Exploring the Design of Context-Aware Widget Recommender System in Mixed Reality

Yiming Sun, Yang Zhan, and Tatsuo Nakajima

Department of Computer Science and Engineering, Waseda University, Tokyo, Japan
{ysun60, yang.zhan, tatsuo}@dcl.cs.waseda.ac.jp

Abstract. Mixed Reality (MR), which overlays virtual user interfaces (UIs) onto physical environments, necessitates UI adaptation to complex scenes and tasks. While prior work focused on adaptive spatial UI layout in MR, functional adaptation such as widget recommendation remains underexplored. We present a user study ($n = 16$) using a Large Language Models (LLMs)-powered widget recommender system (RS) as a technology probe to investigate how context-aware recommendations affect user experience. The system uses LLMs with contextual data (reading text, video transcript and typed data) to suggest MR widgets. Results show that widget recommendations facilitated access to context-relevant functionalities and simplified task workflows, thus enhancing the user experience and reducing workload.However, the usability of the widget RS depends on appropriate widget design and recommendation strategies that enable personal customization. This study serves as an initial step toward MR widget RSs and offers insights for adaptive user-RS interactions.

Keywords: Mixed Reality · Context awareness · Widget Recommendation · Large Language Models

1 Introduction

In recent years, the development of Mixed Reality (MR) has shown its potential to improve users' productivity and become an everyday wearable product. By expanding user interfaces (UIs) into 3D space, MR applications can arbitrarily overlay virtual UIs onto the real world, offering greater mobility compared to conventional laptops and phones. Moreover, MR displays multi-dimensional information simultaneously and enables intuitive interactions [17], which are widely used in programming [29] and education [28]. Although MR systems present novel experiences, their design and implementation are challenging due to complex real-world environments and dynamic user statuses. For instance, it is time-consuming for users to manually place virtual interfaces to avoid occlusion by real-world surroundings and adjust UI layouts when switch physical workspaces [5].

Previous research has emphasized the value of adaptive systems in MR environments. Studies examined how environmental factors affect user experience [12, 23] and how UI placement can align with semantic features of physical objects [5]. Adaptive interfaces have also responded to user states such as mental workload [20]. However, most of these works focus on layout-level adaptation and overlook the recommendation of MR applications based on user needs. In dynamic environments and user activities, repeated app selection can be burdensome. Prior work [6] found that users prefer function-oriented UIs, such as widgets, over application-based ones. Yet, widgets—lightweight tools derived from specific app—may lead to an overwhelming number of options. These findings underscore the need for adaptive widget recommendation to streamline interaction and improve usability in MR.

Motivated by above research gap, this paper takes an initial step toward exploring the user experience and design considerations of widget recommender system (RS) on MR devices. Specifically, we aim to investigate the following two research questions:

- **(RQ1)** How does the recommendation of MR widgets affect users' experience and perception during tasks?
- **(RQ2)** What system features influence the usability of widget RS in MR?

To address our research questions, we conducted a within-subject study ($n = 16$), and designed an open-ended task using an MR widget RS in conjunction with a PC workspace, forming a hybrid workspace [27], which is a common scenario in MR productivity applications [6, 5]. We implemented a technology probe [16] that leverages Large Language Models (LLMs) to recommend task-relevant widgets on MR head-mounted displays (HMDs). Three types of contextual inputs—reading texts, video transcripts, and typed data—were incorporated to investigate how context influences recommendations and user experience. Our main contributions are: (1) extending MR context-aware adaptation from UI layout adjustments to functionality-based adaptation, (2) evaluating its impact on user experience and perception through a user study, and (3) deriving design insights for future RSs and broader MR applications.

2 Related Work

2.1 Adaptive Designs in MR

MR encompassing Virtual Reality (VR) and Augmented Reality (AR), offers high mobility and has been applied across daily scenarios for flexible UIs and efficient interaction. To support dynamic environments, prior studies optimized MR interface placement in offices [22], classrooms [2], and stores [35]. Shin et al. [31] showed that room size affected AR game experience, suggesting space-adaptive UIs for indoor contents. SemanticAdapt [5] aligned UIs with physical objects via contextual cues. For dynamic contexts, Du et al. [10] enabled anchoring UIs to everyday objects, and FingerSwitches [25] introduced UI switching principles across static and dynamic surroundings. While the physical environment is critical in MR interface design, user status and ongoing activities are equally important considerations. Grubert et al. [14] emphasized the necessity of developing context-aware AR systems to support continuous and multi-purpose user experiences. When users transition between contexts, such as tasks or environments, associated changes in cognitive load can be leveraged to optimize both the amount of information displayed and the positioning of MR UIs [20]. To understand UI preferences across activities, MineXR [6] collected personalized UI layouts across four daily XR scenarios, producing 109 layouts tailored to tasks. The findings revealed a preference for functionality UI over complete application, supporting widget-centric designs in XR productivity. Despite the advancement, most existing studies concentrated on adapting UI layouts to environments and users. It remains unclear how adaptive widgets, the functionality of the UI, can enhance user experience in MR.

2.2 App Recommendation

The rapid growth of mobile apps has driven the development of automatic app RS to aid app selection and optimize system operations such as memory management. For example, Baeza-Yates et al. [1] treated app prediction as a classification task, using sensor data and sequential usage patterns to enhance precision. Bayesian and Markov models have been widely used to analyze contextual data and predict subsequent app usage [15, 24]. Despite achieving reasonable accuracy, traditional methods often fail to capture sufficient features from complex contextual data. To address this, Shen et al. [30] applied deep reinforcement learning to reduce app loading time and improve user satisfaction. Likewise, WhatsNextApp [18] used an LSTM model to mitigate the cold-start problem when users set up new devices. The rise of LLMs offers new opportunities for app recommendation by enabling better reasoning over contextual cues and understanding semantic app descriptions. LLMs have proven effective as RSs in various applications, such as open-world recommendation [33] with both fine-tuned [34] and prompt-based approaches [7]. Khaokaew et al. [19] utilized a pre-trained LLM to predict top-k apps using contextual data such as app details and points of interest, outperforming traditional models in standard and cold-start cases.

While prior research has explored app recommendation for real-time task adaptation, few have addressed its application in MR environments, especially regarding user–RS interaction. The distinct characteristics of 3D MR spaces challenge the usability of approaches designed for 2D screens. The most related work focuses on selecting MR apps and adapting UI based on task type and cognitive load [20]. In contrast, we leverages richer contextual information to recommend widgets and explores multiple context-aware mechanisms in everyday tasks, offering new insights into app recommendation in MR.

3 System Overview

We developed a functional system prototype specifically for the user study. Rather than aiming to build a fully mature system, we employed the proto-

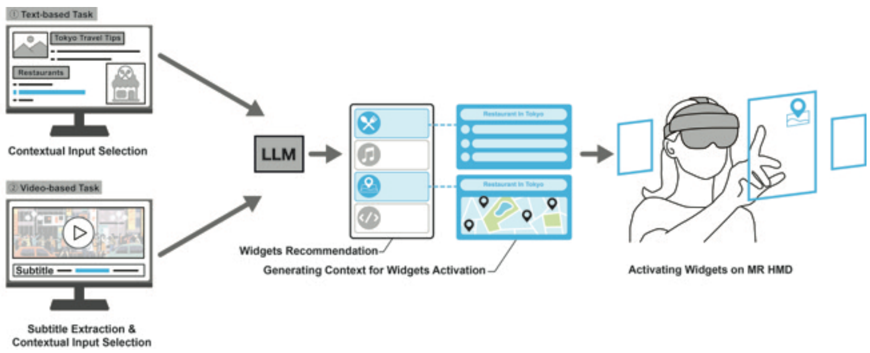


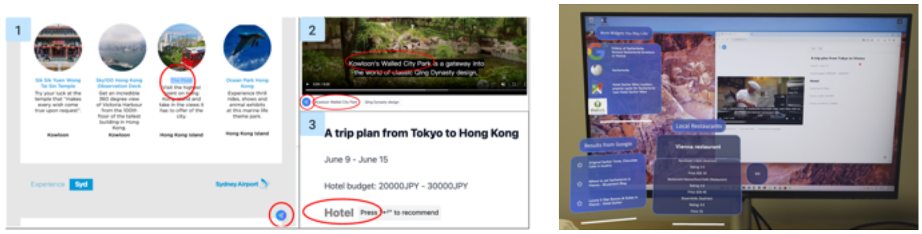
Fig. 1: System pipeline with three main components: context-aware mechanisms (*left*), an LLM-powered RS backend (*middle*), and MR widgets (*right*). *Left*: Passive and proactive mechanisms capture context for the RS (section 3.1). *Middle*: The RS uses WRC to recommend widgets and WAC to adapt their functions (section 3.2). *Right*: Recommended widgets appear in predefined MR areas to support user tasks (section 3.3).

type as a technology probe to gather user insights and design implication for the MR widgets recommendation.

Our system comprises three main components: *context-aware mechanisms* built into computer interfaces that capture on-screen content and treat it as contextual input (see section 3.1), an *LLM-powered RS backend* for widget prediction and activation (see section 3.2), and *MR widgets* for task assistance (see section 3.3). The system pipeline is shown in fig. 1. Three contexts, including reading texts, video transcripts, and typed data, are captured proactively and passively by the context-aware mechanism and fed into the RS. The LLM-powered RS consists of two core features: Widget Recommendation based on Context (WRC) and Widget Activation driven by Context (WAC). WRC focuses on recommending the most suitable widgets based on the current user contexts, whereas WAC enables the selected widgets to autonomously perform relevant functions or guide users to specific functionalities by applying the contextual parameters to the widgets. For example, when viewing a travel video, the WRC recommends a map widget, and the WAC automatically displays information related to the point of interest on the widget.

Our implementation consists of three components that communicate via WebSocket: (1) computer interfaces developed with React[1], (2) a server that hosts a GPT-4o[2]-powered RS backend and widget services, and (3) a MR application developed in Unity with MRTK3[3], integrating 11 widgets. The relevance of widgets is task-dependent and thus are described in section 4.3.

[1] https://react.dev/
[2] https://platform.openai.com/docs/models/gpt-4o
[3] https://github.com/MixedRealityToolkit/MixedRealityToolkit-Unity



(a) Context-aware mechanisms: (1) passive text selection, (2) passive subtitle selection, and (3) proactive title input.

(b) User interacting with MR widgets during the task.

Fig. 2: Context-aware mechanisms and MR interface usage.

3.1 Context-Aware Mechanisms

Although technologies like eye tracking and text recognition may enable MR headsets to autonomously capture context, they risk introducing bias and inconsistency in user studies. Therefore, we extract contextual input on PC. Contextual information is acquired through **passive** and **proactive** mechanisms. The passive mechanism relies on user-selected content. For the text-based contexts (fig. 2a-(1)), users select text content as contextual input, then activates the blue button to send the contexts. For video contexts (fig. 2a-(2)), users extracts interest points from subtitles by pressing the blue button and then selects one to transmit the contextual information. It also incorporates typed input to enrich recommendation context. In contrast, the proactive mechanism automatically captures after title input (fig. 2a-(3)). Our LLM-based RS primarily uses the passive mechanism, as it offers greater precision and fosters user trust [21].

3.2 LLM-powered Widget RS

The LLM-based RS backend analyzes the current context, recommends relevant widgets, and activates their functionalities accordingly. It operates via structured prompts refined through pilot testing, incorporating techniques such as few-shot prompting [3] and task decomposition [32]. Each prompt includes: AI role assignment, task explanation, widget/context descriptions, output format, and a one-shot example (fig. 3). Square-bracketed fields are dynamically replaced with task-specific content during the study.

The RS generates two outputs: (1) WRC results, identifying context-relevant widgets; and (2) refined context parameters ($<$ keyword $>$) for WAC, enabling each selected widget to function with appropriate inputs.

3.3 MR Widgets

MR widgets are supported by the WAC, activating their features based on users' contextual input. To accommodate diverse functionalities, we adopted a Web

Prompt: Now you act as a helpful [assistant]. You will read the user's [contexts], recommend multiple important widgets and fill in the keyword fields for the widgets tagged by < keyword > in the predefined library with the correct information. Using this text, please fill in the following library structure, < library >[...]< /library >. The user's [typed input] is tagged by [< tag >], and the user is currently focusing on the content delimited by < focus >. The focus content could be the [reading context] or [typed input]. You should always recommend 3 to 4 widgets. Your answer must end with a JSON format using the following template: [...]. Below is an example for user's travel description: [...], Your answer could be: [...].
Output: [{"widget": "...", "keywords": "..."}, ...]

Fig. 3: Example of the LLM prompt and output. Task-specific content is enclosed in square brackets. The output includes selected widgets (< library >) for WRC and refined context parameters (< keyword >) for WAC.

Widget deployment strategy: these widgets rely on external web content and do not run natively on the MR headset. Instead, they display simplified information in MR and include navigation buttons for accessing full content via a linked desktop browser. This approach accelerates development and enables rapid deployment of task-relevant widgets. To improve awareness and predictability, users can predefine widget display areas during system initialization.

4 User Study

To answer the research questions, we conducted a within-subject study with two conditions: *Widget RS* and *PC-only*. In the Widget RS condition, the participants wore HoloLens 2 MR headsets and completed the assigned task on a computer with assistance from the RS and MR widgets. In contrast, the PC-only condition required participants to perform the same task merely on the computer without assistance, serving as the baseline. We selected a PC-only condition as the baseline because this familiar paradigm and workflow introduce no additional variables and therefore allow more intuitive comparisons with the MR-based RS. We recruited 16 participants (6 females, 10 males), aged 19–35 ($M = 25.3$, $SD = 3.3$). One participant used MR weekly, 9 less than monthly, and 6 had no prior MR experience. Thirteen participants used LLMs weekly, and three monthly.

4.1 Task

Travel planning, a cognitively demanding task involving various types of contextual information [26], was chosen to evaluate the system across diverse contexts. Its open-ended nature better mirrors real-world MR usage than goal-driven tasks. To limit study duration, participants compiled a travel note rather than a full itinerary. They selected points of interest from text and video materials, searched for related details, recorded supplementary information, and identified a hotel and restaurants within given constraints. Use of conversational LLMs was prohibited to ensure comparability across participants. Two destinations, Vienna and Florence, were assigned in a fixed order while the two system conditions were counterbalanced. We ensured none participants had visited the assigned cities before.

4.2 Procedure

The entire study lasted around two hours. Upon arrival, participants were briefed on the research goals, signed a consent form, and completed a demographic questionnaire. They then performed a training task to familiarize themselves with the system. Next, they proceeded to complete the travel planning task under either the PC-only or Widget RS condition, presented in counterbalanced order. After each task, the participants completed several questionnaires. Finally, a 15-minute semi-structured interview was conducted, and participants were offered a $20 Amazon gift card as compensation. The study was exempted from review by the local IRB.

4.3 Widget Description

In the user study, eleven widgets were designed to support specific types of contextual information or user needs associated with the travel planning task. A list of these widgets is presented in table 1, with visual examples shown in fig. 2b. The selection and design of these widgets were informed by common travel planning requirements, as well as frequently used widget types identified in the MineXR dataset [6]. Each widget features buttons that can be operated via direct finger interaction.

4.4 Measures

We adopted a mixed-method evaluation combining objective and subjective metrics. Objective data included task completion time and the number of passive recommendations across all context types. Subjective evaluation focused on perceived productivity, user experience (measured using the User Experience Questionnaire, UEQ), and workload (measured using the NASA Task Load Index, TLX). Productivity was assessed using four items rated on a 7-point Likert scale: two addressing system usability ("I can quickly find the information I need." and "I am willing to explore more information."), and two adapted from prior research on perceived usefulness [8] ("Using this system in my daily work will improve my productivity." and "...my job performance."). Higher scores indicate better perceived productivity. In the semi-structured interview, we asked participants for general feedback, including preferences, their experiences, and perceptions of the RS and the MR widgets.

Name	Description	Content
Local Restaurant	A list of three restaurants from Google Local	Restaurant's name, cuisine type, price range, and rating
Result from Google	A list of three websites retrieved via Google Search	Website title
Top Hotel Picks	A list of three most-viewed hotels retrieved from Google Hotels	Hotel's name, price, rating, and photos
Google, Google Maps, Wikipedia, YouTube, OpenTable*, Red Note*, The Fork*, Trip*	A context-aware website of the corresponding widget, which is arranged as an item of *More Widgets You May Like*	Widget icon and context-related parameters

Table 1: Overview of widgets used in the study. Asterisks denote widgets without WAC support, functioning only as redirect icons due to API or crawling limitations.

5 Results

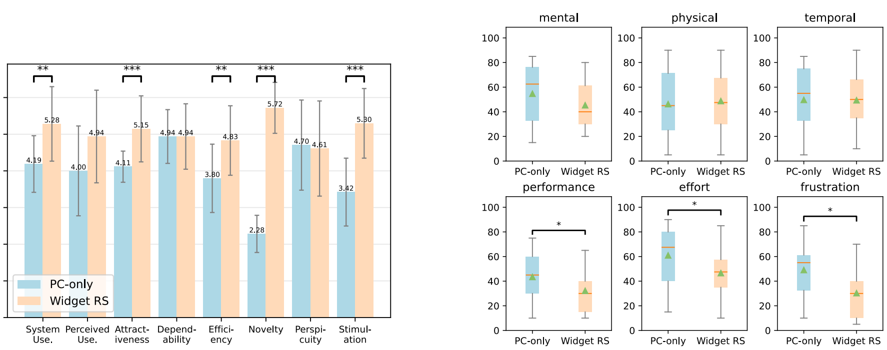
This section presents the quantitative and qualitative findings of our user study. For quantitative data, including objective measures and questionnaire responses, we conducted paired-samples t-tests when the normality assumption was satisfied. When normality was not met, the Wilcoxon signed-rank test was used as a non-parametric alternative. All other statistical assumptions were verified before analysis, and significance levels were set at $\alpha = .05$, .01, and .001. Questionnaire scores were reported as averages of the items within each subscale. For qualitative data, we performed open coding on interview transcripts, followed by affinity mapping to group codes. Key findings were synthesized through team discussions to identify novel and meaningful insights.

5.1 Objective Measures

For task completion time, participants spent an average of 21.66 minutes ($SD = 8.46$) in the PC-only condition and 22.73 minutes ($SD = 9.98$) in the RS condition, with no significant differences confirmed ($t(15) = -0.565$, $p = 0.580$). A Friedman test revealed no significant differences among the three contextual sources ($\chi^2(2) = 0.623$, $p = 0.732$), although more passive recommendations were triggered on average when reading text ($M = 4.25$, $SD = 3.56$), with similar mean values observed in the video ($M = 2.94$, $SD = 1.77$) and the note contexts ($M = 3.00$, $SD = 3.31$).

5.2 Questionnaire Responses

Productivity was measured through system usability and perceived usefulness (fig. 4a). Usability scores were significantly higher in the MR RS condition ($M = 5.28$, $SD = 1.25$) than in the PC-only condition ($M = 4.19$, $SD = 1.15$; $W = 4.0$, $p = 0.004$). Perceived usefulness showed a marginal increase for the MR RS ($M = 4.94$, $SD = 1.29$) over PC-only ($M = 4.00$, $SD = 1.24$; $t(15) = -1.996$, $p = 0.064$). UEQ showed significant improvements in four subscales: *attractiveness* ($t(15) = -4.086$, $p < 0.001$), *efficiency* ($t(15) = -2.948$, $p = 0.009$), *novelty* ($W = 0$, $p < 0.001$), and *stimulation* ($t(15) = -6.141$, $p < 0.001$). No differences were found in *dependability* or *perspicuity*. For the NASA-TLX, overall workload was marginally lower in the MR RS condition ($t(15) = 2.126$, $p = 0.051$). Among subscales (fig. 4b), significant reductions were observed in *performance* ($t(15) = 2.875$, $p = 0.012$), *effort* ($t(15) = 2.375$, $p = 0.031$), and *frustration* ($t(15) = 2.916$, $p = 0.011$). No differences were found in *mental*, *physical*, or *temporal* demand.



(a) Mean and SD of productivity and UEQ. "System Use." denotes System Usability; "Perceived Use." denotes Perceived Usefulness. Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(b) Boxplots of raw NASA-TLX subscales. Green triangles: means; orange lines: medians. * ($p < 0.05$) indicates significance.

Fig. 4: a: Productivity and UEQ, b: NASA-TLX results.

5.3 Interview

User Experience with PC-only vs. MR system Participants compared their task experiences using the Widget RS in MR versus the PC-only condition. Over half ($n = 9$) cited operational familiarity as the main advantage of using a PC. However, the PC-only condition required high cognitive overload when processing abundant information ($n = 4$). P15 noted the lack of task guidance, requiring more mental effort to organize thoughts. In contrast, the MR system was praised for its convenience and time-saving features, with most participants ($n = 7$) noting that recommended widgets reduced search effort. Participants appreciated the centralized layout ($n = 6$), guidance served by RS ($n = 3$), and familiar icons that increased trust and usability (P11, P16). Reported drawbacks of the MR system included its learning curve ($n = 3$), and physical fatigue from interactions and the headset ($n = 7$). A few participants ($n = 2$) also found an excess of widgets overwhelming.

MR Widgets Regarding widget usability, most participants ($n = 13$) preferred widgets that supported WAC, such as Google Map or Local Restaurant. About half ($n = 6$) noted that WAC simplified their workflow and reduced workload. As P1, P6, and P16 remarked, *"Not having to input the search query is convenient."* However, the majority ($n = 13$) found widget functionality insufficient. Participants attributed this to dependence on external webpages (P1, P7, P11) or widgets merely redirecting to main pages (P4, P14, P15). In terms of content richness, most participants ($n = 11$) felt the displayed information was lacking, with several requesting more comprehensive results, especially for hotel ($n = 3$) and restaurant widgets ($n = 2$).

Context-based RS When evaluating RS performance across the context-aware mechanisms, most participants ($n = 11$) favored the video-based passive mechanism, highlighting its convenience and accuracy in reflecting points of interest. P15 emphasized the contrast with the PC-only condition: *"you need to manually type in the names ... but MR can automatically search and provide the key results."* However, some ($n = 3$) noted that pausing videos to extract subtitles disrupted their workflow. As improvements, participants suggested adding support for image frame extraction as context ($n = 9$). The text-based passive mechanism was also well received, with participants appreciating its selection accuracy ($n = 10$) and seamless integration into the reading flow ($n = 4$). P11 noted that selecting and searching in text was more convenient than pausing video. However, some users expected the system to better anticipate their focus and recommend related content automatically (P10). The proactive mechanism, triggered by title input, was considered helpful by most participants ($n = 10$). Still, concerns were raised about mismatches due to abstract or inconsistent titles ($n = 6$). Participants stressed the importance of recommendation timing based on context sequences, suggesting the system should anticipate next steps in the planning process (P1, P13).

User Perception Several participants ($n = 7$) expressed a strong willingness to explore the MR system, citing its interactivity (P5, P15), novelty (P2), and reduced time (P7) and effort (P14). Some ($n = 3$) noted its suitability for open-ended tasks, as it *"suggests things I hadn't originally thought of"* (P3). Most participants ($n = 14$) reported a smooth transition between the MR and PC environments, particularly in interaction and attention switching. Proactive recommendations were generally non-intrusive ($n = 4$), but occasionally ignored ($n = 2$) due to their peripheral positioning around the PC monitor and the need for head movement (P14). As P2 explained, participants could disregard them, especially when *"the timing was wrong."*

6 Discussion

For **RQ1**, participants reported enhanced efficiency, streamlined workflows, and increased exploratory motivation with the MR system. Regarding **RQ2**, key usability features included widget functionality, the intuitiveness of the context-aware mechanisms, and the RS capacity to interpret contextual input. These findings are further discussed in sections 6.1 and 6.2. We also outline potential applications beyond the hybrid workspace scenario explored in this study.

6.1 User Experience and Perception

Although objective performance did not significantly differ, users reported improved usability and experience with the Widget RS. Factors such as task characteristics or prior experience with PC workflows may have influenced the outcome. UEQ results showed that the MR condition was perceived as more attractive, stimulating, and novel, with higher efficiency. However, no significant improvement was observed in completion time or temporal demand, suggesting that RS contributed to a high perceived performance , even if it was not reflected in task completion time, it may be driven by the workflow simplification. Lower effort and frustration scores in TLX further reflected the system's supportive role on inspiration and motivation. For example, P2 described the experience as engaging and curiosity-driven. However, mental demand remained unchanged, We believe that this is due to limitations in widget functionality. Specific tasks like booking still required users to access external websites, which required users to invest additional mental effort.

6.2 Design Implications

Prioritizing Context-Activated Widgets over Icons Participants strongly preferred context-activated MR widgets with WAC support, as these enabled direct access to relevant information without additional navigation. In contrast, icon-type widgets linking to static websites were rarely used, differing from prior findings [6]. This shift may be attributed to the combined cognitive load of MR HMD usage and the travel-planning task since reported workload scores (PC-only: $M = 50.7$; MR RS: $M = 42.1$) fall within cognitively intensive ranges [13]. We suggest future MR systems prioritize widgets with immediately accessible content. Extending prior research [20], our results emphasize context-aware interface design over mere simplification.

Balancing Recommendation Complexity and Workload The RS exploited 3D space to display multiple widgets, but users often prioritized only a few, underscoring the need to avoid redundant recommendations. Some participants also noted limited functionality in individual widgets. However, adding too many features can increase ergonomic burden [11]. Designers should focus on core functionality and reduce unnecessary interaction frequency to balance utility and cognitive load.

Incorporating Users into the Recommendation Loop Low-level contextual input helped narrow recommendations but lacked precision due to limited understanding of user goals. Variations in title writing and user knowledge affected the accuracy of outcomes. Future systems should support preference customization and interactive feedback to better align with user intent. Research in human-AI interaction has underscored the importance of personalization [21] and user autonomy [4] for user experience, which is echoed and extended in our results.

6.3 MR Widget RS in Broader Scenarios

Although the study was conducted on a hybrid workspace, our findings highlight two key advantages of widget RS for broader MR scenarios:

Motivation for continuous exploration. The RS promotes engagement by suggesting context-relevant widgets that simplify tasks and reduce workload (section 5.3, section 6.1), fostering motivation similarly to personalized educational recommendations [9]. For example, the RS could recommend explanatory widgets during MR reading. In such cases, designers should aim to balance recommendation complexity and personalize suggestions (section 6.2).

Inspiration for problem-solving and creativity. The RS guides users during uncertain task stages by suggesting relevant widgets, particularly beneficial for creative or open-ended tasks (section 5.3). For example, in MR shopping, the RS could suggest coupons or alternative products to enhance user decisions. To support this, proactive recommendations should consider task progress and cognitive load, balancing usability and cognitive demands (section 6.2).

6.4 Limitations and Future work

Application scenarios. Due to time constraints, our task design focused on a single scenario and did not fully simulate real-world travel exploration, and also the small sample size limited the generalizability of results. Future research should investigate long-term use of the widget RS across multiple task types. Moreover, the hybrid workspace setting may not reflect the full potential of RS in more immersive MR environments. We encourage applying the system to diverse scenarios as discussed in section 6.3 to support generalizable MR design.

Diverse contexts. Contextual input in our study was extracted via PC rather than MR devices. Although we included text, video, and typed data, other rich inputs—such as gaze tracking and live video—were not explored. Future systems should leverage native MR sensors to enable richer, real-time context capture.

Usability of widget and RS. Some widgets were limited by the lack of public APIs, restricting full WAC support. Integration at the OS level (e.g., visionOS) and more interactive 3D UIs may improve usability. For the RS, proactive mechanisms depend heavily on the quantity and quality of contextual input. Enhancing prompt design or incorporating fine-tuned models [19] may lead to more intelligent and personalized recommendations.

7 Conclusion

This work explored how an MR widget RS supports users in daily tasks and offers design insights for future applications. Through a user study, we examined its impact across different contextual inputs and mechanisms on user experience and perception. Results showed that the RS streamlined workflows and enhanced user satisfaction. Based on qualitative findings, we proposed design implications to improve system usability. Although the study was conducted in a hybrid workspace, its findings extend to broader MR contexts and contribute to advancing adaptive MR system design.

References

1. Baeza-Yates, R., Jiang, D., Silvestri, F., Harrison, B.: Predicting the next app that you are going to use. In: Proceedings of the eighth ACM international conference on web search and data mining. pp. 285–294 (2015)
2. Broussard, D.M., Rahman, Y., Kulshreshth, A.K., Borst, C.W.: An interface for enhanced teacher awareness of student actions and attention in a vr classroom. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). pp. 284–290. IEEE (2021)
3. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al.: Language models are few-shot learners. Advances in neural information processing systems **33**, 1877–1901 (2020)
4. Chen, W., Shi, Y., Wang, Y., Shi, W., Chen, M., Gao, C., Mei, Y., Zhu, Y., Zhang, J., Yu, C.: Investigating context-aware collaborative text entry on smartphones using large language models. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. pp. 1–20 (2025)
5. Cheng, Y., Yan, Y., Yi, X., Shi, Y., Lindlbauer, D.: Semanticadapt: Optimization-based adaptation of mixed reality layouts leveraging virtual-physical semantic connections. In: The 34th Annual ACM Symposium on User Interface Software and Technology. pp. 282–297 (2021)
6. Cho, H., Yan, Y., Todi, K., Parent, M., Smith, M., Jonker, T.R., Benko, H., Lindlbauer, D.: Minexr: mining personalized extended reality interfaces. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–17 (2024)

7. Dai, S., Shao, N., Zhao, H., Yu, W., Si, Z., Xu, C., Sun, Z., Zhang, X., Xu, J.: Uncovering chatgpt's capabilities in recommender systems. In: Proceedings of the 17th ACM Conference on Recommender Systems. pp. 1126–1132 (2023)
8. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly pp. 319–340 (1989)
9. Drissi, S., Chefrour, A., Boussaha, K., Zarzour, H.: Exploring the effects of personalized recommendations on student's motivation and learning achievement in gamified mobile learning framework. Education and Information Technologies **29**(12), 15463–15500 (2024)
10. Du, R., Olwal, A., Le Goc, M., Wu, S., Tang, D., Zhang, Y., Zhang, J., Tan, D.J., Tombari, F., Kim, D.: Opportunistic interfaces for augmented reality: Transforming everyday objects into tangible 6dof interfaces using ad hoc ui. In: CHI Conference on Human Factors in Computing Systems Extended Abstracts. pp. 1–4 (2022)
11. Evangelista Belo, J.a.M., Feit, A.M., Feuchtner, T., Grønbæk, K.: Xrgonomics: Facilitating the creation of ergonomic 3d interfaces. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3411764.3445349
12. Gal, R., Shapira, L., Ofek, E., Kohli, P.: Flare: Fast layout for augmented reality applications. In: 2014 IEEE international symposium on mixed and augmented reality (ISMAR). pp. 207–212. IEEE (2014)
13. Grier, R.A.: How high is high? a meta-analysis of nasa-tlx global workload scores. In: Proceedings of the human factors and ergonomics society annual meeting. vol. 59, pp. 1727–1731. Sage Publications Sage CA: Los Angeles, CA (2015)
14. Grubert, J., Langlotz, T., Zollmann, S., Regenbrecht, H.: Towards pervasive augmented reality: Context-awareness in augmented reality. IEEE transactions on visualization and computer graphics **23**(6), 1706–1724 (2016)
15. Huang, K., Zhang, C., Ma, X., Chen, G.: Predicting mobile application usage using contextual information. In: proceedings of the 2012 ACM conference on ubiquitous computing. pp. 1059–1065 (2012)
16. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B.B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., Eiderbäck, B.: Technology probes: inspiring design for and with families. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. p. 17–24. CHI '03, Association for Computing Machinery, New York, NY, USA (2003). https://doi.org/10.1145/642611.642616
17. Irshad, S., Rambli, D.R.B.A.: User experience of mobile augmented reality: A review of studies. In: 2014 3rd international conference on user science and engineering (i-USEr). pp. 125–130. IEEE (2014)
18. Katsarou, K., Yu, G., Beierle, F.: Whatsnextapp: Lstm-based next-app prediction with app usage sequences. IEEE Access **10**, 18233–18247 (2022)
19. Khaokaew, Y., Xue, H., Salim, F.D.: Maple: mobile app prediction leveraging large language model embeddings. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **8**(1), 1–25 (2024)
20. Lindlbauer, D., Feit, A.M., Hilliges, O.: Context-aware online adaptation of mixed reality interfaces. In: Proceedings of the 32nd annual ACM symposium on user interface software and technology. pp. 147–160 (2019)
21. Lu, F., Xu, Y., Xu, X., Jones, B., Malamed, L.: Exploring the impact of user and system factors on human-ai interactions in head-worn displays. In: 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 109–118. IEEE (2023)

22. Luo, W., Lehmann, A., Widengren, H., Dachselt, R.: Where should we put it? layout and placement strategies of documents in augmented reality for collaborative sensemaking. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. pp. 1–16 (2022)
23. Macedo, M.C., Apolinario, A.L.: Occlusion handling in augmented reality: Past, present and future. IEEE Transactions on Visualization and Computer Graphics **29**(2), 1590–1609 (2021)
24. Natarajan, N., Shin, D., Dhillon, I.S.: Which app will you use next? collaborative filtering with interactional context. In: Proceedings of the 7th ACM Conference on Recommender Systems. pp. 201–208 (2013)
25. Pei, S., Kim, D., Olwal, A., Zhang, Y., Du, R.: Ui mobility control in xr: Switching ui positionings between static, dynamic, and self entities. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–12 (2024)
26. Peng, Y., Qin, X., Zhang, Z., Zhang, J., Lin, Q., Yang, X., Zhang, D., Rajmohan, S., Zhang, Q.: Navigating the unknown: A chat-based collaborative interface for personalized exploratory tasks. In: Proceedings of the 30th International Conference on Intelligent User Interfaces. p. 1048–1063. IUI '25, Association for Computing Machinery, New York, NY, USA (2025). https://doi.org/10.1145/3708359.3712093
27. Plasson, C., Blanch, R., Nigay, L.: Selection techniques for 3d extended desktop workstation with ar hmd. In: 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 460–469 (2022). https://doi.org/10.1109/ISMAR55827.2022.00062
28. Prit Kaur, D., Mantri, A., Horan, B.: Design implications for adaptive augmented reality based interactive learning environment for improved concept comprehension in engineering paradigms. Interactive Learning Environments **30**(4), 589–607 (2022)
29. Segura, R.J., del Pino, F.J., Ogáyar, C.J., Rueda, A.J.: Vr-ocks: A virtual reality game for learning the basic concepts of programming. Computer Applications in Engineering Education **28**(1), 31–41 (2020)
30. Shen, Z., Yang, K., Du, W., Zhao, X., Zou, J.: Deepapp: A deep reinforcement learning framework for mobile application usage prediction. In: Proceedings of the 17th Conference on Embedded Networked Sensor Systems. pp. 153–165 (2019)
31. Shin, J.e., Kim, H., Parker, C., Kim, H.i., Oh, S., Woo, W.: Is any room really ok? the effect of room size and furniture on presence, narrative engagement, and usability during a space-adaptive augmented reality game. In: 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 135–144. IEEE (2019)
32. Wu, T., Terry, M., Cai, C.J.: Ai chains: Transparent and controllable human-ai interaction by chaining large language model prompts. In: Proceedings of the 2022 CHI conference on human factors in computing systems. pp. 1–22 (2022)
33. Xi, Y., Liu, W., Lin, J., Cai, X., Zhu, H., Zhu, J., Chen, B., Tang, R., Zhang, W., Yu, Y.: Towards open-world recommendation with knowledge augmentation from large language models. In: Proceedings of the 18th ACM Conference on Recommender Systems. pp. 12–22 (2024)
34. Yang, F., Chen, Z., Jiang, Z., Cho, E., Huang, X., Lu, Y.: Palr: Personalization aware llms for recommendation. arXiv preprint arXiv:2305.07622 (2023)
35. Zhan, Y., Nakajima, T.: Effect of presentation methods on user experiences and perception in vr shopping recommender systems. In: Proceedings of the 2024 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region. pp. 241–248 (2024)