component may provide a greater comparative advantage.

The current prototype generates 2D textures mapped onto a 3D model. Future work could explore the generation of full 3D packaging models, including custom bottle shapes and structural variations, as well as integration with emerging spatial computing platforms such as mixed reality headsets, which could further enhance the immersive quality of the visualization. We plan to release the HYPAR codebase and prompt corpus as open source to support replication and to enable other researchers to build on this work.

VIII. Conclusion

This paper presented HYPAR, a cross-platform WebAR prototype that combines a hosted generative diffusion model with marker-based augmented reality to allow retail consumers to design and preview personalized packaging on a physical product in real time. The findings from the user study showed preliminary evidence for using such a tool to develop a new pathway for consumer packaging, which can engage with potential consumers by offering an interactive and customizable retail experience.

Acknowledgments

We would like to thank the National Centre of Excellence for Immersive Technology at Algoma University for their support. We also thank Mahnoor Hasan and Simone Jones, from Digital Futures at OCAD University, for their contributions to initial ideation and the prompts for the visual design. Finally, we would like to thank all participants of the user study for providing their time and feedback in the evaluation of this work.

References

[1] J. D. N. Dionisio, W. G. Burns III, and R. Gilbert, "3D virtual worlds and the metaverse: Current status and future possibilities," *ACM Computing Surveys*, vol. 45, no. 3, pp. 1–38, 2013.

[2] M. R. Gleim, H. McCullough, C. Gabler, L. Ferrell, and O. C. Ferrell, "Examining the customer experience in the metaverse retail revolution," *Journal of Business Research*, vol. 186, p. 115045, 2025.

[3] W. Batat, "Phygital customer experience in the metaverse: A study of consumer sensory perception of sight, touch, sound, scent, and taste," *Journal of Retailing and Consumer Services*, vol. 78, p. 103786, 2024.

[4] D. C. Orazi and G. Nyilasy, "Phygital products: Effects and boundaries of metaverse-first retail strategies," *Journal of Retailing*, vol. 102, no. 2, pp. 371–386, 2026.

[5] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695, IEEE, 2022.

[6] M. Rani and M. K. Ramlie, "The utilization of augmented reality (ar) applications as packaging design enhancement," *International Journal of Academic Research in Business & Social Sciences*, vol. 13, no. 7, pp. 251–274, 2023.
[7] C. K. Morewedge and C. E. Giblin, "Explanations of the endowment effect: An integrative review," *Trends in Cognitive Sciences*, vol. 19, no. 6, pp. 339–348, 2015.
[8] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems (NeurIPS '20)*, pp. 6840–6851, Curran Associates, 2020.
[9] Z. Gao and Z. Li, "A tea packaging design method based on artificial intelligence generated content," *Journal of Engineering Design*, vol. 37, no. 2, pp. 339–371, 2025.
[10] T. Chen, D.-B. Luh, and J.-G. Wang, "Analyzing ai-generated packaging's impact on consumer satisfaction with three types of datasets," *International Journal of Data Warehousing and Mining*, vol. 19, no. 1, pp. 1–17, 2023.
[11] A. Wasilewski, Y. Chawla, and E. Pralat, "Enhanced e-commerce personalization through ai-powered content generation tools," *IEEE Access*, vol. 13, pp. 48083–48095, 2025.
[12] L. He, "Ai-driven virtual simulation for packaging customization," *International Journal of Information System Modeling and Design (IJISMD)*, vol. 13, no. 2, pp. 1–10, 2022.
[13] A. Kyguolienė and R. Braziulytė, "Application of augmented reality in product packaging: Challenges and development opportunities.," *Management of Organizations: Systematic Research*, no. 88, 2022.
[14] S. Xia, "Application of augmented reality technology in carton packaging structure design," in *2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 9–13, IEEE, 2021.
[15] S. Accardi, C. Campo, C. Rossi, M. Bilucaglia, C. Pasta, M. Zito, M. Caccamo, V. Russo, *et al.*, "Informative vs. immersive: An exploratory study of augmented reality on packaging labels.," in *NeuroPsychoEconomics Conference Proceedings*, pp. 39–39, Association for NeuroPsychoEconomics, 2025.
[16] L. Dao, "Unfolding the self: Storytelling through interactive packaging with augmented reality," in *Learn X Design 2023 Conference Proceedings*, Design Research Society, 2023.
[17] Y. Hu, D. Zhang, and A. Quigley, "Genair: Exploring design factor of employing generative ai for augmented reality," in *Proceedings of the ACM Symposium on Spatial User Interaction (SUI '23)*, pp. 1–3, ACM, 2023.
[18] J. Shi, R. Jain, S. Chi, H. Doh, H.-g. Chi, A. J. Quinn, and K. Ramani, "Caring-ai: Towards authoring context-aware augmented reality instruction through generative ai," in *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI)*, pp. 1–17, ACM, 2025.
[19] R. K. Kundu, I. Ahmed, and K. A. Hoque, "Pilar: Personalizing augmented reality interactions with llm-based human-centric and trustworthy explanations for daily use cases," in *2025 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 273–280, IEEE, 2025.
[20] Black Forest Labs, "Announcing FLUX1.1 [pro] and the BFL API." https://blackforestlabs.ai/announcing-flux-1-1-pro-and-the-bfl-api/, 2024. Accessed: 2026.
[21] Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le, "Flow matching for generative modeling," in *International Conference on Learning Representations (ICLR)*, 2023.
[22] Black Forest Labs, "FLUX prompting guide." https://docs.bfl.ml/guides/prompting_guide_flux2, 2024. Accessed: 2026.
[23] J. R. Lewis, "The system usability scale: past, present, and future," *International Journal of Human–Computer Interaction*, vol. 34, no. 7, pp. 577–590, 2018.
[24] V. Clarke and V. Braun, "Thematic analysis," *The journal of positive psychology*, vol. 12, no. 3, pp. 297–298, 2017.
[25] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.
[26] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLOv8." https://github.com/ultralytics/ultralytics, 2023. Accessed: 2026.