Bridging Real and Virtual in XR Editing: Color Extraction and Contextual Placement

Ye-Eun Kim, Phil-Joong Kim and Soo-Mi Choi
Department of Computer Science and Engineering and Convergence Engineering for Intelligent Drone,
XR Research Center
Sejong University, Seoul, Korea
kyy1462@sejong.ac.kr, fjfo101@sju.ac.kr, smchoi@sejong.ac.kr

Abstract

This study introduces an XR-based interactive editing framework that allows users to sample colors directly from the physical world, apply them to virtual content, and project results back into the real environment. Existing AR systems often suffer from scanning delays and low color fidelity, while VR systems rely on predefined palettes that limit contextual creativity; recent MR authoring environments improve spatial continuity but still provide limited support for user-driven acquisition and direct editing of real-world colors or materials. To overcome these issues, we designed a cyclical interaction loop—Sense, Edit, Display, and Register—that integrates real-time color extraction, 2D canvas editing, synchronized 3D previews, and spatial alignment, offering a seamless bridge between physical and virtual spaces. We conducted a within-subjects experiment with ten participants who performed the same creative task across AR (paper coloring and scanning), VR (palette-based coloring), and XR (the proposed system). Task completion time and color fidelity (CIEDE2000) were measured as technical indicators, while cognitive load (NASA-TLX), usability (SUS, UEQ-S), and creativity support (CSI) were collected through questionnaires. Semi-structured interviews complemented the quantitative results. Findings show that XR reduced cognitive load compared to AR, while color fidelity did not differ significantly between conditions. Although slower than VR in task completion, XR provided higher ratings in creativity, immersion, and enjoyment. Interview responses emphasized the novelty and intuitiveness of extracting real-world colors and seeing them immediately reflected in physical space. These results suggest that XR's strengths stem from the integration of real-world color sampling and contextual alignment into a unified creative experience. This work suggests XR as more than a coloring tool, establishing a cyclical, bidirectional interaction paradigm that enhances creativity and immersion. The proposed framework informs future XR authoring tools and points toward extensions such as multi-user collaboration, AI-based emotion recognition, and applications in education, art, and therapy.

Keywords: ***Extended Reality (XR), Real-world color sampling, Context-aware placement, Scene understanding, Creativity support***

1. Introduction

With the recent release of new head-mounted displays (HMDs) such as Meta Quest 3, Sony PlayStation VR2, and PICO 4, the accessibility of XR devices has increased rapidly. During the past 3 to 5 years, the average market price of HMDs has decreased by approximately 30%, lowering the entry barrier to adoption. Currently, improvements in network latency and cyber-sickness have expanded the scope of XR applications to diverse domains, including education, art, and industry [12, 39]. XR encompasses virtual reality (VR), augmented reality (AR), and mixed reality (MR), referring to interactive environments where physical and virtual objects coexist. In particular, the alignment between real and virtual objects and their interaction are highlighted as critical factors influencing the presence and immersion of the user [43]. However, some existing systems remain limited to one-way information flow from the physical to the virtual world. The lack of contextual adaptation—such as scanning delays, inconsistent lighting, and scale distortions—often results in visual incoherence and cognitive burden [4].

To address these limitations, this study proposes an XR-based interactive editing framework that cyclically integrates physical and virtual experiences. The framework implements a Sense–Edit–Display–Register loop, wherein users directly extract colors from physical objects (like a digital eyedropper), immediately apply them to a 2D canvas and 3D virtual objects, and then see the results contextually aligned within the physical environment. Through this process, creators can continuously collect inspiration from

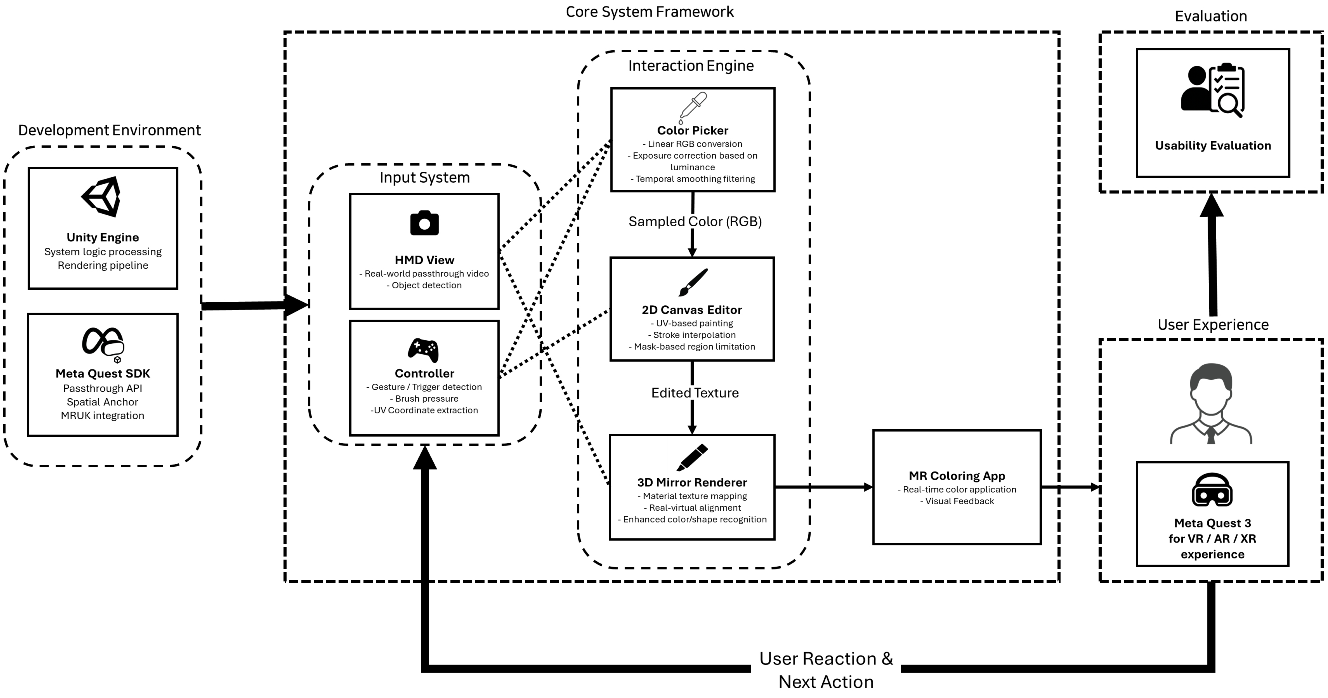


Figure 1: Core System Framework

the real world, edit it digitally, and instantly experience the results projected back into physical space, thus achieving a bidirectional and cyclic creative experience. Unlike VR, which prioritizes speed through palette-based color selection, or AR, which discontinuously digitizes physical canvases, the proposed framework provides a new balance by incorporating real-world stimuli directly into creative materials and reintegrating outcomes into physical context.

To evaluate the effectiveness of the proposed framework, we designed a user study comparing AR, VR, and XR conditions. Using a within-subjects design, ten participants in their twenties with prior XR experience performed the same creative task across all conditions. Technical metrics included task completion time and color fidelity (ΔE_{00}, CIEDE2000), while user experience metrics assessed cognitive load (NASA-TLX), system usability (SUS), user experience quality (UEQ-S), and creativity support (CSI). Semi-structured interviews were additionally conducted after each condition to capture qualitative insights and complement the quantitative findings. Given the scale and composition of the participant sample, the study is intended to identify comparative interaction trends rather than to provide definitive generalizable conclusions.

The contributions of this paper are threefold:

- **Bidirectional and cyclic XR framework.** We present an interaction loop that integrates real-world color extraction, digital editing, and contextual registration in physical space, thereby extending creative experiences beyond the one-way flow of conventional systems.
- **Scene understanding–based contextual alignment.** By leveraging scene understanding techniques to register virtual objects to real-world surfaces, and spatial constraints, the framework enhances visual and cognitive coherence as well as immersion.
- **Exploratory comparative evaluation across AR, VR, and XR.** Through a three-condition comparative study (AR/VR/XR), the results indicate that the proposed framework can provide improvements not only in creativity, immersion, and usability but also in interaction quality and color fidelity, establishing its value beyond a simple coloring tool.

2. Background

2.1. Limitations of Existing Interactive Editing Systems

Creative and editing systems in XR environments have evolved through a wide range of approaches centered on AR, VR, and MR. Early physical–virtual hybrid systems typically relied on workflows in which users colored drawings on paper, scanned or digitized them via a camera, and mapped the results onto virtual 3D objects. Representative examples include TeamLab's Sketch Aquarium [51] and Museum HEI's Puppy Park, which extended physical creative outputs into digital spaces. However, such approaches remain limited to a unidirectional flow of information from the physical to the virtual domain and do not support interactions in which virtual content meaningfully feeds back into the real-world context. As a result, the creative process is interrupted by a separate camera-scanning stage, and final outputs remain confined to screens, undermining immersion and continuity of experience [10, 11].

VR-based creative tools (e.g., Quill, Open Brush) [2, 1] enable freeform drawing and rapid editing in three-dimensional space, offering new expressive possibilities for both professional artists and general users. More recent systems, such as [57], further integrate generative AI into VR authoring workflows. Nevertheless, these tools largely confine creation and editing to closed virtual environments. Color selection typically depends on predefined palettes, making it difficult to directly incorporate contextual cues from the physical world, such as real-world colors, materials, or lighting conditions, as creative resources. This limitation partially conflicts with perspectives from Situated Cognition [14, 23, 25], which emphasize that cognitive processes are deeply coupled with the surrounding environment and tools.

AR-based systems commonly overlay virtual objects onto real scenes and allow users to modify appearance through menus or UI-based interactions [54]. Prior research on color stabilization and illumination compensation [40] and studies emphasizing the importance of context awareness in AR (e.g., Towards Pervasive Augmented Reality) have strengthened the technical foundations of visual coherence. More recent work such as [56] enable users to sketch directly in situ and create expressive, responsive animations anchored in the physical environment, expanding hands-on interaction beyond static overlays. Similarly, systems like [28] leverages scene understanding and AI to generate virtual content on the fly through verbal input while observing the physical environment. Although such systems can automatically adjust scale and placement based on spatial context, content generation is largely driven by AI rather than by direct user creation, which can limit the accurate reflection of user intent and hands-on creative agency. Overall, these approaches tend to focus on improving visual quality and convenience, while offering limited expressive interaction that treats the physical world itself as an active creative material. It has often remained at the level of color substitution or parameter adjustment, constraining exploratory and experimental creative practices [21, 33].

A representative MR authoring environment is CAVE-AR [7], which enables designers to observe and modify users' AR experiences in real time from a VR environment, substantially improving the efficiency of MR content development. More recent MR authoring tools [36] have further explored end-user participation in early-stage creative design, allowing non-expert users to engage directly in spatial ideation and layout within mixed-reality environments. In many cases, creative control continues to be designer-centric, focusing on spatial configuration or object placement rather than on treating real-world sensory attributes as expressive creative inputs.

A closer examination of existing methodologies reveals a critical "workflow disconnection" that prior systems have not fully addressed. In VR, painting systems have advanced physical realism and expressiveness through passive-haptic interaction and multisensory feedback mechanisms [15, 32]. However, these systems still operate within visually isolated virtual environments, where creators typically rely on predefined digital palettes rather than directly sampling and reusing sensory attributes from the physical world. Even recent generative workflows for authoring 3D layouts in VR largely confine creative input within a closed virtual context, remaining disconnected from real-world visual properties such as color and material appearance [57].

Conversely, spatial augmented reality and AR/MR techniques have long demonstrated strong capabilities in projecting or visualizing digital imagery directly onto physical objects and spaces [44]. Yet many AR/MR systems emphasize in-situ visualization or annotation more than bidirectional creative editing. For instance, AR sketching tools enable users to draw responsive animation effects in context [56], and Passthrough MR applications maintain real-world visibility for skill training scenarios [3]. While these approaches preserve visual connectivity, they do not provide an integrated creative loop in which users can capture real-world visual properties and use them to iteratively edit virtual content and re-project results back into the environment.

Moreover, recent AI-assisted in-situ authoring systems can generate and place content from high-level user input, but they do not explicitly center workflows around direct sampling and reuse of physical-world visual attributes [28]. Consequently, the end-to-end workflow remains fragmented: users are isolated in VR creation tools, guided in MR training/overlay experiences, or dependent on AI-driven generation pipelines. This motivates research on workflows that better sustain an in-context creative loop across real-world cue acquisition, editing, and situated evaluation.

2.2. Flow, Cognitive Load, and Interaction Structures

Flow constitutes an optimal psychological state of deep engagement, achieved through a balance between challenge and skill [13]. In XR creative tasks, maintaining this state depends critically on minimizing disruptive transitions between user actions and perceptual feedback. However, conventional digital creation systems often impose extraneous cognitive load [49, 5] through disjointed workflows—such as indirect manipulation, complex scanning procedures, or

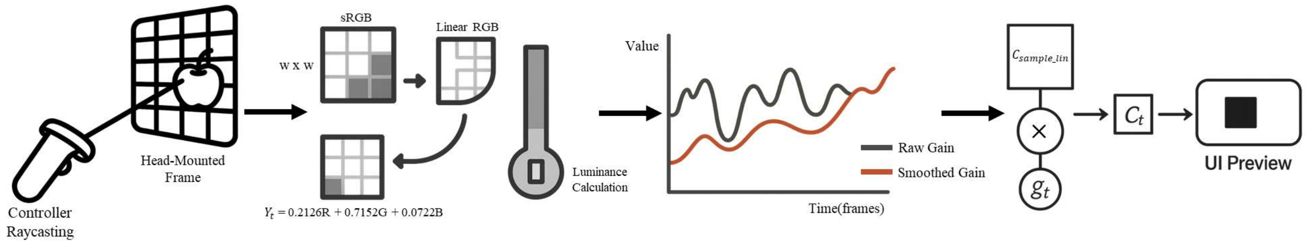


Figure 2: Pipeline for real-world color acquisition and correction

delayed visualization. These discontinuities force users to mentally bridge gaps between intention and outcome, thereby impeding immersion [31, 8, 50].

To address these issues, our framework leverages principles of Direct Manipulation [48, 22] and Embodied Cognition [14, 23]. By enabling users to extract colors directly from physical surroundings and apply them seamlessly to virtual objects, the interface minimizes the distance between user intent and system response. This approach treats real-world colors not as external references but as active creative materials, establishing a continuous perception-action loop that supports situated creativity through the interplay of body, environment, and imagination.

3. Proposed System

3.1. System Overview and Interaction Loop

The proposed system is not a mere aggregation of independent components but a tightly integrated framework. Its architecture centers on a cyclic interaction loop of *Sense* → *Edit* → *Display* → *Register*, which connects creative actions across real and virtual domains. Within this loop, users sense colors from real objects, edit them on a virtual canvas, and display the results immediately via an MR overlay on a 3D model that is spatially registered to the physical workspace.

Each stage is optimized to respond within a few tens of milliseconds, minimizing latency throughout color selection, application, and verification. This low-latency feedback is not merely a technical performance metric but a decisive factor in user experience quality, resolving the discontinuities of "scan-and-recognize"–based AR approaches [52]. By implementing a direct-manipulation loop in XR, the framework reduces the gap between user intent and system response, thereby lowering extraneous cognitive load and allowing creators to focus on the task rather than interface procedures [19, 30](Fig. 2).

Stage 1: Real-World Color Sensing and Acquisition The goal is to extract perceptually consistent color information that reflects user intent despite variations in the HMD's passthrough camera [41]. Beyond pixel copying, the process accounts for color space, exposure, and temporal stability.

(1) Spatial projection for target selection. The user emits a virtual ray using the controller to aim at a real-world object. The system computes the 3D intersection point in world coordinates and back-projects it through the HMD camera model onto the 2D passthrough frame I_t, yielding pixel coordinate $\mathbf{u} = (u_x, u_y)$ used for color sampling.

(2) Linear-space processing for color fidelity. Sampled pixels are converted to a linear RGB space prior to computation. Because camera sensors typically output γ-corrected sRGB data, performing averaging or correction in nonlinear space can cause physical distortion. All computations are performed in linear space and then reconverted to sRGB for display, ensuring both physical accuracy and perceptual consistency.

(3) Adaptive exposure compensation via local luminance. Auto-exposure may change the perceived color with ambient lighting. To stabilize this, the system analyzes a small neighborhood around $\mathbf{u}$, computes the local mean luminance, and applies adaptive correction. For a square window $\Omega_w(\mathbf{u})$ of $w \times w$ pixels, the local mean luminance is

$$\bar{Y}_t = \frac{1}{w^2} \sum_{p \in \Omega_w(\mathbf{u})} (0.2126\, R(p) + 0.7152\, G(p) + 0.0722\, B(p))$$

(4) Temporal smoothing for perceptual stability. The computed local mean $\bar{Y}_t$ is compared with a target luminance Y_{tgt} to derive a gain g_t. Directly applying g_t per frame can induce flicker under small hand tremors or lighting changes; therefore, g_t is smoothed via an exponential moving average, yielding perceptually stable, predictable color transitions.

(5) Final color computation. The smoothed gain g_t is applied to the linear sample color $\mathbf{c}^{\mathrm{lin}}_{\mathrm{sample}}$, and the result is γ-corrected back to sRGB to yield the final sampled color $\mathbf{C}_t$:

$$\mathbf{C}_t \;=\; \gamma\big(g_t \cdot \mathbf{c}^{\mathrm{lin}}_{\mathrm{sample}}\big)\,.$$

Key tunable parameters exposed to the user balance responsiveness and stability, enabling adjustment for different lighting conditions or personal preferences. It is important to note that this local mean luminance-based approach is a strategic design choice intended to prioritize real-time performance (maintaining 90Hz) on mobile HMDs. While this method serves as a simplified approximation that may not fully capture complex lighting dynamics—such as strong specular reflections or extreme contrast—it effectively stabilizes short-term fluctuations caused by the camera's auto-exposure. This optimization is specifically targeted at diffuse surfaces (e.g., paper, canvas), enabling users to perform rapid sampling and editing iterations without latency, rather than aiming for physically perfect exposure reconstruction.

Stage 2: Real-Time 2D Canvas Editing The acquired color is applied to a 2D texture–based virtual canvas. When the user ray-casts to a point on the canvas surface, the system converts it to normalized UV coordinates and performs brush stamping around that point. To prevent stroke discontinuities during fast controller motion, intermediate points are interpolated proportionally to the distance between consecutive frames, automatically constrained below $\Delta_{\max}$ to maintain uniform density and edge quality. Optional mask textures or alpha channels restrict paintable regions, preventing unintended color application. All updates are reflected instantly in the canvas texture, allowing fine-grained adjustment based on immediate visual feedback.

Stages 3 & 4: Synchronized 3D Mirror Preview and Spatial Registration A distinctive feature of the system is the mirror-preview mechanism, which synchronizes the 2D canvas edits with a 3D object in MR space in real time. Users can therefore observe how their chosen color interacts with lighting and geometry directly on the 3D form. To support stable placement of these virtual elements within the physical environment, the system utilizes the Semantic Scene Understanding capabilities provided by the Meta XR SDK. The physical space is pre-scanned and classified into a semantic scene model, identifying entities such as Table, Wall, and Floor.

For spatial registration, the system utilizes this pre-computed semantic data. Upon placement request, the system identifies the appropriate Workable Surface based on its assigned semantic label. The virtual canvas is then instantiated and parented to the coordinate system of the corresponding Semantic Anchor. By attaching content to these pre-classified anchors rather than arbitrary world coordinates, the system helps maintain alignment with the physical environment. This approach leverages the underlying XR runtime to manage spatial stability, minimizing drift and supporting the perceptual feedback loop.

4. Experiments

This section describes the methodology and evaluation process of the user study conducted to validate the effectiveness of the proposed XR-based interactive editing system. The objective of this study is not merely to compare different platforms, but to evaluate the overall user experience package created by the combination of creative methods and output placement strategies provided in each condition. Accordingly, rather than isolating single variables, this experiment adopts a holistic comparison approach that examines the integrated interaction experience across AR, VR, and XR environments, following prior comparative frameworks in immersive media research [34, 37, 26] (Fig. 3).

The evaluation framework included four dimensions: cognitive load (NASA-TLX), usability (SUS), user experience (UEQ-S), and creativity support (CSI) [17]. Technical indicators such as task completion speed and color fidelity (ΔE_{00}) were also measured. Post-condition interviews captured participants' views on feedback immediacy, real-world color sampling, and overall ease of use, providing qualitative context to the quantitative results.

4.1. Experimental Conditions

1. **AR condition:** Participants physically colored a drawing on paper using traditional tools and scanned the completed image with a camera. The extracted color information was then mapped onto virtual 3D objects.
2. **VR condition:** Wearing an HMD, participants worked in a fully immersive virtual environment and painted a predefined 3D object. Colors were selected from a pre-configured virtual palette, representing the traditional VR approach where creation is separated from real-world context.
3. **XR condition:** In the proposed system, participants perceived the real world through the HMD's passthrough function and directly sampled colors from real-world objects. The extracted colors were made available as a brush input, allowing participants to apply them to chosen areas of a 2D virtual canvas. The painted results were then reflected in real time on 3D objects aligned with the physical environment.

4.2. Evaluation Suite

The study involved ten participants (five male and five female) in their twenties, all possessing prior experience with XR applications. All participants completed the same creative task ("beach scene" model). To control for order effects, the sequence of conditions was randomized, following standard counterbalancing procedures used in prior within-subjects experimental designs [53]. Before each task, participants received sufficient training on system operation. After completing each condition, participants filled out questionnaires, and upon completion of all conditions, they provided additional feedback through a comprehensive usability survey and a semi-structured interview.

(a) AR (b) VR (c) XR

Figure 3: User workflows under AR, VR, and XR conditions. AR involves physical coloring and scanning, VR focuses on virtual palette-based coloring, and XR enables real-world color sampling and context-aware virtual placement.

4.2.1 Technical Evaluation

Task Completion Speed: The time required for each participant to complete the assigned task under each condition was measured in seconds [20].

Color Fidelity: The perceptual difference between the original colors and the colors applied via the system was quantified using the CIEDE2000 (ΔE_{00}) color difference formula, grounded in human visual perception [35]. Lower ΔE_{00} values indicate higher fidelity.

4.2.2 Usability Evaluation

Cognitive Load: The NASA Task Load Index (NASA-TLX) was used to measure perceived workload on a 21-point scale, covering physical demand, temporal demand, performance, effort, and frustration [18].

System Usability: The System Usability Scale (SUS) [6] is a widely adopted tool for rapidly assessing usability and user acceptance of interactive systems [29, 38]. It consists of ten items alternating between positive and negative statements, each rated on a 5-point Likert scale.

Interaction Quality: To evaluate overall user experience, the User Experience Questionnaire (UEQ) was adopted in its short form (UEQ-S) to reduce participant fatigue and enhance response quality [27, 47, 58]. UEQ-S employs an 8-item, 7-point semantic differential format, in which users rate their experience along bipolar adjective pairs.

4.2.3 Creativity Evaluation

Creative Outcome: The Creativity Support Index (CSI) [9] was used to measure the extent to which each system supported creativity. CSI evolved from the broader framework of Creativity Support Tools, which aims to evaluate how digital systems facilitate creative processes in HCI [16]. Since this study involved individual users wearing HMDs, the collaboration dimension was excluded.

5. Results

5.1. Technical Evaluation

Task Completion Speed. A Friedman test revealed significant differences in task completion speed across conditions ($\chi^2(2) = 20.00, p < .001$). Post hoc analyses indicated that the VR condition ($M = 146.90$) was the fastest, followed by XR ($M = 220.80$) and AR ($M = 300.00$), with each pairwise difference reaching significance.

Color Fidelity. The analysis of color fidelity (ΔE_{00}) showed no statistically significant difference between conditions ($\chi^2(2) = 1.95$, $p = .38$). Nevertheless, descriptive trends indicated that the VR condition ($M = 0.85$), which enabled exact digital replication, achieved the lowest error values and thus the highest color accuracy. Importantly, the proposed XR system ($M = 3.15$) showed numerically lower color error than the AR condition ($M = 8.45$) in our sample, indicating a trend toward improved fidelity, though the overall test was not significant.

Table 1: Technical performance measures across conditions. Lower values indicate better performance.

Measure	AR	VR	XR
Mean task completion time (s)	300.00	**146.90**	220.80
Mean color difference (ΔE_{00})	8.45	**0.85**	3.15

5.2. Usability Evaluation

NASA-TLX. The NASA Task Load Index demonstrated high internal consistency ($\alpha = .892$), and normality assumptions were met across conditions. A repeated-measures ANOVA revealed significant differences among conditions ($F(2, 18) = 46.09, p < .001, ges = .422$). Adjusted post hoc comparisons indicated that AR ($M = 4.88, 95\% CI[3.76, 6.01]$) imposed the highest workload, VR ($M = 3.20, [2.31, 4.09]$) was intermediate, and XR ($M = 2.03$, $[0.97, 3.09]$) was the lowest. Subscale analyses revealed that AR scored highest in temporal demand and performance pressure, whereas XR consistently reported the lowest levels of effort and frustration.

System Usability Scale (SUS). SUS scores showed that XR ($M = 81.2, SD = 9.5$) and VR ($M = 75.2, SD = 13.5$) exceeded the benchmark score of 68, while AR ($M = 65.8, SD = 18.5$) fell below the threshold. ANOVA results indicated a marginal effect of condition ($F(1.18, 10.64) = 4.54, p = .053$), but a non-parametric Friedman test confirmed significant differences ($\chi^2(2) = 9.95, p = .007, W = .50$). Wilcoxon post hoc tests showed XR significantly outperformed AR ($p = .027$) with a large effect size ($r = .874$). At the item level, XR was rated positively in terms of confidence, integration of functions, and willingness to reuse, while AR was rated negatively in terms of consistency and task flow.

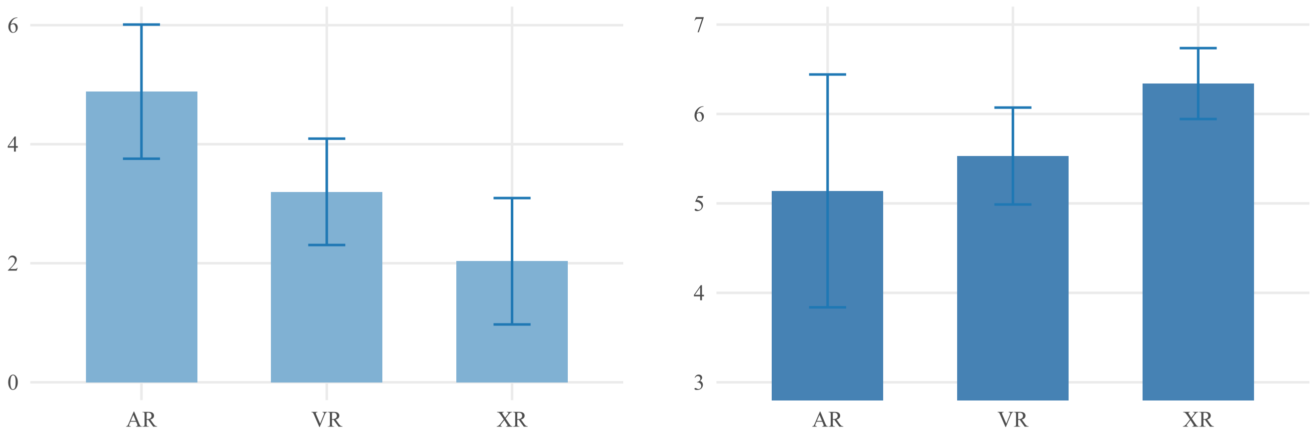
UEQ-S. In the User Experience Questionnaire–Short (UEQ-S), XR achieved the highest ratings for both pragmatic quality ($M = 2.22$) and hedonic quality ($M = 2.42$). Pragmatic quality showed a significant condition effect ($F(2, 18) = 3.99, p = .037$), with XR receiving the highest mean rating compared to VR and AR. Hedonic quality was also significantly higher in XR than in VR ($p = .031$). A Friedman test confirmed overall score differences ($\chi^2(2) = 6.16, p = .046$). Item-level responses showed that XR was consistently associated with descriptors such as “leading edge,” “inventive,” “clear,” and “efficient.”

5.3. Creativity Evaluation

Creativity Support Index (CSI). The CSI demonstrated very high reliability ($\alpha = .951$). A repeated-measures ANOVA revealed significant differences between conditions ($F(2, 18) = 4.71, p = .023, ges = .166$). Adjusted post hoc comparisons indicated that XR was significantly higher than VR ($p = .029$). Friedman test results corroborated this finding ($\chi^2(2) = 11.1, p = .0038, W = .556$), and Wilcoxon tests further showed that XR significantly outperformed AR ($p = .027$). Descriptive statistics showed that XR ($M = 6.34, 95\% CI[5.94, 6.74]$) achieved the highest scores, while VR ($M = 5.53, [4.99, 6.07]$) and AR ($M = 5.14, [3.84, 6.44]$) were comparatively lower and closer to each other.

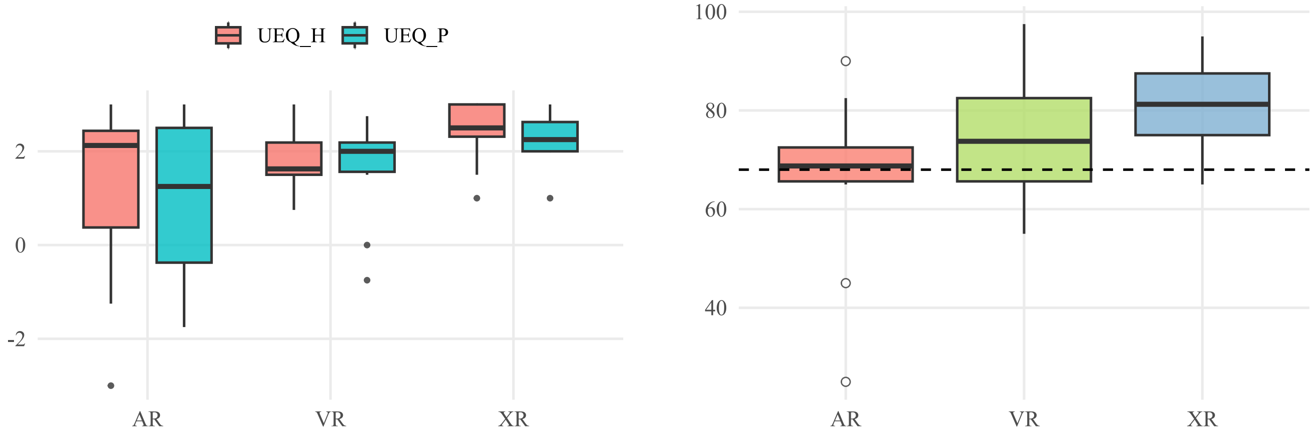
6. Discussion

This study examined the advantages and limitations of an XR-based interactive editing system in comparison with AR and VR approaches. Quantitative analyses revealed that the XR condition yielded the lowest cognitive load in the NASA-TLX. This effect can be attributed to the integration of real-world color sampling and contextual registration in physical space, which created a seamless connection to the creative task. By eliminating the need for palette navigation or scanning procedures, XR reduced extraneous cognitive load and helped maintain immersion. In contrast, AR introduced higher temporal demand(NASA3) and frustration due to disruptive scanning steps(NASA5), consistent with prior findings that AR environments tend to impose greater mental and effort-related workload compared to VR conditions [55]. Although VR enabled faster task execution, its higher NASA-TLX scores suggest that immersive environments can also amplify perceptual and attentional demands,



(a) NASA-TLX mean scores with 95% CI

(b) CSI mean scores



(c) UEQ-S pragmatic vs hedonic quality

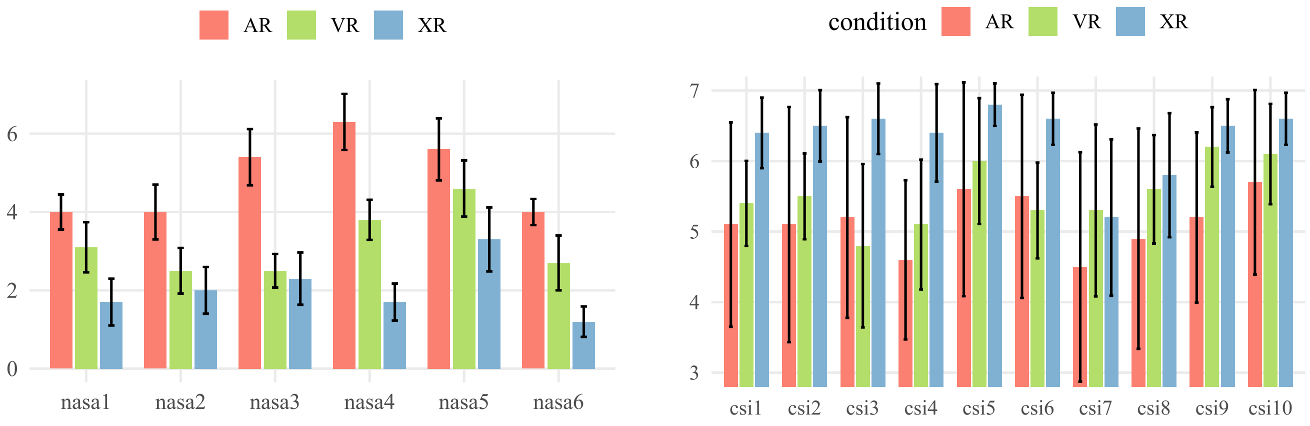
(d) SUS scores by condition

Figure 4: Overall evaluation results across conditions (AR/VR/XR)

rather than uniformly reducing cognitive workload [42].

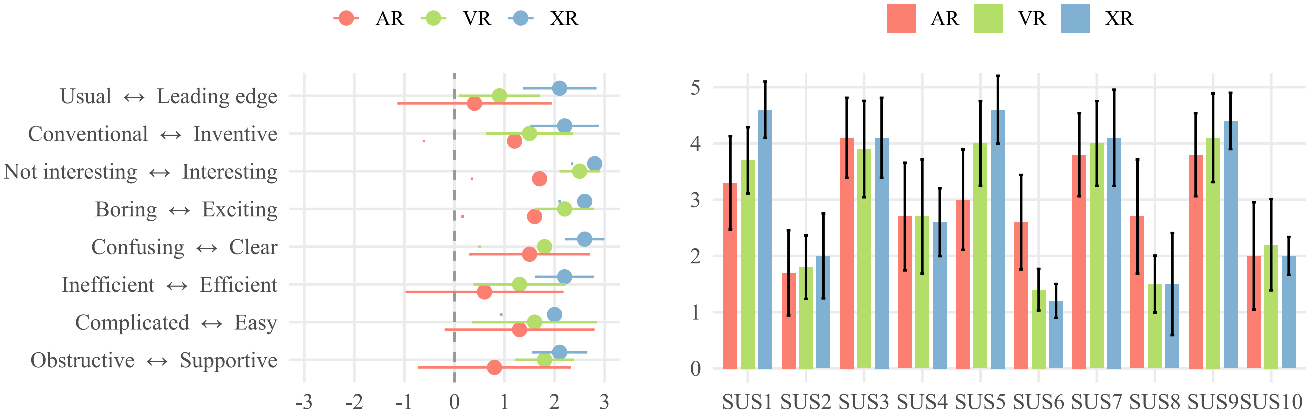
SUS results further substantiated the strengths of XR. Items such as “I would like to use this system frequently” (SUS1), “I thought the functions were well integrated” (SUS5), and “I felt confident using the system” (SUS9) received high ratings, indicating strong usability, consistency, and learnability. Conversely, AR was rated poorly on items such as “I found the system inconsistent” (SUS6) and “I found the system cumbersome to use” (SUS8), highlighting how delayed feedback and perceived color inaccuracies undermine usability(Fig. 5).

XR also outperformed AR and VR in UEQ-S and CSI measures, particularly in creativity and enjoyment. In the UEQ-S, XR was strongly associated with descriptors such as “leading edge,” “inventive,” “clear,” and “efficient.” Interview responses supported this outcome, as participants described the process of sampling colors from real-world objects and immediately applying them to virtual canvases and 3D objects as a novel and stimulating experience. These findings align with prior research showing that XR environments foster enhanced creativity, enjoyment, and innovative user experiences through immersive and contextually grounded interactions [45, 24]. CSI results echoed this trend, with XR achieving the highest scores in items such as “I enjoyed making experimental attempts” (CSI5). Participants engaged playfully by sampling colors from snack wrappers, the sky outside the window, or clothing and applying them directly to their creations. In contrast, AR participants reported frustration that results were not immediate and that unexpected outcomes occurred when attempting to revise colors, which aligned with its lower CSI scores. This finding is consistent with prior research indicating that immediate feedback and freedom for experimental exploration are key drivers of creativity support and enjoyment in immersive environments [46]. VR participants, while efficient, noted that palette-based interaction limited their abil-



(a) NASA-TLX item scores

(b) CSI item scores



(c) UEQ-S item scores (-3 to $+3$ scale)

(d) SUS item mean scores with 95% CI

Figure 5: Item-level evaluation results across conditions (AR/VR/XR)

ity to generate new ideas.

Nevertheless, this study has interpretive limitations. It remains unclear whether XR's superior outcomes can be attributed primarily to real-world color sampling, contextual registration, or the combination of both, as these factors were not independently manipulated. However, repeated interview statements such as "the process of directly exploring colors was enjoyable" and "it was intuitive to see results immediately in the real space" suggest that both factors jointly contributed to enhanced user experiences. While this study cannot disentangle their individual effects, it successfully validates the holistic "experience package" of the XR system, which is itself of both scholarly and practical significance.

6.1. Limitations and future work

This study is exploratory in nature and relies on a relatively small sample of young adults with prior XR experience. While this limits the generalizability of the findings to broader populations—such as older adults, XR novices, or professional artists—the within-subjects design enables meaningful comparison of interaction trends across AR, VR, and XR conditions under controlled settings.

In addition, although the XR condition generally outperformed AR and VR in terms of cognitive load, usability, and creativity support, these results should be interpreted with an important design consideration in mind. The proposed XR system integrates two tightly coupled components: real-world color extraction and contextual MR registration. As such, the present study does not isolate the individual contribution of each component. Rather, the observed benefits reflect the combined effect of a unified interaction loop that connects sensing, editing, and contextual feedback into a single continuous experience. This choice was intentional, as the primary goal of the study was to evaluate holistic interaction quality rather than component-level performance. From a technical perspective, the current exposure compensation algorithm operates within specific hardware and environmental constraints. Since the method relies on a gain-based adjustment derived from local mean luminance, it is inherently vulnerable to sensor clipping and limited dynamic range. In scenarios involving strong specular highlights or lighting intensities that exceed the sensor's capacity, pixel information may be permanently lost, meaning that color cannot be recovered through gain adjustment alone. Consequently, the system's reliability is currently optimized for diffuse surfaces under non-extreme lighting conditions; future iterations could incorporate HDR reconstruction techniques or learning-based estimation to enhance robustness in high-dynamic-range environments.

Future work will extend this evaluation to larger and more diverse participant groups and investigate how expertise level, age, and creative background influence system effectiveness. Moreover, follow-up studies employing factorial or ablation-based experimental designs will be necessary to disentangle the relative contributions of real-world color sampling and spatial contextual alignment, and to clarify their respective roles in reducing cognitive load and supporting creative exploration.

7. Conclusion

This research proposed an XR interactive editing system that integrates real-world color sampling with contextual registration in physical space, and empirically compared it with VR and AR approaches. Quantitative evaluations considered both technical metrics—task completion speed and color fidelity—and user experience metrics, including NASA-TLX, SUS, UEQ-S, and CSI. Results indicated that VR excelled in task speed but restricted creative exploration, while AR involved delayed feedback and showed a numerically higher mean color error in our sample, which may have degraded outcome quality. XR, though slower than VR, showed a numerically lower mean color error than AR in our sample, along with higher immersion and creativity support, demonstrating a balanced overall profile across technical performance and experiential quality. Qualitative analysis further confirmed that the processes of exploring real-world colors and seeing results reflected immediately in physical space provided novelty and playfulness, underscoring that XR's effectiveness derived from the integration of both factors into a unified user experience.

The contributions of this study are threefold. First, it introduces a new creative paradigm that surpasses VR's speed-driven focus and AR's limited realism by enabling a cyclic connection between real and virtual contexts. Second, it expands user experience evaluation beyond usability alone to include immersion, creativity, and enjoyment, thereby broadening the scope of empirical assessment. Third, it proposes a novel design framework for XR creative tools based on a bidirectional interaction loop that integrates real-world stimuli into creation and projects outcomes back into physical context.

Acknowledgement

This work was supported in part by the ITRC Support Program under Grant IITP-2026-RS-2022-00156354 funded by the Ministry of Science and ICT of Korea and the Institute of Information and Communications Technology Planning and Evaluation (IITP).

References

[1] Gesture vr, 2025. Accessed: 2025-10-04. 3

[2] Stencil vr, 2025. Accessed: 2025-10-04. 3

[3] M. Banquiero, G. Valdeolivas, D. Ramón, et al. A color passthrough mixed reality application for learning piano. *Virtual Reality*, 28:67, 2024. 3

[4] J. S. Bazargani, A. Sadeghi-Niaraki, and S. M. Choi. Integrating user input in automated object placement for augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 2025. 1

[5] P. Breves and J. P. Stein. Cognitive load in immersive media settings: the role of spatial presence and cybersickness. *Virtual Reality*, 27(2):1077–1089, 2023. 3

[6] J. Brooke. Sus – a quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194):4–7, 1996. 6

[7] M. Cavallo and A. G. Forbes. Cave-ar: A vr authoring system to interactively design, simulate, and debug multi-user ar experiences. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1724–1729. IEEE, 2019. 3

[8] Y. Chandio, V. Interrante, and F. M. Anwar. Reaction time as a proxy for presence in mixed reality with distraction. *IEEE Transactions on Visualization and Computer Graphics*, 2025. 4

[9] E. Cherry and C. Latulipe. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(4):1–25, 2014. 6

[10] K. Cho, H. Kim, and Y. Lee. Augmented reality coloring book with transitional user interface. *Indian Journal of Science and Technology*, 9(20):1–5, 2016. 3

[11] A. Clark, A. Dünser, and R. Grasset. An interactive augmented reality coloring book. In *SIGGRAPH Asia 2011 Emerging Technologies*, pages 1–1, 2011. 3

[12] Consegic Business Intelligence. Virtual reality market - global market size, trend analysis and forecast 2023-2030. Accessed: 2025-10-04. 1

[13] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper & Row, 1990. 3

[14] P. Dourish. *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press, 2001. 3, 4

[15] A. R. Fender, M. Krichenbauer, C. George, S. S. Wei, and G. Fitzmaurice. Infinitepaint: Painting in virtual reality with

passive haptics using wet brushes and a physical proxy canvas. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 2023. 3

[16] J. Frich, L. MacDonald Vermeulen, C. Remy, M. M. Biskjaer, and P. Dalsgaard. Mapping the landscape of creativity support tools in hci. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–18. ACM, 2019. 7

[17] S. Graser, F. Kirschenlohr, and S. Böhm. User experience evaluation of augmented reality: A systematic literature review. *arXiv preprint*, 2024. 5

[18] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, 1988. 6

[19] X. He, Y. Liu, J. Wang, H. Gao, and Y. Xu. Reducing latency in a collaborative augmented reality service. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 1097–1105. ACM, 2019. 4

[20] K. Hornbæk. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2):79–102, 2006. 6

[21] L. Huang and A. A. Musah. The influence of augmented reality on creativity, student behavior, and pedagogical strategies in technology-infused education management. *Journal of Pedagogical Research*, 8(2):260–275, 2024. 3

[22] E. L. Hutchins, J. D. Hollan, and D. A. Norman. Direct manipulation interfaces. In *User Centered System Design*, pages 87–124. CRC Press, 1986. 4

[23] D. D. Hutto and E. Myin. *Radical Enactivism: Basic Minds without Contents*. MIT Press, 2012. 3, 4

[24] J. Jeon. The impact of xr applications' user experience-based design innovativeness on loyalty. *Cogent Business & Management*, 10, 2022. 8

[25] S. R. Klemmer, B. Hartmann, and L. Takayama. How bodies matter: Five themes for interaction design. In *Proceedings of the 6th Conference on Designing Interactive Systems*, pages 140–149, 2006. 3

[26] N. Kyaw, J. Barbosa, J. Chang, and J. Goncalves. Comparing usability of augmented reality and virtual reality on a mapping task. *Applied Sciences*, 13(21):11693, 2023. 5

[27] B. Laugwitz, T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*, pages 63–76. Springer, 2008. 6

[28] J. Lee, F. Aleotti, D. Mazala, G. Garcia-Hernando, S. Vicente, O. J. Johnston, I. Kraus-Liang, J. Powierza, D. Shin, J. E. Froehlich, G. Brostow, and J. Van Brummelen. Imaginatear: Ai-assisted in-situ authoring in augmented reality. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*, 2025. 3

[29] Y. Lee and K. Jung. Analysis of user experience for the class using metaverse – focus on 'spatial'. *Journal of Practical Engineering Education*, 14(2):367–376, 2022. 6

[30] N. Leins, J. Gonnermann-Müller, and M. Teichmann. Comparing head-mounted and handheld augmented reality for guided assembly. *Journal on Multimodal User Interfaces*, 2024. 4

[31] T. H. Li, H. Suzuki, Y. Ohtake, T. Yatagawa, and S. Matsuda. Efficient evaluation of misalignment between real and virtual objects for hmd-based ar assembly assistance system. *Advanced Engineering Informatics*, 59:102264, 2024. 4

[32] Z. Li, Q. Huo, S. Ding, S. Ye, J. Shi, C. Song, and J. Wang. Painting++: Human–computer collaborative painting in virtual reality with multisensory interaction. *Design and Artificial Intelligence*, 1(2):100019, 2025. 3

[33] A. Liaqat, F. Liu, B. Berengard, J. Cao, and A. Monroy-Hernández. Understanding young people's creative goals with augmented reality. *Proceedings of the ACM on Human-Computer Interaction*, 9(2):1–21, 2025. 3

[34] L. Lisle, K. Davidson, E. J. Gitre, C. North, and D. A. Bowman. Different realities: a comparison of augmented and virtual reality for the sensemaking process. *Frontiers in Virtual Reality*, 4:1177855, 2023. 5

[35] M. R. Luo, G. Cui, and B. Rigg. The development of the cie 2000 colour-difference formula: Ciede2000. *Color Research & Application*, 26(5):340–350, 2001. 6

[36] S. Maurya, K. Arai, K. Moriya, P. A. Arrighi, and C. Mougenot. A mixed reality tool for end-users participation in early creative design tasks. *International Journal on Interactive Design and Manufacturing*, 13(1):163–182, 2019. 3

[37] L. Merino, M. Schwarzl, M. Kraus, M. Sedlmair, D. Schmalstieg, and D. Weiskopf. Evaluating mixed and augmented reality: A systematic literature review (2009–2019). In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 438–451. IEEE, 2020. 5

[38] G. Nam. Analysis of user experience for the metaverse content's production – focused on 'zepeto'. *Journal of Korea Design Forum*, 29(1):173–186, 2024. 6

[39] Omdia. Xr market in 2035 and beyond: Forecast, challenges, and the road to mass adoption, 2023. 1

[40] T. Oskam, A. Hornung, R. W. Sumner, and M. Gross. Fast and stable color balancing for images and augmented reality. In *Proceedings of the Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission (3DIMPVT)*, pages 49–56. IEEE, 2012. 3

[41] T. Oskam, A. Sorkine-Hornung, R. W. Sumner, and M. Gross. Fast and stable color balancing for images and augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 49–57. IEEE, 2012. 4

[42] J. Parong and R. E. Mayer. Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6):785, 2018. 8

[43] M. Radanovic, K. Khoshelham, and C. Fraser. Aligning the real and the virtual world: Mixed reality localisation using learning-based 3d–3d model registration. *Advanced Engineering Informatics*, 56:101960, 2023. 1

[44] R. Raskar, G. Welch, K.-L. Low, and D. Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001 (Eurographics Symposium on Rendering)*, pages 89–102. Springer, 2001. 3

[45] R. Schlagowski, S. Mertes, D. Nazarenko, A. Dauber, and E. André. Xr composition in the wild: the impact of user

environments on creativity, ux and flow during music production in augmented reality. In *Proceedings of the 19th International Audio Mostly Conference: Explorations in Sonic Cultures*, pages 152–161, September 2024. 8

[46] R. Schlagowski, F. Wildgrube, S. Mertes, C. George, and E. André. Flow with the beat! human-centered design of virtual environments for musical creativity support in vr. In *Proceedings of the 14th Conference on Creativity and Cognition*, pages 428–442, June 2022. 8

[47] M. Schrepp, J. Kollmorgen, and J. Thomaschewski. A comparison of sus, umux-lite, and ueq-s. *Journal of User Experience*, 18(2), 2023. 6

[48] B. Shneiderman. Direct manipulation: A step beyond programming languages. *Computer*, 16(8):57–69, 1983. 4

[49] J. Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):257–285, 1988. 3

[50] A. Tang, C. Owen, F. Biocca, and W. Mou. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 73–80, 2003. 4

[51] teamLab. teamlab aquarium. Accessed: 2025-10-04. 2

[52] R. J. Teather, A. Pavlovych, W. Stuerzlinger, and I. S. MacKenzie. Effects of tracking technology, latency, and spatial jitter on object movement. In *IEEE Symposium on 3D User Interfaces (3DUI)*, pages 43–50. IEEE, 2009. 4

[53] S. Villa, T. Kosch, F. Grelka, A. Schmidt, and R. Welsch. The placebo effect of human augmentation: Anticipating cognitive augmentation increases risk-taking behavior. *Computers in Human Behavior*, 146:107787, 2023. 5

[54] M. Whitlock, J. Mitchell, N. Pfeufer, B. Arnot, R. Craig, B. Wilson, B. Chung, and D. A. Szafir. Mrcat: In situ prototyping of interactive ar environments. In *Virtual, Augmented and Mixed Reality. Design and Interaction*, volume 12190 of *Lecture Notes in Computer Science*, pages 235–255. Springer, 2020. 3

[55] N. Xi, J. Chen, F. Gama, M. Riar, and J. Hamari. The challenges of entering the metaverse: An experiment on the effect of extended reality on workload. *Information Systems Frontiers*, 25(2):659–680, 2023. 7

[56] Z. Xia, K. Monteiro, K. Van, C. Rountree, and R. Arora. Realitycanvas: Augmented reality sketching for embedded and responsive scribble animation effects. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*, 2023. 3

[57] L. Zhang, J. Pan, Y. Li, S. Zhang, S. Oney, and A. Hong. Vrcopilot: Authoring 3d layouts with generative ai models in vr. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24)*, 2024. 3

[58] X. Zhang, N. Y. Oh, and J. W. Park. A study on the user experience according to material differences of objects in projection mapping. *The Korean Society of Science & Art*, 41(1):313–322, 2023. 6