2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)

Virtual Roomie: Immersive Layout Co-design with a Virtual Agent

Angela L. Jimenez*
School of Applied and Creative Computing
Purdue University
West Lafayette, IN 47907, USA

Pedro Acevedo†
School of Applied and Creative Computing
Purdue University
West Lafayette, IN 47907, USA

Christos Mousas‡
School of Applied and Creative Computing
Purdue University
West Lafayette, IN 47907, USA

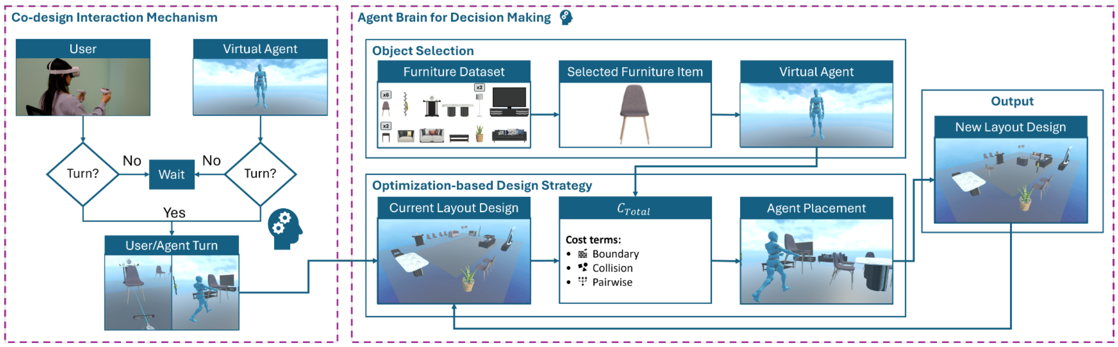


Figure 1: The co-design system pipeline details the interaction mechanism between a virtual agent and a user, as well as the agent's decision-making process. The interaction mechanism utilizes a turn-taking approach for the co-design task, in which both the user and the agent wait for the other to complete actions before editing the environment. The agent's decision-making involves two key actions: (1) selecting objects from a furniture dataset and (2) applying an optimization-based design strategy. This strategy employs a greedy algorithm to update the current layout by minimizing the total cost function (C_{Total}), which evaluates potential furniture placements based on proximity to boundaries, collision costs, and spatial relationships within the living room environment. The optimized placements suggested by the agent, integrated with user input, ultimately yield an updated layout design.

Abstract

We explored human-virtual agent collaboration during a layout design task in a virtual reality environment. Specifically, we developed a human-in-the-loop optimization-based method that drives the decision-making of the virtual agent. Our algorithm accounts for spatial constraints in furniture placement by evaluating boundary proximity, collision costs, and relationships between furniture items in real-time. It also considers the current configuration of the living room, as modified by the user during the co-design process, to guide the virtual agent's furniture placement decisions in the virtual living room. We compared our method (i.e., optimization) against two other co-design strategies (i.e., template and random) following a within-group ($N = 24$) study design. We found the proposed optimization co-design strategy significantly enhanced perceived collaboration compared to the other two co-design strategies. Moreover, our participants attributed higher private and public awareness to the virtual agent in the optimization condition. In addition, the analysis of the logged data showed that participants placed more furniture items and made fewer corrections when co-designing the living room with a virtual agent whose decisions were based on the optimization method. Our results demonstrate that a virtual agent's behavior, which dynamically responds to user actions while maintaining spatial coherence, creates more effective collaborative experiences in an immersive co-design task.

Index Terms: Virtual agent, co-design, layout design, virtual reality, immersive interaction.

*e-mail: jimen198@purdue.edu
†e-mail: paceved@purdue.edu
‡e-mail: cmousas@purdue.edu

1 Introduction

As virtual reality (VR) continues to evolve as an immersive medium, integrating artificial intelligence (AI) and virtual agents into these environments presents promising opportunities for various tasks. These evolving human and virtual agent interactions, such as collaborative design, virtual assistance, personalized training, and interactive storytelling, are transforming how users engage with immersive environments [2, 18, 48]. Motivated by this potential, recent research by Rasch et al. [34] explored how AI representation modes impact 3D object co-creation in VR. They identified that highlighting changes, incremental visualization, and embodiment significantly influence user perception and engagement. Notably, their findings revealed that embodied AI virtual agents affect users' perceived contribution to created models, with users attributing greater creative input to the AI when it was visually represented. The combination of AI interaction and VR spatial capabilities creates an ideal context for exploring effective collaboration between virtual agents and humans [29, 33]. Previous studies have deployed virtual agents in various VR scenarios, including guiding users through environments [48], dynamically sensing and responding to their surroundings [17, 36], engaging in conversations [47], collaborating on task-solving [8], and supporting the development of soft skills [54], showcasing the versatility and utility of virtual agents as valuable tools in VR.

Among the tasks well-suited to human-AI collaboration, layout design stands out as particularly relevant. It involves arranging items or creating virtual environments in a way that reflects users' design goals [41, 49]. Yu et al. [50] and Zhang et al. [52] showed that furniture layouts can accommodate aesthetic, functional, and personal preferences through diverse configurations, all while adhering to spatial constraints. This inherently open-ended problem space offers an ideal setting to explore how humans and AI agents make creative decisions in immersive environments [34, 38]. To support such collaboration, we developed a human-in-

2473-0726/25/$31.00 ©2025 IEEE

DOI 10.1109/ISMAR67309.2025.00026

the-loop optimization-based method that guides the virtual agent's decision-making. Our algorithm (see Figure 1) evaluates spatial constraints, including boundary proximity, collision between furniture items, and relationships between furniture items. To showcase our method, we designed a VR co-design layout (i.e., a living room) experience in which users and the agent alternately place furniture, enabling free navigation and direct contribution from both collaborators.

To evaluate the proposed optimization decision-making of the virtual agents, we conducted a user study to examine the effects of co-design on several variables concerning interaction and perception of the virtual agent, virtual agent awareness, user experience, design satisfaction, and user activity. Our findings offer valuable insights into human-AI co-design, especially in immersive VR environments where AI assists in creative tasks such as layout design.

2 Related works

2.1 AI for Layout Design

Layout design is a well-established problem in computer graphics [25, 30]. In recent years, automation has become increasingly integrated into design tools through AI, parametric modeling, and optimization techniques [16, 37, 42, 50]. Researchers have developed methods to support more effective scene composition, reducing manual effort while improving spatial quality. For instance, Yu et al. [50] developed an automated furniture layout tool that considers accessibility, visibility, and spatial relationships to support layout decisions beyond manual placement. Similarly, Yu et al. [51] proposed an interactive tool for recommending clutter items in indoor scenes, enhancing realism and reducing manual effort. In VR contexts, Wang et al. [41] investigated layout editing through object manipulation techniques that leverage spatial context to enable efficient arrangement in immersive environments. These methods highlight the complexity and creative potential of layout design tasks, which are inherently open-ended and allow for multiple valid solutions. In this study, we adopted the layout design problem as a means to investigate and compare different co-design strategies in a virtual environment, focusing on how human and agent interactions influence the design process.

2.2 Human-Virtual Agent Interaction

Human-virtual agent interactions span diverse contexts, supporting decision making, task execution, and adaptability in collaborative settings [9, 55]. Their effectiveness depends on several factors, including realism, user representation, intelligence, and trust. For instance, agents with realistic facial expressions and movements are often viewed as more credible and engaging [20], while nonverbal cues such as gestures, vocal qualities, and personality traits enhance the quality of interaction [6].

Regarding perceived intelligence, Choi et al. [8] found that users engaged more and performed better when interacting with intelligent agents. Yang et al. [47] showed that deeper agent knowledge enriched conversations, highlighting the value of adaptive and informed agent behavior. Virtual agents can help users learn tasks and pursue shared goals [32], adapting to individual needs, and provide feedback for smooth interaction [10]. However, users tend to cooperate more and risk more with human teammates than with AI ones, which can affect group dynamics and enjoyment [31].

Trust plays a crucial role in shaping human-agent interaction. Daronnat et al. [11] demonstrated that trust can be inferred from user behaviors such as reliance on the agent and perceived task difficulty. Moreover, the type of agent error affects user perception: omission errors tend to be more acceptable than incorrect actions, and users generally prefer moments of silence over inaccurate responses [10]. These findings highlight how various dimensions of agent behavior influence user experience and collaboration quality. Thus, we focus on human-virtual agent collaboration in an immersive design environment, aiming to explore how co-designing in VR with a virtual agent affects users' experiences.

2.3 Designing in VR

Researchers have demonstrated how VR transforms design processes by enhancing spatial understanding and creative exploration [3, 15, 38, 43]. Arora et al. [3] examined factors affecting VR sketching, identifying physical surface absence as a key challenge while establishing principles for improving accuracy through visual guidance. Drey et al. [13] explored pen and tablet interaction for 3D sketching with their VRSketchIn system, demonstrating how combining mid-air drawing with surface-based sketching gives designers greater control in immersive environments.

Freeman et al. [15] found significant advantages of VR-based modeling over traditional desktop CAD interfaces, showing that VR enables users to create more geometric features in the same timeframe and produce designs rated as more creative. In interior design, You et al. [49] developed RedesignUS, a VR system for rebuilding and customizing homes through synthesized layouts and decoration options. Their research demonstrated VR's effectiveness in supporting spatial visualization and furniture arrangement by allowing users to experiment with design alternatives before implementing physical changes. These advances in VR design capabilities form the foundation for our research into human-virtual agent collaboration in immersive environments, focusing on how AI can enhance creative design processes in spatial layout tasks.

2.4 Co-design and Interaction with Virtual Agents

Co-design approaches incorporating AI aim to enhance creativity and improve the efficiency of the design process. Walsh and Wronsky [40] demonstrated how AI integration can create more inclusive design experiences that benefit diverse populations. Vartiainen et al. [39] demonstrated a co-design process using a text-to-image generative AI system, enabling users to produce visual content collaboratively. Zhang et al. [52] introduced a co-design system in which the AI agent offers scene layout suggestions based on the user's mouse position, functioning as an opinionated collaborator. Building on this, Zhang et al. [53] developed a framework using transformable modules to support the synthesis of multipurpose room layouts. Furthermore, Shaoi [28] used Large Language Model (LLM)-driven agents to allow users to design a landscape based on their text input. While these studies have advanced co-design in desktop-based environments, the role of AI in immersive VR settings, particularly when embodied through virtual agents, remains underexplored [34]. Defining the role of AI as a co-designer in VR is critical for understanding human-agent interactions in spatial and embodied contexts. Addressing this gap, our study investigates the dynamics of collaboration with an embodied virtual agent during an open-ended co-design task in VR.

2.5 Research Questions

We identified the following research questions to help us evaluate the proposed method and understand how users interact with a virtual agent during a co-design task in an immersive VR experience:

- **RQ1:** How does the virtual agent's co-design strategy impact participants' interaction and perception of the virtual agent (i.e., co-presence, attentional allocation, perceived collaboration, trust, performance)?
- **RQ2:** How does the virtual agent's co-design strategy impact participants' perceptions of the agent's awareness (i.e., private, public, surrounding)?
- **RQ3:** How does the virtual agent's co-design strategy impact participants' user experience (i.e., enjoyment, task load, system usability, frustration, desire for future interaction)?

- **RQ4:** How does the virtual agent's co-design strategy impact participants' final design satisfaction?
- **RQ5:** How does the virtual agent's co-design strategy impact the users' activity (i.e., time, items placed, items edited)?

2.6 Contributions

AI has demonstrated potential in predicting user intentions and facilitating collaboration, including in VR design contexts [8, 34]. However, existing studies often focus on narrowly defined tasks. Thus, there is a need to explore human-virtual agent interactions in more creative and open-ended scenarios. Our research focused on investigating layout design tasks involving virtual agents, advancing beyond constrained collaborative activities such as puzzle-solving [8, 9], which typically involve single-solution problems (e.g., the correct arrangement of puzzle pieces). Also, prior research highlights the effectiveness of VR for spatial and interior design tasks [3, 13, 15, 34, 44, 49, 53]. Building on these insights, we introduce a human-in-the-loop optimization approach for virtual agent decision-making during interior layout co-design in VR. The agent interacts with the environment by navigating and manipulating objects, adapting its behavior based on user input. This approach integrates real-time layout optimization in an immersive setting, enabling interactive and adaptive co-design experiences. Through a user study, we aim to contribute to the understanding of human-virtual agent collaboration in immersive environments by exploring an open-ended co-design task. We evaluate our method against alternative strategies to assess its impact on user engagement, collaboration quality, and design outcomes. We expect our results to support more natural and effective collaboration between humans and virtual agents in design.

3 Virtual Roomie Framework

We implemented a human-in-the-loop optimization-based method that enables the virtual agent to determine the placement of a randomly selected furniture item based on the current layout of the virtual living room environment. The agent evaluates spatial constraints and the user's previously placed items using a cost function to identify the selected object's most suitable position and rotation. During optimization, the algorithm considers multiple placement options and applies a greedy-based method to find an optimal solution. This optimization approach minimizes computational costs while maintaining coherence with user decisions. Thus, this approach ensures that furniture items are arranged to optimize spatial relationships, supporting a more structured and intelligent co-design process in real-time, as demonstrated in Figure 2.

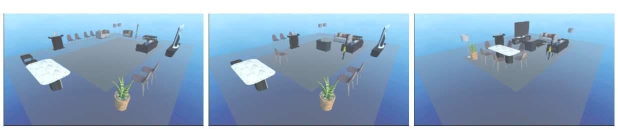


Figure 2: An example of an evolving layout design.

3.1 Preliminaries: Creating Example Layout Designs

We created a desktop-based application in Unity game engine that allows users to create layout configurations of a living room. We asked participants experienced in design to create example configurations of living room furniture to use later as targets in our optimization process (see Section 3.2). Specifically, we invited eight students (five male and three female; age: $M = 25.9$, $SD = 2.53$) from our department, all with backgrounds related to design (e.g., game design, level design). Before starting the task, participants answered questions about their design experience and preferences on a 7-point scale. On average, they reported high experience with design in their daily lives ($M = 5.88$, $SD = 1.73$), moderate professional design experience ($M = 4.75$, $SD = 1.49$), and considered design personally or professionally important ($M = 5.75$, $SD = 1.04$). Based on previously published work [1, 46], and given the variability among participants, we had considered eight participants sufficient to provide reliable data to inform our method.

(a) (b)

Figure 3: (a) The furniture items we used in our application. (b) The initial environment configuration with the furniture items in the outer area.

We instructed our participants to design a living room with 19 furniture items we later used in our user study (see Figure 3a). Once they finished, we saved the designed layouts and extracted information that characterizes their configuration, such as the position $\mathbf{p}_i^T$ and rotation $\mathbf{r}_i^T$ of each furniture item i in the living room environment, and the position $B_{Near}(\mathbf{p}_i^T)$ and rotation $B_{Near}(\mathbf{r}_i^T)$ of each furniture item relative to the nearest boundary. We also defined pairwise relationships between furniture items based on Zhang et al.'s [52] categorization of dominant/subordinate objects, where the dominant object serves as a "parent" and the subordinate is positioned and oriented relative to it. Based on this representation, we also extracted the relative position and rotation between paired furniture items, using the dominant furniture item as a reference point. In our framework, this data served as input to inform the cost terms in our total cost function (see Section 3.2).

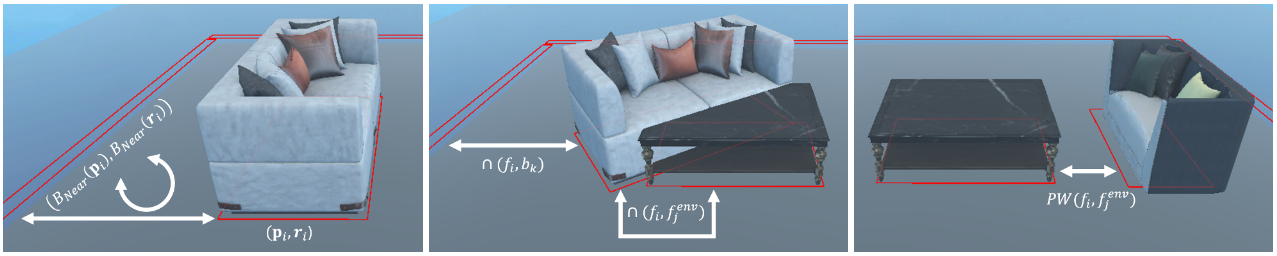
3.2 Virtual Agent's Decision Making

3.2.1 Problem Formulation

Let $F = \{f_1, f_2, ..., f_n\}$ denote a set of n furniture items. Each furniture item is represented as $f_i = \{\mathbf{p}_i, \mathbf{r}_i, \bar{\mathbf{p}}_i^T, \bar{\mathbf{r}}_i^T\}$, where $\mathbf{p}_i$ denotes the position and $\mathbf{r}_i$ denotes the rotation of a furniture item i, and $\bar{\mathbf{p}}_i^T$ denotes the target median position and $\bar{\mathbf{r}}_i^T$ denotes the target median rotation of a furniture item i. Note that the target values for the median position and median rotations were computed from the configurations resulting from our preliminary data collection (see Section 3.1). For the decision-making process of our virtual agent, we considered three costs in the total cost function (C_{Total}) to evaluate a particular design:

$$C_{Total} = w_{Boundary}C_{Boundary} + w_{Collision}C_{Collision} + w_{Pair}C_{Pair}, \quad (1)$$

where $C_{Boundary}$ denotes the boundary cost and $w_{Boundary}$ is its associated weight. $C_{Collision}$ refers the collisions cost and $w_{Collision}$ is its associated weight. Finally, C_{Pair} denotes the pairwise cost between furniture items, with w_{Pair} indicating its weight. The C_{Total} is defined for a selected furniture item f_i being tested in a particular position $\mathbf{p}_i$ and rotation $\mathbf{r}_i$ in the living room environment by following an exhaustive search (brute-force search) through a greedy algorithm; once it tests all the positions and rotations, it will select the minimum cost for the furniture item to be placed (see Section 3.2.5). The $w_{Boundary}$, $w_{Collision}$, and w_{Pair} weights $\in [0,1]$ control the contribution of each cost term in the total cost function. We graphically illustrate the cost terms in Figure 4.



(a) (b) (c)

Figure 4: The (a) boundary, (b) collision, and (c) pairwise costs of our total cost function.

3.2.2 Boundary Cost

For a selected object f_i, the boundary cost, $C_{Boundary}$, computes the alignment between the object's current position and rotation with its target configuration relative to the nearest boundary (see Figure 4a). The $C_{Boundary}$ is computed by:

$$C_{Boundary} = \frac{1}{L_P}||D(B_{Near}(\mathbf{p}_i), \mathbf{p}_i) - \underbrace{D(B_{Near}(\bar{\mathbf{p}}_i^T), \bar{\mathbf{p}}_i^T)}_{target}|| + \frac{1}{L_R}||D(B_{Near}(\mathbf{r}_i), \mathbf{r}_i) - \underbrace{D(B_{Near}(\bar{\mathbf{r}}_i^T), \bar{\mathbf{r}}_i^T)}_{target}||, \quad (2)$$

where $B_{Near}(\cdot)$ returns the relative position or rotation to the nearest boundary and $D(\cdot)$ computes the distance between two vectors. Last, L_P and L_R are normalization factors.

3.2.3 Collision Cost

The collision cost accounts for the degree of overlap between furniture items rather than a simple binary measure. The cost is calculated based on the proportion of an object's colliding area, ensuring more accurate placement evaluation. This cost considers two different types of collision: (1) with furniture items that have been already placed in the living room environment ($f_j^{env} \in \{f_1^{env}, ..., f_J^{env}\}$) and (2) compose the current configuration of the layout and with the boundary ($b_k \in \{b_1, ..., b_K\}$) of the living room (see Figure 4b). We compute the collision cost as follows:

$$C_{Collision} = \sum_{j=1}^{J}\left(\frac{\cap(f_i, f_j^{env})}{A(f_i)}\right) + \sum_{k=1}^{K}\left(\frac{\cap(f_i, b_k)}{A(f_i)}\right), \quad (3)$$

where $\cap(f_i, f_j^{env})$ and $\cap(f_i, b_k)$ calculate the overlapping area between two convex polygons (i.e., between two furniture f_i and f_j^{env}, and between a furniture f_i and a boundary b_k, respectively) using the Sutherland-Hodgman clipping algorithm.[1] Last, the $A(f_i)$ is a function that returns the total areas the furniture f_i occupies.

3.2.4 Pairwise Cost

The pairwise cost considers the relationships between furniture items in the scene, based on a dominant/subordinate hierarchy. For this cost term, we adopted the relationship model proposed by Zhang et al. [52]. Thus, based on our preliminary study (see Section 3.1), we computed the target relative position and rotations. Specifically, given the layouts from our preliminary study, we applied k-means clustering to the observed relative positions and rotations of each dominant-subordinate furniture pair. For each resulting cluster, we computed the median relative position ($\bar{\mathbf{p}}_{ij}^T$) and rotation ($\bar{\mathbf{r}}_{ij}^T$) to identify representative spatial configurations while reducing the influence of outliers in terms of pair relations. This clustering process allowed us to capture common layout patterns from the preliminary designs, which inform the target configurations. Then, we defined a relationship $rl_{ij} = \{\langle \bar{\mathbf{p}}_{ij}^T, \bar{\mathbf{r}}_{ij}^T \rangle, \cdots\}$, as a set of cluster median position and rotation, where f_i is the dominant item and f_j the subordinate one. For each dominant item f_i, we defined a set $RL_i = \{rl_{ij}, rl_{il}, \cdots\}$, which includes all spatial relationships that f_i has with other subordinate items.

Once the virtual agent decides to pick the furniture item f_i, it will check the minimal distance between that furniture item and its dominant/subordinate counterpart in the scene (see Figure 4c). So, the pairwise cost is calculated as:

$$C_{Pair} = \frac{1}{|RL_i|}\sum_{j=1}^{J} PW(f_i, f_j^{env}), \quad (4)$$

where $PW(f_i, f_j^{env})$ represents the cost of the pairwise relationship between f_i and f_j^{env}. PW considers the distance difference between the furniture item position $\mathbf{p}_i$ and any of the target position $\bar{\mathbf{p}}_{ij}^T$ in reference to the related furniture item position $\mathbf{p}_j$ when f_j^{env} is part of relationship RL_i. For the rotation component, we calculate the difference between the current rotation of the furniture item and its ideal rotation relative to the related furniture item. The PW function is defined by the following statement:

$$PW(f_i, f_j^{env}) = \begin{cases} \frac{1}{L_P}D(\mathbf{p}_i, \mathbf{p}_j + \bar{\mathbf{p}}_{ij}^T) & \text{if } RL_i \text{ in relation to } f_j^{env} \\ + \frac{1}{L_R}D(\mathbf{r}_i, \mathbf{r}_j + \bar{\mathbf{r}}_{ij}^T) & \\ 0 & \text{otherwise.} \end{cases}$$

3.2.5 Optimization

Our system updates the virtual agent's decisions by optimizing the placement of furniture items according to our total cost functions. This optimization process takes the cost terms along with their respective weights as input. Rather than using complex sampling techniques, we implemented a simple grid-based approach that balances computational performance with solution quality. The algorithm evaluates 100 positions distributed uniformly across a regular grid spanning the living room floor, with rotation angles considered at 45-degree intervals: $[0°, 45°, 90°, ..., 315°]$. Our controlled testing showed that the regular grid with 100 samples provided sufficient space coverage to generate reasonable furniture arrangements while maintaining real-time responsiveness. By calculating the total cost for each position-rotation combination, the system identified the option with the lowest cost as the optimal placement decision, following the optimization pipeline shown in Figure 1. Last, we should note that we assigned weights to specify the influence of each cost on the total cost function. The decision on the weights was based on a trial-and-error evaluation. We assigned

[1] https://rosettacode.org/wiki/Sutherland-Hodgman_polygon_clipping

$w_{Boundary} = .20$, $w_{Collision} = .80$, and $w_{Pair} = 1.00$ to prioritize avoiding furniture item overlap and emphasize the correct spatial arrangement between dominant/subordinate furniture.

3.3 Virtual Reality Application

Virtual Environment. The virtual environment is divided into two sections: an inner area that represents the living room floor, surrounded by transparent boundaries where the user and virtual agent place furniture, and an outer semi-transparent area that initially holds all furniture items, along with the virtual agent and user (see Figure 3b). The goal is for the virtual agent and user to co-design the living room layout by placing each furniture item from the outer area into the inner one. The inner area measures 7×12 meters. We selected 19 living room-related furniture items to create a cozy layout (see Figure 3a). These items come from Zhang et al.'s [52] 3D scene dataset.

Virtual Agent. The virtual agent is a 3D animated character that navigates the environment using the A* path-finding algorithm [21] (see Figure 5). It can perform two actions: (1) select an object and (2) walk to its destination. After making a decision, the agent moves toward the selected furniture item, grabs it, and then walks toward the computed item's position in the inner area to place it, updating the layout. Once the item is placed, the agent returns to its waiting position.

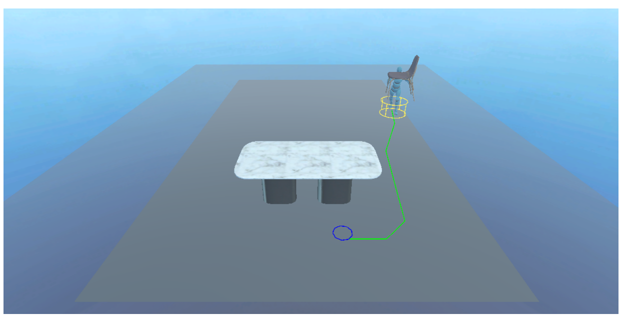
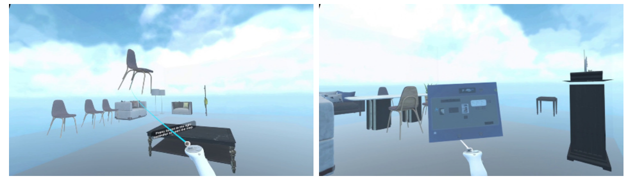


Figure 5: The virtual agent is moving toward placing a furniture item. The green line represents a collision-free path computed using the A* algorithm, and the blue circle represents the target position for the furniture item.

Interaction Mechanism. Users could grab and release furniture items and navigate through the scene. They could manipulate the position of furniture items via distance grabbing using a raycast-based interaction (see Figure 6a). For locomotion, users could teleport or use continuous movement through the controller's joystick. To enhance spatial awareness during the design process, we implemented a top-view map on the user's wrist. This minimap offered a real-time overview of the living room layout, showing the current furniture positions, the user, and the virtual agent (see Figure 6b). This feature addressed a common challenge in VR design where users have a limited field of view while immersed [12].

We established a turn-taking approach between the user and the virtual agent for the co-design task (see Figure 1). During the user's turn, the agent remained stationary in a predefined waiting position. The user could insert a new furniture item or edit an existing one. Once the item was placed in the inner area and no further adjustments were made within a timeout period of five seconds, the user's turn concluded. During the agent's turn, object manipulation was disabled for the user, who could only observe the agent's actions. The virtual agent, following a predefined co-design strategy (see Section 4.3 for experimental conditions), selected and placed a furniture item in the living room. Upon completing its action, the agent returned to the waiting position, allowing the user's turn. This turn-taking process continued until all furniture items had been placed.



(a) (b)

Figure 6: (a) The user grabs a furniture item. (b) The top-view map feature allows for visualization of the living room environment.

4 User Study

4.1 Participants

We calculated the sample size based on an *a priori* power analysis with G*Power v. 3.1 software [14]. For our within-group study, which includes three conditions (i.e., random, template, and optimization), with an effect size $f = .30$, a nonsphericity correction $\varepsilon = .80$, and an $\alpha = .05$, to achieve an 80% power ($1 - \beta$ error probability), the analysis recommended at least 23 participants. We recruited 24 participants (13 male and 11 female; age: $M = 27.25$, $SD = 3.35$) by sending emails and distributing flyers on our university campus.

4.2 Apparatus

We used Unity game engine version 2021.3.34f1 to develop the project for the Meta Quest 2 head-mounted display (HMD). We used an Alienware computer (Intel i7, 32GB RAM, and NVIDIA GeForce RTX 2080) for our implementation and user study.

4.3 Experimental Conditions

We evaluated the proposed optimization against two other co-design strategies described below:

- **Random (RC):** Regardless of where the user positioned the furniture, the virtual agent placed furniture items by randomly selecting the furniture item and its target position in the virtual environment. We included this condition to establish a baseline for agent behavior without spatial reasoning or user input responsiveness.

- **Template (TC):** The virtual agent selected furniture items randomly but placed them according to a predefined layout (see supplementary material document) created by an experienced designer. While users could freely arrange items, the agent did not adapt to these changes and continued following the structured placement pattern. This approach ensured consistent item positioning and allowed us to assess the impact of predefined layouts on collaboration, independent of user-informed design changes.

- **Optimization (OC):** The virtual agent selected furniture items randomly but determined their placement based on the proposed total cost function. Thus, the virtual agent dynamically calculated the best position for each furniture item. This approach allowed the virtual agent to adjust its placement decisions in response to the evolving design.

4.4 Ratings and Measurements

Self-reported Ratings. We collected data from several scales to evaluate the human-virtual agent co-design process. Co-presence and attentional allocation were measured using scales from Biocca et al. [4]. Subjective mental workload was assessed through NASA's task load index (TLX) [22]. We also employed the system usability scale (SUS) [5] to evaluate the system's usability across different conditions. To assess the virtual agent's awareness (i.e., private, public, and surrounding), we adapted scales from Govern and Marsch [19]. For collaboration, we used the perceived collaboration scale by Liu et al. [27], the trust scale by Jian et al. [23], and a perceived contribution measure inspired by Choi et al. [7]. Finally, we measured performance, enjoyment, frustration, and the desire for future interaction using items based on Choi et al. [9].

Application Logs. We recorded interaction data from the VR experience to gain deeper insight into user activity during the co-design task. Specifically, the logged data included: (1) the total time spent in the experience (measured in seconds), calculated from the beginning of furniture placement to the moment users indicated completion; (2) the number of furniture items placed by the user; (3) the total number of edits performed by the user, where an edit is defined as repositioning a furniture item that had already been placed in the inner area; and (4) the total number of edits the user made to items initially placed by the virtual agent. We used this last metric to examine how users responded to the agent's contributions and to determine whether their intended design aligned with the agent's placements.

4.5 Procedure

Each study session included one researcher and one participant. The researcher presented the study's consent form, which our university's Institutional Review Board (IRB) approved. This form provided detailed information about the study, essential points to consider, and participants' rights. We informed each participant that they could report any discomfort they experienced with the VR headset and were free to pause, take a break, or withdraw from the study at any time without consequences. At the start of the study, the researcher asked the participants to complete a demographic questionnaire on the computer. Following this, the researcher explained the study procedure to help the participants understand the experiment and provided instructions on using the HMD. Participants then completed a VR warm-up tutorial, where they practiced interacting with the environment and moving furniture items using the VR controllers. Prior research suggested that tutorials on VR controllers enhanced users' performance, experience, and intrinsic motivation [24]. We exposed participants to all three conditions in an order determined by the Latin square ordering method [45] that balances first-order carry-over (residual) effects across conditions. In each condition, participants co-designed a living room layout with the virtual agent. After each condition, the researcher asked participants to complete a survey on the computer. At the end of the study, the researcher dismissed the participants. Each session lasted no more than one hour.

5 RESULTS

For our statistical analyses, we used the three experimental conditions as independent variables, and the self-reported ratings and logged data as dependent variables. The Q-Q plots of the residuals and the Shapiro-Wilk test at the 5% level confirmed the normality of the collected data. Thus, we performed one-way repeated measures analysis of variance (ANOVA). We used Bonferroni-corrected estimates for pairwise comparisons to assess the statistically significant ($p < .05$) results. We summarize our results in Table 1 and Table 2.

5.1 Self-reported Ratings

5.1.1 Interaction and Perception of the Virtual Agent

Co-presence. We found a statistically significant result (Wilk's $\Lambda = .651$, $F[2,22] = 5.910$, $p = .009$, $\eta_p^2 = .349$). The post hoc pairwise comparison showed that participants rated their co-presence lower when we exposed them to the random ($M = 3.83$, $SD = .88$) than the optimization ($M = 4.81$, $SD = 1.05$) condition at $p = .006$.

Attentional Allocation. We did not find a statistically significant result (Wilk's $\Lambda = .964$, $F[2,22] = .406$, $p = .671$, $\eta_p^2 = .036$).

Perceived Collaboration. We found a statistically significant result (Wilk's $\Lambda = .298$, $F[2,22] = 25.934$, $p < .001$, $\eta_p^2 = .702$). The post hoc pairwise comparison showed that participants rated their perceived collaboration higher in the optimization ($M = 5.11$, $SD = 1.03$) than the template ($M = 3.99$, $SD = 1.46$; $p = .020$) and random ($M = 3.11$, $SD = 1.32$; $p < .001$) conditions.

Trust. We found a statistically significant result (Wilk's $\Lambda = .281$, $F[2,22] = 28.183$, $p < .001$, $\eta_p^2 = .719$). The post hoc pairwise comparisons showed that participants rated their trust in the random ($M = 2.67$, $SD = .64$) lower than in the optimization ($M = 3.79$, $SD = .85$; $p < .001$) and template ($M = 3.66$, $SD = .98$; $p < .001$) conditions.

Performance. We found a statistically significant result (Wilk's $\Lambda = .316$, $F[2,22] = 23.766$, $p < .001$, $\eta_p^2 = .684$). The pairwise comparison showed that participants reported better performed in the optimization ($M = 4.88$, $SD = 1.60$) than random ($M = 2.08$, $SD = 1.50$; $p < .001$) condition, and better in the template ($M = 4.21$, $SD = 1.93$) than the random ($p < .001$) condition.

5.1.2 Virtual Agent Awareness

Private Awareness. The analysis revealed a statistically significant result (Wilk's $\Lambda = .384$, $F[2,22] = 17.648$, $p < .001$, $\eta_p^2 = .616$). The post hoc pairwise comparisons showed that participants rated the virtual agent's private awareness significantly higher in the optimization ($M = 4.40$, $SD = 1.47$) than in the template ($M = 3.29$, $SD = 1.42$; $p = .023$) and random ($M = 2.23$, $SD = 1.24$; $p < .001$) conditions. Participants also rated the template significantly higher than the random ($p = .022$) condition.

Public Awareness. The analysis revealed a statistically significant result (Wilk's $\Lambda = .308$, $F[2,22] = 24.754$, $p < .001$, $\eta_p^2 = .692$). The post hoc pairwise comparison showed that participants rated the virtual agent's public awareness significantly higher in the optimization ($M = 3.98$, $SD = 1.65$) than in the template ($M = 2.88$, $SD = 1.50$; $p = .048$) and random ($M = 1.66$, $SD = .68$; $p < .001$) conditions. Participants also rated the virtual agent's public awareness significantly higher in the template than in the random ($p = .004$) condition.

Surrounding Awareness. The analysis revealed a statistically significant result (Wilk's $\Lambda = .246$, $F[2,22] = 33.627$, $p < .001$, $\eta_p^2 = .754$). The post hoc pairwise comparison showed that participants rated the virtual agent's surrounding awareness significantly lower in the random ($M = 2.10$, $SD = 1.12$) than in the optimization ($M = 4.65$, $SD = 1.48$; $p < .001$) and template ($M = 3.81$, $SD = 1.61$; $p < .001$) conditions.

5.1.3 User Experience

NASA-TLX. The analysis revealed a statistically significant result (Wilk's $\Lambda = .731$, $F[2,22] = 4.055$, $p = .032$, $\eta_p^2 = .269$). Post hoc analysis showed that participants reported a lower subjective mental workload in the optimization ($M = 26.42$, $SD = 8.42$) than in the random ($M = 34.99$, $SD = 14.53$) condition.

SUS. The analysis revealed a statistically significant result (Wilk's $\Lambda = .668$, $F[2,22] = 5.465$, $p = .012$, $\eta_p^2 = .332$). The post hoc pairwise comparison showed that participants rated the optimization ($M = 77.71$, $SD = 13.59$) significantly higher than the random ($M = 58.85$, $SD = 21.30$; $p = .008$) condition.

Table 1: Detailed results of our study for the self-reported ratings (we present significant results with bold font).

	(1)		(2)		(3))		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)		(13)		(14)	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RC	3.83	.88	3.4	.89	3.11	1.32	2.67	.60	2.08	1.5	2.23	1.24	1.66	.68	2.1	1.48	34.99	14.53	58.85	21.3	2.67	1.86	4.29	2.07	2.71	1.81	3.38	2.08
TC	4.44	.99	3.28	.54	3.99	1.45	3.66	.98	4.21	1.93	3.29	1.42	2.88	1.5	3.81	1.61	27.6	11.26	70.21	15.57	4.54	1.77	3.04	1.81	4.29	1.65	5.59	1
OC	4.81	1.05	3.24	.55	5.11	1.03	3.79	.85	4.88	1.6	4.4	1.47	3.98	1.65	4.65	1.12	26.42	8.42	77.71	13.59	5.21	1.47	2.96	1.4	4.96	1.55	5.85	.90
F		**5.910**		.406		**25.934**		**28.183**		**23.766**		**17.648**		**24.754**		**33.627**		**4.055**		**5.465**		**15.478**		3.328		**11.919**		**12.81**
p		**.009**		.671		**<.001**		**<.001**		**<.001**		**<.001**		**<.001**		**<.001**		**.032**		**.012**		**<.001**		.055		**<.001**		**<.001**
η_p^2		**.349**		.036		**.702**		**.719**		**.684**		**.616**		**.692**		**.754**		**.296**		**.332**		**.585**		.232		**.52**		**.538**

$df = 2$, error $df = 22$

(1) Co-presence, (2) Attention, (3) Perceived Collaboration, (4) Trust, (5) Performance, (6) Private Awareness, (7) Public Awareness, (8) Surrounding Awareness, (9) TLX, (10) SUS, (11) Enjoyment, (12) Frustration, (13) Desire for Future Interaction, and (14) Reasonable Design.

Enjoyment. The analysis revealed a statistically significant result (Wilk's $\Lambda = .415$, $F[2,22] = 15.478$, $p < .001$, $\eta_p^2 = .585$). The post hoc pairwise comparisons showed that participants rated their enjoyment significantly higher in the optimization ($M = 5.21$, $SD = 1.47$) than in the random ($M = 2.67$, $SD = 1.86$; $p < .001$) condition and in the template ($M = 4.54$, $SD = 1.77$) than in the random ($p = .004$) condition.

Frustration. The analysis did not yield a statistically significant result (Wilk's $\Lambda = .768$, $F[2,22] = 3.328$, $p = .055$, $\eta_p^2 = .232$). Although the optimization ($M = 2.96$, $SD = 1.40$) and template ($M = 3.04$, $SD = 1.81$) conditions had lower frustration scores compared to the random ($M = 4.29$, $SD = 2.07$) condition.

Desire for Future Interaction. The analysis revealed a statistically significant result (Wilk's $\Lambda = .480$, $F[2,22] = 11.919$, $p < .001$, $\eta_p^2 = .520$). The post hoc comparison showed that participants expressed more desire for future interaction in the optimization ($M = 4.96$, $SD = 1.55$) than in the random ($M = 2.71$, $SD = 1.81$; $p < .001$) condition. Moreover, participants reported significantly higher desire for future interaction in the template ($M = 4.29$, $SD = 1.65$) than in the random ($p = .005$) condition.

5.1.4 Design Satisfaction

Reasonable Design. The analysis revealed a statistically significant result (Wilk's $\Lambda = .462$, $F[2,22] = 12.810$, $p < .001$, $\eta_p^2 = .538$). The post hoc pairwise comparisons showed that participants rated the random ($M = 3.38$, $SD = 2.08$) significantly less reasonable than both the optimization ($M = 5.85$, $SD = .90$; $p < .001$) and the template ($M = 5.59$, $SD = 1.00$; $p < .001$) conditions.

5.2 Application Logs

Table 2: Detailed results of our study for the application logs (we present significant results with bold font).

	(1)		(2)		(3)		(4)	
	M	SD	M	SD	M	SD	M	SD
RC	404.14	168.11	5.42	2.50	7.94	5.11	7.17	5.09
TC	387.75	141.90	6.67	2.16	5.29	4.35	2.33	2.04
OC	371.37	152.99	7.25	1.45	4.13	3.10	2.29	1.73
F		.466		**5.035**		**4.888**		**10.459**
p		.634		**.016**		**.018**		**<.001**
η_p^2		.041		**.314**		**.308**		**.487**

$df = 2$, error $df = 22$

(1) Time, (2) Items Placed by User, (3) Corrections (All Items Edited by User), and (4) Corrections to Virtual Agent (Virtual Agent Items Edited by User).

Time. The analysis did not reveal a statistically significant result (Wilk's $\Lambda = .959$, $F[2,22] = .466$, $p = .634$, $\eta_p^2 = .041$).

Items Placed by User. The analysis revealed a statistically significant result (Wilk's $\Lambda = .686$, $F[2,22] = 5.035$, $p = .016$, $\eta_p^2 = .314$). The post hoc pairwise comparisons showed that participants placed more items in the optimization ($M = 7.25$, $SD = 1.45$) than in the random ($M = 5.42$, $SD = 2.50$; $p = .014$) condition. They also placed more items in the template ($M = 6.67$, $SD = 2.16$) than in the random ($p = .014$) condition.

Corrections (All Items Edited by User). The analysis revealed a statistically significant result (Wilk's $\Lambda = .692$, $F[2,22] = 4.888$, $p = .018$, $\eta_p^2 = .308$). The post hoc pairwise comparison showed that participants edited significantly more items in the random ($M = 7.94$, $SD = 5.11$) than in the optimization ($M = 4.13$, $SD = 3.10$; $p = .014$) condition.

Corrections to Virtual Agent (Virtual Agent Items Edited by User). The analysis revealed a statistically significant result (Wilk's $\Lambda = .513$, $F[2,22] = 10.459$, $p < .001$, $\eta_p^2 = .487$). The post hoc pairwise comparisons showed that participants edited significantly more of the virtual agent's placements in the random condition ($M = 7.17$, $SD = 5.09$) compared to both the optimization condition ($M = 2.29$, $SD = 1.73$; $p < .001$) and the template condition ($M = 2.33$, $SD = 2.04$; $p < .001$).

5.3 Qualitative Data

To complement our findings, we also included an optional open-ended question to collect participants' impressions of the overall experience. Eight participants commented on the varying levels of the virtual agent's intelligence and responsiveness they observed. One participant (P16) noted, "*In some conditions, I can see the virtual agent is aware of my actions and responds to them reasonably.*"

Six participants expressed awareness of the different virtual agent behaviors, with one (P4) remarking: "*I felt that the last AI was aware that it was messing things up...*" Another (P7) stated: "*Some individuals were not the most intelligent, and I felt like I was doing all the work,*" suggesting they perceived varying levels of collaborative engagement across conditions.

Seven participants described the VR interaction in favorable terms, using words like "*fun*" (P2 and P17), "*interesting*" (P6), and "*enjoyable*" (P14). One participant (P19) explicitly ranked their experiences numerically, stating: "*From best to worst system, I would rate them as following: 2* [optimization]*, 1* [template]*, and 3* [random]" suggesting apparent preference differences among the conditions they encountered.

One participant (P18) commented on the virtual agent's communication limitations: "*It's hard to interpret what the virtual character wants to do by just watching,*" suggesting that greater transparency in the virtual agent's decisions could improve collaboration. Last, another (P22) appreciated the spatial awareness aspects, stating: "*I like that I could see the virtual agent through the map. It was easy to capture where it is in the scene.*"

6 Discussion

6.1 RQ1: Interaction and Perception of the Virtual Agent

We examined how co-design strategies impact participants' co-presence, attentional allocation, perceived collaboration, trust, and performance. We found that structured decision-making improved both interaction and perceptions of the agent.

Co-presence was significantly higher in the optimization condition compared to the random condition. When the virtual agent made decisions based on the layout configurations, participants perceived it as more attuned to the environment. They engaged in the interaction, reinforcing the impression of working with an intentional collaborator. The template condition showed intermediate co-presence ratings, suggesting that behavioral consistency contributes to co-presence but not as strongly as adaptive responsiveness. Thus, our results align with Choi et al. [8] who showed that virtual agent intelligence enhances the sense of co-presence during human-virtual agent interaction.

Attentional allocation showed no significant differences across conditions. Interestingly, participants reported a low rating in this metric, indicating that participants focused more on the design task than the virtual agent. Several factors may explain this: the agent was not programmed to follow the user continuously, so it was not always in view; the turn-taking structure naturally divided attention regardless of agent behavior; and the complex spatial nature of the task likely demanded significant cognitive resources. This aligns with Biocca et al. [4], who found that task complexity can redirect attention away from social partners. Future research could employ eye-tracking to understand attentional patterns during collaborative design tasks better.

Perceived collaboration showed that the optimization approach was rated significantly higher than the template and random co-design strategies. This result aligns with Zhou et al. [56], who found that nonlinear collaborative frameworks treating AI as an opinionated collaborator rather than a mere tool significantly enhance user engagement in creative design tasks. The human-in-the-loop nature of the optimization condition exemplifies this balance: the virtual agent makes context-aware design decisions based on the user's choices, contributing meaningfully to the evolving layout. Unlike the template co-design strategy, which follows predetermined patterns regardless of user input, or random, the optimization co-design strategy of the virtual agent demonstrates a "design intention inference" [26], which is the ability to understand and complement human design goals. This decision-making is the critical factor that elevates perceived collaboration in the optimization condition, enabling an effective collaborative interaction in VR layout design tasks.

Participants reported higher trust in the optimization condition than in the random one. The template condition also received significantly higher trust ratings than the random condition. This supports Merritt et al.'s [31] findings on trust development with AI teammates, where predictability plays a crucial role. Virtual agents exhibiting purposeful design (i.e., template and optimization condition) behavior were perceived as more reliable partners, suggesting that apparent intentionality is fundamental for establishing trust during collaborative design.

Performance ratings followed a similar pattern, with participants rating the agent's performance significantly higher in the optimization than in the random condition, and the template condition was also rated significantly higher than the random one. This differential impact of virtual agent behavior on performance perception aligns with previous research that reported higher user confidence when AI acted as an active collaborator [34, 52]. The optimization-based virtual agent created a sense of complementary collaboration that enhanced perceived performance beyond what either the template or random co-design strategies could achieve.

6.2 RQ2: Virtual Agent's Awareness

We also explored how co-design strategies impact perceptions of virtual agent awareness. Participants attributed different awareness levels to the virtual agent across conditions.

Private awareness ratings were significantly higher in the optimization condition than in both the template and random conditions. Participants also rated the template condition significantly higher than the random one, suggesting that agents demonstrating reasonable decision-making were perceived as more self-aware. These findings align with Cerekovic et al. [6], who reported that nonverbal cues and behavioral consistency shape how users infer an agent's internal states. Public awareness ratings showed the most pronounced differences across conditions. The optimization condition received the highest ratings, significantly outperforming the random strategy, with the template strategy also rated higher than random. These results indicate that an agent's ability to place furniture appropriately strongly influences how users perceive its understanding of the environment, supporting findings by Ye et al. [48] regarding situational awareness and environmental responsiveness.

Finally, participants rated the random condition significantly lower in terms of surrounding awareness compared to the optimization and template conditions. As expected, users perceived the agents as unaware of their surroundings when they made unreasonable or random placements. This finding aligns with previous research showing that participants perceived virtual agents as more aware when they completed tasks correctly [9].

6.3 RQ3: User Experience

We also examined how co-design strategies impact user experience. Enjoyment, frustration, willingness for future interaction, task load, and design evaluation all showed significant differences favoring structured (i.e., template or optimized) virtual agent behavior.

Enjoyment ratings revealed a clear distinction between conditions, with both the optimization and template conditions rated significantly more enjoyable than the random condition. This pattern suggests that enjoyment increases with behavioral structure and user input responsiveness. This finding supports research by Walsh and Wronsky [40], who found that adaptive AI integration creates more engaging and satisfying design experiences. The optimization-based virtual agent's ability to respond to the evolving design context induced a more dynamic and rewarding collaborative experience than less responsive approaches.

Task load revealed significantly lower subjective mental workload in the optimization condition compared to the random condition, with the template condition showing intermediate workload levels not significantly different from either extreme. This reduced workload can be explained by the fact that predictable and supportive virtual agent behavior minimizes the mental workload resources users must allocate to managing the virtual agent interactions [32]. Similarly, Daronnat et al. [11] found that virtual agents with predictable behaviors improved performance and reduced cognitive load in real-time collaborative tasks.

System usability ratings were significantly higher for the optimization condition than the random condition, with the template condition receiving intermediate ratings not significantly different from either extreme. Optimizing-based decision-making induced more intuitive interactions, making the system feel more natural and easier to work with than the template or random co-design strategies. This aligns with You et al. [49], who found that responsive design tools significantly enhance usability perceptions in VR.

According to our qualitative analysis, several participants expressed difficulty interpreting the virtual agent's intentions, suggesting that making the decision-making process more transparent might enhance collaboration. Furthermore, our findings also indicated that participants valued the spatial awareness features in the interface, which helped them track the virtual agent's position throughout the collaborative design process.

6.4 RQ4: Design Satisfaction

Ratings on design satisfaction were significantly higher for the optimization and template conditions compared to the random condition. This finding suggests that structured virtual agent behavior, whether adaptive or predetermined, leads to more satisfactory design outcomes. This supports research by Yu et al. [50] on automated furniture layout, demonstrating that algorithmic approaches considering spatial relationships produce more acceptable designs. The optimization condition's slightly higher ratings suggest potential advantages for adaptive strategies in achieving design coherence, since structured approaches yielded satisfactory results.

6.5 RQ5: User Activity

We also examined how co-design strategies influence user activity. The logged data revealed distinct interaction patterns that objectively show how different co-design strategies affect collaboration.

Time spent in the environment showed no significant differences across conditions, suggesting that overall task completion time remained consistent regardless of virtual agent behavior. This indicates that the virtual agent's co-design strategy primarily affected the quality rather than duration of the interaction, as VR's spatial affordances for layout design were similar across conditions [15].

Participants placed more items in the optimization and template conditions than in the random condition. This pattern suggests that working with purposefully behaving virtual agents encourages more item additions from users. This aligns with research showing that intentional virtual agents boost human engagement in co-design tasks [35]. This relates to user corrections, which shows that participants made significantly fewer overall corrections in the optimization condition than in the random condition, with the template condition showing an intermediate number not significantly different from either extreme. Participants also modified fewer of the agent's placements in both the optimization and template conditions compared to the random one. These outcomes suggest that purpose-driven agent behaviors yielded more acceptable contributions, reducing the need for user intervention. This aligns with findings from Zhou et al. [56], showing that adaptive virtual agents contribute more effectively and require fewer adjustments.

Interestingly, the number of corrections to agent-placed items was similar in the optimization and template conditions, indicating that both strategies led to comparably acceptable outputs. However, the lower overall correction count in the optimization condition suggests that this strategy created a more coherent and satisfactory layout. Participants made fewer edits overall, suggesting that the optimization strategy supported a more balanced co-design process in which both the user and virtual agent made meaningful contributions to the intended layout. This reduction in participant corrections may indicate smoother collaboration with the agent during the co-design task, as fewer adjustments imply the agent's choices were more closely aligned with the user's design intentions.

6.6 Implications

Our findings offer several implications for the design of collaborative virtual agents in VR design environments. First, the positive user responses to the optimization-based virtual agent suggest that adaptive behavior can meaningfully enhance collaboration when grounded in a coherent and transparent design rationale. This highlights the need for virtual agents to exhibit algorithmic sophistication in a way that aligns with the user's creative intent, thereby supporting a shared understanding and trust. Second, the consistent patterns of attentional allocation across conditions point to the value of structured interaction frameworks, such as turn-taking, in facilitating collaborative engagement. Structured protocols help establish clear expectations and coordination between human users and virtual agents, even when agent behavior varies. Third, although agent support improved the overall experience, users continued to adjust the agent's contributions. This indicates that systems should be designed to balance automation with user control, allowing for refinement or override of agent actions without breaking the collaborative flow. Finally, the strong relationship between early agent behavior and user trust highlights the importance of first impressions in human-agent collaboration. Designers should prioritize reliable, competent early interactions to establish a foundation for continued engagement and cooperation.

6.7 Limitations

Our study had several limitations that should be considered when interpreting the results. However, these limitations do not invalidate our findings but rather highlight important directions for future research. First, our simplified grid-based optimization approach constrained the virtual agent's decision-making, which evaluated only 100 positions distributed uniformly across the living room floor, with rotation angles of 45-degree intervals. This discrete sampling method limits the problem's search space, reducing the virtual agent's ability to discover the optimal furniture item placement. While this approach was computationally efficient and maintained real-time performance in VR, it could have limited our framework's results. Second, while our virtual agent worked and helped us understand collaborative interactions, other decision-making methods could result in a more adaptive and responsive virtual agent. Third, our study utilized a single-viewing perspective, where participants interacted with the environment at a first-person view level. This limited viewpoint may affect how users perceive the environment and make spatial decisions during the collaborative task. Finally, our study focused on a living room layout task with specific furniture items and constraints. This setting may not generalize to other design contexts that involve different spatial relationships, aesthetic considerations, or functional requirements.

7 Conclusions and Future Work

We proposed a human-in-the-loop optimization-based method for virtual agent decision-making during layout co-design in VR. A user study compared our proposed optimization method against random and template co-design strategies and demonstrated that the optimization approach significantly improved perceived collaboration, trust, and co-presence compared to the random condition. Moreover, the virtual agent's ability to adapt to user actions and improvise during the evolving design appears to have created stronger impressions of awareness and intentionality. The optimization-based approach also reduced subjective mental workload while increasing enjoyment and satisfaction with the final design. Logged data showed that participants made fewer corrections to optimization-based virtual agent placements than those in the random condition. Based on qualitative feedback, participants noted varying levels of virtual agent awareness and responsiveness aligned with our experimental conditions. They described more positive collaborative experiences with virtual agents demonstrating purposeful behavior and responsiveness to user actions.

In future work, we argue that more advanced methods incorporating intention recognition should be explored to improve the virtual agent's decision-making. By understanding the user's underlying intentions, the system could better guide the agent in making design choices aligned with user goals. Furthermore, identifying user perception of the agent's role in the design (e.g., leading or subordinate) can be relevant for human-agent collaborations. Additionally, integrating LLMs could enable natural language communication between the user and the virtual agent, allowing them to discuss and decide on design decisions collaboratively.

REFERENCES

[1] P. Acevedo, M. Choi, H. Liu, D. Kao, and C. Mousas. Procedural game level design to trigger spatial exploration. In *Proceedings of the 17th International Conference on the Foundations of Digital Games*, FDG '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3555858.3563272 3

[2] A. Arntz, A. Di Dia, T. Riebner, C. Straßmann, and S. C. Eimler. Teamwork makes the dream work: A virtual reality-based human-robot collaboration sandbox simulating multiple teams. In *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, pp. 335–339, 2024. doi: 10.1109/AIxVR59861.2024.00057 1

[3] R. Arora, R. H. Kazi, F. Anderson, T. Grossman, K. Singh, and G. Fitzmaurice. Experimental evaluation of sketching on surfaces in vr. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, p. 5643–5654. ACM, May 2017. doi: 10.1145/3025453.3025474 2, 3

[4] F. Biocca, C. Harms, and J. Gregg. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA*, pp. 1–9, 2001. 6, 8

[5] J. Brooke and Others. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996. 6

[6] A. Cerekovic, O. Aran, and D. Gatica-Perez. *How Do You Like Your Virtual Agent?: Human-Agent Interaction Experience through Nonverbal Features and Personality Traits*, p. 1–15. Springer International Publishing, 2014. doi: 10.1007/978-3-319-11839-0_1 2, 8

[7] M. Choi, D. Cui, A. Koilias, and C. Mousas. The effects of virtual character's intelligence and task's complexity during an immersive jigsaw puzzle co-solving task. In *The 17th ACM SIGGRAPH Conference on Motion, Interaction, and Games*, MIG '24, p. 1–12. ACM, Nov. 2024. doi: 10.1145/3677388.3696324 6

[8] M. Choi, D. Cui, M. Volonte, A. Koilias, D. Kao, and C. Mousas. Toward understanding the effects of intelligence of a virtual character during an immersive jigsaw puzzle co-solving task. *ACM Transactions on Applied Perception*, 22(2):1–28, Jan. 2025. doi: 10.1145/3700822 1, 2, 3, 8

[9] M. Choi, S. Guo, A. Koilias, M. Volonte, D. Kao, and C. Mousas. Exploring the effects of self-correction behavior of an intelligent virtual character during a jigsaw puzzle co-solving task. *ACM Transactions on Interactive Intelligent Systems*, 14(3):1–33, Sept. 2024. doi: 10.1145/3688006 2, 3, 6, 8

[10] S. Daronnat, L. Azzopardi, and M. Halvey. Impact of agents' errors on performance, reliance and trust in human-agent collaboration. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64(1):405–409, Dec. 2020. doi: 10.1177/1071181320641091 2

[11] S. Daronnat, L. Azzopardi, M. Halvey, and M. Dubiel. Inferring trust from users' behaviours; agents' predictability positively affects trust, task performance and cognitive load in human-agent real-time collaboration. *Frontiers in Robotics and AI*, 8, July 2021. doi: 10.3389/frobt.2021.642201 2, 8

[12] G. D. Domenico, F. D. C. Becker, B. B. Schirmer, N. L. P. Berwaldt, G. M. d. Freitas, A. C. Neto, C. T. Pozzer, L. M. Fontoura, and R. C. Nunes. Evaluation of navigation in immersive large-scale vr environments using minimap-assisted teleportation. In *Proceedings of the 26th Symposium on Virtual and Augmented Reality*, pp. 270–274, 2024. 5

[13] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio. Vrsketchin: Exploring the design space of pen and tablet interaction for 3d sketching in virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–14. ACM, Apr. 2020. doi: 10.1145/3313831.3376628 2, 3

[14] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner. G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2):175–191, May 2007. doi: 10.3758/bf03193146 5

[15] S. M. Feeman, L. B. Wright, and J. L. Salmon. Exploration and evaluation of cad modeling in virtual reality. *Computer-Aided Design and Applications*, 15(6):892–904, Apr. 2018. doi: 10.1080/16864360.2018.1462570 2, 3, 9

[16] M. Fisher, D. Ritchie, M. Savva, T. Funkhouser, and P. Hanrahan. Example-based synthesis of 3d object arrangements. *ACM Trans. Graph.*, 31(6), Nov. 2012. doi: 10.1145/2366145.2366154 2

[17] R. Gao, H. Li, G. Dharan, Z. Wang, C. Li, F. Xia, S. Savarese, L. Fei-Fei, and J. Wu. Sonicverse: A multisensory simulation platform for embodied household agents that see and hear. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 704–711, 2023. doi: 10.1109/ICRA48891.2023.10160461 1

[18] E. Gatti, D. Giunchi, N. Numan, and A. Steed. Aisop: Exploring immersive vr storytelling leveraging generative ai. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 865–866, 2024. doi: 10.1109/VRW62533.2024.00229 1

[19] J. M. Govern and L. A. Marsch. Development and validation of the situational self-awareness scale. *Consciousness and Cognition*, 10(3):366–378, Sept. 2001. doi: 10.1006/ccog.2001.0506 6

[20] R. E. Guadagno, K. R. Swinth, and J. Blascovich. Social evaluations of embodied agents and avatars. *Computers in Human Behavior*, 27(6):2380–2385, nov 2011. doi: 10.1016/j.chb.2011.07.017 2

[21] P. E. Hart, N. J. Nilsson, and B. Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968. doi: 10.1109/TSSC.1968.300136 5

[22] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, Oct. 2006. doi: 10.1177/154193120605000909 6

[23] J.-Y. Jian, A. M. Bisantz, and C. G. Drury. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, 4(1):53–71, mar 2000. doi: 10.1207/s15327566ijce0401_04 6

[24] D. Kao, A. J. Magana, and C. Mousas. Evaluating tutorial-based instructions for controllers in virtual reality games. *Proceedings of the ACM on Human-Computer Interaction*, 5(CHI PLAY):1–28, Oct. 2021. doi: 10.1145/3474661 6

[25] J. Krarup and P. M. Pruzan. Computer-aided layout design. In *Mathematical programming in use*, pp. 75–94. Springer, 2009. 2

[26] M. V. Law, A. Kwatra, N. Dhawan, M. Einhorn, A. Rajesh, and G. Hoffman. Design intention inference for virtual co-design agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, IVA '20, p. 1–8. ACM, Oct. 2020. doi: 10.1145/3383652.3423861 8

[27] H. Liu, M. Choi, D. Kao, and C. Mousas. Synthesizing game levels for collaborative gameplay in a shared virtual environment. *ACM Transactions on Interactive Intelligent Systems*, 13(1):1–36, Mar. 2023. doi: 10.1145/3558773 6

[28] J.-H. Liu, S.-K. Zhang, C. Zhang, and S.-H. Zhang. Controllable procedural generation of landscapes. In *Proceedings of the 32nd ACM International Conference on Multimedia*, MM '24, p. 6394–6403. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3664647.3681129 2

[29] K.-Y. Liu, S.-K. Wong, M. Volonte, E. Ebrahimi, and S. V. Babu. Investigating the effects of leading and following behaviors of virtual humans in collaborative fine motor tasks in virtual reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 330–339, 2022. doi: 10.1109/VR51125.2022.00052 1

[30] P. Merrell, E. Schkufza, Z. Li, M. Agrawala, and V. Koltun. Interactive furniture layout using interior design guidelines. *ACM Trans. Graph.*, 30(4), July 2011. doi: 10.1145/2010324.1964982 2

[31] T. Merritt, C. Ong, T. L. Chuah, and K. McGee. *Did You Notice? Artificial Team-Mates Take Risks for Players*, p. 338–349. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-23974-8_37 2, 8

[32] D. A. Norman. How might people interact with agents. *Communications of the ACM*, 37(7):68–71, July 1994. doi: 10.1145/176789.176796 2, 8

[33] Y. Pan and A. Steed. The impact of self-avatars on trust and collaboration in shared virtual environments. *PLOS ONE*, 12(12):e0189078, Dec. 2017. doi: 10.1371/journal.pone.0189078 1

[34] J. Rasch, J. Töws, T. Hirzle, F. Müller, and M. Schmitz. Creepycocreator? investigating ai representation modes for 3d object co-creation

in virtual reality. 2025. doi: 10.1145/3706598.3713720 1, 2, 3, 8

[35] A. Romani, F. Casnati, and A. Ianniello. Codesign with more-than-humans: toward a meta co-design tool for human-non-human collaborations. *European Journal of Futures Research*, 10(1), July 2022. doi: 10.1186/s40309-022-00205-7 9

[36] T. Steel, D. Kuiper, and R. Z. Wenkstern. Context-aware virtual agents in open environments. In *2010 Sixth International Conference on Autonomic and Autonomous Systems*, p. 90–96. IEEE, Mar. 2010. doi: 10.1109/icas.2010.36 1

[37] H. I. I. Tam, H. I. D. Pun, A. T. Wang, A. X. Chang, and M. Savva. Sceneeval: Evaluating semantic coherence in text-conditioned 3d indoor scene synthesis, 2025. doi: 10.48550/ARXIV.2503.14756 2

[38] J. Urban Davis, F. Anderson, M. Stroetzel, T. Grossman, and G. Fitzmaurice. Designing co-creative ai for virtual environments. In *Proceedings of the 13th Conference on Creativity and Cognition*, CC '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3450741.3465260 1, 2

[39] H. Vartiainen, P. Liukkonen, and M. Tedre. The mosaic of human-ai co-creation: Emerging human-technology relationships in a co-design process with generative ai. Nov. 2023. doi: 10.35542/osf.io/g5fb8 2

[40] G. Walsh and E. Wronsky. Ai + co-design: Developing a novel computer-supported approach to inclusive design. In *Companion Publication of the 2019 Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '19, p. 408–412. ACM, Nov. 2019. doi: 10.1145/3311957.3359456 2, 8

[41] M. Wang, Z.-M. Ye, J.-C. Shi, and Y.-L. Yang. Scene-context-aware indoor object selection and movement in vr. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, p. 235–244. IEEE, Mar. 2021. doi: 10.1109/vr50410.2021.00045 1, 2

[42] Y. Wang, X. Qiu, J. Liu, Z. Chen, J. Cai, Y. Wang, T.-H. Wang, Z. Xian, and C. Gan. Architect: Generating vivid and interactive 3d scenes with hierarchical 2d inpainting. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, eds., *Advances in Neural Information Processing Systems*, vol. 37, pp. 67575–67603. Curran Associates, Inc., 2024. 2

[43] A. Weingarten, B. Lafreniere, G. W. Fitzmaurice, and T. Grossman. Dreamrooms: Prototyping rooms in collaboration with a generative process. In *Graphics Interface*, pp. 19–1, 2019. 2

[44] J. Whyte, N. Bouchlaghem, A. Thorpe, and R. McCaffer. From cad to virtual reality: modelling approaches, data exchange and interactive 3d building design tools. *Automation in Construction*, 10(1):43–55, Nov. 2000. doi: 10.1016/s0926-5805(99)00012-6 3

[45] E. J. Williams. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2):149–168, 1949. 6

[46] B. Xie, Y. Zhang, H. Huang, E. Ogawa, T. You, and L.-F. Yu. Exercise intensity-driven level design. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1661–1670, 2018. doi: 10.1109/TVCG.2018.2793618 3

[47] F.-C. Yang, K. Duque, and C. Mousas. The effects of depth of knowledge of a virtual agent. *IEEE Transactions on Visualization and Computer Graphics*, 30(11):7140–7151, Nov. 2024. doi: 10.1109/tvcg.2024.3456148 1, 2

[48] Z.-M. Ye, J.-L. Chen, M. Wang, and Y.-L. Yang. Paval: Position-aware virtual agent locomotion for assisted virtual reality navigation. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, p. 239–247. IEEE, Oct. 2021. doi: 10.1109/ismar52148.2021.00039 1, 8

[49] W. You, Y. Shao, Z. Zheng, Y. Lu, C. Yang, Z. Zhou, and L. Sun. Old house new home: Facilitating interior design with redesignus in virtual reality. *Displays*, 80:102555, Dec. 2023. doi: 10.1016/j.displa.2023.102555 1, 2, 3, 8

[50] L.-F. Yu, S.-K. Yeung, C.-K. Tang, D. Terzopoulos, T. F. Chan, and S. J. Osher. Make it home: automatic optimization of furniture arrangement. In *ACM SIGGRAPH 2011 papers*, SIGGRAPH '11, p. 1–12. ACM, July 2011. doi: 10.1145/1964921.1964981 1, 2, 9

[51] L.-F. Yu, S.-K. Yeung, and D. Terzopoulos. The clutterpalette: An interactive tool for detailing indoor scenes. *IEEE Transactions on Visualization and Computer Graphics*, 22(2):1138–1148, Feb. 2016. doi: 10.1109/tvcg.2015.2417575 2

[52] S.-K. Zhang, Y.-X. Li, Y. He, Y.-L. Yang, and S.-H. Zhang. Mageadd: Real-time interaction simulation for scene synthesis. In *Proceedings of the 29th ACM International Conference on Multimedia*, MM '21, p. 965–973. ACM, Oct. 2021. doi: 10.1145/3474085.3475194 1, 2, 3, 4, 5, 8

[53] S.-K. Zhang, J.-H. Liu, J. Huang, Z.-W. Chi, H. Tam, Y.-L. Yang, and S.-H. Zhang. Sceneexplorer: An interactive system for expanding, scheduling, and organizing transformable layouts. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2024. doi: 10.1109/TVCG.2024.3488744 2, 3

[54] Z. Zhao and X. Ma. Situated learning of soft skills with an interactive agent in virtual reality via multimodal feedback. In *Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20 Adjunct, p. 25–27. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3379350.3416174 1

[55] C. Zhou, Y. Bian, S. Zhang, Z. Zhang, Y. Wang, and Y.-J. Liu. Exploring user experience and performance of a tedious task through human–agent relationship. *Scientific Reports*, 13(1), Feb. 2023. doi: 10.1038/s41598-023-29874-5 2

[56] J. Zhou, R. Li, J. Tang, T. Tang, H. Li, W. Cui, and Y. Wu. Understanding nonlinear collaboration between human and ai agents: A co-design framework for creative design. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, p. 1–16. ACM, May 2024. doi: 10.1145/3613904.3642812 8, 9