A Comparison of Virtual Reality Menu Archetypes: Raycasting, Direct Input, and Marking Menus

Johann Wentzel*, Matthew Lakier*, Jeremy Hartmann*, Falah Shazib*, Géry Casiez*†‡, Daniel Vogel*

***Abstract*—**
We contribute an analysis of the prevalence and relative performance of archetypal VR menu techniques. An initial survey of 108 menu interfaces in 84 popular commercial VR applications establishes common design characteristics. These characteristics motivate the design of raycast, direct, and marking menu archetypes, and a two-experiment comparison of their relative performance with one and two levels of hierarchy using 8 or 24 items. With a single-level menu, direct input is the fastest interaction technique in general, and is unaffected by number of items. With a two-level hierarchical menu, marking is fastest regardless of item number. Menus using raycasting, the most common menu interaction technique, were among the slowest of the tested menus but were rated most consistently usable. Using the combined results, we provide design and implementation recommendations with applications to general VR menu design.

***Index Terms*—virtual reality, interaction techniques, menus, direct input, raycasting, marking menu**

I. Introduction

SPATIAL menus, meaning visual menus requiring movement in space to interact, are a fundamental part of virtual reality (VR) interfaces. The design, development, and evaluation of novel spatial menu styles has been a research topic for decades (e.g. [14], [27], [35]).

However, not all demonstrated concepts in academia make the jump to real-world implementations. Moreover, since VR design and development on a consumer scale is still in its relative infancy, designers lack informed insight into the relative performance, accuracy, comfort, and preference between current common implementations. Some previous work compares various subsets of spatial menus [18], [19], or proposes quick and effective new menus [38], neither of which capture performance comparisons based on consumer reality. As such, there remains a gap in understanding the relative performance and preference of the most common menu characteristics in spatial menus today.

Motivated by this, our work explores three research questions. First, we determine the most common characteristics by exploring: (**RQ1**) *What are the menu characteristics of commercial VR applications, and how are these characteristics commonly combined?* Next, we evaluate these menu characteristics in their most common configurations, by exploring: (**RQ2**) *How do common archetypes of VR menus perform relative to each other in typical usage configurations?* Finally, to cross the academia-to-product gap, we generate real-world design recommendations by exploring: (**RQ3**) *How can designers use these empirical performance results to inform the design of future VR systems?*

We present an analysis of the prevalence and relative performance of VR spatial menu techniques. Our survey of 108 menus from 84 popular VR applications finds that menus positioned in world space using raycasting interaction were the most common, with many examples using direct input as well.

Radial marking menus were less common in the survey, despite being common and performant in academic work [38].

This survey informed the characterization and evaluation of three archetypal menu techniques: raycasting, direct input, and marking menus. We included marking menus as an archetype to represent radial layouts in the survey, and in recognition of their performance benefits demonstrated in the context of traditional, screen-based interfaces [38], which may be extensible to VR. The survey found that many non-radial menus display a large number of items at once. However, the planar radial design of a 2D marking menu is impractical for many items per level. To balance our experiments and explore archetypes that enable more items, we created a straightforward extension for a 3D marking menu layout, by dispersing items over the surface of a sphere. An initial experiment compares the three archetypal menus with regard to performance, accuracy, and usability across two menu sizes: 8 items and 24 items. We found that marking menus were fastest with few items, but slower with many items. Direct interaction was consistently fast. A second experiment compares the same menus with 2 levels of hierarchy and a refined focus on expert motor performance. In this setting, we found that marking menus were consistently fast at both menu sizes but, at a large number of items, had lower perceived usability than direct interaction. Across both experiments, raycast menus were consistently slow but rated as highly usable.

In sum, we make 3 main contributions: (1) an analysis of common menu characteristics from popular VR applications; (2) a detailed comparison of two common types of VR menus with a VR marking menu; and (3) empirically-informed design guidelines for choosing menu designs for consumer VR.

II. Background and Related Work

Our first focus is to understand the gap between controlled studies in academic work and real-world implementations. To better understand which techniques from academic literature have made it to real-world products, as well as better inform our evaluation methods, we first review the most relevant

* University of Waterloo, Ontario, Canada

†Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL, Lille, France

‡Institut Universitaire de France, Paris, France

TABLE I
SUMMARY AND COMPARISON OF THE MOST RELEVANT PREVIOUS WORKS THAT ALSO COMPARED MIDAIR VR MENUS. SEE TEXT FOR MORE DETAILS.

	Platform	Number of Items	Item Layouts	Interaction Techniques	Hierarchy	Anchoring	Measures
Monteiro et al. [52]	HMD VR	4 items	linear, radial	raycast, touchpad	2-level	world, hand	time, error, questionnaires
Lediaeva and LaViola [40]	HMD VR	6 items	linear, radial	raycast, head gaze, eye gaze	1-level	world, hand, body (*arm, waist*)	time, error, questionnaire
Wall [67]	HMD VR	5 items	linear, radial	raycast, direct, joystick	2-level	world, hand	time, error, questionnaire
Pourmemar and Poullis [60]	HMD AR	6 items	linear, radial	raycast, direct, voice, head gaze	3-level	world	time, error
Das and Borst [21]	CAVE VR	4-10 items	linear, radial	raycast	1-level, 2-level	world	time
Pfeuffer et al. [58]	HMD VR	4-16 items	linear, grid	raycast, direct, gaze	1-level	hand	time, error, movement, questionnaires
Mundt and Mathew [54]	HMD VR	8 items	radial	raycast, direct, wrist rotation, joystick	1-level	hand	time, error, questionnaire
Our Work	HMD VR	8 items, 24 items	linear, grid, radial	raycast, direct, marking	1-level, 2-level	world, hand	time, error, movement, questionnaire

spatial menu surveys, designs, and comparisons in the literature. Table I shows a summary of selected previous work, organized by the categories defined in Section III.

A. Surveys of 3D Pointing Techniques and Menus

Previous literature provides surveys and classifications of 3D pointing techniques and menus. Dang [20] classified 3D pointing techniques based on how targets are selected. In point-based techniques, such as Go-Go [59], the user controls a single selection point in 3D space. In line- or curve-based techniques, such as raycasting [49], the user controls the angle of a line or curve in 3D space to select a target. Argelaguet and Andujar [1] surveyed 3D pointing techniques and classified them based on criteria including the selection approach, input degrees of freedom, and the approach for disambiguating selections. Bailly et al. [5] developed a general taxonomy of menu properties, not specific to VR. This included vocabulary for menus such as "menu structure" and "super-menu", and properties such as number of items, geometry, visual versus motor space, and feedback.

Dachselt and Hübner [18], [19] surveyed VR and AR menu techniques proposed by researchers, and produced a taxonomy of 3D menus, including aspects like number of displayed items, structural layout, and orientation. Surveys of 3D pointing and menu techniques are important as they supply initial structure for discussing menu design characteristics. However, these surveys focus on 3D menus in academic literature with little focus on consumer applications. We use a taxonomy inspired by Dachselt and Hübner [18], [19] as a tool to describe consumer VR menu implementations.

B. Comparing Novel AR and VR Menus to Existing Work

Several past works focused on designing and implementing novel AR and VR menu designs, including those using raycast selection, direct input, radial designs, and marking menus. Raycast menus include early work on gaze and hand-tracked gestures [35], [57]. Direct input menus include those with physical object metaphors [4], body-anchored menu widgets [2], [3], or tangible props [70].

Some previous work on direct input also examines passive haptics, which can improve item selection compared to mid-air [72], [73]. However, passive haptic feedback can be situational depending on the user's environment and specific application context.

Radial layouts and marking menus include several approaches that use hand and wrist rotation [27], [44], [63], or contextually-overloaded marking for additional specificity [25]. Most relevant to the 3D marking menu we include in our evaluation, is work on depth-based selection on a 3D grid layout [30], [71], and freehand 3D gestures on both a vertical and horizontal plane [61]. These works either have no evaluation, or evaluate different internal parameters of the novel menu's design (e.g. type of feedback, item size) with little comparison to previous VR menus.

We focus instead on work that evaluates novel menus using common or pre-existing spatial menu designs as comparative baselines. Bowman and Wingrave [14] designed the TULIP menu for VR, which uses fingers to anchor items and provide input, and compared it with a pen-and-tablet-based menu. Gebhardt et al. [26] investigated using a smartphone as a secondary menu in a bimanual raycasting technique for CAVE-based VR. They compared it to one of their own earlier menu designs [25] and found it to have low usability for novice users. Wang et al. [68] compared VR menus fixed in world-space versus held in the non-dominant hand, using selection techniques based on hand or head pointing. Similar work by Pfeuffer et al. [58] investigated three gaze-based selection techniques, and compared them to raycasting and direct input as baselines. Perrault et al. [56] proposed using in-room "Physical Loci" for AR, and compared this approach with mid-air 2D marking menus. Lim et al. [45] designed a loop-based AR marking menu for see-through HMDs. They compared the menu to a standard marking menu and a hybrid of the two.

These works evaluate novel techniques using pre-existing designs as baselines, but their focus remains primarily on the novel techniques instead of the comparison of established designs. These works inform our experiment methodology, but we focus instead on comparing existing menu types to inform consumer application design choices.

C. Comparing Spatial Menu Configurations

Previous work has investigated the performance and subjective preference differences between some menu configurations in AR and VR environments. Das and Borst [21] evaluated various

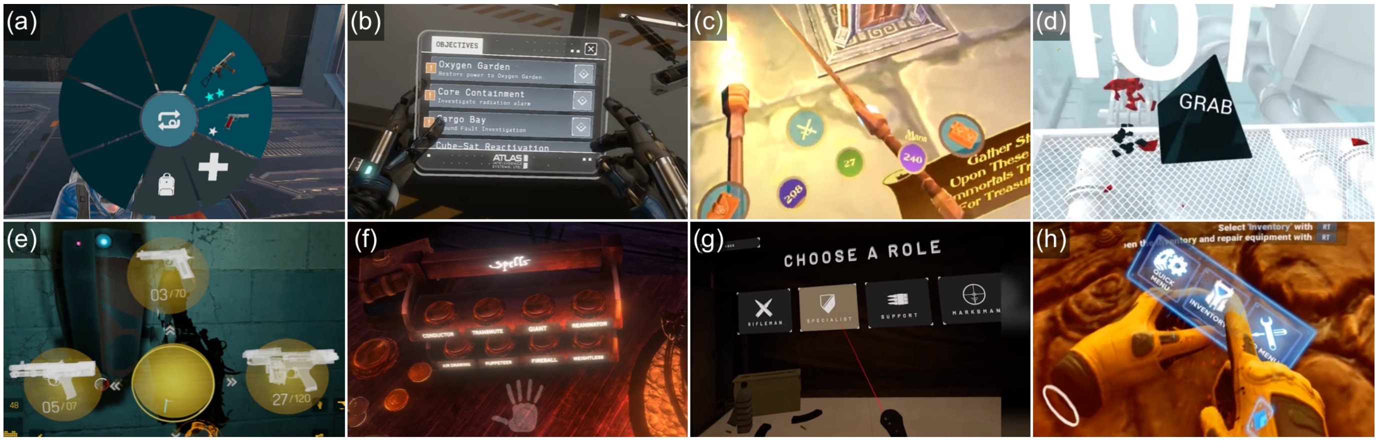

Fig. 1. A sample of spatial menus from our survey: (a) *radial raycast* menu in *POPULATION: ONE* [64]; (b) *linear direct* menu in *Lone Echo* [65]; (c) *radial marking* menu in *Vanishing Realms* [10]; (d) *scattered direct* menu in *SUPERHOT VR* [41]; (e) *radial marking* menu in *Half Life: Alyx* [34]; (f) *grid direct* menu in *Waltz of the Wizard* [17]; (g) *linear raycast* menu in *Onward* [12]; (h) *linear direct* menu in *No Man's Sky* [11].

menu types in a CAVE VR environment, comparing pie versus linear menu layouts, conventional raycasting versus a novel variation, a variable number of items from four to ten, one or two levels of hierarchy, and contextual menu placement versus fixed placement. Pie menus were typically faster and more accurate than linear, raycasting was faster than their novel method, and smaller menus were faster. However, CAVE-based VR evaluations are meaningfully distinct from HMD-based VR, making direct comparisons inappropriate. For example, CAVE menus preclude the use of arm-mounted menus due to depth differences between the arm and the visualization [23], and the user's hands may not be visible in HMD-based VR [50].

Monteiro et al. [52] compared all 4 combinations of linear raycast or radial touchpad menus, and world-anchoring or body-anchoring. At two levels of hierarchy with four items per level, linear world-anchored menus were most preferred despite all performing similarly. Lediaeva and LaViola [40] investigated different placements for VR menus around the body and compared these to a world-anchored placement. They also compared a linear versus radial menu shape, and selection using controller raycasting, head pointing, or eye-gaze pointing. At one level of hierarchy with six items per level, participants preferred world-anchored, hand, and waist placements. Mundt and Mathew [54] compared four interaction techniques for hand-anchored VR pie menus: raycasting, direct input, wrist rotation, or controller joystick input. Direct input and raycasting were the fastest and most preferred. Wall [67] compared a vertical linear menu with raycast selection, a hand-anchored joystick-controlled radial menu, and a radial direct-input menu, finding that the vertical linear menu had the best performance and joystick movements were rated the least usable. In a pilot test investigating AR menus, Pourmemar and Poullis [60] compared a vertical linear menu with a hierarchical radial menu, using three different input modalities: hand pointing plus air tapping, voice input, and head pointing. The hierarchical radial menu with head pointing performed the fastest and was the most preferred.

D. Summary

Despite the variety of contexts explored, previous studies do not address key combinations of characteristics found in simple, common VR menus. For example, in the context of HMD-based VR, past works have not investigated the effect of adding a hierarchy, have not compared marking menus to more standard menu types, and have not compared the combined effects of menu layouts, item counts, and interaction techniques. Further, most past works only evaluate menus with fewer than 10 items per level, despite the prevalence of menus with large numbers of items in common VR applications. Our work fills this gap in understanding the effects of VR menu characteristics.

III. Survey of Common VR Menus (**RQ1**)

In addition to understanding past academic work, providing real-world design recommendations demands knowledge of real-world implementations. Our work draws menu archetypes directly from the most common characteristics of menus found in popular VR games, warranting a deeper understanding of real implementations. To inform our selection of archetypal menus and answer **RQ1**, we survey and summarize menu design characteristics used in popular consumer VR applications.

We first selected the top 50 applications based on purchase and usage statistics from Steamcharts [29] at two time periods: June 2019 and February 2021. We selected games at two time periods to respond to a large influx of PC VR titles being published over the course of carrying out this work. Steam is a primary marketplace for game and non-game VR applications. Spatial menus in these applications were selected using three criteria. They had to: (1) be designed for use with a VR headset; (2) be visible and usable within the application's 3D environment; and (3) require input from VR hand controllers. The third condition specifically excludes menus designed for non-spatial controllers (like an Xbox controller). Additionally, while acknowledging recent advances in hand tracking technology like that on the Meta Quest 2, we focus on controller-based input as it remains the most widely-implemented cross-platform input configuration.

TABLE II
COUNTS AND PROPORTIONS OF VARIATIONS IN DESIGN CHARACTERISTICS FOR 108 SPATIAL MENUS USED IN POPULAR VR APPLICATIONS.

Characteristic	Variation with Count (N = 108)				
Item Layout Spatial arrangement of menu items. κ = 0.68	*Linear* (single horizontal or vertical row) 45 (42%)	*Grid* (multiple rows) 23 (21%)	*Radial* (circular layout) 10 (9%)	*Scatter* (non-uniform, e.g. scattered desk) 3 (3%)	*Mixed* (compound with linear and grid parts) 27 (25%)
Interaction Technique Method for invoking menu and highlighting items. κ = 0.67	*Raycast* (controller direction intersects distant item) 66 (61%)	*Direct* (controller position intersects with item) 26 (24%)	*Hardware* (controller button or joystick) 10 (9%)	*Mixed* (support both raycast and direct) 6 (6%)	
Selection Technique Method to confirm item and dismiss menu. κ = 0.69	*Hardware Button* (button press) 84 (78%)	*Immediate* (as item highlighted with interaction tech.) 23 (21%)	*Dwell* (hold still for time period) 1 (1%)		
Number of Items Items typically visible in first menu level. κ = 0.57	*1-5* 45 (42%)	*6-10* 39 (36%)	*11-20* 22 (20%)	*21 or more* 2 (2%)	
Anchoring Where menu is positioned. κ = 0.69	*World* (independent of user) 80 (74%)	*Hand* (relative to user's hand) 23 (21%)	*Body* (relative to user's arm, head, hips, etc.) 5 (5%)		
Hierarchy Number of nested layouts of items. κ = 0.83	*1 level* 44 (41%)	*2 levels* 50 (46%)	*3 or more levels* 14 (13%)		
Hand Usage Number of hands required. κ = 0.64	*Unimanual* (one hand) 90 (83%)	*Bimanual* (both hands) 18 (17%)			

TABLE III
THE TOP 10 COMBINATIONS OF ANCHORING, ITEM LAYOUT, NUMBER OF ITEMS, AND INTERACTION TECHNIQUE FOUND IN THE GAME SURVEY.

Anchoring	Layout	Items	Interaction	Count
World	Linear	1-5	Raycast	**12**
World	Mixed	11-20	Raycast	**9**
World	Mixed	6-10	Raycast	**7**
World	Linear	6-10	Raycast	**6**
World	Grid	6-10	Raycast	**5**
World	Linear	1-5	Direct	**5**
World	Radial	1-5	Raycast	**4**
Hand	Linear	1-5	Raycast	**3**
Hand	Radial	1-5	Direct	**3**
World	Grid	11-20	Raycast	**3**

The final dataset consisted of 108 unique menus from 84 applications. Figure 1 shows a sample, and a full list of applications with links to recordings of menu interactions is in supplementary materials.

Two reviewers each independently categorized all 108 menus by either watching online videos or through direct experience (using HTC Vive Pro or Oculus Rift S system). Informed by Dachselt et al.'s menu taxonomies [18], [19], we used open coding to converge on seven menu design characteristics: item layout, interaction technique, selection technique, number of items, anchoring, levels of hierarchy, and hand usage. To classify a menu, each characteristic has multiple variations, such as "Unimanual" and "Bimanual" variations for the "Hand Usage" characteristic. Definitions of the seven characteristics and their variations are in Table II. Inter-reviewer agreement for primary categorizations was verified using Cohen's Kappa [39].

A. Results

Table II summarizes the results for individual characteristics. There was a dominance of *Linear* and *Grid* Item Layouts, including menus that dynamically *Mixed* the two depending on number of items. The Interaction Technique to point at items was dominated by *Raycast* and *Direct* methods, and items were most often Selected *Immediately* when first highlighted or with a *Hardware Button*. We found only 10 examples of *Radial* layouts used for pie or marking menus. Most menus used a *1* or *2 level* Hierarchy, with the Number of Items per level ranging from less than 6, to more than 20. Menus were Anchored most often in the *World* away from the user's body, or on the user's *Hand*. The latter typically required *Bimanual* hand usage, but most menus were designed to be *Unimanual*.

We also examined frequent combinations of characteristics. The most common combination of *Item Layout* and *Number of Items* was linear menus with 1–5 items (27 menus total), while the next most common was mixed-layout menus with 6–10 items. Menus with item counts above 10 most commonly used mixed or grid layouts (16 menus total), and no menus with more than 10 items used radial or scattered layouts. Both menus that used more than 21 items used grid layouts. Radial menus were most commonly hand-anchored, whereas all other menus were most commonly world-anchored. Mixed-layout menus had the highest proportion of world-anchoring (89%). Radial menus had the highest proportion of 2-level hierarchies (60%), but menus with 3 or more levels of hierarchy only had linear or mixed layouts.

Hand-anchored menus most commonly used *Direct* interaction (43% of all hand-anchored menus) while world-anchored menus most commonly used *Raycast* selection (70% of all world-anchored menus). Of the 10 *Radial* menus, five were marking menus (marking is categorized as an *Immediate* selection technique).

Menus using *Direct* interaction most frequently used hand-anchored *Grid* and *Linear* layouts, with 1–2 levels of hierarchy. Menus using *Raycast* interaction most commonly used world-anchored *Linear* or *Mixed* layouts with a wider variety of hierarchy than those using *Direct*. Menus that used marking as a selection technique most commonly used a hand-anchored radial layout, with a number of items below 10 and multiple levels of hierarchy.

Menus in the survey adopted a wide variety of overall configurations. Table III shows the 10 most frequent combinations of Layout, Anchoring, Number of Items, and Interaction Technique. Note that the most frequent *combinations* of characteristics is not representative of the rate these characteristics occur individually.

B. Discussion

This survey illustrates the variety of VR menus, but also how design characteristics can influence each other. For example, if the context of an application requires a menu to be suspended in the VR scene far away, designers often choose to have the user raycast rather than interact with the menu directly. Likewise, menus shown on the hand typically employ direct interaction, likely because raycasting toward a hand-anchored menu involves the dual challenge of short-range pointing as well as holding the hand anchoring the menu sufficiently steady.

The survey also revealed that the vast majority of menus used raycasting or direct interaction, in linear or grid layouts. Radial menus (five of which were marking menus) were the next most common non-compound layout. All menu layouts placed all items within the same 2D plane. Why did this survey not find more examples of novel menu designs found in HCI research, such as those using 3 dimensions, or other more elaborate methods? We believe this could be related to consumer VR development constraints for software and hardware, and possibly a more conservative approach to VR interface design overall. Designers working with unfamiliar technologies are prone to legacy bias, and could default to interface designs used for devices with which they are more familiar. Linear lists and grids are historically commonplace arrangements in desktop UI design. Similarly, raycasting is well-established in HCI [35], [57], and is common in consumer devices like television remotes, the Nintendo Wii, and other general pointing devices, making the popularity of *Raycast* menus in VR applications expected. There were many occurrences of all four combinations of linear or grid layouts with raycasting or direct interaction, and for that reason we include all four combinations in our menu experiments.

Developers use UI interaction toolkits like MRTK [47] or the Meta Interaction SDK [46] to speed up the implementation of common interaction techniques and interface layouts. The most common interaction toolkits include pre-made implementations of linear or grid menus, using raycasting or direct interaction, with little emphasis on marking or radial menus, which could further explain the commonality of these menu designs.

IV. Menu Archetypes

The results of the survey informed the creation of six menu archetypes,[1] whose configurations represent combinations of the most frequent menu configuration states. We describe our construction process for these archetypes, as well as their common features below.

The menu archetypes are built based on the most common configuration states encountered in the survey. First, the three most common item layouts were *Linear*, *Grid* and *Radial*, forming the basis of our three archetypal layouts. Note that while *Radial* menus were much less common than *Linear* and *Grid*, they were included enough to warrant detailed comparison. Next, based on the survey, we chose *Raycast* and *Direct* as representative interaction techniques.

All menus require an invocation method. Based on the prevalence of button-based interaction techniques in the survey, as well as to establish a common interaction method between all archetypes, we choose a single trigger button with a “press-hold-highlight-release” method for selection. Because the survey found both unimanual and bimanual menus, we implement two unimanual techniques (raycasting, marking) and one bimanual technique (direct input). All interactions are done with the dominant hand, and similarly to the surveyed menus. The non-dominant hand is used as an anchor for the menu items in the direct technique.

Considering that comparing all combinations of all design characteristic variations would make studies inaccessibly long, we narrowed down the number of menu archetypes based on combinations that were commonly represented in the survey and distinct enough to enable explicit or implicit comparisons of all common menu configuration characteristics. As a result, we selected three menu styles: *Hand*-anchored bimanual *Direct* interaction, *World*-anchored unimanual *Raycast* interaction, and *Hand*-anchored unimanual *Marking* interaction. Note that *Marking* was the most common *Immediate* selection technique for *Radial* menus. For more accessible study lengths, we include *Anchoring* as a secondary independent variable by spreading the anchoring properties across the three menu archetypes, instead of explicitly testing all combinations.

Each menu archetype has a small and large version. Based on the *Number of Items* characteristic, we chose 8 items for the small version since this captures a reasonable number within the common *1-5* and *6-10* range variations. We chose 24 items for the larger menu since it is a multiple of 8 which makes linear and grid layouts more comparable. Additionally, 24 items also serves as a conservative upper limit for menu capacity, considering that almost all menus in the survey showed fewer than 21 items at once.

Menu items for all archetypes are icons contained in a white square. The icons are an unambiguous set of pictures of simple objects (e.g. a cat, a boot, a kiwi fruit) curated by Lewis et al. [42] (based on an original set by Grossman et al. [31]). To ensure visual and interaction consistency across menus, icon sizes are dynamically adjusted to maintain similar visual angular width. When the menu is invoked, the distance between the headset and the user's hand is used to scale all icons equally, such that an icon at the centre of the menu has a 5° angular width when viewed straight-on. In practice, because all icons are scaled by an equal amount, there is slight variation in actual visual icon size due to perspective and item depth. This allows for visual depth cues, important in spatial interaction [13].

[1] The accompanying video demonstrates all variations of menu archetypes.

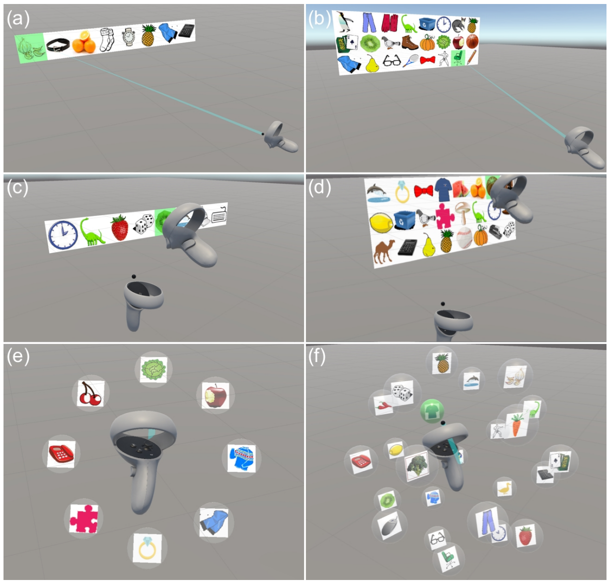

Fig. 2. Menu archetypes: (a) *raycasting* with 8 items in a linear layout; (b) *raycasting* with 24 items in a grid layout; (c) *direct input* with 8 items in a linear layout; (d) *direct input* with 24 items in a grid layout; (e) *marking* with 8 items in a 2D radial layout; and (f) *marking* with 24 items in a 3D spherical layout.

A. Raycasting Menu Archetype

Raycasting menus are visualized as a fixed panel of items positioned in world space 3 metres away from the user's location at moment of invocation. This placement is in line with the survey, which found that raycast menus are primarily anchored in world space. With 8 items, the layout is a single linear row of 8 icons (Figure 2a). With 24 items, the layout is a 3×8 grid of icons (Figure 2b). When the menu is invoked, a visible ray extends from the controller with the intersecting menu item rendered with a green outline. The selection ray is rendered as a 5 mm cyan cylinder projecting from the top of the controller model. The ray is stabilized using the 1€ filter [16] to reduce error rate [6].

For single-level menus, the user presses and holds the controller trigger to invoke the menu, points the ray at the intended menu item, and releases the trigger to select it. Several menus in the survey used an *Immediate* selection technique, including this "press-hold-highlight-release" technique for one-button menu invocation and selection. We include it for single-level menus for ecological validity and to maintain consistency between menu techniques. For multiple levels of hierarchy, the user presses and releases the controller trigger to invoke the menu, points the ray at the intended menu item, and presses and releases the trigger to select it. The first level of items is hidden and the second level of items is shown as a new panel slightly offset from the first to make it distinct. The second-level selection follows the same pattern as the first.

B. Direct Input Menu Archetype

Direct input menus are visualized as a panel of items anchored to the user's non-dominant hand, displayed as a single linear row for 8 items (Figure 2c) and a 3×8 grid for 24 items (Figure 2d). This placement is in line with the survey, which found hand-anchored menus are commonly used with *Direct* input. When invoked, the user brings the "tip" of their controller model into the desired item to select it. The tip is defined as a 5×80 mm cyan cone projecting out from the top of the controller. We include the tip as a way to control for different VR controller models in our remote experiments. When any part of the tip intersects the icon square, the item is outlined in green to indicate that it is currently selected. If the tip intersects multiple menu items, the item closest to the controller is selected.

For single-level menus, the controller trigger is pressed and held to invoke the menu, the tip is moved to directly intersect the intended menu item, and the trigger is released to complete the selection and close the menu. For multiple levels of hierarchy, the user presses and releases the trigger to invoke the menu, intersects the tip with the menu item, and presses and releases the trigger to select it. The second level is revealed as in raycast.

C. Marking Menu Archetype

Based on the survey, the *Radial* layout was the third most common menu layout (not counting *Mixed* compound linear and grid layouts), of which half were marking menus. Although direct or raycast selection could be used with a radial layout as a pie menu, research has shown that selection with a swift movement from the menu centre towards the desired item (a "mark") is better in non-VR contexts [38]. For example, the index of difficulty to select an item is lower and experts can select items eyes-free. For these reasons, we include marking menus as the third archetype.

Our survey found that all *Radial* layouts had 10 items or fewer, but multiple examples of other layouts with more than 10 items. To balance our experiment by filling out all combinations of item number and layout, we also explore 24-item marking menus. However, this number of items per level presents a design challenge. As more items are added to a radial layout, the effective item selection angle tolerance (for a marking menu) or item sector area (for a pie menu) shrinks, increasing time and errors [37]. Introducing a hierarchy is a common remedy, but this means multiple selection actions for each item. To scale a radial layout with marking selection to support more items, we extend the 2D planar circle layout to a 3D sphere, widening the effective range of marking angles for each item. We argue this is a simple extension for supporting more items.

Based on the above considerations, the marking menu archetype uses two styles of marking menu: a planar 2D style designed for up to 8 items and a spherical 3D style tested with 24 items. To validate the choice of a 2D circular layout for 8 items, we conducted a pilot experiment to compare it with an 8-item 3D layout. The 2D layout was slightly advantageous, so we use it for 8 items and the 3D menu for 24 items[2].

For both the 2D and 3D marking menus, the interaction is identical. The user presses and holds the controller trigger to invoke the menu, which then appears in world space centred on the controller's position. The user then moves their hand in the direction of the desired item. After moving 5 cm from

[2]See supplementary materials for a technical report of this experiment.

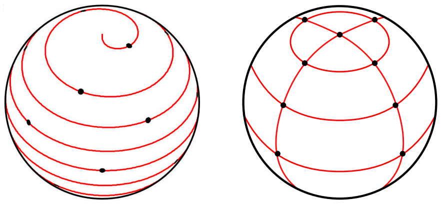

Fig. 3. Techniques for distributing items (represented by black dots) along the surface of a sphere. (a) Fibonacci lattice, including its placement spiral in red (illustration from González et al. [28]); (b) latitude-longitude lattice with placement grid in red.

the menu centre, the item to be selected is highlighted in semi-transparent green. The selected item can be fine-tuned by changing the position of the controller relative to the menu centre. For a single-level menu, the user can then release the trigger to make their final selection. For multiple levels of hierarchy, the user moves 2.5 cm past the radius of the menu to make the first-level selection, at which point the first-level menu disappears and the second-level menu appears at the new controller position. The user moves the controller in the direction of the desired second-level item, it is highlighted as above, and the user can release the trigger to select it. There is no delay between when the user presses the trigger and the menu items appearing, which previous work found does not affect performance [33].

The small (8-item) marking menu is visualized in 2D, similar to those found in the survey. The 8 items are spaced equally around the circumference of a 30 cm diameter circle (Figure 2e). The circle plane is oriented to face the user upon invocation and remains fixed until dismissed. The large 3D marking menu distributes icons over a 30 cm diameter sphere, with each icon rotating to face the user (Figure 2f). Like the 2D marking menu, the positions of icons remain fixed from menu invocation to dismissal. The spiral pattern of a Fibonacci lattice [28] is used to disperse items uniformly over the surface of the sphere (Figure 3a). Compared to previous 3D selection techniques [30], [61], [71] or traditional latitude-longitude lattices (Figure 3b), a Fibonacci lattice has lower axial symmetry, creating fewer opportunities for menu item occlusion [28], a known problem with depth-based menus [1]. The 3D menu extends the 2D menu metaphor directly; just as all items in the 2D menu occupy an equal amount of circumference, the Fibonacci placement ensures all items in the 3D menu occupy an equal amount of surface area [28].

V. Experiment 1: Single-Level Menus (**RQ2**)

Our survey of VR games revealed a collection of menu characteristics from real-world implementations. This experiment answers **RQ2** by comparing the most common real-world variations of these menu characteristics, as expressed through our archetypal menu designs. This first experiment explores single-level menus, using randomized ordering of menu item selection to balance visual search and motor performance.

We compare the three menu archetypes introduced above (raycast, direct input, and marking) for menus of two different sizes (8 and 24 items). Primary measures are trial time, error rate, hand movement distance, and perceived usability via the System Usability Scale (SUS) [15]. This experiment was administered remotely due to the COVID-19 pandemic.

Previous work found that raycasting and direct menus are often preferable based on familiarity [40], [54], and marking menus often gain performance after learning due to eyes-free interaction [38], [60]. As a result, we hypothesize that: (**H1**) *Marking* menus will have the best performance, but *Raycast* menus will be the most preferred.

A. Protocol

Participants: We recruited 18 right-handed participants (ages 21 to 62, 10 male, 8 female). Because the experiment was remote, they had to have access to a 6DOF VR headset with hand controllers, running either SteamVR or Oculus systems. Remuneration was a $25 e-gift card.

Apparatus: Participants used a variety of devices: Oculus Quest (6), Oculus Rift (4), Valve Index (3), HTC Vive (2), HTC Vive Pro (2), and Oculus Rift S (1). The study was developed and compiled for final distribution using Unity and SteamVR for input cross-compatibility. Upon experiment completion, log files were uploaded to AWS.

Procedure: Each trial began with a word prompt (e.g. "cat", "boot", "kiwi") anchored to the participant's view. The participant pressed the controller trigger to invoke the menu, selected the icon corresponding to the prompt, and released the trigger to confirm their selection and end the trial. If the selection was incomplete or an incorrect selection was made, a buzzer sound was played and the trial would restart. After all trials were completed for a menu condition, the participant completed a SUS questionnaire. This study was approved by our research ethics board, and study sessions lasted approximately 45 minutes including rest breaks.

Design: This is a within-subjects design with two main independent variables: ITEMS with two conditions (8-ITEMS, 24-ITEMS) and TECHNIQUE representing the three archetypal menu conditions (RAYCAST, DIRECT, MARKING). We refer to combinations of ITEMS and TECHNIQUE with this shorthand: RAYCAST-8, RAYCAST-24, DIRECT-8, etc. We use the secondary independent variable BLOCK to facilitate measuring learning.

For each combination of ITEMS and TECHNIQUE, 8 icons were randomly chosen as selection prompts. Participants completed 8 blocks of these 8 icon prompts in randomized order. Combinations of ITEMS and TECHNIQUE were counterbalanced by participant using a Latin square. The icon positions within each menu were randomized at the start of each condition, but remained the same across all blocks in a condition. Using the same icons at the same positions across all blocks increases ecological validity and encourages learning.

To evaluate and compare VR menus, we use dependent variables drawn from previous work on VR selection and controlled interaction studies [51], [69] as well as common industry practices [36], [62]. Dependent measures are computed from logs. *Time* is the time from the invocation of the menu to the final selection of an icon. Time for incorrect selections is not included. *Error Rate* is the proportion of trials that had one or more incorrect selections before selecting correctly. *Travel Distance* is the distance travelled by the dominant-hand controller between invoking the menu and making a selection. We measure Travel Distance as a proxy for physical movement and overall comfort [51], [69]. At the end of each condition, participants answered the ten-item SUS questionnaire, common in UX studies [36], [62]. Each question evaluated different usability aspects on a scale from 1 ("strongly disagree") to 5 ("strongly agree"). A *Combined SUS* is between 1 and 100, with 68 considered an "average" score [66]. Participants were also encouraged to provide informal feedback via email.

In summary: 2 ITEMS × 3 TECHNIQUES × 8 BLOCKS × 8 selection prompts = 384 data points per participant.

B. Results

For each combination of participant and TECHNIQUE, trials with *Time* or *Travel Distance* more than 3 standard deviations from the mean time were excluded as outliers. In total, 237 trials (3.4%) were removed. In the analysis to follow, if residuals did not follow a normal distribution, we applied an aligned rank transform (ART) [22] before running an ANOVA. All ANOVA preconditions were met. Pairwise comparisons used Tukey HSD or Holm corrections. A post-hoc power analysis with 20% β error and 5% α error suggests the number of participants was reasonable to detect a medium effect, but we recognize that post-hoc power analyses are less effective than a priori [24], [32]. We supplement with effect sizes and 95% confidence intervals where appropriate.

We examine several different menu characteristics, whose effects when combined might differ from their separate individual effects. As such, we also explore and analyze interaction effects for a more thorough description and more complete understanding of menu performance.

1) Learning Effect: We are interested in practised performance, so we remove initial slower blocks due to learning effects. Figure 4d shows *Time* by BLOCK for combinations of TECHNIQUE and ITEMS. For analysis, we create a 6-level TECH-ITEMS factor representing these combinations. This reduces the number of factors to better understand learning. There was a main effect of BLOCK on *Time* ($F_{7,799} = 42.4$, $p < .001$), and a TECH-ITEMS × BLOCK interaction ($F_{35,799} = 3.3$, $p < .001$). Pairwise comparisons showed block 1 was slower than blocks 4 – 8 for MARKING-8, MARKING-24 and RAYCAST-24; and block 1 was slower than blocks 5 – 8 for MARKING-8, DIRECT-8, RAYCAST-8 and DIRECT-24. In subsequent analysis, we use blocks 5 through 8 as they represent practised performance.

2) Time: Direct input was the fastest technique in general, and was unaffected by number of items (Figure 4a). MARKING-8 and DIRECT-8 were the fastest combinations overall, but the marking technique was slower than direct when selecting from 24 items. Raycasting was consistently slow.

Residuals were normally distributed, so ART was not used. There was a main effect of TECHNIQUE on *Time* ($F_{2,34} = 22.6$, $p < .001$, $\eta^2_G = .32$). Pairwise comparisons revealed differences between all techniques (all $p < 0.01$), showing that DIRECT input was fastest overall (1176 ms), followed by MARKING (1360 ms) and RAYCAST (1603 ms). There was a TECHNIQUE × ITEMS interaction effect ($F_{2,34} = 16.9$, $p < .001$, $\eta^2_G = .14$), showing that in the 8-ITEM menus, RAYCAST (1536 ms) was significantly slower than MARKING (1108 ms, $p < .001$) and DIRECT (1167 ms, $p < .001$). The MARKING-8 and DIRECT-8 menus were not significantly different. In the 24-ITEM menus, DIRECT (1186 ms) was faster than MARKING (1609 ms, $p < .004$) and RAYCAST (1671 ms, $p < .004$). The MARKING-24 and RAYCAST-24 menus were not significantly different. Post-hoc tests by TECHNIQUE did not find differences between RAYCAST-8 and RAYCAST-24 or between DIRECT-8 and DIRECT-24, but MARKING-8 was faster than MARKING-24 ($p < 0.001$).

3) Error Rate: The mean error rate was low, only 1.1%, with a single pairwise difference between the 8-item marking menu (M = 0.2%) and the 24-item raycast menu (M = 2.4%). There was a main effect of TECHNIQUE on *Error Rate* ($F_{2,85} = 3.5$, $p < .03$) but no post-hoc differences. There was also a TECHNIQUE × ITEMS interaction ($F_{2,85} = 4.7$, $p < .01$) revealing a post-hoc difference between MARKING-8 and RAYCAST-24 ($p < 0.03$).

4) Travel Distance: Participants moved their hands the most while using marking menus, followed by direct input, and finally raycasting (Figure 4b). There was a main effect of TECHNIQUE on *Travel Distance* ($F_{2,85} = 71.5$, $p < .001$). Pairwise comparisons showed MARKING had the most (13.0 cm), followed by DIRECT (9.1 cm), then RAYCAST (5.4 cm) (all $p < 0.001$). There was also a main effect of ITEMS on *Travel Distance* ($F_{1,85} = 8.8$, $p < .001$), showing that 8-ITEM menus resulted in less travel than 24-ITEM menus (8.3 vs. 10.0 cm). There was a TECHNIQUE × ITEMS interaction ($F_{2,85} = 3.0$, $p < .001$) with post-hoc tests finding differences between most conditions ($p < 0.02$), except between DIRECT-24 and (DIRECT-8, MARKING-8) and between MARKING-24 and (MARKING-8, RAYCAST-24, RAYCAST-8).

5) Combined SUS: Participants found raycasting the most usable for 24-item menus, and had no preference between techniques for 8-item menus (Figure 4c). There was a main effect of TECHNIQUE ($F_{2,85} = 14.5$, $p < .001$), ITEMS ($F_{1,85} = 19.1$, $p < .001$) and a TECHNIQUE × ITEMS interaction ($F_{2,85} = 6.7$, $p < .001$) on *Combined SUS*. There were no differences in *Combined SUS* between TECHNIQUES for 8-item menus. For 24-ITEM menus, participants found RAYCAST (M = 87) more usable than MARKING (M = 58) and DIRECT (M = 64) ($p < .001$ for both). There was no difference between DIRECT-24 and MARKING-24. Separate post-hoc tests for each TECHNIQUE showed that participants found MARKING-8 more usable than MARKING-24 ($p < 0.001$). There were no significant effects of ITEMS with other TECHNIQUES.

6) Participant Feedback: Some participants provided feedback via email. P1 preferred RAYCAST due to depth perception: "*[raycasting] felt easiest to use because in situations where I struggle with depth, [raycast] felt like depth perception mattered the least.*" P5 preferred RAYCAST based on movement: "*I liked that I didn't have to move my hands as much, I could just keep my hands at my hips and move my wrist.*" P5 commented on how menu layout affected visual search: "*I felt like [the 24-item marking menu] was disorganized, and like it was so much harder to look through quickly. I liked the [linear or*

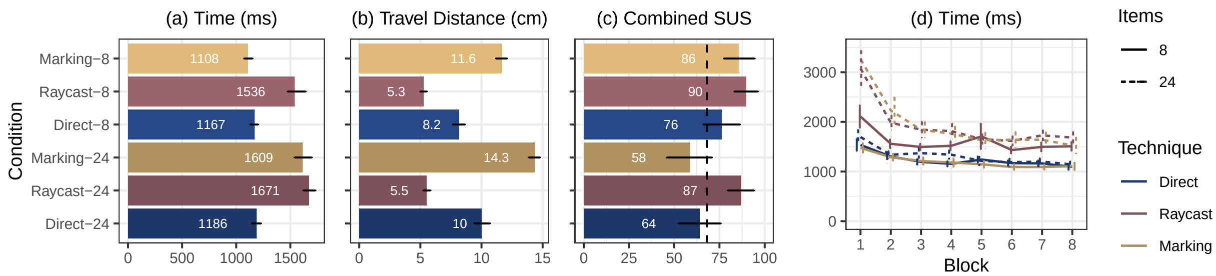

Fig. 4. Experiment 1 results: (a) *Time* after learning; (b) *Travel Distance*; (c) *Combined SUS* (dashed line is average SUS score of 68); and (d) *Time* (ms) by BLOCK. Solid lines represent 8-ITEM conditions, and dashed lines represent 24-ITEM conditions. Error bars are 95% CI.

grid] menus because they had distinct rows that were easier to skim." P4 also commented on issues with MARKING-24 layout: "*sometimes I would bring up the menu and there would be things a few inches from my face. I could still select them fine, but it was disorienting at first.*"

C. Discussion

Overall, our results illustrate how different design characteristics affect performance, accuracy, and perceived usability for single-level menus. Interaction technique and number of items strongly affect performance. Marking menus were most affected by an increase in items, possibly due to factors affecting the speed of the marking gestures. Because the marking menu appears at the position of the user's hand when invoked, selections in some directions may be harder to make depending on invocation position. For example, invoking the menu while the arm is fully extended forward may make selecting an item even further forward more difficult. Additionally, while pilot testing showed that the 3D marking menu mitigates occlusion, the experiment showed that it does not eliminate occlusion. With a high number of items, participants sometimes had to move their head to see some items. This may have impacted visual search time, and as a result, selection time.

Visual search time may have increased when items in the menu appeared beyond the headset -of-view. Although our experiment controlled the visual angular width of individual menu items, the overall width of the direct and raycast menu layouts may have caused some items to be outside the field of view, requiring more head movement during visual search. The 8-item marking menu was fastest overall, possibly because it was narrow enough to fit inside the field of view and not affected by occlusion since it is planar. Additional testing is necessary to draw strong conclusions regarding the effect of field of view on menu performance.

All menu techniques have acceptably low error rates with little variation, considering all remain below 2.5%. This suggests all tested menu types are reliable for selecting a specific item and that our experiment task captures a practical performance balance between accuracy and time.

Decreased physical movement does not translate into decreased selection time. Raycasting primarily uses the rotation of the hand rather than position, so a very low travel distance at either number of items follows. Yet, this conservation of movement did not translate into faster selection times. Hand movement is a critical part of marking interaction, so having the highest travel distance follows. Yet, the selection time was comparable to direct in the 8-item condition. In general, direct input requires more movement, but the bimanual design of the direct menu reduces this. The anchoring hand brings the menu partway toward the selecting hand, thus reducing its movement.

The number of items affects the perceived usability of a menus in different ways. Participants felt that the menus in the 8-item condition were equally usable, but preferred raycasting in the 24-item conditions. Participants may have preferred raycasting due to the relatively stable amount of movement between menu sizes, as shown by the results for *Travel Distance*. The reduced need for depth perception, as mentioned by P1, may have contributed to the higher raycasting score than other techniques. The number of items in the menu could also be seen as more frustrating, as there are additional items to search. In particular, some participants noted that the 24-item direct and marking menus felt overwhelming to use at first. This, combined with the additional proximity to the user's view (despite controlling for visual angular icon width), could account for the more negative SUS scores.

The qualitative feedback, along with the SUS scores from the 3D marking menu, suggest some readability issues with the Fibonacci layout used by the 3D marking. The spiral pattern has lower axial symmetry [28] which reduces occlusion. Yet, this lack of symmetry, likely compounded by the non-planar layout, may lower perceived usability. Users tend to parse an interface in horizontal lines [55] which is supported with direct and raycast, but not possible with the spherical 24-item marking menu layout.

Overall, the results lead us to partially reject **H1**: *Raycast* was the most preferable, but the inconsistency in *Marking* performance meant that *Direct* was the fastest overall.

There are two aspects to understanding menu performance that this experiment did not capture. First, although a single menu level represents the large number of single-level menus in the survey (44%), made it simpler for participants to learn the menus and experiment task, and made it possible for participants to familiarize themselves with menu item locations over time, the survey also found a large number of menus with 2 or more levels. Second, using a fully randomized order for menu item prompts per block is a common experiment design and has good external validity, but it does not measure pure motor performance. It would take many more blocks for participants to fully memorize item locations. Even after the learning effect levelled out, it is very likely there was some hesitation to recall the item location and some visual search time to find the target item.

VI. Experiment 2: Two-Level Menus (**RQ2**)

The single-level menu comparison in Experiment 1 partially answers **RQ2**, but recall 56% of menus we surveyed were hierarchical. The additional complexity of hierarchical selection requires more controlled isolation of motor performance, independent of recall and visual search. Addressing this additional complexity enables a more thorough understanding of relative performance. We more thoroughly answer **RQ2** using a second experiment with two key changes: (1) comparing two-level hierarchical versions of the menu archetypes; and (2) repeating item prompts to better isolate motor performance in more complex selection tasks. As before, the experiment was administered remotely and primary measures were trial time, error rate, hand movement distance, and SUS.

Our previous experiment showed that *Direct* was the fastest in general, primarily because MARKING-24 was much slower than MARKING-8. *Raycast* menus were the most preferable in general. We hypothesize that these trends from Experiment 1 will continue, namely that (**H2**) *Direct* will remain the fastest overall, and *Raycast* will be the most preferred.

A. Protocol

We recruited 18 right-handed people (ages 23 to 64, 10 male, 6 female, 2 non-binary). Note 6 participants also completed the first experiment, but there was a 12-month gap. Remuneration was a $25 e-gift card. VR headsets used by participants were: Oculus Quest (8), Oculus Quest 2 (5), HTC Vive (2), Valve Index (1), Oculus Rift (1), and HP Reverb G2 (1). The compilation and distribution process of the experiment application was the same as the previous experiment. The overall procedure and task were the same as in the previous experiment except the prompt describes which icons to select in both levels of the menu (e.g. "Pigeon > Hat"). This study was approved by our research ethics board, and the entire study session took about one hour, including break times.

Design: This is also a within subjects design with two independent variables: ITEMS (8-ITEMS, 24-ITEMS) and TECHNIQUE (RAYCAST, DIRECT, MARKING). The same shorthand is used: RAYCAST-8, DIRECT-24, etc. Due to the more mentally demanding 2-level interaction and to reach practised motor performance faster, the block and trial design was modified. For each combination of ITEMS and TECHNIQUE, there were 2 BLOCKS of 32 trials. The trials were divided into 8 menu prompts, each prompt repeated four times in sequence. This means the participant completed 4 repetitions of the same menu prompt before moving to the next. First and second level menu icons and positions were maintained throughout a condition. Icons were unique between the two levels. Each two-level selection is considered one trial. The primary computed measures remained: *Time*, *Error Rate*, *Travel Distance*, and *Combined SUS*. In summary: 2 ITEMS $\times$ 3 TECHNIQUES $\times$ 2 BLOCKS $\times$ 8 selection prompts $\times$ 4 consecutive repetitions = 384 data points per participant.

B. Results

For each combination of participant and TECHNIQUE, trials with *Time* or *Travel Distance* more than 3 standard deviations from the mean were excluded as outliers: 177 trials (2.6%) were removed. As before, analysis uses ANOVA with Tukey HSD or Holm post-hoc comparisons, with aligned rank transform if residuals were not normally distributed. All ANOVA preconditions were met. As before, we also explore and report interaction effects to describe how variables have different effects when combined.

1) Learning Effect: As before, we create a combined 6-level factor TECH-ITEMS representing each combination of TECHNIQUE $\times$ ITEMS. We also created a 8-level factor BLOCK-REP, corresponding to each combination of BLOCK $\times$ CONSECUTIVE REPETITION. Figure 5d shows *Time* by ITEMS, TECHNIQUE, and BLOCK-REP with no blocks removed.

There was a main effect for TECH-ITEMS ($F_{5,799} = 133.1$, $p < .001$) and BLOCK-REP ($F_{7,799} = 299.6$, $p < .001$), and a TECH-ITEMS $\times$ BLOCK-REP interaction ($F_{35,799} = 12.4$, $p < .001$) on *Time*. Pairwise comparisons show the first repetition of block 2 was slower than the last 3, and this was consistent for each TECH-ITEMS (all $p < 0.01$). No significant differences were observed between these last 3 repetitions for any TECH-ITEMS. To focus on practised motor performance, subsequent analysis uses only the last 3 prompt repetitions of the second block.

2) Time: Marking menus were the fastest, followed by direct and raycast, with 8-item marking menus faster than all others (Figure 5a). Surprisingly, raycast was consistently slow. There was a main effect of TECHNIQUE on *Time* ($F_{2,85} = 17.7$, $p < .001$). Pairwise comparisons revealed MARKING (1513 ms) was faster than DIRECT (1826 ms) and RAYCAST (1836 ms) (all $p < 0.001$). There was no significant difference between DIRECT and RAYCAST. A main effect of ITEMS ($F_{1,85} = 67.4$, $p < .001$) revealed 8-ITEM menus were faster than 24-ITEM menus overall (1527 ms vs 1923 ms). A TECHNIQUE $\times$ ITEMS interaction effect ($F_{2,85} = 14.6$, $p < .001$) with pairwise comparisons showing MARKING-8 (897 ms) was faster ($p < 0.001$) than RAYCAST-8 (1476 ms) and DIRECT-8 (1500 ms). No significant difference was found between RAYCAST-8 and DIRECT-8. MARKING-8 was also faster than all 24-item techniques ($p < 0.001$ for all). For 24 items, MARKING-24 (1297 ms) was faster than RAYCAST-24 (1674 ms) ($p < 0.001$) but not significantly different from DIRECT-24 (1511 ms).

3) Error Rate: Direct interaction was more accurate than marking, and error rate increased with the number of items, especially for the marking menu. There was a main effect of TECHNIQUE ($F_{2,85} = 7.3$, $p < .001$), ITEMS ($F_{1,85} = 11.5$, $p < .001$) and a TECHNIQUE $\times$ ITEMS ($F_{2,85} = 3.4$, $p < .04$) interaction on *Error Rate*. Pairwise comparisons revealed DIRECT (2.5%) was lower than MARKING (7.1%) ($p < 0.001$), and 8-ITEMS (3.5%) was lower than 24-ITEMS (5.6%) ($p < 0.001$). For specific combinations, DIRECT-24 (3.5%) was lower than MARKING-24 (9.5%) ($p < 0.02$) and MARKING-8 (4.6%) was lower than MARKING-24 ($p < 0.001$). Menus using DIRECT and RAYCAST were not significantly different.

4) Travel Distance: Raycast required the least movement, followed by direct, and then marking (Figure 5b). There was a main effect of TECHNIQUE ($F_{2,85} = 210.6$, $p < .001$), ITEMS

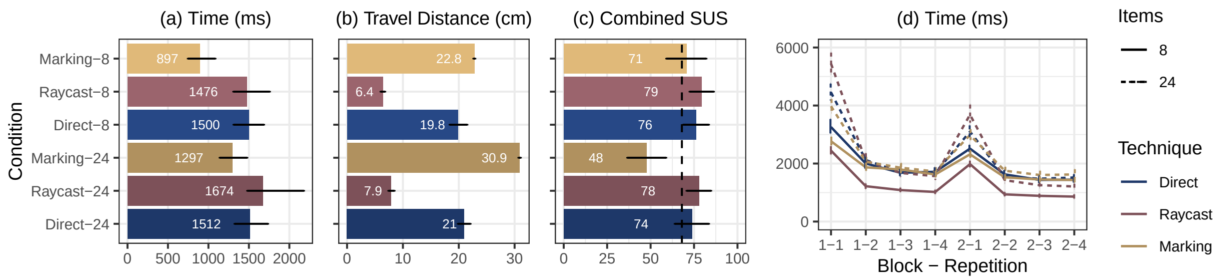

Fig. 5. Experiment 2 results: (a) *Time* after learning; (b) *Travel Distance*; (c) *Combined SUS* (dashed line is average SUS score of 68); and (d) *Time* (ms) by BLOCK-REP. Solid lines represent 8-ITEM conditions, and dashed lines represent 24-ITEM conditions. Error bars are 95% CI.

($F_{1,85} = 56.5$, $p < .001$), and a TECHNIQUE × ITEMS ($F_{2,85} = 21.8$, $p < .001$) interaction on *Travel Distance*. Pairwise comparisons found differences between all TECHNIQUES (RAYCAST: 7.1 cm, DIRECT: 20.4 cm, MARKING: 26.8 cm) (all $p < 0.001$). There was an increase in *Travel Distance* between 8-ITEMS (16.3 cm) and 24-ITEMS (19.9 cm) ($p < 0.001$). For specific combinations, MARKING-8 (22.8 cm) used less movement than MARKING-24 (30.9 cm) ($p < 0.001$), with no differences between DIRECT and RAYCAST for different levels of ITEMS.

5) Combined SUS: All techniques were rated similarly, except for MARKING-24, which was rated lower (Figure 5c). Residuals were normally distributed. There was a main effect of TECHNIQUE ($F_{2,34} = 9.8$, $p < .001$, $\eta^2_G = 0.18$), ITEMS ($F_{1,17} = 10.2$, $p < .001$, $\eta^2_G = 0.06$) and a TECHNIQUE × ITEMS ($F_{2,34} = 9.1$, $p < .07$, $\eta^2_G = 0.07$) interaction on *Combined SUS*. Pairwise comparisons reveal that MARKING (59) had a lower score compared to DIRECT (75) and RAYCAST (79) ($p < 0.004$) and the score decreased from 75 with *8-items* to 66 with *24-items* ($p < 0.01$). For specific combinations of technique and menu, MARKING-8 (71) was much higher than MARKING-24 (48) ($p < 0.002$) with no other significant differences between techniques for 8-items. MARKING-24 was lower than other 24-item techniques (both $p < 0.001$).

6) Participant Feedback: Specific to hierarchical selection, participants noted that MARKING-24 could place items in uncomfortable positions. P4 elaborates: *"sometimes I'll have to bring my hand toward the headset to make the first selection and then half the menu is all around or behind my head."* Similarly, participants found that if menus were invoked too close to their head position, the icon scaling (to maintain visual angular width) could result in icons that are hard to discern – an issue exacerbated by the hierarchical selection conditions.

C. Discussion

The results show how hierarchy and expert-level performance alter the relative usability of direct, raycast, and marking-based menus. The BLOCK by REPETITION analysis (Figure 5d) reveals an interesting effect of immediate rehearsal. For a single prompt, the two sets of four repetitions between blocks 1 and 2 were separated by a significant time gap, while other prompts were completed. The clear "spike" in results for *Time* demonstrates the effect of this time gap, and illustrates why our analysis used only the last 3 repetitions of block 2 for each prompt.

Our results show that hierarchy and expert performance can affect the relative ranking of menu interaction time. In the first experiment, MARKING-24 was among the slowest menus tested, but with the addition of hierarchy and a focus on motor performance, it was only slower than MARKING-8, which was the fastest technique. Previous work shows how fast conventional 2D desktop marking menus are for hierarchical selection [37], [38], and now our results extend this to a 3D spherical marking menu layout. Raycast was the least affected by the addition of hierarchy or expert performance, it remained the slowest menu technique overall.

To understand how hierarchy and expert performance affected absolute selection times, we compare *Selection Time* between our two experiments. To do this, we introduce an EXPERIMENT between-subjects factor, and use the 6-level TECH-ITEMS factor (representing all combinations of TECHNIQUE and ITEMS). Keeping only blocks with practised performance in each experiment, we applied an ART and ran an ANOVA revealing an EXPERIMENT × TECH-ITEMS interaction ($F_{5,170} = 12.5$, $p < .001$). Pairwise comparisons found only one significant difference, MARKING-24 was *faster* in Experiment 2. Even though Experiment 1 had only one level, the random menu item prompt order likely caused some time penalty for visual search and recall. But with the focus on motor performance enabled by the repeated menu item design in Experiment 2, MARKING-24 is much faster, even for more complex 2-level menu interactions.

To understand how expert performance affects menu selection times for a single level, we conducted the same statistical analysis using only the selection time for the first level in Experiment 2. Again, we find an EXPERIMENT × TECH-ITEMS interaction ($F_{5,170} = 12.4$, $p < .001$), with comparisons between pairs of the same menu technique and item number (i.e. same level of TECH-ITEMS) revealing that each menu type is faster in Experiment 2 than in Experiment 1 (all $p < 0.001$). This confirms practised motor performance in Experiment 2.

The results for *Error Rate* demonstrate that hierarchy provides additional opportunities for incorrect selection, especially at a higher number of items. Whereas almost all pairs of techniques in Experiment 1 had similar error rates, differences became more pronounced in Experiment 2. This is especially true for the marking menu, which had the highest error rate once hierarchy was added. The more levels a marking menu has, the greater space it needs for error-free marking [37]. Our marking menu centres the second level at the end of a 17.5 cm movement for a level 1 selection. If both item directions face away, this means a user must reach 35 cm from where they first invoked the menu. This can be beyond arms reach if the menu is first invoked too far from the body. Likewise, invoking the menu near the body when both item directions point towards the user can result in item directions obstructed by their torso. These situations can cause more errors with the 3D marking menus despite its faster selection time.

Introducing added menu complexity with hierarchy highlights usability issues that may be less prominent for simpler single-level menus. Raycast and direct interaction were rated highly, similar to earlier work comparing these techniques in a radial layout [54]. Despite an improvement in selection time, the 3D marking menu was still rated as least usable. In addition to accuracy issues, participant feedback suggests challenges for 3D menu layouts. If invoked too close to the headset, items can spread beyond the field-of-view, even behind the headset. If invoked too far from the headset, dynamically scaling icons for consistent visual size can increase occlusion, since the icon size increases but the menu radius remains constant. This is exacerbated by adding hierarchy, because the second-level position is determined by first-level selection position. A possible solution is to also dynamically scale the menu radius and tune 3D directions to remain in arms-reach.

Overall, adding hierarchy changes the relationship between design characteristics, and the resulting performance and preference. The addition of hierarchy alleviated much of the performance penalty of MARKING-24 relative to other techniques, making *Marking* overall the fastest. All but MARKING-24 were rated similarly preferable. These results lead us to reject **H2**.

VII. General Discussion

We summarize our findings for the first two research questions. For **RQ1**, our survey shows the variety of characteristics found in VR menus today, and the experiments show the effect that these different characteristics can have on overall performance. Menu configurations in consumer VR applications still mostly favour the linear and grid layouts from 2D computing, despite the proven benefits of radial marking selection in research. For **RQ2**, we found that at a single level of hierarchy, direct menus offer the most consistent speed. Marking menus are fast for a small number of items, but slower with a large number of items. With two levels and expert performance, marking menus are consistently fast but are rated less usable than other menu techniques at a high number of items. Raycast menus are consistently the slowest, but have higher perceived usability.

A. Design Recommendations *(RQ3)*

We tested fundamental performance characteristics which are largely context-independent. However, choosing menus for an application requires consideration. For example, a settings menu in a VR 3D modelling application has less need for selection speed, and prioritizes a higher number of items as well as selection accuracy. Conversely, a weapon selection menu in a first-person shooter game requires quick selection for fast-paced gameplay. The importance of different aspects of VR menu performance depends on the application context. For that reason, this section answers **RQ3** by recommending how our empirical results can inform real-world implementations, based on the priorities of individual designers. Table IV shows condensed recommendations by design priority.

Marking Menus for Speed, Direct Menus for Capacity: Designers optimizing for speed should consider directional marking style of interaction for increased selection speed. Our results show that for 8 items, marking and direct techniques have a clear advantage over raycast. Moreover, marking has the advantage of gestural learning and potential for eyes-free operation, making it an especially good choice for fast-paced applications like games. However, at a single level of hierarchy and a higher number of items, the advantage of marking menu gestural learning is reduced. As such, for larger single-level menus, we recommend direct interaction.

This guideline presents novel insight into how marking menus compare to conventional menu types in VR, where the user is making menu selections when holding a controller. This is in contrast with past work, primarily work on mid-air input for 2D displays and AR, which compared gestural pointing selection to other selection approaches [48] or gestural marking menus to novel AR menu types [56].

Single Level Menus for Usability and Accuracy: If subjective usability is paramount, raycast menus offer the most consistent high ratings regardless of number of items. The depth-based selection of MARKING-24 was the lowest rated technique despite being among the fastest with hierarchical selection, suggesting that depth-based selection still requires additional design work to overcome barriers to perceived usability. Similarly, if prioritizing accurate selection, flatter menus like the two-dimensional layouts of direct and raycast offer lower levels of error rate.

While past work on HMD-based VR has not investigated the effect of menu hierarchy, this guideline aligns with past work on hierarchy in other contexts like traditional 2D interfaces [37] and mid-air pie menus [43] which show that deeper hierarchies generally result in comparable or higher error rates.

Raycast Menus to Minimize Movement: Our results show raycast menus require the least controller movement. Discomfort or difficulty when manipulating VR controllers is a common accessibility barrier to VR for people with limited mobility [53]. Designers prioritizing increased accessibility should use raycast selection. The bimanual selection used by our direct menu archetype may also be difficult or even impossible for those with limited mobility, in which case unimanual techniques like raycast or marking may be advantageous.

This guideline is in line with past work on 4-item and 16-item single-level grid menus in VR [58], which found that direct interaction required more movement than raycast; however, our results expand these findings by comparing hierarchical menus, and linear and radial menu layouts.

Marking Menus for Hierarchy when Few Items Per Level: Hierarchy is often more a consequence of the categorization of the menu items themselves, rather than an intentional menu design choice. However, menu item layouts can have significant downstream effects on design priorities, performance, and usability. Direct interaction offers high performance in large, single-level menus. However, with multiple levels of hierarchy, especially with few items per level, marking menus offer the same information capacity with similar performance and ges-

TABLE IV
MENU CHARACTERISTIC RECOMMENDATIONS BASED ON DESIGN PRIORITY. LINKED SECTIONS PROVIDE SUPPORT FOR EACH RECOMMENDATION.

Design Priority		Recommended Menu Characteristics
Performance	Selection time	Marking or Direct techniques; fewer items; no hierarchy [VI-B2]
	Accuracy	Flatter menu layouts (linear or grid) [VI-B3]
	Subjective usability	Raycast technique; flatter menu layouts (linear or grid); fewer items [VI-B5]
	Minimal movement	Raycast technique; no hierarchy; unimanual techniques [VI-B4]
Functionality	Capacity	Direct technique; flatter menu layouts (linear or grid) [V-B2]
	Hierarchical selection	Marking technique; fewer items [VI-C]

tural learning potential. Marking menus also offer consistently high speed at a lower number of items, regardless of hierarchy.

B. Limitations

The wide variety of VR menus' in-game placements and scenarios complicates the accurate measurement of their size and distance from video reviews or gameplay. Future studies should explore these aspects for more detailed design insights.

Two factors introduce external validity limitations. First, the ordering and frequency of tasks are representative of previous work examining practised menu performance [54], [58], [67], but such concentrated repetition of menu item selection is not typical of real-world usage. Real-world performance will depend on user expertise including motor ability as well as locating or recalling item positions. Second, real-world menu items often use text. We used icons only, which is representative of past work examining menu performance [31], [42]. Using icons introduces some interpretation ambiguity, but reduces the irregularity and occlusion introduced by text.

Because our experiment conditions compared both closer hand-anchored and further world-anchored menus, one could consider depth from the user as an implicit secondary independent variable. Depth can affect menu selection, considering that the vergence-accommodation conflict [9] and stereo display deficiencies [7], [8] can slow down target selection. Our menu evaluation did not consider depth movement. In addition to some effect on relative menu performance, this could also affect 3D menus' internal performance, like the 3D marking menu archetype depending on item placement. Future work should examine other 3D menu designs with regard to depth.

We chose to use SUS over NASA-TLX due to its emphasis on overall user experience, as well as being faster and easier for remote participants. This gave us a more holistic self-reported measure of each menu technique's perceived usability, which we thought would be more useful for developers deciding between menu techniques. Previous work [36], [62] shows links between the SUS questions and various types of cognitive load, allowing for indirect workload measurement in addition to a more holistic focus. Future work should more deeply examine cognitive workload and how it affects menu characteristics.

Remote experiments allowed participants to use their own VR headsets, enhancing ecological validity but skewing the sample towards experienced users. While we represented novices by measuring learning and only using stable performance blocks, usability perceptions could vary for total beginners. Participants' differing hardware may introduce variances in tracking accuracy, but we address this by filtering input [16]. At-home participants may face more disruptions than those in-lab. To mitigate this, we designed our experiments and log files to handle interruptions smoothly and included explicit prompts for users to take breaks.

Our experiments aimed to fairly compare techniques by standardizing visual icon size and stabilizing selection rays. However, certain internal parameters might affect performance and usability more than anticipated, suggesting a need for a more systematic study to clarify their influence.

VIII. CONCLUSION

Previous work evaluated various sets of menu techniques and design characteristics drawn from academia, but little work has directly investigated the relative performance of common VR menu techniques directly drawn from real-life configurations.

Our survey of spatial menus in popular VR applications surfaced trends in menu design characteristics which helped inform our three menu archetypes. Raycast and direct methods were obvious choices, and our implementation of radial menu layouts as marking menus is based on decades of work from conventional computing. Our two experiments explore the relative performance of these menu archetypes with small and large numbers of items, with 1 and 2 levels of hierarchy, and with and without expert motor performance. To our knowledge, this is the first VR study to formalize these three archetypal menu types and test them in a controlled setting with consideration for key factors like number of items, hierarchy, and level of expert performance. Our most surprising results are: (1) the relatively poor performance of world-anchored raycast menus, even though they are very common in commercial VR applications; and (2) a 3D spherical marking menu is a viable choice for a relatively large number of items, especially when users reach expert-level motor performance. These results and general recommendations allow designers to make more informed choices for VR interfaces.

REFERENCES

[1] F. Argelaguet and C. Andujar. A Survey of 3D Object Selection Techniques for Virtual Environments. *Computers & Graphics*, 37(3):121–136, 2013.

[2] T. Azai, S. Ogawa, M. Otsuki, F. Shibata, and A. Kimura. Selection and Manipulation Methods for a Menu Widget on the Human Forearm. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*, pp. 357–360. ACM Press, New York, New York, USA, 2017. doi: 10.1145/3027063.3052959

[3] T. Azai, M. Otsuki, F. Shibata, and A. Kimura. Open Palm Menu: a Virtual Menu Placed in Front of the Palm. In *Proceedings of the 9th Augmented Human International Conference*, AH '18. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3174910.3174929

[4] T. Babic, H. Reiterer, and M. Haller. GestureDrawer: One-Handed Interaction Technique for Spatial User-Defined Imaginary Interfaces. In *Proceedings of the 5th Symposium on Spatial User Interaction*, SUI '17, p. 128–137. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3131277.3132185

[5] G. Bailly, E. Lecolinet, and L. Nigay. Visual Menu Techniques. *ACM Comput. Surv.*, 49(4), dec 2016. doi: 10.1145/3002171

[6] M. Baloup, T. Pietrzak, and G. Casiez. RayCursor: a 3D Pointing Facilitation Technique Based on Raycasting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–12. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300331

[7] M. D. Barrera Machuca and W. Stuerzlinger. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–14. ACM, Glasgow Scotland Uk, May 2019. doi: 10.1145/3290605.3300437

[8] A. U. Batmaz, M. D. B. Machuca, D. M. Pham, and W. Stuerzlinger. Do Head-Mounted Display Stereo Deficiencies Affect 3D Pointing Tasks in AR and VR? In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 585–592, Mar. 2019. ISSN: 2642-5254. doi: 10.1109/VR.2019.8797975

[9] A. U. Batmaz, R. Turkmen, M. Sarac, M. D. Barrera Machuca, and W. Stuerzlinger. Re-investigating the Effect of the Vergence-Accommodation Conflict on 3D Pointing. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, VRST '23, pp. 1–10. Association for Computing Machinery, New York, NY, USA, Oct. 2023. doi: 10.1145/3611659.3615686

[10] Ben Plays VR. Let's Play Vanishing Realms - Part 1 - YouTube. https://www.youtube.com/watch?v=zLFU6n5_37A, Aug 2017. (Accessed on 09/13/2020).

[11] Ben Plays VR. Let's Play: No Man's Sky VR - YouTube. https://www.youtube.com/watch?v=cD53j9CQg4s&t=228s, Aug 2019. (Accessed on 09/11/2022).

[12] Ben Plays VR. Let's Play Onward on Oculus Quest! - YouTube. https://www.youtube.com/watch?v=txl52iudDnw&t=86s, Aug 2020. (Accessed on 09/11/2022).

[13] J. Boritz and K. S. Booth. A Study of Interactive 3D Point Location in a Computer Simulated Virtual Environment. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology - VRST '97*, pp. 181–187, 1997. doi: 10.1145/261135.261168

[14] D. Bowman and C. Wingrave. Design and Evaluation of Menu Systems for Immersive Virtual Environments. In *Proceedings IEEE Virtual Reality 2001*, pp. 149–156. IEEE Comput. Soc, 2001. doi: 10.1109/VR.2001.913781

[15] J. Brooke. SUS: A 'Quick and Dirty' Usability Scale. *Usability Evaluation in Industry*, p. 189, 1996.

[16] G. Casiez, N. Roussel, and D. Vogel. 1€ Filter: a Simple Speed-Based Low-Pass Filter for Noisy Input in Interactive Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, p. 2527–2530. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2207676.2208639

[17] Ctop. Making Potions and Casting Spells in VR! - Waltz of the Wizard Gameplay - VR HTC Vive - YouTube. https://www.youtube.com/watch?v=NSlkpNN5-4o&t=92s, Oct 2018. (Accessed on 09/11/2022).

[18] R. Dachselt and A. Hübner. A Survey and Taxonomy of 3D Menu Techniques. In M. Lin and R. Hubbold, eds., *Eurographics Symposium on Virtual Environments*. The Eurographics Association, 2006. doi: 10.2312/EGVE/EGVE06/089-099

[19] R. Dachselt and A. Hübner. Three-dimensional Menus: A Survey and Taxonomy. *Computers & Graphics*, 31(1):53–65, 2007. doi: 10.1016/j.cag.2006.09.006

[20] N.-T. Dang. A Survey and Classification of 3D Pointing Techniques. In *2007 IEEE International Conference on Research, Innovation and Vision for the Future*, pp. 71–80. IEEE, Mar. 2007. doi: 10.1109/RIVF.2007.369138

[21] K. Das and C. W. Borst. An Evaluation of Menu Properties and Pointing Techniques in a Projection-based VR Environment. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 47–50. IEEE, Mar. 2010. doi: 10.1109/3DUI.2010.5444721

[22] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, p. 754–768. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474784

[23] A. Elor, M. Powell, E. Mahmoodi, N. Hawthorne, M. Teodorescu, and S. Kurniawan. On Shooting Stars: Comparing CAVE and HMD Immersive Virtual Reality Exergaming for Adults with Mixed Ability. *ACM Transactions on Computing for Healthcare*, 1(4):22:1–22:22, Sept. 2020. doi: 10.1145/3396249

[24] C. Ford. Post Hoc Power Calculations Are Not Useful | UVA Library, Aug. 2021.

[25] S. Gebhardt, S. Pick, F. Leithold, B. Hentschel, and T. Kuhlen. Extended Pie Menus for Immersive Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):644–651, Apr. 2013. doi: 10.1109/TVCG.2013.31

[26] S. Gebhardt, S. Pick, T. Oster, B. Hentschel, and T. Kuhlen. An Evaluation of a Smart-Phone-Based Menu System for Immersive Virtual Environments. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 31–34. IEEE, Mar. 2014. doi: 10.1109/3DUI.2014.6798837

[27] D. Gerber and D. Bechmann. The Spin Menu: A Menu System for Virtual Environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, vol. 2005, pp. 271–272. IEEE, 2005. doi: 10.1109/VR.2005.1492790

[28] Á. González. Measurement of Areas on a Sphere Using Fibonacci and Latitude–Longitude Lattices. *Mathematical Geosciences*, 42(1):49, 2010.

[29] J. Gray. Steam Charts – Tracking What's Played. https://steamcharts.com/, Jun 2019.

[30] J. Grosjean, J.-M. Burkhardt, S. Coquillart, and P. Richard. Evaluation of the Command and Control Cube. In *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*, pp. 473–478, 2002. doi: 10.1109/ICMI.2002.1167041

[31] T. Grossman, P. Dragicevic, and R. Balakrishnan. Strategies for Accelerating On-line Learning of Hotkeys. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1591–1600, 2007.

[32] M. G. Heckman, J. M. Davis, and C. S. Crowson. Post Hoc Power Calculations: An Inappropriate Method for Interpreting the Findings of a Research Study. *The Journal of Rheumatology*, 49(8):867–870, Aug. 2022. doi: 10.3899/jrheum.211115

[33] J. Henderson, S. Malacria, M. Nancel, and E. Lank. Investigating the Necessity of Delay in Marking Menu Invocation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376296

[34] Honosuseri. Half Life: Alyx - Weapons Showcase & Gameplay with Tips (no Plot Spoilers) - Rift S & 980ti - YouTube. https://www.youtube.com/watch?v=mvmT23RP7fE&t=293s, Mar 2020. (Accessed on 09/11/2022).

[35] R. H. Jacoby and S. R. Ellis. Using Virtual Menus in a Virtual Environment. *Proceedings of SPIE – The International Society for Optical Engineering*, pp. 39–48, June 1992. doi: 10.1117/12.59654

[36] T. Kosch, J. Karolus, J. Zagermann, H. Reiterer, A. Schmidt, and P. W. Woźniak. A Survey on Measuring Cognitive Workload in Human-Computer Interaction. *ACM Comput. Surv.*, 55(13s), July 2023. doi: 10.1145/3582272

[37] G. Kurtenbach and W. Buxton. The Limits of Expert Performance Using Hierarchic Marking Menus. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pp. 482–487. Association for Computing Machinery, New York, NY, USA, 1993. doi: 10.1145/169059.169426

[38] G. Kurtenbach and W. Buxton. User Learning and Performance with Marking Menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Celebrating Interdependence - CHI '94*, pp. 258–264, 1994. doi: 10.1145/191666.191759

[39] J. R. Landis and G. G. Koch. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159–174, 1977.

[40] I. Lediaeva and J. LaViola. Evaluation of Body-referenced Graphical Menus in Virtual Environments. In *Proceedings of Graphics Interface 2020*, GI 2020, pp. 308 – 316. Canadian Human-Computer Communications Society / Société Canadienne du Dialogue Humain-Machine, 2020. doi: 10.20380/GI2020.31

[41] Let's STFU and Play. SUPERHOT VR (Full Playthrough, No Commentary) - YouTube. https://www.youtube.com/watch?v=iIQdbU8vSJc&t=37s, Aug 2017. (Accessed on 09/11/2022).

[42] B. Lewis, G. d'Eon, A. Cockburn, and D. Vogel. Keymap: Improving Keyboard Shortcut Vocabulary Using Norman's Mapping. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–10. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376483

[43] W. Li, X. Wan, Y. Shi, N. Yao, C. Wang, and Z. Gao. Depth and Breadth of Pie Menus for Mid-air Gesture Interaction. *International*

Journal of Human–Computer Interaction, 37(2):131–140, 2021. doi: 10.1080/10447318.2020.1809245

[44] J. Liang and M. Green. JDCAD: A Highly Interactive 3D Modeling System. *Computers & Graphics*, 18(4):499–506, 1994. doi: 10.1016/0097-8493(94)90062-0

[45] Z. H. Lim and P. O. Kristensson. An Evaluation of Discrete and Continuous Mid-Air Loop and Marking Menu Selection in Optical See-Through HMDs. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3338286.3340127

[46] Meta. Interaction SDK Overview — Oculus Developers. https://developer.oculus.com/documentation/unity/unity-isdk-interaction-sdk-overview/. (Accessed on 12/06/2022).

[47] Microsoft. Github - Microsoft/MixedRealityToolkit-Unity: Mixed Reality Toolkit. https://github.com/microsoft/MixedRealityToolkit-Unity. (Accessed on 12/06/2022).

[48] D. M. Mifsud, A. S. Williams, F. Ortega, and R. J. Teather. Augmented Reality Fitts' Law Input Comparison Between Touchpad, Pointing Gesture, and Raycast. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 590–591, 2022. doi: 10.1109/VRW55335.2022.00146

[49] M. R. Mine. Virtual Environment Interaction Techniques. Technical report, Department of Computer Science, University of North Carolina, 1995.

[50] G. Molina, J. Gimeno, C. Portalés, and S. Casas. A Comparative Analysis of Two Immersive Virtual Reality Systems in the Integration and Visualization of Natural Hand Interaction. *Multimedia Tools and Applications*, 81(6):7733–7758, Mar. 2022. doi: 10.1007/s11042-021-11760-9

[51] R. A. Montano Murillo, S. Subramanian, and D. Martinez Plasencia. Erg-O: Ergonomic Optimization of Immersive Virtual Environments. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 759–771. Association for Computing Machinery, New York, NY, USA, Oct. 2017. doi: 10.1145/3126594.3126605

[52] P. Monteiro, H. Coelho, G. Gonçalves, M. Melo, and M. Bessa. Comparison of Radial and Panel Menus in Virtual Reality. *IEEE Access*, 7:116370–116379, 2019. doi: 10.1109/ACCESS.2019.2933055

[53] M. Mott, J. Tang, S. Kane, E. Cutrell, and M. Ringel Morris. "I Just Went Into It Assuming That I Wouldn't Be Able to Have the Full Experience": Understanding the Accessibility of Virtual Reality for People with Limited Mobility. In *The 22nd International ACM Sigaccess Conference on Computers and Accessibility*, ASSETS '20. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3373625.3416998

[54] M. Mundt and T. Mathew. An Evaluation of Pie Menus for System Control in Virtual Reality. In *Proceedings of the 11th Nordic Conference on Human-computer Interaction: Shaping Experiences, Shaping Society*, NordiCHI '20. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3419249.3420146

[55] K. Pernice. Exhaustive Review or "I Can't Believe It's Not There" Phenomenon. https://www.nngroup.com/articles/exhaustive-review-eyetracking/, Oct 2017. (Accessed on 09/13/2022).

[56] S. T. Perrault, E. Lecolinet, Y. P. Bourse, S. Zhao, and Y. Guiard. Physical Loci: Leveraging Spatial, Object and Semantic Memory for Command Selection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 299–308, 2015.

[57] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen. Gaze + Pinch Interaction in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, SUI '17, pp. 99–108. ACM, New York, NY, USA, 2017. doi: 10.1145/3131277.3132180

[58] K. Pfeuffer, L. Mecke, S. Delgado Rodriguez, M. Hassib, H. Maier, and F. Alt. Empirical Evaluation of Gaze-Enhanced Menus in Virtual Reality. In *26th ACM Symposium on Virtual Reality Software and Technology*, VRST '20. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3385956.3418962

[59] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, UIST '96, p. 79–80. Association for Computing Machinery, New York, NY, USA, 1996. doi: 10.1145/237091.237102

[60] M. Pourmemar and C. Poullis. Visualizing and Interacting with Hierarchical Menus in Immersive Augmented Reality. In *The 17th International Conference on Virtual-reality Continuum and Its Applications in Industry*, VRCAI '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359997.3365693

[61] G. Ren and E. O'Neill. 3D Marking Menu Selection with Freehand Gestures. *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 61–68, 2012.

[62] R. Rohani Ghahari, J. Palilonis, G. Hossain, L. Kaser, and D. Bolchini. Semi-Aural Interfaces: Investigating Voice-Controlled Aural Flows. *Interacting with Computers*, 28(6):826–842, 10 2016. doi: 10.1093/iwc/iww004

[63] T. N. Ryan, P. McMahan, and D. A. Bowman. Tech-Note: rapMenu: Remote Menu Selection Using Freehand Gestural Input. In *2008 IEEE Symposium on 3D User Interfaces*, pp. 55–58. IEEE, Mar. 2008. doi: 10.1109/3DUI.2008.4476592

[64] Seeker_HCA. Population One Tutorial (and Small Tips) - YouTube. https://www.youtube.com/watch?v=S5wHNu74l00&t=308s, Oct 2020. (Accessed on 09/11/2022).

[65] Shaken Martini. Lone Echo - The Complete Cinematic Walkthrough at 60fps - YouTube. https://www.youtube.com/watch?v=zxPuZYMIzuQ&t=1514s, Aug 2017. (Accessed on 09/11/2022).

[66] U.S. General Services Administration, Technology Transformation Services. System Usability Scale (SUS) — Usability.gov. https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html, September 2020. (Accessed on 09/15/2020).

[67] E. S. Wall. *An Empirical Study of Virtual Reality Menu Interaction and Design*. PhD thesis, 2021.

[68] Y. Wang, Y. Hu, and Y. Chen. An Experimental Investigation of Menu Selection for Immersive Virtual Environments: Fixed Versus Handheld Menus. *Virtual Reality*, 25(2):409–419, June 2021. doi: 10.1007/s10055-020-00464-4

[69] J. Wentzel, G. d'Eon, and D. Vogel. Improving Virtual Reality Ergonomics Through Reach-Bounded Non-Linear Input Amplification. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, Honolulu HI USA, Apr. 2020. doi: 10.1145/3313831.3376687

[70] S. White, D. Feng, and S. Feiner. Interaction and Presentation Techniques for Shake Menus in Tangible Augmented Reality. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, ISMAR '09, pp. 39–48. IEEE, Washington, DC, USA, Oct. 2009. doi: 10.1109/ISMAR.2009.5336500

[71] N. Yanagihara and B. Shizuki. Cubic Keyboard for Virtual Reality. In *Proceedings of the Symposium on Spatial User Interaction - SUI '18*, pp. 170–170. ACM Press, New York, New York, USA, 2018. doi: 10.1145/3267782.3274687

[72] D. Zielasko, M. Krüger, B. Weyers, and T. Kuhlen. Passive Haptic Menus for Desk-Based and HMD-Projected Virtual Reality. In *2019 IEEE 5th Workshop on Everyday Virtual Reality (WEVR)*. Osaka, Japan, Mar. 2019. doi: 10.1109/WEVR.2019.8809589

[73] D. Zielasko, B. Weyers, M. Bellgardt, S. Pick, A. Meibner, T. Vierjahn, and T. W. Kuhlen. Remain Seated: Towards Fully-Immersive Desktop VR. In *2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)*, pp. 1–6, Mar. 2017. doi: 10.1109/WEVR.2017.7957707