

Towards Context-Aware Adaptation in Extended Reality: A Design Space for XR Interfaces and an Adaptive Placement Strategy

Shakiba Davari  *Doug. A. Bowman  †

Center for Human-Computer Interaction
Department of Computer Science, Virginia Tech, Blacksburg, VA, USA

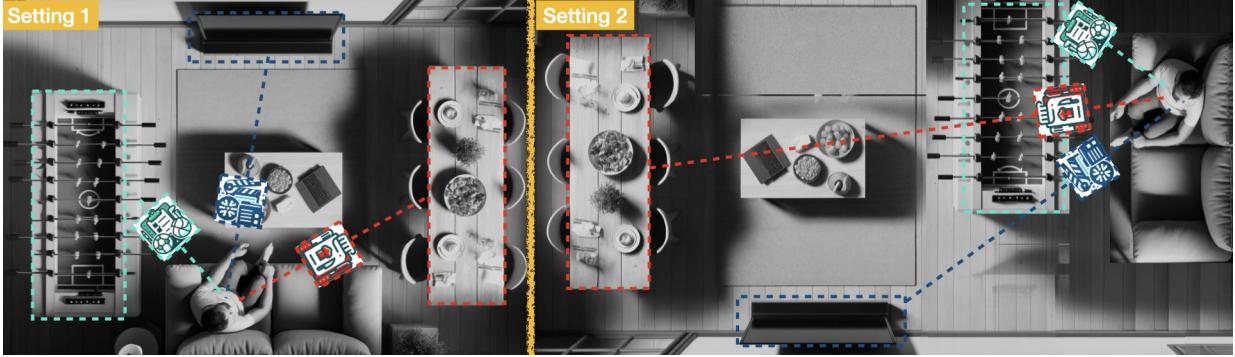


Fig. 1: Environment-referenced placement strategy: to adaptively place an XR object within varying contexts, this strategy identifies a relevant intermediary from the environment to be utilized within the object's frame of reference.

Abstract—By converting the entire 3D space around the user into a screen, Extended Reality (XR) can ameliorate traditional displays' space limitations and facilitate the consumption of multiple pieces of information at a time. However, if designed inappropriately, these XR interfaces can overwhelm the user and complicate information access. In this work, we explored the design dimensions that can be adapted to enable suitable presentation and interaction within an XR interface. To investigate a specific use case of context-aware adaptations within our proposed design space, we concentrated on the spatial layout of the XR content and investigated non-adaptive and adaptive placement strategies. In this paper, we (1) present a comprehensive design space for XR interfaces, (2) propose *Environment-referenced*, an adaptive placement strategy that uses a relevant intermediary from the environment within a Hybrid Frame of Reference (FoR) for each XR object, and (3) evaluate the effectiveness of this adaptive placement strategy and a non-adaptive *Body-Fixed* placement strategy in four contextual scenarios varying in terms of social setting and user mobility in the environment. The performance of these placement strategies from our within-subjects user study emphasized the importance of intermediaries' relevance to the user's focus. These findings underscore the importance of context-aware interfaces, indicating that the appropriate use of an adaptive content placement strategy in a context can significantly improve task efficiency, accuracy, and usability.

Index Terms—Extended Reality, Design Space, Augmented Reality, Context-Aware, Adaptation, Interface Design, Intelligent Interface, XR/MR/AR/VR.

1 INTRODUCTION

Extended Reality (XR) is envisioned to become a primary modality for information access, providing continuous, effortless, and hands-free access to multiple pieces of information through a glance, anywhere, and at any time [11, 21, 38, 40]. For instance, consider a conversation about scheduling a hike on a sunny day. Using a mobile phone involves multiple steps to 1) retrieve the device, 2) unlock the screen, 3) locate the calendar and weather apps, and 4) frequently switch between them. Furthermore, this process diverts the user's attention away from their conversation partner and social cues. In this scenario, Augmented Reality (AR) offers simultaneous access to both apps, eliminating the need for frequent switching, simplifying the process of locating the apps. Moreover, placing the apps near the interlocutor's face can enhance social awareness and engagement in the conversation [12].

However, these XR interfaces may also increase intrusiveness, environmental occlusion, and information overload, posing challenges to

the user's perception and awareness of their surroundings. XR's limitless possibilities for providing information pose a unique challenge for designers, who must create interfaces that seamlessly integrate into users' environments while maintaining usability and clarity. For instance, when overlaying content onto a busy environment, maintaining content visibility while avoiding intrusions to the user's interactions with other real-world or physical objects can be challenging. Failure to design a suitable interface can diminish performance (e.g., illegible presentation of instructions in a virtual training simulation), or even environment awareness and safety (e.g., occlusion of a car by virtual navigation cues when crossing the street).

Identifying the XR interface's design elements that can adapt is of utmost significance in designing efficient yet non-intrusive XR interfaces. Numerous adaptation strategies can be deployed within an interface to utilize various contextual information. Consider a scenario in which Miles and Alex meet at the library to review the course material together and prepare for an exam. Overlaying documents on Alex's face can disrupt Miles' social interactions with Alex. In this context, an optimal interface tracks Alex's face and adapts to prevent its occlusion.

However, the mere presence of such adaptations does not necessarily ensure optimal effectiveness. The efficiency of specific adaptive and non-adaptive interfaces is contingent upon the specific context of their usage [10, 26, 37]. For instance, in the scenario above, after reviewing

*e-mail: sdavari@vt.edu

†e-mail: dbowman@vt.edu

the material, Miles and Alex decide to study the next chapter separately. While the overall scenario remains unchanged, preventing occlusion of Alex's face can now lead to suboptimal placement, inconsistency, and distraction, as Miles is no longer interacting with him. Conversely, a non-adaptive placement may occlude Alex's face, yet consistently offer efficient access and minimize distractions. Designing an effective XR interface requires identifying the relevant XR content and their optimal presentation and interaction within various contexts [10].

Towards context-aware adaptation in XR, we explore three fundamental Research Questions (**RQs**) investigating the XR design space (**RQ1**), utilizing this design space to propose an adaptive spatial placement strategy (**RQ2**), and exploring context-awareness through empirical evaluation of a non-adaptive and an adaptive placement strategy in various contexts (**RQ3**).

RQ1: What are the **design elements** of an XR interface?

An XR interface is composed of a set of interactable XR objects that present information in various modalities. Adapting the design elements of these XR objects enables context-aware interfaces that optimize the provided XR content and its presentation within a specific context. In [Section 3](#) we examine potential XR adaptations and outline a design space for XR.

RQ2: How can we adapt the **spatial placement** of XR objects utilizing context?

The inappropriate spatial layout of XR objects can lead to distractions and occlusions, overwhelm the user, and hinder their interactions with the environment. [Section 4](#) proposes an adaptive placement strategy utilizing an intermediary object from the environment within a hybrid frame of reference. This placement strategy leverages contextual information to maintain efficiency and unobtrusiveness within context-switching scenarios.

RQ3: Does adaptive XR placement *always* enhance performance?

While an adaptive placement strategy adjusts to specific contextual information, relying solely on one or a few context components may be insufficient for optimal performance. The people, objects, events, and elements within the environment can affect the user's behavior, information access needs, and interests. Similarly, one's physical, cognitive, and emotional states influence perception and interactions within the environment, changing perspective, spatial awareness, directional perception, and distance from different objects. Varying the setting and the user state can influence the effectiveness and performance of an adaptive placement strategy within a scenario. [Section 5](#) designs an experiment to explore the nuanced interplay between these contextual components and the effectiveness of placement strategy. This experimental design evaluates the accuracy, efficiency, and usability of our proposed adaptive placement strategy and a traditional body-fixed placement across four distinct contexts. [Section 6](#) presents the results from our user study within *Dynamic Mobile*, *Static Mobile*, *Dynamic Stationary*, and *Static Stationary* contexts.

2 RELATED WORK

Previous work explored various innovative applications and design dimensions of XR to spatially place 2D & 3D objects and transition between them, visualize hierarchies, and provide persistent and portable presentation of the personal information [14, 16, 20, 34, 38]. Morrison et al. highlighted unique design elements within AR for enhancing accessibility for visually impaired children [45]. These studies underscore the broad design space of XR interfaces and the versatile and transformative applications that XR enables across various contexts. This work, investigates previous work and the design elements they utilized, providing a comprehensive XR design space.

XR offers the potential to enable efficient information access, reduce cognitive load, and enhance user convenience compared to traditional methods such as mobile phones [11, 12, 40, 43, 55]. However, intrusive XR interfaces may result in challenges such as information overload

and occlusion of important cues within the environment [4, 31], increase cognitive load and discomfort, and reduce the user's situational awareness and performance [17, 25, 54]. Various approaches for intuitive and seamless integration of XR content into the environment have been extensively explored. For instance, to enhance efficiency and minimize intrusiveness, numerous designs adapt the XR content's availability, transparency, placement, and Level of Detail (LoD) [3, 15, 46], as well as spatial layout and size [9, 19]. Lages & Bowman highlighted the significance of adapting the AR content placement strategy to avoid occlusions and accommodate activities like walking [33]. For adaptations to the XR content placement, the concept of the frame of reference, also referred to as fixation was introduced [20].

User-triggered adaptation through gaze, hand, and head-based inputs such as finger taps and handheld controllers are extensively explored for adjustments to the transparency, LoD, and spatial layout of XR content [19, 32, 39, 47, 48, 50]. In AR, for instance, many applications were designed to prioritize the real-world [11] by initially keeping the XR content hidden, in the peripheral, or at a lower LoD, and granting access to them through explicit interactions [42, 46, 49]. However, AR's definition emphasizes the *integration of the digital content into the real world* [1], underscoring the significance of context awareness.

While user-triggered adaptations offer control and predictability, they increase the user's physical and mental workload of deciding when, what, and how to apply the adaptations [51]. Automatic XR adaptation can enhance efficiency and reduce workload compared to the user-triggered ones [11]. Numerous studies suggest rule-based approaches for XR adaptations. The significance of such rule-based adaptations in meeting the XR task requirements within various applications such as driving and conversation have been highlighted [4, 12]. To prevent occlusion issues, Ens et al. suggested a rule-based adaptive design to exclusively place the XR objects on empty surfaces [18]. Constraints, explicitly imposed by the users, were utilized as guidelines to group related XR objects together and prevent their occlusion within a rule-based view management [3]. Such rule-based adaptive approaches are highly tailored to specific use cases and applications. Even within the same application or use case, slight contextual deviations can cause a rule to fail, making it suitable only within unchanging contexts. This work proposes an adaptive placement strategy, applicable within changing contexts, to extract and utilize contextual information from the environment and user state to spatially place the XR content.

Context refers to the external components that influence or relate to the user's interactions with the interface [13]. In recent years, context-aware XR has become a focal point of research, promising the potential for "ubiquitous" and "pervasive" computing through AR [26, 56]. Contextual aspects such as user preferences, cognitive load, device profiles, task environment, semantic changes, and task-specific security parameters have been utilized for adaptations to the XR content's appearance, LoD, frame of reference, and spatial layout [9, 36, 37].

3 DESIGN SPACE OF XR INTERFACES

An intelligent XR interface dynamically adapts the content, presentation, and interaction of each XR object to contextual changes. Prior research extensively investigates design elements and adaptation techniques for XR interfaces [1, 7, 24, 52]. We draw from the existing literature to identify these design elements and propose a comprehensive design space for XR interfaces (see [Figure 2](#)).

We regard an XR object as the fundamental component of any XR interface, irrespective of its modality for presenting digital information or its potential for interactivity. Each XR object may comprise multiple other sub-objects and can provide applications, auditory notifications, 3D models, 2D images, system recommendations, menus, and the like. We divide the design elements of XR objects into two high-level categories: content design and presentation design ([Figure 2](#)).

3.1 Content Design

Previous research explored adapting the **availability** of XR content to provide the right information at the right time, enhancing performance and preventing potential obstructions and distractions in various contexts [4, 12, 37]. Effectively adapting in-use XR objects' **availability**

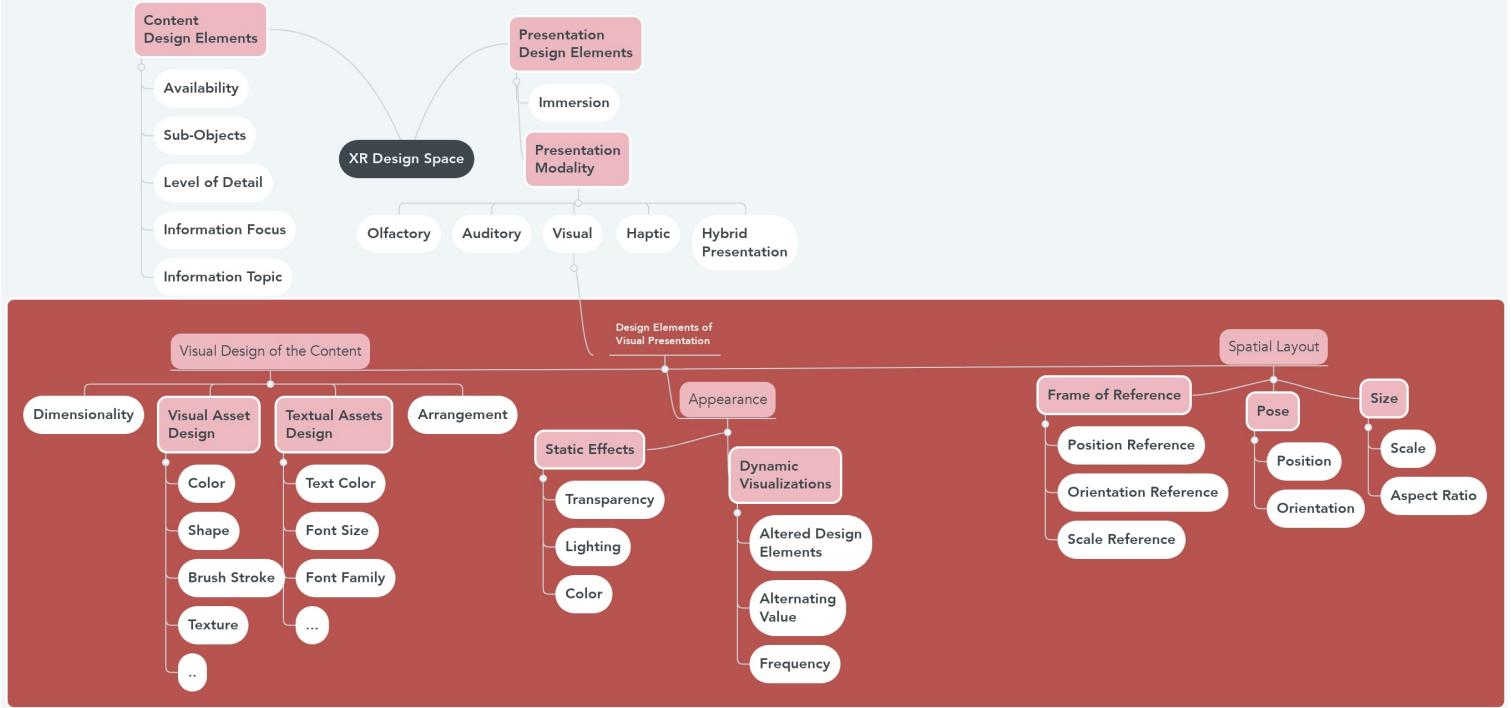


Fig. 2: Design Space for XR interfaces

allows the creation of interfaces that “close”, “open”, or “minimize” apps tailored to the user. Within the design of an XR interface, the availability of each XR object can be modified either manually by the user or through context-aware adaptations, or be immutable (e.g., a constantly “opened” system-level XR object). For instance, in educational XR aimed at facilitating student learning experiences, the availability of specific XR objects may be contingent upon the student’s location, fulfillment of predetermined criteria, or their class and exam schedule.

Each XR object is dedicated to a specific subject inherent to its essence and remains immutable. For instance, one XR object functions as a weather app while another provides recommendations or notifications to the user. However, within an interface, adaptations can also be made to the content of an XR object, varying its **sub-objects**, **Level of Detail (LoD)**, **information focus**, and **topic** [14, 15, 37]. For instance, the content within the weather app could display data focused on either the hourly forecast or the 10-day forecast, exclusively feature precipitation forecasts, or offer minimal LoD, showcasing solely visualizations of the current temperature and conditions. Adapting the content within an XR object allows for the creation of a personalized interface, offering only the intended content. Imagine an XR object that provides recommendations to the user. Context-aware adaptations to this XR object’s content facilitate the recommendation of various apps and information tailored to the user’s needs, thereby enhancing their performance. The objective of an effective XR interface is to present pertinent information in a timely and seamlessly integrated manner, ensuring it remains intuitive, easily accessible, understandable, and unobtrusive to the user. This underscores the pivotal role of content presentation in the design process of XR objects.

3.2 Presentation Design

The presentation design of an XR object involves its immersion and the modality in which it is presented (refer to Figure 2). The **immersion** design element of an XR object determines whether it is immersive or non-immersive. A non-immersive XR object allows simultaneous interaction with other XR objects. However, an immersive XR object restricts user engagement solely to itself and its sub-objects, limiting interaction with external XR objects.

Presentation Modality

Skarbez et al. propose that all technology-mediated experiences fall under the term mixed reality, suggesting that an XR object can present information through various modalities [53]. The efficacy of an XR

interface and the information it provides also relies on its presentation modality. In modern XR interfaces, XR objects are predominantly showcased as 2D and 3D overlays within the XR environment, utilizing the *visual modality*. However, XR objects can possess the capacity to convey content through *audio*, *haptic*, and *olfactory modalities*, as well as through *hybrid presentations*.

3.2.1 Visual Presentation

Extensive research in the literature has investigated the visual design of XR interfaces and their design elements. We propose a taxonomy for the visual design elements of XR objects, categorizing them into the visual design of the content of an XR object, and the XR object’s appearance, and spatial layout within the environment (Figure 2).

Visual Design of the Content

An XR object’s visual content involves a structured arrangement of various 2D or 3D visual and textual assets, impacting the clarity, readability, and comprehensibility of the XR object. The **dimensionality** design element indicates whether the XR object employs 2D or 3D visualizations to present its content, while the **arrangement** design element determines how these assets are organized and integrated to create a coherent visualization.

Visual Asset Design: Design elements such as **texture**, **brushstroke**, **size**, and **color** are instrumental in designing 2D or 3D images, models, lines, and shapes [35]. These visual assets represent either the entirety or specific aspects of an XR object’s content.

Textual Asset Design: Design elements pertinent to the typography of the text influence the legibility, readability, and aesthetic appeal of textual assets, such as text boxes [2]. These typography features encompass various design elements such as **typeface selection**, **text color**, **font and point size**, **line length**, **line and letter spacing**. Typography design elements can serve readability purposes and creative endeavors, such as conveying specific messages and establishing hierarchy, tone, and visual rhythm. For instance, context-aware adaptations to typography enable XR to accommodate low-vision or dyslexic users.

Appearance

A multitude of static and dynamic design elements collectively influence the realism, perceptibility, and noticeability of an XR object’s appearance. Previous work proposed guidelines for limiting unnecessary or non-conforming dynamic visuals and effectively using visual cues, color, and animation to enhance performance and prevent potential obstructions and distractions [4, 52]. These elements contribute to

the natural integration and alignment of visual XR objects within the XR environment.

Static Effects: Various visual effects can be applied to manipulate the **transparency**, **lighting**, and **colors** of the 2D or 3D visualization of an XR object. Some common static visual effects include night/day mode color themes and adaptations to color saturation, warmth, tint, brightness, and contrast. These adaptations significantly impact the opacity, legibility, and intuitiveness of the XR interface within its context. For example, adjusting an XR object's transparency and colors enables its seamless integration into the XR environment [3]. Similarly, aligning the shadow and brightness of an XR object with the environmental lighting conditions ensures its noticeability without overwhelming brightness, enhancing the natural appearance of the scene.

Dynamic Visualizations: Adapting at least one design element of an XR object at a specified frequency results in dynamic 2D or 3D animated visualizations within the XR object. For example, an animated XR object that dynamically moves at speed v visualizes pose adaptations to its position design element at the frequency of v . While dynamic visualizations do not necessarily follow a periodic pattern, their design elements can consist of the **altered design elements** and their **alternating values** and **alteration frequencies**. Such dynamic visualizations can enhance the overall realism, noticability, and interactivity of the XR interface, and the user's awareness by highlighting relevant information or events within specific XR objects. Consider a dynamic visualization that alters the content compositing of an XR object by altering its line colors between green and transparent at frequency f . Such blinking outlines can provide notifications about new mission assignments in an immersive VR game or indicate the occlusion of an important object in a context-aware AR [11].

Spatial Layout

The spatial integration of a visual XR object plays a crucial role in the interface's efficiency and the user's engagement and safety [19,32]. The proper placement and alignment of XR objects can facilitate and speed up access to XR content, prevent occlusions, and influence the user's perception and awareness of the virtual and real-world environments. To ensure that an XR object is displayed in a non-intrusive, realistic, and easily accessible manner within the user's field of view, it is essential to give careful consideration to the design of the frame of reference, the object's pose, and size.

Frame of Reference (FoR): A coordinate system is used as the FoR to determine the XR object's position, orientation, and scale in relation to the environment [5–7,19,20,30]. The literature has identified various FoRs for registering XR objects within the environment as world-fixed, object-fixed, head-fixed, body-fixed, or device-fixed [7]. *World-fixed FoR* registers the XR object in a fixed position relative to stationary real-world locations and objects. *Object-fixed FoR* is defined relative to another XR object or real-world object (including input devices), so that the content moves along with that object. *Head-fixed FoR* registers XR objects relative to the virtual camera. This FoR is also referred to as display-fixed, Heads-Up-Display (HUD), or fixed display. *Body-fixed FoR* registers the XR objects relative to the user's body. World-fixed, display-fixed, and body-fixed settings have been commonly used as the FoR for XR objects in AR interfaces [20]. However, in all these scenarios, the coordinate system of the world or another object is uniformly employed as the FoR for position, orientation, and scale. For example, by utilizing ref as the reference object for $O1$, any alterations to its position, orientation, and scale would apply to $O1$.

While previous research studies commonly apply a **Unified FoR**, we propose that an XR object's FoR can be **Hybrid** and adhere to a distinct **position reference**, **orientation reference**, and **scale reference**. Recent research by Manakhov et al. touches on the concept of hybrid FoRs, specifically discussing the *HeadDelay* and *Path* FoRs in the context of walking scenarios [41]. For instance, consider N as an XR object positioned on the north side of the user and displaying a pointer towards the north direction in an AR navigation interface. As its FoR, this XR object can utilize a custom coordinate system that employs the user as its $reference_{position}$ and scale reference and the world as its

orientation reference.

$$N_{FoR} = (ref_{position} : user, ref_{scale} : user, ref_{orientation} : world)$$

This FoR enables the XR object to be consistently positioned on the north side of the user and point towards the north regardless of the user's orientation while simultaneously following the user's position and maintaining a fixed distance from the user in mobile contexts.

Pose: The pose of an XR object in 3D space is specified through its **position** and its **orientation** relative to its FoR. Cartesian (x, y, z) or spherical (r, θ, ϕ) coordinates can denote the **position**, while the facing or pointing **orientation** can be indicated through Euler angles, quaternions, or rotation matrices.

Size: The **scale** of an XR object's size relative to its FoR, along with its **aspect ratio**, represents its size within the environment. Scale can be depicted by a singular value, such as a uniform scale factor, or by distinct scale factors for each dimension. The size of the XR object, including its length, width, and height within the environment, indicates the space it occupies, contributes to a sense of depth, and affects the realism and occlusion of the XR environment. Maintaining the correct aspect ratio can be essential in specific XR objects, as deviations may lead to distortion or misalignment, detrimentally affecting its integration within the environment and the user experience.

3.2.2 Non-Visual Presentation

XR objects can be presented through *audio*, *haptic*, and *olfactory modalities*, as well as through *hybrid presentations*. As XR technology advances, utilizing the full potential of non-visual outputs is crucial for enhancing realism and interaction in virtual and augmented environments. For instance, delivering content through olfactory sensations enhances the presence and plausibility of immersive XR experiences more effectively than instructing users to "imagine smelling grass" through visual text. However, employing many of these modalities necessitates specialized hardware and software, and their implementation relies on the specifications of the XR system.

This research primarily concentrates on visual presentation in AR. However, comprehending the design elements of non-visual modalities allows AR designers to develop more immersive, intuitive, interactive, and engaging experiences for users across various applications and contexts. We briefly examine the common design elements for delivering information through non-visual modalities and encourage researchers to further investigate and expand upon this line of inquiry (See Figure 3).

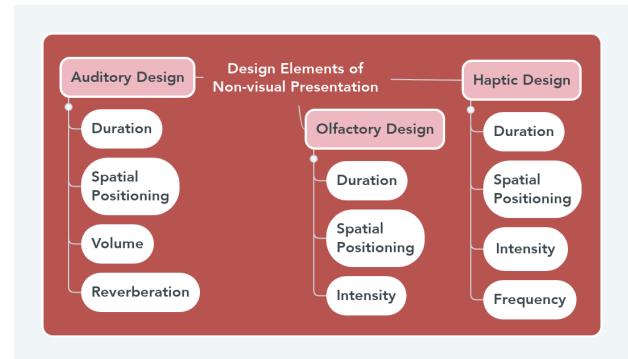


Fig. 3: XR design elements for non-visual modalities

Audio Presentation

Delivering XR object content through audio outputs, such as speech synthesis or sound effects, via the *audio modality* is relatively commonplace. Utilizing speech specifically enables intuitive communication of complex information, reduces ambiguity, and enhances accessibility for visually impaired users. We consider the design elements of audio output in XR, focusing on its duration, volume, spatialization, and

reverberation. The **duration** of audio output in an XR interface can significantly influence distractions, disturbances, and conflicts. Prolonged audio presentations may divert the user's attention away from their task and influence their awareness of the XR environment. Additionally, repetitive or non-verbal audio played for extended periods can become annoying and disruptive. Furthermore, aligning the **spatial positioning** of the audio output in 3D space with its associated XR object and surroundings enhances realism.

Matching the audio output's **volume** and intensity to the XR environment's noise level ensures audibility without disruption. Moreover, the audio output can employ reverb effects to authentically mimic the acoustic properties of the surrounding XR environment. **Reverberation** artificially creates a spatial perception by imitating how sound bounces off surfaces in different environments, thereby shaping the acoustic ambiance.

Haptic Presentation

The *haptic modality* employs tactile sensations such as vibration, force, and thermal feedback to convey diverse information. For instance, force feedback can signify pressure and weight, while subtle variations can convey texture, shape, and deformation. We explore the design elements of haptic output in XR, focusing on its duration, spatial positioning, intensity, and frequency. The **duration** of haptic sensations during XR interaction varies based on context and purpose. Brief vibrations may signal events or interactions, while prolonged feedback can simulate ongoing physical interactions or environmental cues. Moreover, the **spatial positioning** of haptic sensations in the XR environment, relative to the user's body or XR objects, affects the realism and efficacy of haptic output. Aligning haptic output with visual or auditory cues enhances immersion and intuitiveness, enabling users to perceive virtual interactions more accurately.

User engagement and immersion in XR can be influenced by the **intensity** or amplitude of haptic sensations experienced by the user. These sensations simulate a broad spectrum of tactile experiences, ranging from intense vibrations, temperatures, or weights conveying a heightened sense of impact or force to gentle taps. Similarly, the **frequency** of haptic vibrations and output can shape user perception of the content and evoke various sensory experiences. High frequencies may signal urgency, while low frequencies can mimic gentle movements or subtle textures.

Olfactory Presentation

The *olfactory modality* utilizes scent to convey information or evoke a sense of smell. It reproduces the natural scents of the virtual environment through diffusers, fans, scented materials, electronic devices, and computational methods. Due to technical complexities, such olfactory outputs remain uncommon in XR and are primarily reserved for short, stationary immersive experiences. We discuss the main design elements of olfactory output in XR, focusing on duration, spatial positioning, and intensity. Considering the **duration** of a scent within the design of an XR environment is crucial to ensuring its alignment with user interactions and the experience's narrative. On the other hand, strategic **spatial positioning** of scent sources relative to the user's position within the XR environment can heighten the sense of presence and realism. For instance, positioning floral scent emitters near XR objects depicting flowers enriches the immersive experience by reinforcing visual and olfactory cues. Additionally, a scent's **intensity** / potency, from subtle hint to powerful odor, can evoke various emotional responses or amplify specific narrative elements.

Hybrid Presentation

The hybrid presentation involves multimodal XR objects that integrate sub-objects of various modalities to present the content. Visual and auditory modalities are frequently integrated to enhance the efficacy of an XR object. For example, combining sound effects with visual notifications or background music with visual animations enhances their capacity to convey intended information to the user. Similarly, haptic and olfactory modalities are not commonly utilized individually, as the information they convey tends to be ambiguous and limited in range. However, they can be effectively employed in conjunction with visual or auditory modalities to enhance presence and plausibility in

tailored immersive XR experiences. For example, an XR object in an immersive experience can integrate haptic and visual modalities to simulate various temperatures that match the XR object's 3D visual presentation. In more general-purpose scenarios, these modalities can also provide cues or signals to the user within a multimodal XR object and increase the efficacy of information presentation.

3.3 Input Design

The efficiency and functionality of an XR interface depend on the input design and user interactions as well. Key input design elements of XR encompass the interactivity of each XR object and the selected input modality and interaction design for individual XR objects.

The **interactivity** of an XR object represents users' active engagement to provide input and adapt its design elements, including actions like clicking buttons, entering text, repositioning, or adjusting audio levels. An interactive XR interface enhances efficiency by enabling users to customize it, control information flow, and access relevant data according to their needs. An XR object can offer an adaptive level of interactivity for specific design elements. For instance, typically, the availability of XR objects is interactive, allowing the user to open and close them. However, some XR objects, for instance, static visual assets like walls in a VR game, may be designed with no interactivity even in terms of their availability, limiting the user to "close" and remove them from the environment. Additionally, intelligent XR interfaces can adapt the level of interactivity for individual XR objects and particular design elements depending on the context. For instance, they can limit the availability of specific XR objects to restrict students' content access during an exam.

Input modality and interaction design encompass the methods used to capture and interpret user input. These are crucial elements in XR interface design influencing ease of use, intuitiveness, accessibility, and efficiency. Inappropriate selection of input modality and poorly crafted interactions can result in user frustration and challenges.

Input modalities range from natural inputs like eye-based, hand-based, and voice-based methods to specialized equipment like controllers and touchpads. The interplay between input modality and user context shapes the XR experience, affecting users' perception of their virtual and physical surroundings and their ability to interact with the interface. Even within a chosen input modality, selecting the right **interaction technique** significantly impacts users' ability to navigate and utilize the XR interface, influencing its learnability, usability, and overall effectiveness.

4 ADAPTIVE ENVIRONMENT-REFERENCED PLACEMENT

A placement strategy determines the *reference_{position}* and *reference_{orientation}* of an XR object's FoR, as well as the pose of the object within its FoR. The literature has investigated different world-fixed or user-fixed placement strategies for XR objects, including surround-fixed, object-fixed, body-fixed, head-fixed, and device-fixed [7, 11, 20]. These placement strategies utilize empirical data from previous use cases to either statically pose the XR content within a uniform FoR or propose a context-specific adaptive placement. In this work, we propose an adaptive strategy for utilizing each XR object's association with the environment to dynamically place it without prior knowledge of the specific context or environment.

We propose *Environment-referenced* placement, an adaptive placement strategy that employs a hybrid FoR tailored to each XR object utilizing an *reference_{orientation}* distinct from the XR object's *reference_{position}*. Bell et al. investigated placing relevant objects in close proximity as a strategy for view management [3]. This approach intelligently identifies a relevant object or area within the environment as an intermediary for each specific XR object by leveraging machine learning (ML) and large language models (LLMs). Using each intermediary's relative position to the XR object's *reference_{position}* our placement strategy determines a distinct *reference_{orientation}* for each XR object. In essence, while the XR content moves along with and remains anchored at a consistent distance from its *reference_{position}*, it continuously rotates to maintain a persistent pose between its *reference_{position}* and intermediary (See Figure 4: **1-b** and **2-b**).

Implementing this adaptive placement strategy within a context-aware interface that can identify relevant intermediaries for the user's task and specific context allows for placing the XR objects precisely where needed, thereby enhancing the interface's efficiency. By leveraging the user's spatial awareness and focusing their attention on nearby RW and virtual objects within their visual range, this placement strategy seamlessly integrates XR objects into the environment, simplifying the interface complexity [31]. Furthermore, the intelligent detection of intermediaries within new environments and tracking their relative position to the XR object's *reference_{position}* offers an adaptive placement strategy for portable and dynamic use cases. This is particularly useful when users move to new or unfamiliar environments, or when the position of the user, *reference_{position}*, or intermediaries change within the same environment. However, in such contexts, especially when the selected intermediaries are irrelevant to the setting and the user's task, this placement strategy can be unintuitive and introduce additional mental and physical load to locate the intermediaries. Envision utilizing various placement strategies for name tags at a conference. A non-adaptive object-fixed placement strategy might use each attendee as the uniform FoR for their name tag. In this interface, the pose of each name tag depends on the specific attendee's position and orientation. This can result in the inconsistent appearance of name tags, creating a confusing interface. For instance, depending on people's relative orientation, one person's name tag might appear to their right while another's name tag appears to their left in the user's view. Additionally, an attendee's name tag might appear too small or at an angle, depending on their distance from the user or whether they are facing the user, affecting its visibility and readability. On the other hand, using the user as the uniform FoR ensures consistency in the distance and angle at which name tags appear in the user's view. However, without additional contextual information and adaptations to clearly associate each name tag with the corresponding attendee, a non-adaptive body-fixed placement strategy can result in visual clutter and overwhelm the user with excessive information.

Our placement strategy enables a solution to consider the user as the *reference_{position}* and each attendee as their intermediary to their name tag. This enables each person's name tag to appear in a uniform pose around them in the user's view. Using the user as the *reference_{position}* ensures consistency in the distance and angle at which the name tags appear in the user's view. While these XR name tags consistently follow the user, their distinct *reference_{orientation}* ensures they are visible solely when the specific attendee, i.e., intermediary, is within the user's field of view, minimizing visual clutter and information overload.

5 EXPERIMENT

To address *RQ3.*, we designed an experiment to assess information acquisition using our proposed adaptive placement strategy compared to a non-adaptive placement strategy across four distinct contexts: *Dynamic Mobile*, *Dynamic Stationary*, *Static Stationary*, and *Static Mobile*.

This experimental design aims to investigate the nuanced interplay between placement strategy and contextual components. It is important to note that, while our proposed adaptive strategy intelligently identifies intermediaries, this work does not aim to propose, implement, or evaluate the performance of ML models in identifying optimal intermediaries. Consequently, to ensure that the reliability and performance of the intermediary detection do not influence our results, within our experimental design we manually identify relevant intermediaries for the XR content within each context.

5.1 Experimental Design

Our experiment incorporated three independent variables each with two levels: *Placement Strategy* (adaptive *Environment-referenced* or non-adaptive *Body-fixed*), *Real-World (RW) Setting* (either *Dynamic* or *Static*), and *User State (Mobile or Stationary)*. We conducted a within-subjects user-study to evaluate the performance of each *Placement Strategy* within each combination of *RW Setting* and *User State*. Thus, each participant's performance was evaluated across eight sessions.

5.1.1 Study Task

The challenges associated with information acquisition and improper placement of XR content are exacerbated when dealing with a large amount of content. We devised a trivia game about 28 different countries to assess information acquisition within a large collection of AR documents. The trivia questions covered three categories: international food, movies, and sports. Our AR interface displayed a total of 84 documents, one for each category and country. The documents were displayed across three panels in the interface, one for each category. The user's objective was to identify the category and country mentioned in the question and open the correct document. All questions explicitly stated both the category name and the country name. The question was repeated every 30 seconds until the correct document was opened.

5.1.2 RW Setting

The user was positioned in an empty room featuring three designated intermediary objects, each representing one of our categories.

Static Individual Setting: In the *Static* contexts, the user participated in a trivia game alone in the room, where questions were displayed on an AR screen in front of them. In these *Static* contexts, the RW environment involved three virtual posters, each displaying images relevant to a specific category, and remained static.

Dynamic Social Setting: In the *Dynamic* contexts, three virtual humans acted as trivia hosts. These trivia hosts engaged in dialogue with the user and each other, moved around the room, and switched positions throughout the session. The virtual avatars were each attired in clothing pertinent to one of our categories: a chef, a soccer player, and a director. Each host presented questions aligned with their respective category.

5.1.3 User State

At the beginning of each session, the user was standing at the *Start Point*, facing the sports intermediary.

Stationary User: In the *Stationary* contexts, the participant remained consistently positioned at the *Start Point*.

Mobile User: In contrast, in the *Dynamic* contexts, participants were prompted to move near a different intermediary at specific times during each session.

5.1.4 Placement Strategy

We designed a non-adaptive and adaptive placement strategy to display these panels (Refer to [Figure 4](#)). Variations in the spatial layout design elements of the panels influence not only their placement but also their appearance, occlusion, and ease of interaction. This emphasizes the importance of consistently presenting the panels at a uniform size, height, and distance within the user's view to ensure that our results are solely due to placement. Additionally, to access the panels in *Mobile* contexts, they must follow the user. To address these concerns, in both placement strategies, we used the user as the *reference_{position}* and designed all category panels in a uniform *pose* and *size*.

Environment-referenced Placement: As described in [Section 4](#), we implement an adaptive placement strategy to determine an individual *reference_{orientation}* for each category panel utilizing its intermediary's relative position to the user. In our design, we placed each panel directly between the user and its intermediary (See [Figure 6](#)). To ensure that intermediary detection's performance does not influence our results, within each setting, we manually assigned each panel's relevant intermediary. In *Static* contexts, the virtual poster, displaying images relevant to each category, served as that category's intermediary.

In social interactions, visual engagement with conversation partners is crucial. Research by Hood et al. demonstrates that individuals tend to gaze at their conversation partners for a significant portion of the interaction [28]. Miller et al. revealed that the mere presence of virtual agents influences the user's non-verbal and social behavior, with participants avoiding avertion of their gaze from avatars and treating

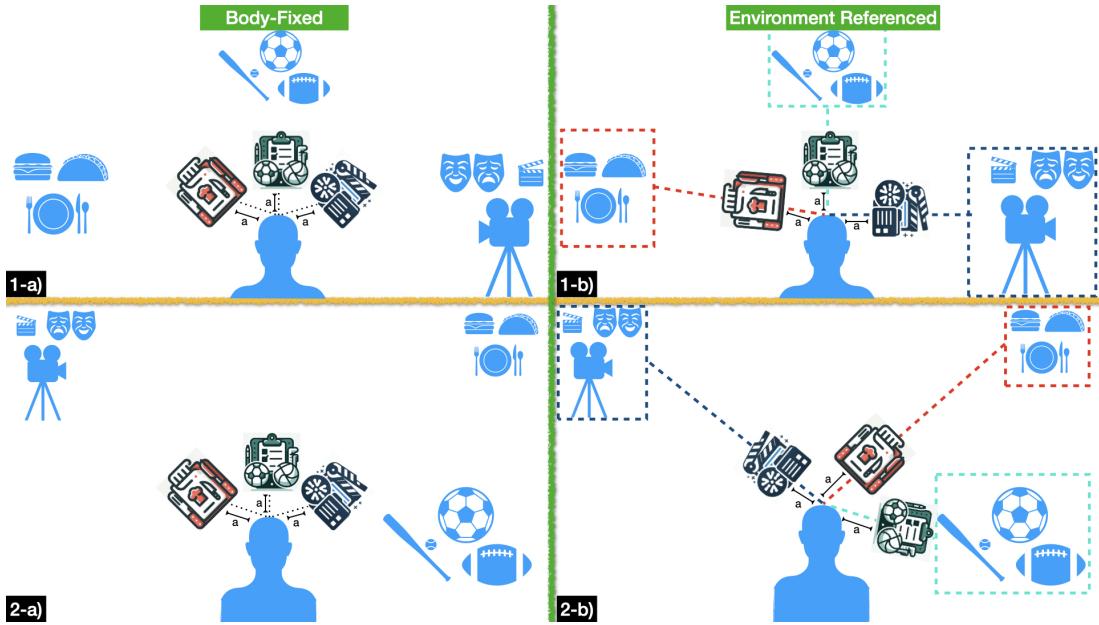


Fig. 4: Illustration of our two placement strategies. (1-a, 2-a): Body-fixed placement: each XR object's pose remains consistent based on the orientation of the user, i.e., frame of reference. (1-b, 2-b): Environment-referenced placement: each XR object's pose remains consistent based on the orientation of its intermediary relative to the frame of reference, maintaining a consistent distance from the user. (1-a, 1-b): The user is at the session start point, positioned at an equal distance from the intermediaries. (2-a, 2-b): The user has moved near the Sports category's intermediary.



Fig. 5: Each document displays the answer to a trivia question

them as real individuals [44]. Consequently, in *Dynamic* contexts, we considered each trivia host as the intermediary for their category panel.

Body-fixed Placement: As a baseline non-adaptive placement strategy, we implemented a commonly-used *Body-fixed Placement*, in which the user is considered as the uniform FoR. By using the orientation of the user's body as the *reference orientation* for all panels, within this non-adaptive placement, a particular panel is consistently placed on the same side of the user's body. In our implementation, we

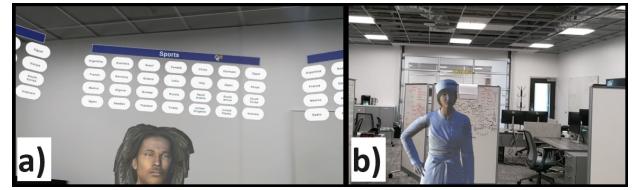


Fig. 7: Body-fixed placement in dynamic contexts: a) All three category panels in front of the user's body. b) No AR content is placed near the trivia hosts, not directly in front of the user.

situated the sports panel directly in front of the user, the food panel to their right side, and the Movies panel to their left (See Figure 7).

5.2 Hypotheses

Both adaptive and non-adaptive placement strategies have unique advantages that influence their effectiveness for the same task. Regarding *RQ3*, we hypothesize that neither of the two placement strategies would be singularly superior for information acquisition in all contexts, as their advantages and disadvantages can vary depending on the specific context. To explore the impact of different context components on the effectiveness of these placement strategies for information acquisition, we formulated and tested three hypotheses denoted as **H1, H2, and H3**.

H1. Placement strategy effectiveness varies across different contexts.

When utilizing the appropriate contextual information for adaptation, the *Environment-referenced* placement facilitates rapid access to information, by placing each XR object where it is needed. Within this experiment, this adaptive strategy can lead to intuitive placement and optimal access to information with no added cognitive load for navigation between category panels. On the other hand, *Body-fixed* placement facilitates constant and rapid access to all category panels at all times, regardless of the *RW Setting* or *User State*, or the suitability of the intermediaries. However, this strategy's non-adaptive placement increases the decision-making and mental load associated with

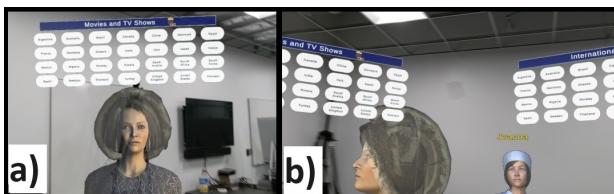


Fig. 6: Environment-referenced placement in dynamic contexts: a) The category panel relevant to each trivia host is displayed above their head. b) the panels move along with their associated trivia host.

navigating the correct category panel.

H2. The effectiveness of Environment-referenced placement strategy is influenced by RW Setting.

The people, objects, events, and elements within the RW setting can affect user behavior, information needs, and interests.

We evaluated the performance of adaptive and non-adaptive placement strategies within various *RW Settings*, hypothesizing that *Environment-referenced*:

H2-a outperform Body-fixed in Dynamic contexts.

In social contexts, information acquisitions are usually associated with their conversation and interlocutor with whom the users typically maintain eye contact [28]. Similarly, in both *Dynamic* contexts within this experiment, the information acquisition inquiries are associated with the same trivia host with whom the user is conversing and visually engaged. Additionally, the user's engagement in social and conversational interactions, along with the dynamic interactions between the trivia hosts, can increase the user's mental load in our *Dynamic* contexts. By considering the conversation partner as the intermediary in these contexts, *Environment-referenced* intuitively displays solely the required category panel where needed. We hypothesize that in these *Dynamic* contexts, the *Environment-referenced* strategy facilitates the cognitive and physical load of finding and accessing the relevant panels and enhances information access.

H2-b underperforms Body-fixed in Static contexts.

The advantages of *Environment-referenced* rely on its potential to place the AR content precisely where needed. On the other hand, the suboptimal selection of intermediaries can make this placement unintuitive and lead to increased cognitive load and physical effort for information access. Unlike the *Dynamic* contexts, in our *Static* ones, the user's visual attention within the environment varies, posing challenges in anticipating the optimal intermediaries. Additionally, due to the unchanging environment and the *solo* social setting, our *Static* contexts impose a lower mental load on the user. Considering the limited number of category panels, this higher cognitive capacity can facilitate decision-making associated with the *Body-fixed* placement. We speculate that the *Body-fixed* placement's consistency and effortless access to all documents through a glance, outweigh its cognitive decision-making in *Static* contexts. Consequently, we hypothesize that in *Static* context, the *Body-fixed* placement's consistent glanceable access enhances information acquisition, compared to *Environment-referenced*.

H3. Environment-referenced underperforms Body-fixed in Mobile contexts.

The user's physical, cognitive, and emotional states influence their capability to interact with and perceive the RW environment, including their perspective, spatial awareness, directional perception, and distance from different objects. While within our *Stationary* contexts, the user's perception and interaction remain unchanged, in our *Mobile* contexts, users move across different locations within the environment.

When utilizing *Environment-referenced* placement, depending on the user's position, the relative position of each category panel changes, potentially appearing on their left side at one time and on their right side at another. The user's varying perspective and directional perception of the environment can lead to inconsistency and confuse the user. In our *Mobile* contexts, the *Environment-referenced* placement requires spatial recall for the users to determine their current perspective and the relative position of each intermediary, increasing their cognitive load, especially within unfamiliar environments. Conversely, regardless of the user's mobility and location, *Body-fixed* placement ensures a persistent approach to position each category panel relative to the user. We evaluate the effectiveness of each placement strategy and hypothesize that the consistency of the non-adaptive *Body-fixed* placement enhances information acquisition in *Mobile* contexts.

5.3 Apparatus

Throughout the experiment, participants utilized a Microsoft HoloLens (2nd gen) Augmented Reality Head-Worn Display (AR HWD). This

wireless apparatus has a display density of 47 pixels per degree and a diagonal field of view spanning 52 degrees. Our virtual environment and task-specific panels and documents were developed using the Unity 2022.3.4f1 game engine in conjunction with Microsoft's MRTK toolkit.

All sessions started with the user positioned at the *Start Point*, situated at an equal distance from the intermediaries (**Figure 4: 1-a and 1-b**). The participant was directly facing the sports intermediary, with the food intermediary to their left, and the movies intermediary to their right. We designed eight unique trial sets, each serving as the script for one session. Each *Stationary* session comprised three trials, with one trial for each category. In all *Stationary* trials the user was far (farther than 1.5 m) from all intermediaries. In specific *Mobile* trials, the user was positioned near the intermediary relevant to the trial's question. To ensure the independence of our trial results from this distance, the *Mobile* sessions consisted of six trials, with two trials for each category, one when the user was near its intermediary, and another one when far.

The interface provided three panels, each containing 28 documents corresponding to one of our three trivia categories. These documents were organized alphabetically in a grid format, with four rows and seven columns within each panel. All three panels shared a uniform appearance, except for the panel header, which showcased the category title along with an icon representing that category. Participants utilized hand interactions facilitated by the HoloLens device to select and interact with documents, enabling them to open and close them as needed. Each document displayed a single word providing the answer to the question, closing once the user clicked anywhere on it.

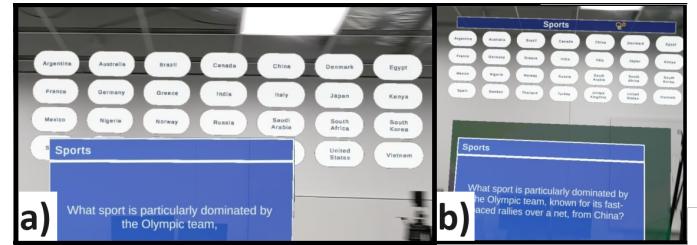


Fig. 8: a) The questions are presented word-by-word. b) The header for each category panel reappears after the question is asked.

In the *Static* contexts, a screen was displayed in front of the user at eye level to present the trivia question (**Figure 8**). In the *Dynamic* contexts, a trivia host engaged in a conversation with the user and verbally asked the question. Consequently, the questions were presented differently across contexts, providing text in the *Static* contexts while audio in the *Dynamic* ones. We employed multiple methods to prevent participants from initiating document searches during the question in specific contexts and ensure consistency across sessions. First, all questions were formulated to reveal the name of the country as the concluding word of the question. Second, in *Dynamic* contexts, participants were instructed to maintain eye contact during questions and to respond only after the entire question was asked. Third, we matched the rate at which the question was presented across various *RW Settings*. Using the duration of each question's audio within the *Dynamic* contexts, its equivalent question was revealed one word at a time in the *Static* contexts. Fourth, while presenting the question, the panel headers were rendered transparent.

Additionally, within *Dynamic* contexts, to enhance the ecological validity of social interactions and alleviate any biases arising from individual differences and social behavior among participants, we instructed all participants to engage with the trivia hosts as if they were real people and to maintain eye contact during conversations. These trivia hosts were designed using avatars and animations sourced from the Microsoft Rocketbox library and Maximo [22, 23].

5.4 Procedure

Following approval from our university's Institutional Review Board, we enlisted participants from our local university. The experiment lasted 120 minutes and was divided into four parts, each session corresponding

to a distinct context. Within each context, each participant underwent two sessions, each employing one of our placement strategies. Overall, each participant underwent eight sessions.

The presentation order of the sessions was counterbalanced to mitigate any potential bias arising from a specific trial set or the order in which participants experienced the session. To counterbalance the ordering of the sessions, participants were presented with the four contexts in a Latin Square ordering. Within each context ordering, half of the participants began with *Body-fixed* placement, while the other half started with *Environment-referenced*. We recruited 24 subjects, ensuring three participants per session-ordering ($3 \times 4 \times 2$).

Upon their arrival, participants were asked to carefully review and sign the consent form. Subsequently, we gathered demographic details and information regarding their prior experience with XR. Once this process was completed, we provided a concise overview of the study, acquainted participants with the device, and instructed them on its operation and interaction methods. They then proceeded to engage in the four parts of the experiment.

We started each part by introducing the particular task associated with that context. Within each part, participants experienced each *Placement Strategy* in a separate session. Prior to each session, they underwent a training session, completing a minimum of six trials to acquaint themselves with the environment, tasks, and specific content arrangement offered by the *Placement Strategy* within that context. Following each session, participants were directed to complete the post-session survey on a computer. During the post-context interview conducted after trying both placement strategies within a context, participants were prompted to compare and rank the two *Placement Strategy* options based on speed, ease of use, intrusiveness, and preference.

At the conclusion of the study, all participants underwent interviews regarding their overall experiences across all contexts, as well as the advantages and drawbacks of each *Placement Strategy* within each combination of *RW Setting* and *User State*. Participants had the freedom to take breaks between sessions as needed.

5.5 Evaluation Metrics

The impact of XR on mental workload and task performance is frequently assessed through objective metrics like task completion time and error rate, as well as subjective evaluations like NASA TLX [29]. To explore information acquisition using our proposed placement strategies within these contexts, we evaluated the Accuracy, Efficiency, and Usability of each *Placement Strategy* for information acquisition. To assess these metrics, we collected objective task-specific data from each trial and subjective data from survey and interview responses.

5.5.1 Accuracy

A critical aspect in determining the effectiveness of an AR placement strategy for information acquisition is its potential to maintain task performance accuracy. In this experiment, a trial continued until the participant opened the correct document. We regard every instance of opening the wrong document as an error, adversely affecting accuracy. To assess accuracy, we evaluated *Error Rate* based on the total number of trials in which any errors occurred.

5.5.2 Efficiency

An effective placement strategy facilitates efficient and swift information acquisition through a large quantity of AR documents and minimizes the associated physical and mental load. We assessed information acquisition efficiency through various metrics evaluating the time and difficulty of finding the correct document within each session.

Navigation Time: utilized the logged time-stamped gaze data to measure the duration of time for answering a question. This time was measured from the moment a question was completely presented to the user for the first time until the user's gaze fixated on the correct document. Unfortunately, we observed that the accuracy of the device in recognizing hand gestures was significantly influenced by the participant's race and skin color, resulting in longer times to open a document for certain participants. To remove this bias from our data,

we measured *Navigation Time* using the gaze data rather than the time at which a document was opened.

Gaze Switch: measured the frequency of gaze switches among the category panels within each trial. An increase in the frequency of gaze switches suggests a higher head-turn rate and physical load for accessing the document. Additionally, this measurement can be indicative of the unintuitiveness of the placement and the user's difficulty in navigating the correct category panel.

Perceived Speed: quantified the participants' response to the post-context interview question about ranking the fastest placement strategy for information acquisition within that context.

Perceived Ease: quantified the participants' post-context ranking of the easiest placement strategy for information acquisition.

NASA TLX: collected the post-session response to the short version of the Standard NASA Task Load Index (TLX) questionnaire [27].

5.5.3 Usability

To assess the usability of our placement strategies within these contexts:

Perceived Intrusiveness: quantified the post-context ranking of the most intrusive placement strategy within each context.

Preferred Placement: quantified the ranking of the most-preferred strategy within each context.

SUS: collected participants' responses to the short version of the Standard System Usability Survey (SUS) [8] in our post-session survey.

5.6 Participants

We enrolled 24 participants. This population aged between 18 and 45 years ($M = 25.62, SD = 5.24$), and included eight female and thirteen graduate students. 13 participants had little to no experience with XR, and only 5 had used XR more than ten times.

6 USER STUDY RESULTS

To test our hypotheses, we conducted a series of analyses on our quantitative data to assess the effects of our three independent variables: *RW Setting*, *User State*, and *Placement Strategy*. Our data, collected from 24 participants in a within-subjects study comprising eight sessions, represents a sizeable sample size. Considering our large sample size ($n > 50$), we conducted Kolmogorov-Smirnov (KS) goodness-of-fit tests on our measurements, revealing the non-normal distribution of our data.

We opted to conduct nonparametric statistical analysis using Aligned Rank Transform (ART) [57]. ART enables parametric statistical analyses, such as Repeated Measures Analysis of Variances (RM-ANOVA), for both continuous and ordinal responses, without the assumption of normality or limiting the number of involved factors. For all metrics, we applied ART to the data and conducted a factorial non-parametric ART RM-ANOVA to test the significant effect of our independent variables. For all pairwise comparisons on the interactions that showed a significant effect, we applied the ART-C algorithm for multifactor post-hoc contrast tests with Bonferroni adjustments. We used an α level of 0.05 in all significance tests.

6.1 Accuracy

Our analyses indicate that *Error Rate* is significantly influenced by *RW Setting* X *Placement Strategy* ($F_{(1,161)} = 5.25, p \sim 0.02$). Pairwise comparisons on the effect of *RW Setting* on *Placement Strategy Error Rates* are illustrated in Figure 9. These results indicate significantly lower *Error Rate* when using *Environment-referenced*, compared to *Body-fixed* in both **a**) in *Dynamic* contexts ($p \sim 0.01$, supporting **H2-a**), and **b**) in *Static* contexts ($p < 0.0001$, rejecting **H2-b**).

6.2 Efficiency

Overall analysis of our quantitative and qualitative data on the *Placement Strategy* efficiency within various contexts, points toward a significant influence of the interaction between *RW Setting* and *Placement Strategy*, supporting **H2**.

Navigation Time:

Information acquisition within correct trials, where no errors occurred, significantly affect *RW Setting* X *Placement Strategy* on *Navigation*

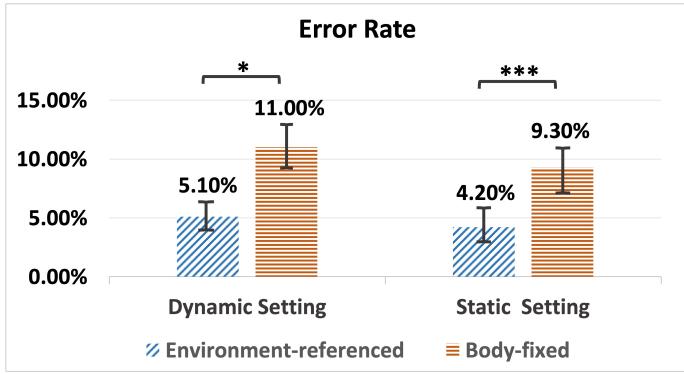


Fig. 9: Accuracy within various RW settings

Time ($F_{(1,769.53)} = 22.32, p < 0.0001$). Our pairwise comparison on this interaction indicated that *Environment-referenced* performed **a)** statistically faster than *Body-fixed* in *Dynamic* contexts ($p < 0.0001$, supporting **H2-a**). **b)** significantly faster in *Dynamic* contexts compared to *Static* ones ($p < 0.0001$ supporting **H1**) (See Figure 10).

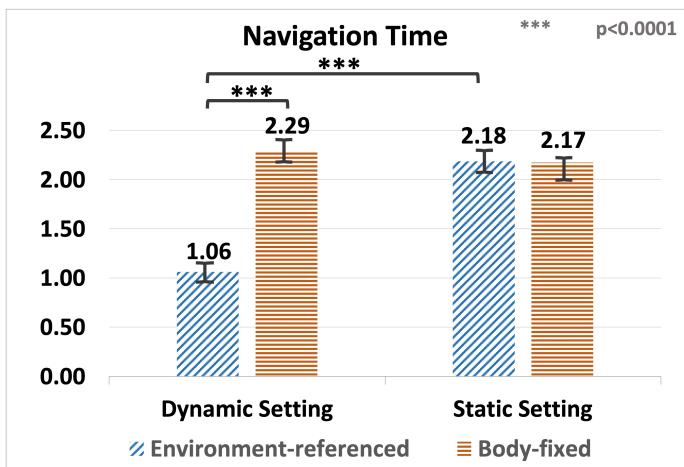


Fig. 10: Time efficiency within various RW Settings

Gaze Switch:

Figure 11 indicates that *Gaze Switch* fluctuations within correct trials, is significantly affected by *RW Setting X User State X Placement Strategy* ($F_{(1,770.32)} = 18.6, p < 0.0001$), *RW Setting X Placement Strategy* ($F_{(1,770.91)} = 154.24, p < 0.0001$), and *User State X Placement Strategy* ($F_{(1,769.8)} = 56.69, p < 0.0001$).

Pairwise comparison on *RW Setting X User State X Placement Strategy* revealed that gaze switches between the category panels when using *Environment-referenced* placement were significantly fewer **a)** compared to *Body-fixed* in the *Dynamic Mobile* ($p < 0.0001$), *Static Mobile* ($p \sim 0.001$), and *Dynamic Stationary* ($p \sim 0.009$) contexts, and **b)** in the *Dynamic Mobile* context compared to the *Static Mobile* one ($p < 0.0001$). Additionally, this analysis indicated that gaze fluctuations when using *Body-fixed* placement were significantly fewer in **c)** the *Dynamic Stationary* context rather than the *Dynamic Mobile* ($p \sim 0.005$), **d)** the *Static Stationary* context rather than the *Static Mobile* ($p \sim 0.035$), **e)** in the *Static Mobile* context compared to the *Dynamic Mobile* ($p \sim 0.006$).

These results revealed that compared to *Body-fixed*, the *Environment-referenced* placement reduces gaze switches in *Dynamic* contexts ($p < 0.0001$, supporting **H2-a**), *Static* ones ($p < 0.0001$, rejecting **H2-b**), and *Mobile* contexts ($p < 0.0001$, rejecting **H3**).

Perceived Speed:

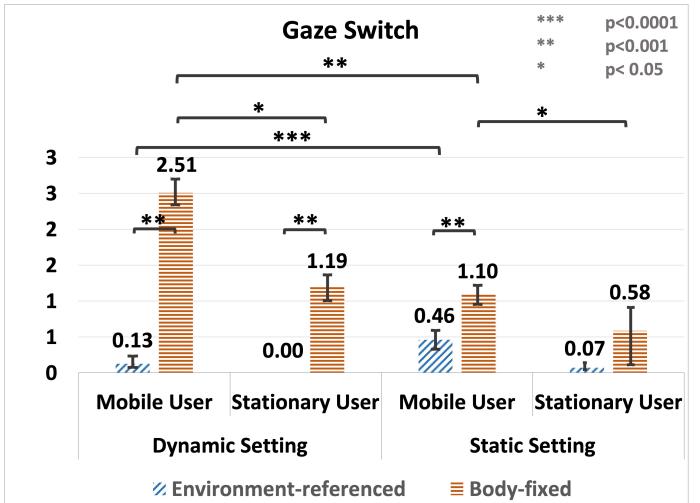


Fig. 11: Gaze switch efficiency within various contexts

Participants' post-context ranking of the fastest placement strategy revealed a significant interaction effect of *RW Setting X Placement Strategy* ($F_{(1,161)} = 135.01, p < 0.0001$). Pairwise comparison on the effect of these interactions revealed that *Environment-referenced* was ranked **a)** faster than *Body-fixed* in the *Dynamic* contexts by 86% (supporting **H2-a**), **b)** slower than *Body-fixed* in the *Static* contexts by 84% ($p < 0.0001$) (supporting **H2-b**), and **c)** faster in the *Dynamic* contexts compared *Static* ones ($p < 0.0001$, supporting **H1**). Additionally, **d)** *Body-fixed* was ranked higher in *Static* contexts compared to *Dynamic* ones ($p < 0.0001$, supporting **H1**).

Perceived Ease:

Participants' post-context ranking of the easiest placement strategies revealed a significant interaction effect of *RW Setting X Placement Strategy* ($F_{(1,161)} = 231.4, p < 0.0001$). *Environment-referenced* was ranked **a)** easier than *Body-fixed* in the *Dynamic* contexts by 93% (supporting **H2-a**), **b)** less easy than *Body-fixed* in the *Static* contexts by 80% ($p < 0.0001$) (supporting **H2-b**). Indicating that **c)** *Environment-referenced* placement was perceived as easier in *Dynamic* contexts compared to *Static* ones ($p < 0.0001$, supporting **H1**), while **d)** *Body-fixed* placement was perceived significantly more difficult in *Static* contexts compared to *Dynamic* ones ($p < 0.0001$, supporting **H1**).

NASA TLX:

Our pairwise comparison on the effect of *RW Setting X Placement Strategy* on the **TLX Score** ($F_{(1,161.06)} = 11.68, p < 0.001$) indicates significantly lower overall task loads, **a)** in *Dynamic* contexts, when using *Environment-referenced* compares to *Body-fixed* ($p \sim 0.001$, supporting **H2-a**), and **b)** when utilizing *Body-fixed* in *Static* contexts rather than *Dynamic* ones ($p \sim 0.002$, supporting **H2-b**).

Analyzing the metrics from within the TLX questionnaire identified that **Effort** significantly varied dependent on the specific *RW Setting X User State X Placement Strategy* ($F_{(1,161.08)} = 4.51, p \sim 0.03$, providing support for **H1** and **H2**), and *RW Setting X Placement Strategy* significantly impact **Physical Demand** ($F_{(1,161.08)} = 6.25, p \sim 0.01$), **Mental Demand** ($F_{(1,161.10)} = 6.24, p \sim 0.01$), and **Effort** ($F_{(1,161.08)} = 10.4, p \sim 0.001$). Pairwise comparison on these metrics shows that:

Significantly less **Effort** was required **a)** in *Dynamic* contexts, when using *Environment-referenced* compared to *Body-fixed* ($p < 0.001$, supporting **H2-a**), **b)** when utilizing *Body-fixed* in *Dynamic* contexts rather than *Static* ones ($p \sim 0.04$, supporting **H1**), and **c)** when using *Environment-referenced* in *Static Stationary* contexts, rather than *Static Mobile* ones ($p \sim 0.02$, providing partial support for **H3**).

Mental Demand is significantly decreased, **a)** in *Dynamic* contexts, when using *Environment-referenced* compared to *Body-fixed* ($p \sim 0.001$, supporting **H2-a**), and **b)** when using *Body-fixed* in *Static* contexts rather than *Dynamic* ones ($p \sim 0.01$, supporting **H1**).

Physical Demand is significantly lower, **a)** in *Static* contexts, when using *Body-fixed* compared to *Environment-referenced* ($p \sim 0.02$, supporting **H2-b**), and **b)** when utilizing *Body-fixed* in *Static* contexts rather than *Dynamic* ones ($p \sim 0.01$, supporting **H1**).

6.3 Usability

Our Analysis indicates a significant interaction effect of *RW Setting X Placement Strategy* for *Perceived Intrusiveness* ($F_{(1,161)} = 34.82$, $p < 0.0001$), *Preferred Placement* ($F_{(1,161)} = 246.45$, $p < 0.0001$), and *Overall SUS Scores* ($F_{(1,161.05)} = 11.6$, $p < 0.001$).

Perceived Intrusiveness:

RW Setting X Placement Strategy significantly influenced *Perceived Intrusiveness* ($F_{(1,161)} = 34.82$, $p < 0.0001$). The majority of our participants (79.2% within the *Static* contexts, and 56.3% within the *Dynamic* ones), insisted that both placement strategies were non-intrusive and refused to choose any one as the most intrusive. In *Static* contexts, only 6.3% ranked *Body-fixed* as intrusive. On the contrary, only 4.2% ranked *Environment-referenced* as intrusive for *Dynamic* contexts. Our pairwise comparison indicates that **a)** in *Dynamic* contexts, *Environment-referenced* is significantly less intrusive than *Body-fixed* ($p < 0.0001$, supporting **H2-a**), **b)** *Environment-referenced* is significantly less intrusive in *Dynamic* contexts compared to the *Static* ones ($p < 0.0001$, supporting **H1**), and **c)** *Body-fixed* is significantly less intrusive in *Static* contexts compared to the *Dynamic* ones ($p < 0.0001$, supporting **H2-b**).

Preferred Placement:

ART RM-ANOVA results on *Perceived Intrusiveness* indicate a significant interaction effect of *RW Setting X Placement Strategy* ($F_{(1,161)} = 246.45$, $p < 0.0001$). 90% of our participants ranked *Environment-referenced* as the most preferred placement strategy for information acquisition in *Dynamic* contexts, and 82% ranked *Body-fixed* as the most preferred for *Static* contexts. Our pairwise comparison indicates that **a)** in *Dynamic* contexts, significantly more participants preferred *Environment-referenced* over *Body-fixed* ($p < 0.0001$, supporting **H2-a**), **b)** in *Static* contexts, significantly more participants preferred *Body-fixed* over *Environment-referenced* ($p < 0.0001$, supporting **H2-b**),

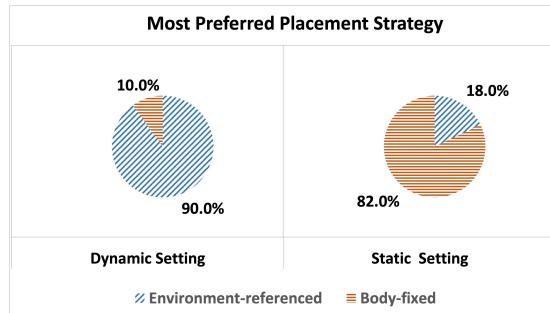


Fig. 12: User preference on placement strategy in various RW settings

c) *Environment-referenced* is significantly more preferred for *Dynamic* contexts compared to *Static* ones ($p < 0.0001$, supporting **H1**), and **d)** using *Body-fixed* is significantly more preferred for *Static* contexts compared to *Dynamic* ones ($p < 0.0001$, supporting **H1**).

SUS:

Our Analysis of the SUS responses indicate significant effect of *RW Setting X Placement Strategy* on the **overall Usability Score** ($F_{(1,161.05)} = 11.6$, $p < 0.001$). Delving deeper into the metrics from within SUS, found that other than the *overall Usability Score*, more specifically, *User Satisfaction* ($F_{(1,161.12)} = 16.37$, $p < 0.0001$), *Synergy* ($F_{(1,161.1)} = 8.62$, $p \sim 0.003$), *Consistency* ($F_{(1,161.07)} = 7.4$, $p \sim 0.007$), *Learnability* ($F_{(1,161.09)} = 3.62$, $p \sim 0.05$), and *Effortlessness* ($F_{(1,161.13)} = 5.55$, $p \sim 0.02$) were significantly affected by *RW*

Setting X Placement Strategy. Pairwise comparisons on these measurements provide support for **H2-a**, indicating that in *Dynamic* contexts, *Environment-referenced* significantly improves overall usability score ($p < 0.0001$) and provides a significantly more satisfactory ($p < 0.0001$), well-integrated ($p < 0.0001$), and consistent ($p \sim 0.04$) interface that better facilitates learning ($p < 0.0001$) and effortless document navigation ($p \sim 0.004$) compared to *Body-fixed*. These analyses also support **H1** by showing that **a)** utilizing *Environment-referenced* resulted in significantly higher user satisfaction in *Dynamic* contexts compared to the *Static* ones ($p \sim 0.03$), and **b)** utilizing *Body-fixed* results in significantly higher user satisfaction in *Static* contexts compared to the *Dynamic* ones ($p \sim 0.005$).

7 DISCUSSION

In this section, we evaluate the user study results, examining their alignment with our hypotheses. We explore alternative hypotheses to explain any discrepancies between our initial hypotheses and the findings, offering insights into the study's limitations and proposing directions for future research.

7.1 Evaluation of Our Hypotheses

Our results revealed that the effectiveness of *Environment-referenced* varies across different contexts, excelling in certain scenarios while falling short compared to *Body-fixed* in others. For each of our measurements, the factorial analysis revealed at least one significant interaction effect involving *Placement Strategy* and one or both of our contextual components, i.e., *RW Setting* and *User State*. Considering *Environment-referenced* placement, the participants switched their gaze between panels less often, navigated the documents faster, and expressed overall better feedback on the usability and efficiency of the placement strategy, when in *Dynamic* contexts, compared to *Static* ones. *Body-fixed* placement increased the user's gaze switches within the *Dynamic* contexts compared to the *Static* contexts. Additionally, considering the *User State*, when *Mobile*, this placement increased gaze switches within a *Dynamic* setting but reduced it within the *Static* setting. This underscores the impact of contextual scenario on the efficacy of placement strategy, thereby supporting **H1**.

To evaluate **H2**, we investigated the performance of *Environment-referenced* information acquisition, compared to *Body-fixed* in our *Dynamic* (**H2-a**) and *Static* (**H2-b**) contexts. All of our measurements consistently support **H2-a**, indicating that in *Dynamic* contexts *Environment-referenced* offers a similar level of physical efficiency as *Body-fixed*, with significantly lower cognitive load, optimizing accuracy, speed, ease, and usability of information acquisition. This underscores the importance of context awareness, suggesting that in contexts where the relevant intermediary objects are dynamic, applying *Environment-referenced* placement in a context-aware interface could enhance the user experience.

Our data yielded inconclusive outcomes regarding **H2-b** and **H3**. While participants' post-context ranking of the placement strategies and the Overall TLX score support **H2-b**, indicating that users perceive *Body-fixed* to outperform *Environment-referenced* in *Static* contexts, the *Error Rate* and *Gaze Switch* measures contradict this hypothesis, indicating lower accuracy and efficiency of *Body-fixed* placement. Similarly, according to our results, the *User State* only influences placement strategy performance in terms of *Gaze Switch* and the *Effort* subscale of the TLX. Both these metrics indicate the inferior performance of *Body-fixed* in the *Mobile* contexts and refute **H3**.

7.2 Influence of Info-focus Relevance

Our results strongly confirm **H1** and **H2-a** while offering partial support for **H2-b** and partially rejecting **H3**. The inconsistency in our results can be attributed to a subtle interplay among our contextual components, which influences the suitability of the intermediaries. Across these contexts and even within specific ones, the user's focus on specific objects within the room varied. We used the user's primary conversation partner or location within each trial to extract the user's primary focus. While the choice of intermediaries ensures that the information access needs are relevant to the intermediaries, within each trial, the relevance

of the information being accessed to the user's primary focus differs. This discrepancy impacted the efficacy of the *Environment-referenced* placement strategy placement strategy across diverse contexts. We use the term *Info-focus Relevance* to refer to the degree of relevance between information access need and the user's current focus within each trial. The following discussion delves deeper into *Info-focus Relevance*'s effect on placement strategy performance in our contexts.

Dynamic Contexts: Within our *Dynamic* contexts, the user was always focused on their conversation partner. Additionally, in these contexts, regardless of the *User State*, the information needs within a trial were always *Relevant* to the same trivia host the user was interacting with. Due to the high *Info-focus Relevance*, the intermediaries are suitable for all trials in the *Dynamic* contexts. This enables *Environment-referenced* placement to always provide effortless access to the required document through a forward glance and eliminate the cognitive load associated with navigating to the correct category panel. This can explain our results regarding **H3**, as the *Environment-referenced* performs significantly better than *Body-fixed* in *Dynamic* contexts, even when the *User State* was *Mobile*.

Static Contexts: In the *Static* contexts, the user's focus and, therefore, *Info-focus Relevance* vary based on the *User State*. The *Static Mobile* context involved the user moving to one of the intermediary objects and maintaining visual attention to it. In this context, half of the trials require information relevant to the same category as the nearby intermediary object. Consequently, only half of the trials in this context were *Relevant* in terms of *Info-focus Relevance*. On the other hand,

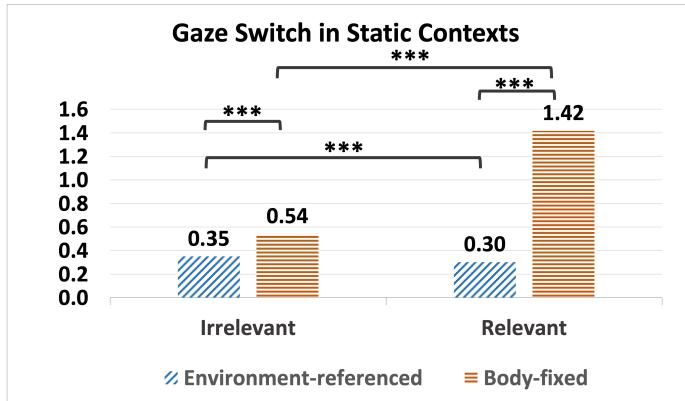


Fig. 13: The influence of Info-focus Relevance on gaze switch efficiency within static contexts

within all the *Static Stationary* trials, the user was facing forward and focused on the Sports category's poster. In this context, only 1/3 of the trials require information relevant to the Sports category and were considered as *Relevant* in terms of *Info-focus Relevance*.

To better understand *Info-focus Relevance*, we conducted a factorial non-parametric ART RM-ANOVA to test the interaction effects of *User State*, *Placement Strategy* and *Info-focus Relevance* on the quantitative trial data within the *Static* contexts. The results indicate significant influence of *Placement Strategy* and *Info-focus Relevance* on *Gaze Switch* ($F_{1,372.59} = 40.02, p < 0.0001$) (See Figure 13) and significant effect of *User State X Placement Strategy X Info-focus Relevance* on *Navigation Time* ($F_{1,372.39} = 11.65, p < 0.001$) (See Figure 14).

Static Stationary: The *Static Stationary* context entails a stationary user positioned between the three posters (intermediaries), maintaining a consistent orientation and location. Consequently, *Environment-referenced* appears almost identical to *Body-fixed* and imposes equivalent physical and mental load levels. This is contradictory to our **H2-b** and partially explains the lack of support for this hypothesis in our findings. Additional nonparametric analyses using Wilcoxon Rank Sums tests on our objective data, NASA TLX, and SUS measurements within this context indicate no significant effect of *Placement Strategy* on any of our metrics.

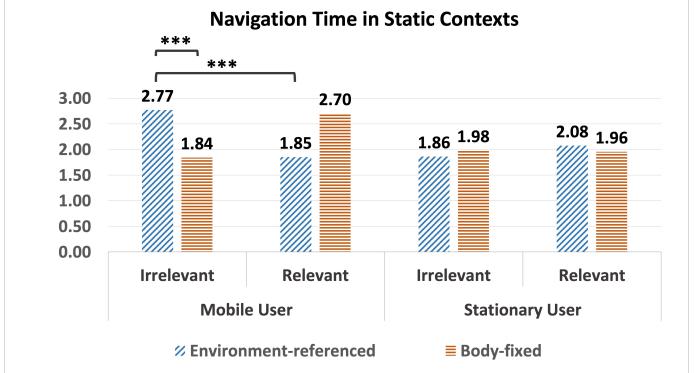


Fig. 14: The influence of Info-focus Relevance on time efficiency within static contexts

Static Mobile: We conducted nonparametric analyses using Wilcoxon Rank Sums tests to examine the effect of *Placement Strategy* within this context. The results indicated that *Environment-referenced* placement in this context resulted in significantly lower *Gaze Switch* compared to *Body-fixed* ($\chi^2 = 11.84, Z = 3.44, p \sim 0.0006$). This helps further explain the contradictory results to **H2-b** and **H3**.

8 LIMITATIONS & FUTURE WORK

In each combination of our context components (*RW Setting* and *User State*), the user's focus within the environment varied. This resulted in the involvement of *Info-focus Relevance* as an independent variable, influencing the effectiveness of *Environment-referenced* placement strategy. On the other hand, *Body-fixed* offers reliable and effortless access to all documents through a glance, irrespective of the particular context and *Info-focus Relevance*. While this placement strategy imposes an additional mental load to select the required information, this mental load is consistent across different contexts. Additionally, as mentioned by our participants, due to the limited number of category panels within this study, this mental load is relatively easy.

Our results indicate that contingent upon the user's focus and interactions with their setting, the choice of optimal placement strategy within each context varies. In a context-switching scenario, it is essential for a context-aware interface to evaluate the cognitive and physical load associated with various placement strategies and determine the optimal placement for the specific context. Such intelligent interfaces must infer information such as *Info-focus Relevance* from the context. In addition, future work must explore various ways for such interfaces to provide feedback to users when the placement strategy changes.

9 CONCLUSIONS

Towards context-aware adaptations in XR, we explored various XR design elements that can be adapted, providing a comprehensive design space for XR (Section 3). Concentrating on the content's spatial layout from our design space, we proposed utilizing a relevant intermediary from the environment within a Hybrid Frame of Reference (FoR) for each XR object (Section 4). We investigated the effectiveness of this adaptive placement strategy and a non-adaptive *Body-Fixed* placement strategy in four contextual scenarios varying in *RW Setting* and *User State* (Section 5).

Within our experiment, the *Environment-referenced* strategy utilized relevant people and objects from the environment as intermediaries to position the XR content. On the other hand, the *Body-fixed* strategy provided a consistent glanceable interface, offering effortless access to all content, regardless of the specific context. Our results indicate the significance of context in the choice of optimal placement strategy and highlight the significance of *Info-focus Relevance* on the effectiveness of *Environment-referenced* placement strategy across different contexts. In contexts with more *Relevant Info-focus* trials, *Environment-referenced* intelligently positioned each piece of AR content precisely

where needed, thereby enhancing efficiency while mitigating information overload. However, with *Irrelevant Info-focus*, *Environment-referenced* diminished performance as it necessitated excessive head and body movements towards the intermediaries and was inconsistent when the user or the environment was dynamic. These observations indicate the significance of *RW Setting* and *User State* in determining *Info-focus Relevance*. More specifically, within our contexts, the interlocutor, the user's proximity to the intermediary objects, and the mobility of the user influenced the user's focus and information needs. Our findings underscore the significance of context-aware interfaces that determine the optimal *Placement Strategy* through inferring information such as *Info-focus Relevance* from the context.

ACKNOWLEDGMENTS

Special thanks to our study participants and our colleagues at the 3DI Lab and Adobe for their contributions in shaping the direction of this paper. AI tools were used only in the final round of polishing of this paper.

REFERENCES

- [1] R. T. Azuma. A survey of augmented reality. *Presence: teleoperators & virtual environments*, 6(4):355–385, 1997. 2
- [2] P. Baines and A. Haslam. *Type & typography*. Laurence King Publishing, 2005. 3
- [3] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pp. 101–110, 2001. 2, 4, 5
- [4] K. Bengler and R. Passaro. Augmented reality in cars: Requirements and constraints. *ISMAR'06 industrial track*, 2006. 2, 3
- [5] M. Billinghurst, A. Clark, and G. Lee. A survey of augmented reality. *Foundations and Trends® in Human–Computer Interaction*, 8(2-3):73–272, 2015. doi: 10.1561/1100000049 4
- [6] M. Billinghurst and T. Starner. Wearable devices: new ways to manage information. *Computer*, 32(1):57–64, 1999. 4
- [7] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. An introduction to 3-d user interface design. *Presence*, 10(1):96–108, 2001. doi: 10.1162/105474601750182342 2, 4, 5
- [8] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996. 9
- [9] Y. Cheng, Y. Yan, X. Yi, Y. Shi, and D. Lindlbauer. 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. 2
- [10] S. Davari. [dc] context-aware inference and adaptation in augmented reality. In *2022 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 938–939. IEEE, 2022. doi: 10.1109/VRW55335.2022.00320 1, 2
- [11] S. Davari, F. Lu, and D. A. Bowman. Occlusion management techniques for everyday glanceable ar interfaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 324–330. IEEE, 2020. doi: 10.1109/VRW50115.2020.00072 1, 2, 4, 5
- [12] S. Davari, F. Lu, and D. A. Bowman. Validating the benefits of glanceable and context-aware augmented reality for everyday information access tasks. In *2022 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 336–444. IEEE, 2022. doi: 10.1109/VR51125.2022.00063 1, 2
- [13] A. K. Dey, G. D. Abowd, and D. Salber. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction*, 16(2-4):97–166, 2001. 2
- [14] S. Di Verdi, D. Nurmi, and T. Höllerer. Arwin-a desktop augmented reality window manager. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, pp. 298–299. IEEE, 2003. 2, 3
- [15] S. DiVerdi, T. Hollerer, and R. Schreyer. Level of detail interfaces. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 300–301. IEEE, 2004. 2, 3
- [16] P. S. Dunston et al. Identification of application areas for augmented reality in industrial construction based on technology suitability. *Automation in Construction*, 17(7):882–894, 2008. 2
- [17] M. R. Endsley. Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society annual meeting*, 32(2):97–101, 1988. 2
- [18] B. Ens, E. Ofek, N. Bruce, and P. Irani. Spatial constancy of surface-embedded layouts across multiple environments. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, pp. 65–68, 2015. 2
- [19] B. M. Ens, R. Finnegan, and P. P. Irani. The personal cockpit: a spatial interface for effective task switching on head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3171–3180, 2014. 2, 4
- [20] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the world: 2d windows for 3d augmented reality. In *Proceedings of the 6th annual ACM symposium on User interface software and technology*, pp. 145–155, 1993. 2, 4, 5
- [21] S. K. Feiner. Augmented reality: A new way of seeing. *Scientific American*, 286(4):48–55, 2002. 1
- [22] A. Giovannelli, J. Thomas, L. Lane, F. Rodrigues, and D. A. Bowman. Gestures vs. emojis: Comparing non-verbal reaction visualizations for immersive collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 8
- [23] M. Gonzalez-Franco, E. Ofek, Y. Pan, A. Antley, A. Steed, B. Spanlang, A. Maselli, D. Banakou, N. Pelechano, S. Orts-Escolano, et al. The rocketbox library and the utility of freely available rigged avatars. *Frontiers in virtual reality*, 1:20, 2020. 8
- [24] J. Grubert. Mixed reality interaction techniques. *arXiv preprint arXiv:2103.05984*, 2021. 2
- [25] J. Grubert, D. Hamacher, R. Mecke, I. Böckelmann, L. Schega, A. Huckauf, M. Urbina, M. Schenkl, F. Doil, and J. Tümler. Extended investigations of user-related issues in mobile industrial ar. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 229–230, 2010. doi: 10.1109/ISMAR.2010.5643581 2
- [26] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE transactions on visualization and computer graphics*, 23(6):1706–1724, 2016. 1, 2
- [27] S. G. Hart, S. G. Hart, L. E. Staveland, and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 1988. doi: 10.1016/s0166-4115(08)62386-9 9
- [28] S. A. Hood, B. M. Beauchesne, T. A. Fahmie, and A. Go. Descriptive assessment of conversational skills: towards benchmarks for young adults with social deficits. *Journal of Applied Behavior Analysis*, 54:1075–1094, 2021. doi: 10.1002/jaba.831 6, 8
- [29] N. F. S. Jeffri and S. Sulaiman. A review of augmented reality systems and their effects on mental workload and task performance. *Heliyon*, 7:e06277, 2021. doi: 10.1016/j.heliyon.2021.e06277 9
- [30] R. Jose, G. A. Lee, and M. Billinghurst. A comparative study of simulated augmented reality displays for vehicle navigation. In *Proceedings of the 28th Australian conference on computer-human interaction*, pp. 40–48, 2016. 4
- [31] R. v. Krevelen and R. Poelman. A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, 9:1–20, 2010. doi: 10.20870/ijvr.2010.9.2.2767 2, 6
- [32] W. Lages and D. Bowman. Adjustable adaptation for spatial augmented reality workspaces. In *Symposium on Spatial User Interaction*, p. 20. ACM, 2019. 2, 4
- [33] W. S. Lages and D. A. Bowman. Walking with adaptive augmented reality workspaces: design and usage patterns. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, pp. 356–366, 2019. 2
- [34] T. Langlotz, T. Nguyen, D. Schmalstieg, and R. Grasset. Next-generation augmented reality browsers: rich, seamless, and adaptive. *Proceedings of the IEEE*, 102(2):155–169, 2014. 2
- [35] E. Laviola, M. Gattullo, and V. e. a. Manghisi. Minimal ar: visual asset optimization for the authoring of augmented reality work instructions in manufacturing. In *The International Journal of Advanced Manufacturing Technology*, vol. 119, p. 1769–1784, 2022. doi: doi.org/10.1007/s00170-021-08449-6 3
- [36] J. Y. Lee, D. W. Seo, and G. Rhee. Visualization and interaction of pervasive services using context-aware augmented reality. *Expert Systems with Applications*, 35(4):1873–1882, 2008. 2
- [37] D. Lindlbauer, A. M. Feit, and O. Hilliges. 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. 1, 2, 3
- [38] F. Lu and D. A. Bowman. Evaluating the potential of glanceable AR interfaces for authentic everyday uses. In *2021 IEEE Virtual Reality and*

- 3D User Interfaces (VR)*, pp. 768–777. IEEE, 2021. 1, 2
- [39] F. Lu, S. Davari, and D. Bowman. Exploration of techniques for rapid activation of glanceable information in head-worn augmented reality. In *Symposium on Spatial User Interaction, SUI ’21*. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3485279.3485286 2
- [40] F. Lu, S. Davari, L. Lisle, Y. Li, and D. A. Bowman. Glanceable AR: Evaluating information access methods for head-worn augmented reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 930–939, 2020. doi: 10.1109/VR46266.2020.00113 1, 2
- [41] P. Manakhov, L. Sidenmark, K. Pfeuffer, and H. Gellersen. Gaze on the go: Effect of spatial reference frame on visual target acquisition during physical locomotion in extended reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI ’24*. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3613904.3642915 4
- [42] T. Matthews, J. Forlizzi, and S. Rohrbach. Designing glanceable peripheral displays. *UC Berkeley*. Retrieved September, 23:2015, 2006. 2
- [43] T. L. Matthews, J. Mankoff, and J. Canny. *Designing and evaluating glanceable peripheral displays*. University of California, Berkeley, 2007. 2
- [44] M. R. Miller, H. Jun, F. Herrera, J. Y. Villa, G. Welch, and J. N. Bailenson. Social interaction in augmented reality. *Plos One*, 14:e0216290, 2019. doi: 10.1371/journal.pone.0216290 7
- [45] C. Morrison, E. Cutrell, M. Grayson, A. Thieme, A. Taylor, G. Roumen, C. Longden, S. Tschiatschek, R. Faia Marques, and A. Sellen. Social sensemaking with ai: Designing an open-ended ai experience with a blind child. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2021. 2
- [46] T. Pfeiffer, M. E. Latoschik, and I. Wachsmuth. Evaluation of binocular eye trackers and algorithms for 3d gaze interaction in virtual reality environments. *JVRB-Journal of Virtual Reality and Broadcasting*, 5(16), 2008. 2
- [47] K. Pfeuffer, Y. Abdabou, A. Esteves, R. Rivu, Y. Abdelrahman, S. Meitner, A. Saadi, and F. Alt. Artention: A design space for gaze-adaptive user interfaces in augmented reality. *Computers & Graphics*, 95:1–12, 2021. 2
- [48] R. Piening, K. Pfeuffer, A. Esteves, T. Mittermeier, S. Prange, P. Schröder, and F. Alt. Looking for info: Evaluation of gaze based information retrieval in augmented reality. In C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen, eds., *Human-Computer Interaction – INTERACT 2021*, pp. 544–565. Springer International Publishing, Cham, 2021. 2
- [49] R. Piening, K. Pfeuffer, A. Esteves, T. Mittermeier, S. Prange, P. Schröder, and F. Alt. Looking for info: Evaluation of gaze based information retrieval in augmented reality. In *IFIP Conference on Human-Computer Interaction*, pp. 544–565. Springer, 2021. 2
- [50] R. Rivu, Y. Abdabou, K. Pfeuffer, A. Esteves, S. Meitner, and F. Alt. Stare: Gaze-assisted face-to-face communication in augmented reality. In *ACM Symposium on Eye Tracking Research and Applications*, pp. 1–5, 2020. 2
- [51] Q. Roy, F. Zhang, and D. Vogel. Automation accuracy is good, but high controllability may be better. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2019. 2
- [52] D. Schmalstieg and T. Hollerer. Augmented reality: Principles and practice. In *2017 IEEE Virtual Reality (VR)*, pp. 425–426. IEEE Computer Society, 2017. 2, 3
- [53] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino’s reality-virtuality continuum. *Frontiers in Virtual Reality*, 2:647997, 2021. 3
- [54] M.-H. Stoltz, V. Giannikas, D. McFarlane, J. Strachan, J. Um, and R. Srinivasan. Augmented reality in warehouse operations: opportunities and barriers. *IFAC-PapersOnLine*, 50(1):12979–12984, 2017. 2
- [55] R. B. Valimont, D. A. Vincenzi, S. N. Gangadharan, and A. Majoros. The effectiveness of augmented reality as a facilitator of information acquisition. In *Proceedings. The 21st Digital Avionics Systems Conference*, vol. 2, pp. 7C5–7C5. IEEE, 2002. 2
- [56] M. Weiser. The computer for the 21st century. *ACM SIGMOBILE mobile computing and communications review*, 3(3):3–11, 1999. 2
- [57] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011. doi: 10.1145/1978942.1978963 9



Shakiba Davari

Shakiba Davari received her Ph.D. from Virginia Tech with a focus on intelligent XR interfaces. Throughout her Ph.D., she has contributed methodological guidelines and frameworks for developing intelligent XR and investigated context-aware AR interfaces that address AR challenges such as real-world occlusion and social intrusiveness.



Doug A. Bowman

Doug A. Bowman is the Frank J. Maher Professor of Computer Science and Director of the Center for Human-Computer Interaction at Virginia Tech. He is the principal investigator of the 3D Interaction Group, focusing on the topics of three-dimensional user interfaces, VR/AR user experience, and the benefits of immersion in virtual environments