Layerable Apps: Comparing Concurrent and Exclusive Display of Augmented Reality Applications

Brandon Huynh*
University of California, Santa Barbara

Abby Wysopal†
University of California, Santa Barbara

Vivian Ross‡
University of California, Santa Barbara

Jason Orlosky§
Augusta University and Osaka University

Tobias Höllerer¶
University of California, Santa Barbara

Figure 1: Images of the Layerable Apps paradigm showing (left to right) Code Entry and the Application Switcher, Atlas, Item Inspector, and Device Groups, all of which can be displayed concurrently to facilitate passing of information among them. In experiments, users were required to switch between and integrate information from all apps using different methods of application selection.

Abstract

Current augmented reality (AR) interfaces are often designed for interacting with one application at a time, significantly limiting a user's ability to concurrently interact with and switch between multiple applications or modalities that could run in parallel. In this work, we introduce an application model called *Layerable Apps*, which supports a variety of AR application types while enabling multitasking through concurrent execution, fast application switching, and the ability to layer application views to adjust the degree of augmentation to the user's preference. We evaluated Layerable Apps through a within-subjects user study (n=44), compared against a traditional single-focus application model on a split-information task involving the simultaneous use of multiple applications. We report the results of our study, where we found differences in quantitative task performance, favoring Layerable mode. We also analyzed app usage patterns, spatial awareness, and overall preferences between both modes as well as between experienced and novice AR users.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Interaction design—Interaction design process and methods—User interface design

1 Introduction

Today's AR systems are often operated in a single-application paradigm, in which users switch between one active application at a time. Though this model is good for interacting with individual pieces of content, it is not suitable for interaction with and viewing of multiple applications that might be displaying content using different modalities and might need to be cross-referenced with each other. For example, one application may require the use of a pin-pad for text entry, whereas another may augment existing waypoints with annotations. Current devices require the user to switch from one application to another, despite the fact that the applications may be used together, such as note taking during navigation.

In this work, we propose and evaluate the concept of Layerable apps: applications which can be quickly and easily layered on top of each other by the user. Layerable apps provide an increased degree of control and granularity, allowing the user to decide how much of their world is augmented, while still being able to perform tasks that integrate information between multiple applications or require simultaneous interaction between them. One of the goals of this paradigm is to provide a more consistent user experience in which interaction is seamless and application switching is less noticeable. Our primary research questions include:

- Do Layerable Apps provide advantages for multitasking performance?
- What effect does the use of Layerable Apps have on users' application usage and spatial awareness?
- What do user preferences look like when presented with Layerable Apps vs. traditional approaches?

In service of these questions, we implemented a prototype system consisting of an application switcher and four example applications, which are shown in Figure 1. This system allows for application switching via exclusive display (i.e., the currently adopted application switching scheme in most AR operating systems) and concurrent display (our Layerable Apps approach). We designed an experimental task with 44 participants that required users to actively engage with each application, and we used quantitative and qualitative methods to examine how users interact with multiple AR applications under the layerable application model.

2 Related Work

Related work primarily falls into two categories, including research that seeks to develop Augmented Reality as a personal computing paradigm, and view management systems that deal with menu placement and interaction.

2.1 Augmented Reality for Personal Computing

Throughout the development of AR technologies, one goal has been to integrate AR systems into everyday life as a type of personal computing device. For example, Starner et al.'s conceptualization

*e-mail: bhuynh@cs.ucsb.edu
†e-mail: abbywysopal@ucsb.edu
‡e-mail: vivianross@ucsb.edu
§e-mail: jorlosky@augusta.edu
¶e-mail: holl@cs.ucsb.edu

978-1-6654-5325-7/22/$31.00 ©2022 IEEE
DOI 10.1109/ISMAR55827.2022.00104

of an augmented reality wearable interface [31] focused on the use of wearable AR as an assistive technology, acting as a kind of extended memory for the user, capable of storing and retrieving timely information.

A recent survey by Merino et al. [25] provided a comprehensive review of Mixed and Augmented Reality research and identified pervasive and always-on AR as a growing and important topic. Grubert et al. laid a foundation and taxonomy for describing this type of work, termed pervasive augmented reality [11]. Many works have examined individual application scenarios targeting everyday consumers, for example interior design [18], cooking [13, 27], and retail shopping [1, 30].

One such application by Knierim et al. utilized a technology probe to explore the potential of augmented reality usage in the home [20]. They found that most domestic participants were very accepting of AR as a personal technology to be used in domestic spaces, although they had some concerns about privacy and transparency. They identified potential use cases, including the use of AR to support everyday activities like grocery shopping, and the enhancement of everyday objects with new AR functionality.

Our work builds on these usage scenarios by investigating user behaviours and expectations for how to switch between these applications, and how to design the interfaces such that they can operate in seemingly seamless and non-obtrusive ways.

2.2 Information Placement and AR App Management

More recently, researchers have begun to more thoroughly explore different interfaces and paradigms for interacting with multiple information sources, including the simultaneous integration of menus, annotations, and augmentations in the same environment. One of the early attempts at managing a user's view was the work by Bell et al. [5], which allowed for improved placement of text and images such that all content was viewable. Hoang et al. developed a similar system for interacting with in-situ 3D objects from world-relative and head-relative in-situ menus [14]. Probably one of the most comprehensive menu systems was that of Brudy et al, who came up with a number of different menu styles that allowed for in-situ selection and manipulation of menu items [7]. Though not a menu system, Ubii provides for interaction with and selection of icons or other widgets in-situ [15]. Pourmemar took this a step further and developed heirarchical menus that could be used to select from multi-level lists as well as conduct manipulations [29].

In addition to menu-based interaction, context sensitivity has often been integrated into information presentation in AR. Integration of context or context awareness is present in many applications, such as location detection for relevant content placement [23], activity detection for AR video instruction [9], face detection for conversation-based AR [28], and object detection for in-situ language learning [16]. Other interfaces such as Glanceable AR allow for a combination of context and natural glance-based interaction for easy information access [24].

On the commercial front, both the Microsoft HoloLens 2 and the Magic Leap AR headsets have implemented limited forms of multi-app management. On the HoloLens 2 users are limited to certain combinations of a single 'mixed reality app' and a single '2D view' app alongside it [26]. Magic Leap has a 'Landscape' experience which allows multiple apps to display simple 2D content only [21].

While these systems provide a variety of different ways to interact with and view individual applications or specific groups of applications, the management of multiple applications that may be constantly available to the user is still not well explored. Lebeck et al. identified the problem space of multi-app AR laying the foundation for our work [22]. They suggested user-managed application output as a potential solution to the challenges of multi-app AR, which is our central focus.

Our work seeks to address this problem by determining what methods of application activation are most effective for dealing with multiple AR paradigms that are simultaneously available to the user. Simply put, we ask if it is better to manage applications through currently available menu systems that launch apps that take over the user environment exclusively, or if an in-situ layered approach may be more effective?

3 Layerable Applications

Augmented reality can often be described as the layering of digital information on top of the real world. Many futurists envision this digital layer to be a monolithic application that services all the needs of a user. For example, the concept of the Metaverse, where an AR user would engage with a single shared digital layer for all their entertainment and productivity needs, as is the current vision of companies such as Meta and Epic Games. While such efforts are necessary, they are also susceptible to privacy and security implications and could significantly limit users' technology choices, while giving an unprecedented degree of personal access to the companies and stakeholders who own the Metaverse platforms.

In this work, our goal was to explore AR applications not as monolithic do-everything systems, but as smaller, single-purpose, modular elements, with the goal of empowering the user to decide to what extent they want to engage with an augmented world. For this purpose, we came up with Layerable applications, which treat content as a series of "layers" on top of the physical world that can be toggled quickly and seamlessly. Multiple application layers can be used at the same time. This encourages the creation of applications that are still singular in scope, but allow the user to mix and match preferred functions depending on the situation.

When approaching the design and evaluation of Layerable applications, our goals were to (1) create a working prototype system capable of simulating the experience of using Layerable apps, (2) create a set of example applications to implement within the prototype system, and (3) develop an experimental task that required users to engage with each application modality to solve tasks.

3.1 System Design

We implemented a prototype of Layerable applications using Unity and deployed it to the Microsoft HoloLens 2. The system features an application menu that is brought up by looking at the palm of your hands. With one hand, users can bring up the application menu, and with their other hand, they can tap the application icons to toggle the respective application layer on and off. Currently open layers are indicated with a green underline, as shown in leftmost image in Figure 1. The menu is ambidextrous and can be viewed on either hand.

When the system is in *Layerable* mode, application layers can be toggled on and off based on user preference. Users may prefer to use more or fewer applications, or to activate certain applications which have higher or lower amounts of augmentation, depending on their goals and physical situation.

Our system also features an implementation of the single-focus app model for the purposes of comparison in our user study. This model, which we call *Immersive* mode, imitates the behaviour of applications in most contemporary AR headsets. In this mode, apps are launched one at a time, and opening an app will suspend any other currently open app. We chose to re-implement this behaviour within our system instead of using HoloLens' default application launcher to provide a fairer comparison, as a) the HoloLens performs other operating system tasks that dramatically increase the time it takes to open an application, and b) this choice allowed us to use matching visual identities for the UI design of either mode.

Table 1: Application Categories

	Context-aware	Context-free
2D Presentation	Item Inspector	Code Entry
3D Presentation	Device Groups	Atlas

3.2 Representative Applications

The visual presentation and interaction capabilities of an AR application can vary significantly depending on the intent of the application designer. There is no widely accepted standard by which AR apps should look and feel. This makes it difficult to implement meaningful exemplar applications for testing an application switching system. With the Layerable Apps prototype, we identified a minimal taxonomy for the most common styles of AR applications that the system should support. We identified 2D and 3D as modalities of application presentation. 2D applications are those where all of their graphics are rendered within the confines of a 2D plane, though the plane itself may exist in 3D space. Notably, this encompasses all instances of traditional applications found on desktops and touchscreen devices, making it plausible to port those applications into our Layerable App system (a pathway that Microsoft has outlined for their 2D windows universal platform (UWP) apps and Windows Mixed Reality). 3D applications are all other applications that render graphics at multiple 3D positions. We found these categories to be representative of nearly all types of AR applications found on current commercially available HMDs.

Additionally, we wanted to incorporate some element of context awareness into our design. This was inspired by Grubert et. al's work on Pervasive Augmented Reality [11], which suggests that future AR applications are likely to feature context-sensitive functionality. We chose to further categorize applications by whether or not they utilize context. Thus, our final design includes 4 example applications, with each app representing one of the 4 possible combinations of context-awareness and spatiality (cf. Table 1).

After enumerating the desired application types to support in our system, we implemented representative applications for each category. When conceptualizing the design of these apps, we tried to think of functionality that would be desirable to users in a real world setting. Images showing the contents of each application are shown in Figure 1. The applications we arrived at over many iterations and pilot evaluations are the following:

Item Inspector. Inspired by Internet of Things (IoT) applications, Item Inspector allows you to visually inspect the status and associated technical information about objects and devices in your home, such as battery life, model number, and manufacturing date. Information about each object is displayed within a 2D plane fixated above the object itself as shown in the 2nd and 3rd images in Figure 1. Object locations were tracked using manually placed spatial anchors in our controlled testing environment.

Device Groups. Using device groups, participants can group physical objects together in their space in order to perform aggregate actions such as turning all devices in a group on or off. Device groups are visually represented with colored lines connecting every object in a particular group to every other object in that group. These lines are rendered in 3D space, allowing the user to quickly grasp which objects are part of a group and where their locations are in the space.

Code Entry. This app enables users to virtually enter passcodes and pin numbers in place of traditional keypads on door locks, ATMs, and other security systems. In practice, the app functions similar to a calculator, displaying a number pad on a 2D plane, but not with any spatial dependency on any specific objects in the environment. This makes it suitable as a representative for a context-free and 2D application.

Atlas. Atlas displays a large 3D model of planet Earth that users can explore, displaying geographical information about cities and landmarks around the world. The model is rendered intentionally large – it can be scaled within certain limits but maintains a minimum size so as to 'fill' the space and require users to walk around when looking for a particular location.

3.3 Experimental Task

Our goal was to design an experimental task that would require the user to engage with all 4 applications in order to complete the task. We chose to employ a split-information task where the necessary pieces of information needed to complete the task are split up and distributed to each representative application.

In the study, users were tasked with finding pieces of a 6 digit code. Each code was split into 3 code fragments and each fragment was embedded into random 'flavor text' within the Item Inspector, Device Groups, and Atlas applications. Each code fragment also featured 2 leading alphabetical characters to help identify which fragments belonged to the same code. For instance, the user might encounter the fragment SC-12 in one app, SC-23 in another app, and SC-89 in a third app. After finding all 3 code fragments for a corresponding code, users could enter the digit pairs into the Code Entry application in any order. Participants were scored by the number of correct codes entered within a fixed amount of time. There were no penalties for incorrect codes (apart from the elapsed time used to enter them).

4 Experiment Design

We conducted a within-subjects user study with 44 participants over the course of 2 weeks. The study was conducted primarily with students and affiliates at a university campus and included students from different departments as well as local community participants signed up with a human subjects pool managed by the university. Study sessions took approximately 1.5 hours to complete. The average age of participants was 22.9, with 16 male, 25 female, and 3 identifying as non-binary.

4.1 Procedure

Participants filled out a demographics questionnaire and consent form prior to arriving for the experiment. Upon arriving, participants were trained on how to perform hand gestures within the HoloLens. Specifically, they were taught how to press a button and how to air tap on buttons that were far away. Participants used a training application that provided multiple opportunities to test their ability to execute the gestures correctly. Participants were asked if they were confident in their ability to execute the gesture before proceeding.

Following training, participants were placed into the Layerable Apps prototype and provided with a guided tutorial on how to complete the experimental task. The tutorial provided step-by-step instructions with text bubbles and text-to-speech voiceover, demonstrating how to open application layers, how to find codes hidden in each application, what the structure of the codes were, and how to input them. Participants were required to find and enter a code successfully to complete the tutorial. Afterwards, participants were asked to verbally describe to the experimenter, in their own words, what the task was and how to complete it.

After completing the training and tutorial, participants performed 4 task sessions of 7 minutes each, alternating between *layerable* and *immersive* modes. Participants were counter-balanced with respect to their starting application mode. After the 1st task session, participants were asked to fill out a post-task questionnaire to capture their thoughts on the usability of that mode, as well as an object recognition quiz and object placement quiz where they were asked to recall information about objects in the scene. After the 2nd task session, participants were administered another post-task questionnaire to capture their thoughts on the alternative app switching mode. Finally, a post-study questionnaire and semi-structured interview was administered following the 4th and final task session. Throughout the study, modes were coded to "Mode A" for layerable and "Mode B" for immersive to avoid name bias.

To validate the sufficiency of our training procedures, we asked participants to rate their understanding of the tutorial, the task, and ease of use of the HoloLens, on a 7-point Likert scale.

4.2 Metrics

For task metrics, we measured participants score after each task session. We also tracked the number of mistakes made during each session. We measured the number of times participants opened and closed applications, as well as the average time they spent using each application. For layerable modes, we calculated the average number of application layers open on a per-frame basis.

To measure the usability of each mode, we employed the System Usability Scale [6], which was part of the post-task questionnaires. We also employed a single ease question in the post-task questionnaire to assess how difficult users found the task under each mode.

To measure spatial awareness, we employed an object recognition quiz as well as an object placement quiz. We used the same testing and scoring methodology as Suma et al. [32] in their previous work evaluating cognitive effects of exploration in mixed reality spaces. For the object recognition quiz, participants were given a list of 30 objects, with half of the objects actually being present in the experiment space, and the other half being absent. Participants were asked to answer true or false for each object. The number of false positives was subtracted from the number of true positives, yielding a score between 0 and 15. Following the object recall quiz, participants took an object placement quiz in which they were given the correct list of 15 present objects, and asked to mark their locations on a 2D top-down view floor plan of the space to the best of their memory. Participants were scored based on the number of objects that were correctly placed relative to other objects, for a max score of 15.

In our post-study questionnaire we focus on overall preferences. Participants were asked to rank in which mode they felt the most productive, fastest, focused, distracted, spatially aware, or tired. We asked participants which mode they preferred the most, which was the easiest to use, and which was the most enjoyable.

5 Results

We examine differences between users in Layerable and Immersive modes, with a focus on evaluating Layerable apps in the context of personal computing. Additionally during the course of piloting the user study, we also noticed a trend where users who had prior AR experience tended to score higher overall. Recognizing the importance of application switching in the context of productivity tasks, we decided to examine differences between experienced AR users vs. novice users. Expertise was determined based on subjective responses from the pre-study questionnaire, where participants were asked on a scale of 1 to 5 how familiar they were with Augmented Reality. Those who answered 4 or 5 were considered as experienced AR users. Using this criteria resulted in 23 users categorized as experienced and 21 as novice.

5.1 Tutorial Adequacy and HoloLens Usability

We assessed the suitability of our training procedures in post-hoc questionnaires employing a 7-point Likert scale (higher numbers indicating higher amounts of understanding of the tutorial, the task, and ease of use of the HoloLens). 86.4% of participants rated highly (5 or higher) for tutorial understanding, 95.5% of users rated highly for task understanding, and 97.7% of participants rated highly for ease of use.

5.2 Task Performance

Participants were scored based on the number of codes they were able to find and enter successfully within a 7-minute task session.

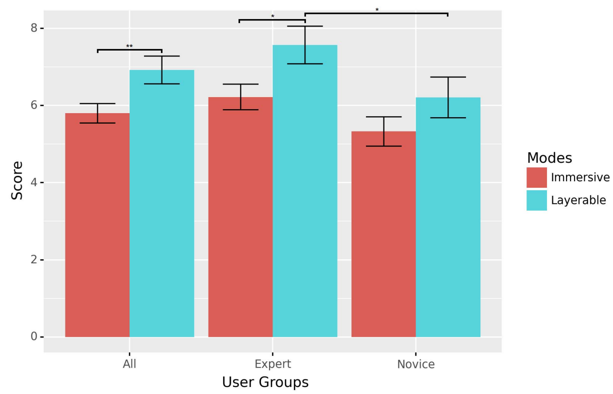

Figure 2: Average task score in terms of number of codes successfully found. These are shown according to user group and application mode.

We averaged scores for both Immersive and Layerable modes for all participants, as well as for the subgroups of Expert and Novice, shown in Figure 2. We compared scores between modes using the Wilcoxon signed-rank test with Bonferroni correction to account for multiple comparisons error. We report effect size r adopting Cohen's classification [8] of small (0.1 to 0.3), medium (0.3 to 0.5) and large (>0.5) effect sizes. For significance tests, we used α of 0.05.

When looking at all participants as a whole, we found significantly higher scores ($p = .002, r = .365$) when completing the task under Layerable mode, averaging 6.92 ($SD = 3.39$) compared to 5.8 ($SD = 2.39$) under Immersive mode. We also found significance among experienced AR users ($p = .012, r = .425$), with an average score of 7.57 ($SD = 3.29$) compared to 6.22 ($SD = 2.27$) in Immersive. We did not find significance between modes for the novice group ($p = .178, r = .291$), with an average score of 6.21 ($SD = 3.4$) under Layerable and 5.33 ($SD = 2.46$) under Immersive.

We used the Mann-Whitney U test to analyze differences between experienced and novice users. When aggregating across both modes, experienced users scored significantly higher ($p = .005, r = .242$) in the experimental task ($M = 6.89, SD = 2.89$) compared to novice users ($M = 5.77, SD = 2.98$). They also scored significantly higher ($p = .035, r = .26$) when using Layerable mode compared to novice users. Results were inconclusive ($p = .068, r = .225$) when comparing Immersive scores between expert and novice users, with a trend to higher performance by expert users.

We also looked at the number of mistakes participants made during each task session but did not find any significant differences between modes or experience levels. Mistakes were defined as incorrect code entries. On average, participants made 1.59 ($SD = 1.7$) mistakes with Layerable and 1.61 ($SD = 1.97$) with Immersive. Experienced AR users averaged 1.78 ($SD = 1.78$) mistakes compared to 1.4 ($SD = 1.89$) for novices.

5.3 Application Usage

We examined application usage behaviours by looking at time spent in each app as well as app switching actions. We measured the duration of time applications were kept open during a task session. Applications were automatically closed at the end of each task session, so the maximum length of time is 7 minutes. Among the total sample population, participants kept apps open significantly longer ($p < .001, r = .867$) in layerable mode ($M = 130, SD = 156.22$), compared to immersive mode ($M = 9.01, SD = 21.58$). While this is unsurprising, it is noteworthy that applicants did not simply leave apps open continuously, as one might expect that to be an optimal strategy. We discuss this further in Section 6.

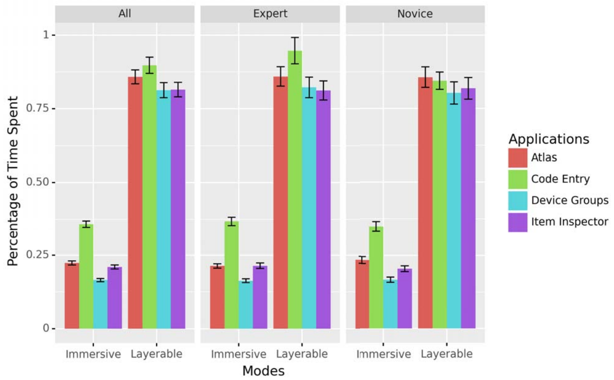

Figure 3: Average percentage of time spent in each app. Note that this includes all time that the application was open, not necessarily time spend directly interacting with content.

Additionally, we calculated the average proportion of time spent in each app within a single task session. A breakdown of the mean proportion of time spent in each application during a task session can be found in Figure 3. Though we did not find any significant differences, we can see interesting trends in usage that may warrant further exploration. For instance, users appear to engage with context-free applications (Atlas and Code Entry) slightly more often than context-aware applications (Device Groups and Item Inspector). This may be due to users perceiving the space as more cluttered with context-aware applications. We also see that in general, Code Entry is used more often than any other app, with the exception of novice users under layerable mode. This is surprising and may by an indication of choice overload amongst novice users who are inexperienced with the different capabilities of AR.

For application switching behaviour, we measured application open and close actions performed by each user. It should be noted that since apps can be layered and open simultaneously, opening an application does not necessarily mean a user has switched their attention to that application's contents. The average number of application open actions in layerable mode were 21.3 ($SD = 12.96$) and 90.34 ($SD = 26.19$) in immersive mode. A Wilcoxon Signed-Rank test showed these to be significantly different ($p < .001, r = .859$). Average counts for the close app action were 14.8 ($SD = 13.49$) for layerable and 88.64 ($SD = 26.1$) for immersive modes. We found these means to also be be significantly different ($p < 0.001, r = .851$).

When comparing between user groups, we did not find differences between experts and novices on the amount of open actions in layerable ($p = .541$) or immersive ($p = .125$), nor did we find differences for close actions in layerable ($p = .663$) or immersive modes ($p = .126$) modes.

Specifically for layerable mode, we were interested in the number of applications participants kept open at any given time. We averaged the number of apps open on a per frame basis for each task session when using layerable mode. Mean apps open was 3.39 ($SD = .694$), with that number being slightly higher amongst AR experts ($M = 3.43, SD = .610$) and slightly lower ($M = 3.35, SD = .781$) amongst AR novices. We wanted to know if participants cycled between different applications or if they chose to keep all applications open simultaneously to provide themselves with the most available information for completing the task. We used a one-sided Wilcoxon Signed-Rank test, subtracting the expected mean of 4 from all our sample observations to fit the test hypothesis. We found average apps open to be significantly less ($p < .001$) than our expected mean, suggesting users don't opt to use all apps simultaneously even though that is, in our opinion, the more efficient strategy.

5.4 Usability

We employed the System Usabiliy Scale (SUS) [6] after the first use of each mode, as well as a Single Ease Question (SEQ) rating the ease of task completion from 1 (easy) to 7 (difficult). A Wilcoxon Signed-Rank test showed no differences in SUS score ($p = .607$) between modes with a mean layerable mode score of 67.27 ($SD = 13.91$) and mean immersive mode score of 68.47 ($SD = 15.58$). There were also no differences found between modes when examining the scores of the experienced user groups ($p = .425$) and novice user groups ($p = .708$). When interpreting SUS scores, an 'OK' score is generally 51-71 and a 'Good' score is generally 72-85, so both layerable and immersive modes fall somewhere between 'OK' and 'Good' [4]. We also found no differences in SEQ score between modes ($p = .549$), nor amongst experienced users ($p = .773$) or novices ($p = .598$).

5.5 Spatial Awareness

We tested the effects of each mode on spatial awareness with an object recognition and object placement quiz. Quizzes were administered after the first task session only, as we did not want to influence task performance by having participants divert attention to memorizing parts of the space in later trials. As we counterbalanced the starting modes for each participant, we can effectively treat these results as coming from independent groups, but with spatial awareness results for only n=22 (half our user population) participants for each mode. We use a Two-Way ANOVA to analyze the quiz scores, and confirmed normality using Shapiro-Wilk test as well as homogeneity of variances using Levene's test. We used system mode (layerable vs. immersive) and AR experience (expert vs. novice) as our independent factors, using quiz score as our dependent variable for the ANOVA model. Post-hoc analysis was performed using Tukey's HSD test for all pairwise comparisons.

Table 2: Two-Way ANOVA of Object Placement Scores

effect	sum sq	df	F	PR(>F)
Mode	50.62	1	5.64	**0.022**
Experience	16.97	1	1.89	0.177
Mode x Experience	8.58	1	0.96	0.334
Residual	358.99	40		

Both quizzes had a max score of 15. Please refer to section 4.2 for details on how the quizzes were scored. In the object placement quiz, we found a significant main effect of system mode on the object score ($p = .026, r = .665$), with users averaging a score of 8.64 ($SD = 2.5$) in layerable compared to 6.55 ($SD = 3.47$) in immersive. We did not find any effect for user experience level ($p = .212$), nor did we find any significant interaction effects between mode used and user experience ($p = .334$). Summary statistics for the ANOVA model are shown in Table 2.

For the object recognition quiz, we did not find any statistical significance for either system mode ($p = .497$) or experience level ($p = .245$), nor did we find any significance for the interaction between independent factors ($p = .692$). The average score was 6.5 ($SD = 3.57$) for layerable and 5.73 ($SD = 3.87$) for immersive.

The significantly higher object placement score when using layerable is notable, as in the next section we will see that most users rated themselves as more aware of their surroundings in the immersive mode rather than the layerable mode. While immersive mode may give the feeling of greater spatial awareness due to increased visibility of the physical scene, users' actual spatial awareness performance may be better facilitated by the increased context-related content and visual stimuli in the layerable mode. Similar results have been found in other works regarding AR and memory [17].

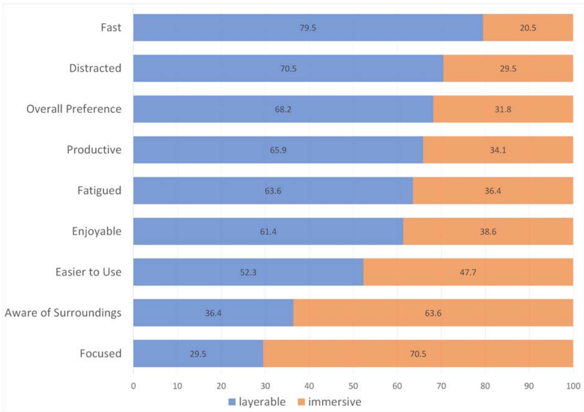

Figure 4: Proportion of User's Mode Preferences

5.6 Overall Preferences

In the post-study questionnaire, we asked users to rank their preferred modes based on several different criteria, including which mode they felt faster, more productive, more distracted, more fatigued, more focused, and more aware of their surroundings in. We also asked which mode users enjoyed the most, found easiest to use, and preferred overall. Figure 4 shows the proportion of user responses for each ranking criteria.

Concerning attitudes around multitasking, 79.5% of users felt faster in layerable, and 65.9% of users felt more productive in layerable. These self-reported rankings fall in line with our task performance results, suggesting that from a task efficiency standpoint, layerable appears to be better.

However, a majority (63.6%) of users also found layerable more tiring to use. We had designed layerable apps with the goal of reducing context switching fatigue, but that does not appear to be the outcome. We believe that while context switching fatigue may be reduced compared to immersive mode, overall fatigue is increased due to increased visual demands or eye strain.

Only 36.4% of users felt more "aware of their surroundings" in layerable and only 29.5% of users felt more focused. 70.5% of users ranked layerable as the more distracting option. These results are counter-intuitive considering spatial awareness quiz scores were generally higher and in some cases significantly higher for layerable mode. We believe these results are due to participants attributing the quality of being "aware of surroundings" to their visibility of the real world. When self-reporting on their spatial understanding, users appear to be biased towards consciously perceived visual cues, which may not be indicative of their actual spatial understanding. Even in AR/VR settings [3, 12], spatial awareness is additionally facilitated by other unconscious non-visual inferences such as orientation processing.

6 Discussion

Reviewing our initial research questions, our results show Layerable Applications to be a promising application model for Augmented Reality. Layerable was ranked as the more preferred mode to use and was also ranked as more enjoyable by the majority of users in our study.

Statistical analysis showed significant improvements in performance on our tasks (which necessitated cross-referencing) when using layerable applications, compared to a traditional single-application model, on average higher by 1.12 points. We also found significant improvements on layerable apps task performance for experienced AR users compared to novice users, suggesting its suitability as an application paradigm for 'power users' who have more technical knowledge or are willing to overcome the initial learning curve. We designed Layerable Apps to increase the degree of control users have on the augmented world. One of our main research questions was to determine user preferences around augmentation control, as such information could be used to inform future application designs. We were concerned due to the nature of the task that users would open all apps all the time, but that was not the case. Rather, our results show that users do frequently choose to switch between applications in layerable mode, switching apps an average of 21.3 times and using app instances an average of 130 seconds.

We found that the number of apps they kept open at any given time was significantly less than the total number of apps available, even though opening all apps may have provided a potentially faster pathway for the task (if one were to discount negative effects from clutter and information overload). It looks like users self-regulated the amount of information display they were willing to take in at a time, shielding against higher levels of clutter and information overload. Additionally, we found some evidence of users being more spatially aware in layerable AR. When analyzing object placement scores, where users had to position objects on a 2D floor plan of the experiment space, those who started in layerable scoring significantly higher. However in contrast to that result, participants also ranked layerable as causing them to be less 'aware of their surroundings'. These results are interesting, and more work needs to be done to find the 'sweet spot' of number of applications and degree of augmentation that users prefer to use.

7 Future Work

This work focuses on a simple implementation of the Layerable Apps paradigm, where the onus on view management is strictly on the user. While this form may be appealing to power users and early adopters, it may not be appropriate for mass adoption. In future iterations, we would like to explore how to incorporate view management [5,10,33] and information filtering [19,34] as an element, while preserving the degree of user control that helps distinguish Layerable Apps from other application paradigms. For instance, it may be possible to define a standard set of rules for the presentation and layout of AR app elements, similar to HTML and CSS for web design. Such a system could alleviate the issues of visual fatigue while maintaining the productivity benefits of Layerable Apps.

There is also a mental load involved in determining which applications are appropriate to use in which context and a related challenge for app developers in testing their application to work well in a variety of contexts, as reported in recent developer surveys [2]. One potential solution that we would like to incorporate into Layerable Apps is the inclusion of a "target scene description" with each application, indicating the types of spaces that are appropriate for the application, perhaps in the form of a hierarchical description of objects and surfaces in the scene or similar spatial representation structure. This would provide the user with a quantifiable indicator of how appropriate an application is based on how closely their current space matches the target scene description. Developers would also benefit by being able to narrowly scope their application's operational context and having concrete test cases that they could evaluate their app on.

8 Conclusion

Simultaneous usage of multiple applications in Augmented Reality is a challenging but important problem to solve. In this work, we set forth and evaluated one application model that supports concurrent display of application content, which we call Layerable Apps. We compared a prototype implementation against the commonly used single-application display paradigm through a within-subjects user study with 44 participants. We found significantly higher task performance and demonstrated spatial awareness when using layerable apps, and a majority of users preferred this mode overall. We also analyzed results between experienced and novice AR users and found that experienced users had significantly higher task performance in Layerable as well, suggesting an additional benefit of the system for 'power users.' We documented our design process for the system prototype, experiment task, and choice of sample applications, and analyzed application usage during the study to provide insight towards the design of future multi-app AR interfaces.

Acknowledgments

This work was supported in part by the Office of Naval Research, under grants #N00014-19-1-2553, #N00174-19-1-0024, and #N62909-18-1-2036, and by grant #21H03482 from the Japan Society for the Promotion of Science. Additional funding came from NSF award IIS-1911230.

References

[1] J. O. Álvarez Márquez and J. Ziegler. In-store augmented reality-enabled product comparison and recommendation. In *Fourteenth ACM Conference on Recommender Systems*, pp. 180–189, 2020.

[2] N. Ashtari, A. Bunt, J. McGrenere, M. Nebeling, and P. K. Chilana. Creating augmented and virtual reality applications: Current practices, challenges, and opportunities. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pp. 1–13, 2020.

[3] B. Bahrami, D. Carmel, V. Walsh, G. Rees, and N. Lavie. Spatial attention can modulate unconscious orientation processing. *Perception*, 37(10):1520–1528, 2008.

[4] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[5] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pp. 101–110, 2001.

[6] J. Brooke. Sus: a "quick and dirty'usability. *Usability evaluation in industry*, 189(3), 1996.

[7] F. Brudy. Interactive menus in augmented reality environments. *Beyond the Desktop*, p. 1, 2013.

[8] J. Cohen. *Statistical power analysis for the behavioral sciences*. Routledge, 2013.

[9] D. Damen, T. Leelasawassuk, O. Haines, A. Calway, and W. W. Mayol-Cuevas. You-do, i-learn: Discovering task relevant objects and their modes of interaction from multi-user egocentric video. In *BMVC*, vol. 2, p. 3, 2014.

[10] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Image-driven view management for augmented reality browsers. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 177–186. IEEE, 2012.

[11] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE transactions on visualization and computer graphics*, 23(6):1706–1724, 2016.

[12] L. R. Harris, M. Jenkin, and D. C. Zikovitz. Visual and non-visual cues in the perception of linear self motion. *Experimental brain research*, 135(1):12–21, 2000.

[13] H. Hasada, J. Zhang, K. Yamamoto, B. Ryskeldiev, and Y. Ochiai. Ar cooking: comparing display methods for the instructions of cookwares on ar goggles. In *International Conference on Human-Computer Interaction*, pp. 127–140. Springer, 2019.

[14] T. N. Hoang and B. H. Thomas. Augmented reality in-situ 3d model menu for outdoors. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 185–186. IEEE, 2008.

[15] Z. Huang, W. Li, and P. Hui. Ubii: Towards seamless interaction between digital and physical worlds. In *Proceedings of the 23rd ACM international conference on Multimedia*, pp. 341–350, 2015.

[16] B. Huynh, J. Orlosky, and T. Höllerer. In-situ labeling for augmented reality language learning. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1606–1611. IEEE, 2019.

[17] A. Ibrahim, B. Huynh, J. Downey, T. Höllerer, D. Chun, and J. O'donovan. Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE transactions on visualization and computer graphics*, 24(11):2867–2874, 2018.

[18] S. Irawati, S. Green, M. Billinghurst, A. Duenser, and H. Ko. " move the couch where?": developing an augmented reality multimodal interface. In *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 183–186. IEEE, 2006.

[19] S. Julier, M. Lanzagorta, Y. Baillot, L. Rosenblum, S. Feiner, T. Hollerer, and S. Sestito. Information filtering for mobile augmented reality. In *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, pp. 3–11, 2000. doi: 10.1109/ISAR.2000.880917

[20] P. Knierim, P. W. Woźniak, Y. Abdelrahman, and A. Schmidt. Exploring the potential of augmented reality in domestic environments. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1–12, 2019.

[21] M. Leap. Landscape design. Article, May 2019. Retrieved June 4, 2022 from `https://developer.magicleap.com/en-us/learn/guides/design-landscape`.

[22] K. Lebeck, T. Kohno, and F. Roesner. Enabling multiple applications to simultaneously augment reality: Challenges and directions. In *Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications*, pp. 81–86, 2019.

[23] D. Lindlbauer, A. M. Feit, and O. Hilliges. Context-aware online adaptation of mixed reality interfaces. In *Proceedings of the 32nd annual ACM symposium on user interface software and technology*, pp. 147–160, 2019.

[24] F. Lu, S. Davari, L. Lisle, Y. Li, and D. A. Bowman. Glanceable ar: Evaluating information access methods for head-worn augmented reality. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 930–939. IEEE, 2020.

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

[26] Microsoft. App views. Article, March 2022. Retrieved June 4, 2022 from `https://docs.microsoft.com/en-us/windows/mixed-reality/design/app-views`.

[27] T. Narumi, Y. Ban, T. Kajinami, T. Tanikawa, and M. Hirose. Augmented perception of satiety: controlling food consumption by changing apparent size of food with augmented reality. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 109–118, 2012.

[28] J. Orlosky, K. Kiyokawa, T. Toyama, and D. Sonntag. Halo content: Context-aware viewspace management for non-invasive augmented reality. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, pp. 369–373, 2015.

[29] 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*, pp. 1–9, 2019.

[30] M. Riar, J. J. Korbel, N. Xi, R. Zarnekow, and J. Hamari. The use of augmented reality in retail: a review of literature. In *Proceedings of the 54th Hawaii International Conference on System Sciences*, p. 638, 2021.

[31] T. Starner. Human-powered wearable computing. *IBM systems Journal*, 35(3.4):618–629, 1996.

[32] E. Suma, S. Finkelstein, M. Reid, S. Babu, A. Ulinski, and L. F. Hodges. Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 16(4):690–702, 2009.

[33] M. Tatzgern, D. Kalkofen, R. Grasset, and D. Schmalstieg. Hedgehog labeling: View management techniques for external labels in 3d space. In *2014 IEEE Virtual Reality (VR)*, pp. 27–32. IEEE, 2014.

[34] M. Tatzgern, V. Orso, D. Kalkofen, G. Jacucci, L. Gamberini, and D. Schmalstieg. Adaptive information density for augmented reality displays. In *2016 IEEE Virtual Reality (VR)*, pp. 83–92, 2016. doi: 10.1109/VR.2016.7504691