User-Centered Evaluation of Different Configurations of a Touchless Gestural Interface for Interactive Displays

Vito Gentile[1], Habiba Farzand[2], Simona Bonaccorso[3], Davide Rocchesso[3], Alessio Malizia[4], Mohamed Khamis[2], and Salvatore Sorce[5]

[1] synbrAIn srl, Milano, Italy
[2] University of Glasgow, United Kingdom
[3] Università degli Studi di Palermo, Italy
[4] Università degli Studi di Pisa, Italy
[5] Università degli Studi di Enna "Kore", Italy
salvatore.sorce@unikore.it

Fig. 1. Interactive displays are now commonly used in various contexts such as airports, train stations, and much more. The figure shows an example of interacting with a display in the presence of an avatar.

Abstract. Approaches for improving the user experience when interacting with touchless displays have been proposed, such as using activation gestures and representing users as avatars in real-time. However, the novelty of such approaches may hinder users' natural interaction behavior bringing challenges such as ease of use. In this paper, we investigate how the presence of avatars and their configurations, the usage of activation gestures, and the arrangement of interactive tiles in a touchless visual interface impact users' experience, usability and task performance. We also

compare users' willingness to promote the interaction setup, perceived task difficulty, and time consumed to perform four different tasks in each configuration. We found that using a squared arrangement of elements, adopting activation gestures to trigger actions, and showing a moving avatar, resulted in the highest perceived usability and user experience, also reducing errors, task completion time, and perceived task difficulty. Our findings support the design of interactive displays to ensure high usability and user experience.

Keywords: Touchless Gestural Interfaces · Public Displays · Interface Evaluation

1 Introduction

Hands-free interaction with displays is becoming more and more pervasive. Interactive displays are being deployed at various public places such as airports, train stations, and alike [9]. While they are being adopted at a fast speed, novel interaction styles have also been proposed such as using avatars to represent the user and facilitate interaction [13]. However, these novel methods bring along user experience and usability challenges.

Avatars and activation gestures have been studied in many contexts in terms of perceived cognitive load [13] or their effectiveness in communicating touchless interactivity [29]. Prior work focused on exploring perceived cognitive load with the presence and absence of avatars and activation gestures [13]. It was found that the use of avatars may reduce perceived cognitive load by increasing performance and reducing user efforts. However, usability and user experience are yet to be explored with respect to the presence of activation gestures and avatars. Exploring user experience and usability is crucial because, without prime usability and user experience, users are more likely to make errors and not utilize the system to its full features. Usability and user experience are the two important factors for determining user acceptance and appreciation of a system.

In this paper, we explore how user experience and usability are impacted by the presence, use, and behavior of the avatar, by the presence of activation gestures, and by the layout of the visual interface (i.e. squared and middle of the screen arrangement of icons). We report comparative results from a within-subject study with 19 participants.

Our results show that usability was perceived as highest in a squared layout with a moving avatar and activation gestures. User experience was optimum when using a squared layout with a fixed avatar and a required activation gesture and in a squared layout with an activation gesture but without an avatar.

Contribution Statement: In the context of touchless gesture interaction with displays, this study investigates the impact of the avatar presence, the usage of activation gestures, and the visual interface layout on usability and user experience.

2 Related Work

In this section, we present previous work around interactive displays, interaction media and paradigms, and evaluation of user experience and usability.

2.1 Interactive and Public Displays

Many of today's public displays are interactive. Examples include ticket vending machines, ATMs, and info stands in malls and airports. While the predominant interaction modality used on these devices is touch, more displays today support touchless interaction. Touchless interaction can come in several forms. Some displays are gaze-enabled [20], allowing users to interact using their eye movements. Several works proposed leveraging personal mobile devices to interact with displays [31]. Closer to our work, displays can also support interaction using mid-air gestures [11, 36]. Researchers have also proposed techniques that combine gaze and mid-air gestures. For example, in their implementation of Pocket Transfers [23] Mäkelä et al. deployed a display where users can transfer content from the display to their phones by gazing at items on the screen and using mid-air gestures to indicate which items they wish to transfer.

Our work focuses on interaction using mid-air gestures. This interaction modality is more common for large displays and has been argued to be particularly useful for public displays as they allow hygienic interactions that do not require touching the display. Its playful nature also contributes to extending interaction durations [2]. There are also downsides to interaction using gestures. For example, they often require the use of user representations to communicate interactivity [18, 2]. Also to distinguish natural body movements from interactions, the system needs to either teach its users very specific gestures [35, 1] or require the use of activation gestures to trigger the interaction [26]. The presented study investigates the impact of multiple configurations on usability and user experience, including configurations that involve an avatar to represent users and activation gestures.

2.2 Evaluating & Supporting Interaction

Bystanders often miss or ignore interactive public displays thinking they are advertisements [9]. This can also be due to passersby being overwhelmed by the amount of information on the screen or when the interactive features are not noticeable or unintuitive. This results in a phenomenon referred to as interaction blindness [4, 29]. Apart from this, even after users have noticed that the system is interactive, they still encounter difficulty in understanding how to interact with it. This problem, known as affordance blindness, is due to the novel and uncommon media that is often used to interact. This is particularly true for public displays for which common mouse/keyboard or even touch-screens are not often feasible due to being deployed in public places. Therefore, these unique circumstances necessitate innovative interaction paradigms that cater to distinct user behaviors and require specialized study methods. In this section, we will

describe said evaluation methods and users' behaviors in the context of public displays.

Evaluation Methods for Public Displays Observational studies, surveys, and experiments all have their place in evaluating user experience and systems usability in general. Still, their effectiveness can be significantly enhanced when combined into a mixed-method approach. The complexity and variety of interactions with public displays necessitate this diverse approach to capture the full range of user behaviors and experiences [4]. Observational studies are widely used in the field of public display research. They involve researchers closely observing and recording the behavior of individuals or groups interacting with the display. For instance, a researcher might note the time a person interacts with a display or the number of errors during the interaction [6]. Sometimes observation in the real world might be so difficult or expensive that it requires the use of models or simulations [24]. Surveys, either self-administered or interviewer-led, are another common method for evaluating public displays. These may be used to gather subjective information on user perceptions and experiences, such as satisfaction, understanding, and intention to interact again. Müller et al. [30] used a survey method in conjunction with video observation to assess whether and why people pay attention to public displays. Of course one could use custom questionnaires. Still, to allow for comparisons and thus reliable evaluations, there are some well-known questionnaires or surveys whose scores can be measured and then compared with established thresholds. Among them, particularly in the context of public displays, researchers often use NASA Task Load Index (NASA-TLX) [15], System Usability Scale (SUS) [7], User Experience Questionnaire (UEQ) [22], and Net Promoter Score (NPS) [37], which are widely used tools for evaluating various aspects of human interaction with technology. They each serve a specific purpose: NASA-TLX for perceived workload, SUS for overall usability, UEQ for user experience, and NPS for customer satisfaction and loyalty. These all contribute to evaluating usability and user experience in a comprehensive, objective, and unbiased way. Experimental designs are employed when researchers seek to establish causal relationships. This involves manipulating one or more variables (independent variables) and measuring their effect on other variables (dependent variables). For example, an experiment could manipulate the content of a public display to determine its impact on user engagement, such as in [34].

Display and Interaction Blindness A major contributor to interaction blindness is display blindness [25]. If passersby cannot see or notice the display, they will not consider interacting with it. Dalton et al. [8] studied display blindness in an eye-tracking study to find that passersby often look at displays. However, the study could not conclude whether passersby actually notice the displays, as people may gaze at a target without paying attention to it [19]. Display blindness was tackled by using a stimulus that attracts attention to the display [27] such as curiosity-provoking artifacts [16], or an animatronic hand [17]. More

commonly, avatars and silhouettes that mirror passersby's behavior were found to be effective in tackling display blindness [18].

Even if display blindness is overcome, passersby may never realize that the display they noticed is interactive [29]. Many similar approaches were deployed to address interaction blindness. For example, in Looking Glass [29], passersby's movements were mirrored on displays and combined with call-to-action labels to explain to users how to interact. Once users start interacting with a display, the honeypot effect results in attracting other passersby to the display [6, 24].

Activation gestures Representing users using silhouettes and avatars helps overcome display and interaction blindness. A byproduct of representing users using silhouettes and avatars is that their arms and hands are visible on the screen; this makes interaction using mid-air gestures easier to communicate and to give feedback on. However, a drawback of touchless gestural interaction is that it lacks an equivalent of a "mouse click". As a result, the system needs a mechanism by which users can indicate whether they are simply pointing, or activating. One way to address this is by requiring users to move their hands on top of the desired target and "dwell" on it i.e., keep the hand steady on the target for a short period, typically in seconds [38]. An alternative approach is to require an activation gesture. For example, Yoo et al. [38] compared point-and-dwell to activation gestures such as push and grab-and-pull to find that the former is more accurate while push was preferred for selection and grab-and-pull was preferred for navigation. Gentile et al. [14] found that activation gestures may discourage users from continuing to interact and also that they may require a steep learning curve compared to point-and-dwell.

2.3 Research Gap

In summary, previous work on interactive displays shows that deploying avatars that mirror user movements has a positive impact on display and interaction blindness and that requiring activation gestures is sometimes necessary to avoid unintentional selections. Our work studies the impact of the presence of an avatar, its configuration, the layout of the screen, and the configuration of activation gestures on usability and user experience.

3 User study

In this study, we aim to understand how usability and user experience are affected by the layout of the visual interface and by the presence and behavior of an avatar that is displayed in the middle of the screen. We also explore the use and impact of activation gestures on the above-mentioned factors. To this end, we used different metrics, including task difficulty, task completion time, number of errors, and well-known questionnaires to evaluate perceived usability and user experience as discussed in the previous Section.

3.1 Apparatus

The system used for our study consisted of a 65" projected display placed at eye level, showing the interface being tested. The projector was driven by a computer, to which a Kinect for Xbox One was connected. The Kinect was placed below the screen, and gathered information on users' body gestures, using the Microsoft Kinect SDK v2. The study was conducted in the lab at our institute.

Visual Interface Layout (L) We selected two interface layouts: (1) square (SQ) and (2) middle (MID). In the square layout, the icons were arranged in a square style along the edges of the interface (see Figure 2B and 2D), whereas, in the middle layout, the icons were arranged all over the interface, including the middle part of it (see Figure 2A and 2C). The interface layouts were chosen based on previous work [13].

Activation Gesture (AG) In our interfaces, we used the "two hands icon" for interactive tile selection purposes as shown in Figure 2. The hand icons moved along with the hands of the users. For the purposes of our study, we tried two different solutions for interactive tile activation: (1) no activation gesture: in this case, the users can trigger the interaction events just by driving the hand icons and keeping them on top of the available tile-shaped components (i.e., point and dwell). (2) With an activation gesture: in this case, the user must execute a gesture that, if executed when a cursor (hand icon) overlays an interactive tile, triggers the corresponding event. We used a "push-to-press" gesture, which emulates a mid-air pushing action in accordance with Microsoft's Human Interface Guidelines (HIG) [26].

Avatar Design (AV) Avatars have been intensively studied for their use in interactive displays [28, 35, 29]. An avatar is a user representation on the interface that mimics the user's movements. We adopted an avatar design from prior work by Gentile et al. [13]. In our study, we experimented with three possible solutions for the user's avatar: (1) avatar present and fixed (FIX) in the middle of the screen (see Figure 2C and D); (2) avatar present and moves on the x-axis (MV), mirroring the user's body position in front of the display (see Figure 2E); (3) avatar absent (NO) (see Figure 2A and B). In both MV and FX, the avatar was always mimicking the user's arm movements.

3.2 Study Design

Our study was designed as a repeated measures experiment with three independent variables: IV1) visual interface layout (L), which had two conditions: a) *square* (SQ) and b) *middle* (MID); IV2) Activation Gesture (AG), which had two conditions: a) *absent* and b) *present*; and finally IV3) Avatar design (AV), which had three conditions: a) *present and fixed* in the middle of the screen

(FIX), b) *present and movable* on the x-axis (MV), and c) *absent* (NO). Note that in IV3a and IV3b, the avatar mimicked the user's arm movements.

While the total number of conditions was 2 L × 2 AG × 3 AV = 12, we discarded the two conditions that had an avatar present (one condition had the moving avatar whereas one had the fixed avatar), a middle-positioned layout, and did not feature an activation gesture. These two conditions were excluded because a) they are subject to involuntary activations of tiles which are located at the bottom of the interface, and b) they were reported to be disliked by users in prior work [14]. The tested conditions in this study are summarized in Table 1.

The dependent variables of interest for our study, and the methods and metrics used to evaluate them, were:

- **perceived usability**, evaluated using the System Usability Scale (SUS) questionnaire [7], which provides a score ranging from 1 to 100;
- **user experience**, evaluated using the User Experience Questionnaire (UEQ) [22] and the Net Promoter Score (NPS) [37];
- **task difficulty**, evaluated using the Single Ease Question (SEQ) [33] which provides a 7-point rating scale to assess how difficult users find a task, where 1 = "very difficult" and 7 = "very easy";
- **task completion time**, evaluated counting the time in seconds elapsed from the moment each task was started by the participant until successful completion;
- **error rates**, evaluated counting the number of times users perform an error. We counted an error whenever the user did any of the following:
  - activated the wrong interactive tile;
  - tried to activate a non-interactive tile;
  - assessed that they have finished the task when they actually had to continue;
  - asked for help;
- **two-handed interactions**, evaluated counting the number of times the user:
  - switched from using one arm to another arm;
  - used both arms simultaneously.

3.3 Procedure

We welcomed participants with an information sheet that provided details of the study. The participants were then presented with a consent form. Upon receiving the participants' consent, they were presented with the display and were provided with a training session of five minutes. This was to ensure adequate interaction and familiarity with the display. During the training session, the experimenter debriefed participants about the available features and interactive tiles. After the training session, the participants were asked to perform the following tasks by driving the Avatar's hands (in cases where an avatar was present) or the hand-shaped cursor (in cases where there was no avatar): (1) find specific news,

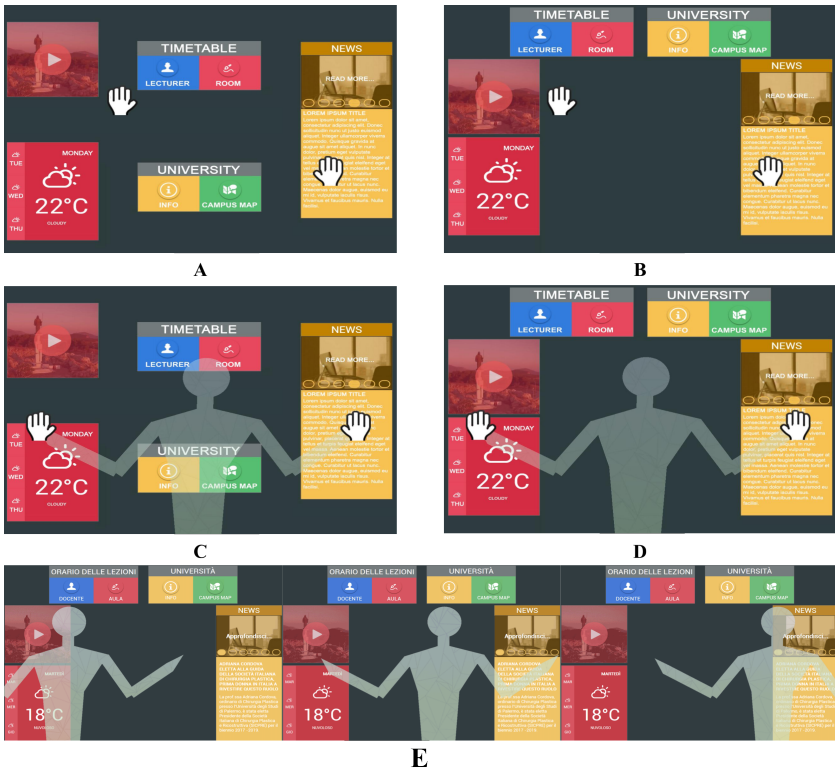

Fig. 2. The investigated configurations of a Touchless Gestural Interface: avatar absence (A, B) vs. presence (C, D, E); squared (B, D) vs. middle (A, C) tile layout; fixed (C, D) vs. moving (E) avatar.

Condition ID	L (Layout)	AV (Activation Avatar)	AG (Activation Gesture)
CN1	Square	with moving avatar	with activation gesture
CN2		with moving avatar	without activation gesture
CN3		with fixed avatar	with activation gesture
CN4		with fixed avatar	without activation gesture
CN5		without avatar	with activation gesture
CN6		without avatar	without activation gesture
CN7	Middle	with moving avatar	with activation gesture
CN8		with fixed avatar	with activation gesture
CN9		without avatar	with activation gesture
CN10		without avatar	without activation gesture

Table 1. The table shows the conditions investigated in this paper.

(2) access university information, (3) find the timetable for a specific class, and (4) play a video. All the visual components that allow the above-mentioned tasks were accessible from the main page of the interface.

Participants were instructed to perform the tasks as fast as they can and as accurately as possible. We used a within-group setup i.e., all participants experienced all the conditions. In each condition, participants were required to perform all four tasks. The order of the conditions was counterbalanced to reduce the biases and to level out any learning effects. After the completion of each task, participants were presented with a questionnaire comprising the System Usability Scale (SUS), the User Experience Questionnaire (UEQ), the Net Promoter Score, and the Single Ease Question (SEQ). During the study, the experimenter manually took notes of errors made, time consumed to perform each task, and if participants used one or both hands for interaction. After the completion of all tasks, participants were presented with an exit questionnaire asking questions about the different conditions such as preference for the layout, avatar preference, and use of activation gestures for selecting the interactive tiles. The followed procedure is visualized in Figure 3.

Our institute has no ethics board, therefore we followed the best practices and ethics guidelines when conducting our study. The participants were briefed before and after the study, were made aware they can withdraw anytime, had to read an information sheet and sign a consent form before participating, and were made aware that their data was stored securely to protect sensitive information.

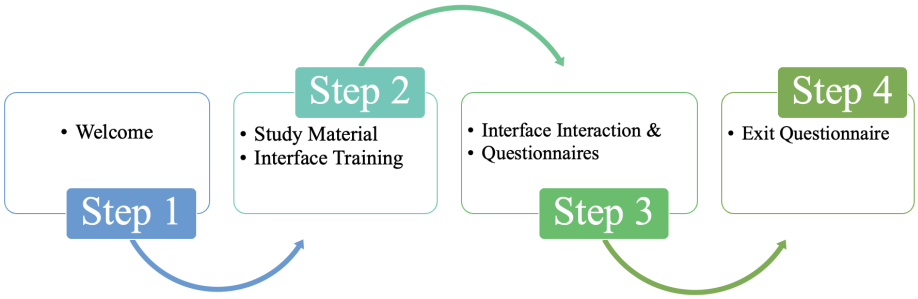

Fig. 3. The figure shows the followed procedure in our study.

3.4 Participants

We recruited N=19 participants (M=7, Female=12; self-identified) through word of mouth and snowball sampling. This number of participants was found to be appropriate in consideration of the number of participants in similar studies such as [12, 29, 5]. Participants were on average 30.89 years old (SD=12.90). Sixteen participants were right-handed and three participants reported to be equally good with left and right-handedness. None of the participants had issues that

could limit their movements. A few participants (N=3) reported having used a similar gesture-based interaction system before such as in training sessions for disabled children, on PlayStation, and in museums. Whereas, the majority of the participants (N=16) had no prior experience.

3.5 Data Analysis

We use participant IDs to refer to participants, such as P2 throughout the data collection and reporting to ensure anonymity. Where necessary, we use participants' quotes to support the results of the study but they cannot be traced back to the participants' identities. We ran one-way repeated measures ANOVA and used Bonferroni correction to correct for multiple comparisons in post-hoc tests. The results of the standardized questionnaires (i.e., SUS, UEQ and NPS and SEQ) were analyzed using the corresponding standard analysis for each. Qualitative responses were analyzed using open coding to translate them into meaningful snippets.

3.6 Limitations & Future Work

Although our study was carefully designed, like all studies it has some limitations. We acknowledge the following limitations. This study was conducted in a controlled lab environment. While in-the-lab studies impact ecological validity, they also allow controlling numerous confounding variables such as distraction caused by the presence of bystanders that may impact users' task performance, which is commonplace in public display interactions [12]. Second, this study was conducted at our academic institute where there are many computer science students who are knowledgeable about technology. However, the users' experience may vary as their knowledge and technology experience change.

Third, we acknowledge that the small number of left-handed participants (N=3) may have impacted the results. Further, the effect of habituation was absent from the user study as the only way the habituation effect would have been possible would have been during the training session where the repeated aspect of the interaction from one setup to another was the task set. However, the participants were only knowledgeable of what the tiles do but not how to interact with them as this is the only information we provided about them in the training session. Another way habituation could have happened can be due to repeating the task. Even though the task was the same, the way it was carried out by the participants was different in every condition. Thus, we expect there is little to no effect of habituation. We counterbalanced the conditions to avoid learning effect.

Lastly, the sample size chosen for this study is appropriate in line with guidelines for usability studies [3, 32, 10], however, we realise this sample size may appear low from some perspectives therefore, we acknowledge and propose that future studies with more diversity and large sample size should be conducted.

4 Results

The main goal of this study was to evaluate the impact of the avatar presence, the usage of activation gestures, and of the visual interface layout on perceived usability and user experience, task difficulty, task completion time, and use of two-handed interactions. To this end, we evaluated the dependent variables in ten different conditions corresponding to the ten relevant combinations of the independent variables, as described in the previous section. In this section, we present the findings gathered from N=19 participants, which are detailed in the following subsections and summarized in Tables 2 and 3.

4.1 Usability

System Usability Scale (SUS) The System Usability Scale (SUS) [7] is a standard metric for evaluating usability. We used it to check for usability for all conditions. The perceived usability was the highest in CN1 i.e., square arrangement with a moving avatar and an activation gesture, among all conditions and CN8 i.e., middle arrangement with a fixed avatar and with an activation gesture, received the lowest SUS score.

A Friedman test for the SUS scores revealed significant differences ($\chi^2(9) = 107.920$, $p < 0.0005$). Pairwise comparisons revealed significant differences between CN8 and CN7 (p=0.005), CN8 and CN9 (p=0.003), CN8 and CN10 (p=0.002), CN8 and CN6 (p=0.002), CN8 and CN4($p < 0.0005$), CN8 and CN3($p < 0.0005$), CN8 and CN5($p < 0.0005$), CN8 and CN1(p=0.0005), CN2 and CN1 ($p < 0.0005$), CN7 and CN1($p < 0.0005$), CN9 and CN1($p < 0.0005$), CN10 and CN1($p < 0.0005$), CN6 and CN1($p < 0.0005$), CN4 and CN1 (p=0.001), CN3 and CN1 ((p=0.001), and CN5 and CN1 (p=0.033).

Task Difficulty & Duration After each condition, participants were asked to rate their perceived difficulty of the interaction on a scale from 1 to 7 where 1 represented the least difficulty and 7 represented the highest difficulty. We then checked for significant differences in task difficulty of each task in each condition. A Friedman test was run (with a Bonferroni correction for multiple comparisons where appropriate) to determine if there are significant differences between the task difficulty in each condition. No significant differences were found for all tasks' perceived difficulty across all conditions. Next, we checked which task took longer to perform in each condition. Significant differences were only found for Task 3 ($\chi^2(9) = 26.505$, p=0.002). Pairwise comparisons revealed significant differences between CN10 and CN2 (p=0.003) and CN10 and CN3 (p=0.027).

4.2 User Experience & Willingness to Promote the Interface

The user experience was recorded using the short version of the User Experience Questionnaire [22]. It assesses user experience on two aspects; pragmatic and hedonic quality. Values between -0.8 and 0.8 represent a neutral evaluation whereas values greater than 0.8 represent a positive evaluation and values less

than -0.8 represent a negative evaluation. The range of the scales is between -3 (”horribly bad”) and +3 (”extremely good”). Overall, CN3 and CN5 provided the best user experience both with a score of 0.618.

Next, we measured participants’ willingness to promote the interface system through Net Promoter Score (NPS) [37]. NPS is a frequently used market research metric that assesses how likely users are to promote a product or a service. Participants rated their likelihood of promoting the interaction on a scale of one to ten. We then calculated the NPS score using the dedicated score calculation method [37]. CN5 received the highest NPS score (47.37) while CN10 received the lowest NPS score (-57.89). The results of UEQ, NPS, and SUS scores for all conditions are summarized in Table 2, where: Layout L can be square (SQ) or middle (MID); Avatar AV can be movable (MV), fixed (FIX), or absent (NO); Activation Gesture AG can be present (Y) or absent (N).

CONDITIONS				**RESULTS**				
				UEQ			**NPS**	**SUS**
	L	AV	AG	Pragmatic Quality	Hedonic Quality	Overall		
CN1	SQ	MV	Y	0.434	0.618	0.526	42.11	91.45
CN2			N	0.33	0.47	0.428	5.26	33.74
CN3		FIX	Y	0.5	0.74	0.618	31.58	37.95
CN4			N	0.54	0.5	0.519	21.05	36.47
CN5		NO	Y	0.51	0.72	0.618	47.37	38.32
CN6			N	0.49	0.55	0.52	5.26	36.84
CN7	MID	MV	Y	0.38	0.63	0.507	31.58	36.63
CN8		FIX		0.32	0.54	0.428	21.05	28.11
CN9		NO		0.43	0.68	0.559	31.58	37.42
CN10			N	0.41	0.5	0.454	-57.89	36.26

Table 2. The Table shows User Experience Questionnaire (UEQ), Net Promoter Score (NPS), and System Usability Score (SUS) for the conditions investigated in the study.

	Conditions			**Task Difficulty**				**Task Duration**				**Errors Committed**			
	L	AV	AG	Task 1	Task 2	Task 3	Task 4	Task 1	Task 2	Task 3	Task 4	Task 1	Task 2	Task 3	Task 4
CN1	SQ	MV	Y	6.63	6.68	6.63	6.79	10.68	8.58	13.21	7	0	0.11	0.32	0
CN2			N	6.79	6.68	5.58	6.42	6.89	7.05	13.68	5.74	0.11	0.21	1.11	0
CN3		FIX	Y	6.85	6.8	6.8	6.7	7.58	7.05	11.79	5.89	0.1	0.1	0.45	0.05
CN4			N	6.95	6.84	6.32	6.89	7.95	5.58	11.58	5.26	0.11	0	0.42	0
CN5		NO	Y	6.89	7	6.79	6.89	7.84	5.95	11.05	5.21	0.11	0.05	0.21	0.05
CN6			N	6.63	6.47	6.32	6.89	6.74	6.68	12.21	4.79	0.26	0.32	0.68	0
CN7	MID	MV	Y	6.74	6.53	6.84	6.89	7	6.47	10.58	5.68	0.16	0.26	0.05	0.05
CN8		FIX		6.89	6.21	6.63	6.84	7	6	9.32	6	0.21	0.05	0.11	0.11
CN9		NO		6.53	6.53	6.53	6.74	6.63	6.42	12.05	5.84	0.32	0.16	0.32	0
CN10			N	6.68	6.84	6.84	6.84	5.42	4.89	8.79	4.79	0.16	0.11	0.21	0

Table 3. The Table shows the participants' perceived average task difficulty, time consumed to perform each task, and errors made in each condition while performing tasks.

4.3 Overall Evaluation

To evaluate the layout design, use of avatars and activation gestures, we asked participants to rate their experience of each on a scale of 1 to 10 where 1 represented the least rating and 10 represented the highest rating. We also asked participants to provide reasoning for their choice. In this section, we report the descriptive and statistical analysis of the ratings and results of the inductive coding of the qualitative insights. To analyse the qualitative data, one researcher went through the data and assigned codes. Then, the second researcher revisited the qualitative data and assigned codes. The two researchers then discussed the disagreements and finalized the codes. This iterative refinement of the coding process ensured the validity of the results. In the following results, we report the frequency of codes appearing in the dataset to give the readers an impression of how many times each category appeared.

Layout (L): Square layout received a mean rating of 8.89 out of 10 (Min=6, Max=10, SD=1.37) for participants' experience while the middle layout received a mean rating of 7.89 out of 10 (Min=1, Max=10, SD=2.30). Nine participants preferred the square layout whereas ten participants preferred the middle layout. An exact sign test was run to uncover significant differences between the ratings of the two layouts but no significant differences were found (p=1.000).

The qualitative insights reveal that the square layout was preferred mainly for two reasons; the arrangement of icons (6), ease of selection (2), and visual appearance (1). The arrangement of icons gave freedom of movement to the participants.Other reasons included the fact that information is easier to reach with hands and the possibility to freely move hands without activating the central icons. One participant mentioned *"...comfortable arrangement of windows"* (P1) as the reason for preferring the square layout. Participants did not prefer the middle layout because it offered less freedom of movement, and close placement of the icons, and some participants even found it *"..confusing.."* (P16). On the contrary, the middle layout was preferred mainly for four reasons; the arrangement of icons (2), visual appearance (4), and ease of selection (5). Less empty space, the centre positioning of the interface, and the orderly position of icons were the reasons to opt for the middle layout.

Avatar (AV): Participants rated their experience of interaction without the avatar present with a mean of 8.15 out of 10 (Min=4, Max=10, SD=1.92) and a mean of 8.36 out of 10 for interaction with the avatar present (Min=5, Max=10, SD=1.53). A Wilcoxon signed-rank test did not reveal significant differences (z=0.516, p=606). Avatar was not preferred by participants (N=9) due to visual cumbersome (4), inconvenience (1) and because participants felt that its presence was unnecessary (1) and it made the view complicated to deal with (1). However, some participants (N=10) preferred to have an avatar with a predilection towards having the avatar mobile (N=8) as compared to a fixed position (N=2). The mobile positioning of the avatar was liked by participants because it bought

convenience (5), better engagement (1), and design (1). Participants were able to move around better and mirror themselves in the avatar. The mobile state of the avatar made it easier reach to icons and facilitated movement. It was also seen as letting the users be more involved in the interaction and being able to understand the interaction through the avatar. The fixed position of the avatar was favoured as participants felt that the accuracy of the hands was better (1) and due to convenience (1).

Activation Gestures (AG): The experience with activation gestures (AG) was rated a mean of 8.84 out of 10 (Min=4, Max=10, SD=1.46). An exact sign test revealed significant differences (p=0.035). Participants were inclined towards preferring activation gestures (N=15). This was mainly because of four reasons; controlled interaction (6), accuracy (3), ease of use (3), and fewer chances of error (1). Participants voiced that activation gestures appeared as *"..safer in use"* (P1). Participants felt that they could safely move around the screen without activating unneeded functions. On the contrary, some participants (N=4) also felt that the greater ease that is provided by activation gestures is not really needed (1), could result in system failure because sometimes the system did not recognize the activation gesture (1), and that the system is faster (1) and simpler without AG (1).

One-Handed vs Two-Handed Interaction: All participants (N=19) reported having used and preferred both hands for interaction. Participants felt that the use of both hands was convenient (6), and allowed free (6) and fast movement (3). Participants also mentioned that use of both hands was a source of comfort while interacting (3) and therefore was necessary (1). Participants also voiced that using two hands while interacting made them pay more attention to the task.

Lastly, we asked participants what they perceive are the advantages of using one-handed interaction. The participants voiced that using one hand for interaction is easy (1) and gives greater freedom (1). It also leaves the other hand to perform some other task (2) such as holding a jacket. Further, they also mentioned that single-handed interaction could assist people with disabilities (1). For two-handed interaction, participants mentioned that using two hands appears more naturalistic (2) and focused (1). Use of two hands was perceived as providing faster interaction (6), free movement (5), convenient (1), ease of selection (1), and giving more opportunities to interact (1).

5 Discussion

Our study revealed how usability and user experience are influenced by modifications in the interface style i.e. one and two-dimensional and with avatar and activation gestures. Based on our findings, we present the following general observations and key takeaways.

5.1 General Observations

Prior work [13] evaluated whether and how the presence of an Avatar that replays the user's movements may decrease the perceived cognitive workload during interactions. In this paper, we focus on the user experience and usability. From the users' perspective, usability and user experience are highest in the square interface layout. On the other hand, middle-layout has the lowest usability. Considering the design perspective, the arrangement of icons on the layout should be placed carefully as users look for "neatness" in design and "freedom of movement". The avatars can support the users' communication with the touchless display if they are in a mobile position. The activation gestures further facilitate the selection process, assisting in making fewer mistakes and making the user more involved.

The use of two hands was highly favoured by participants as it makes the interaction a lot easier than one-handed interaction. On the contrary, in a study by Walter et al. [36], 80% of users only used their right hand to perform the interaction while they were given the option of both hands. However, participants in our study favoured using two hands. This was especially the case for complex tasks such as using a map. One of the reasons our results differ from those of Walter et al. [36] could be because people usually carry something such as a bag or a jacket in public places and therefore prefer to use one hand for the interaction. However, the use of two hands is perceived to make interaction easier and quicker. Second, Walter et al. [36] state that a possible reason their study participants preferred using one hand was that they were asked to register a gesture, and thus the participants continued to use the same hand used for registration believing this was the only way to interact. We expect the reason we had a different result is that our participants were free to use one or both hands. Considering the results of our study, it can be implied that simple tasks requiring only a few interactions could be designed for one-handed interaction as they give the freedom to the other hand to perform other tasks. This is also beneficial from the perspective of designing accessible interaction as voiced by the participants of our study.

On the other hand, some participants found that the use of two hands is problematic when carrying other objects. This was also reported in prior work on two-handed interaction using mid-air gestures with public displays [21]. This suggests that interactive displays deployed in public spaces should provide users with the option of interacting using one hand.

5.2 Key Takeaways

Based on the findings of our user study, we pen down the following key takeaways.

Usability is perceived as highest in a square layout with moving avatar and activation gestures. Participants in our study were slightly inclined towards the square layout as opposed to the middle layout. This was so because the square layout allowed participants to move freely with an easier

reach to icons. The moving avatar helped the participants in mirroring themselves in the interface making the interaction closer to real. It is likely that this supported participants in reaching farther targets e.g., to activate a target on the left, the user could not only extend their arm but also move to the left to reach the target. The activation gestures assisted in giving a sense of confirmation of the selection of icons to the participants which in turn was perceived as helping in making fewer interaction errors.

User experience is optimum in a square layout with activation gestures irrespective of whether the avatar was moving or fixed. The User Experience Questionnaire focuses on two aspects; pragmatic (i.e., efficiency, dependability, and perspicuity) and hedonic quality (i.e., stimulation and novelty). Square layout with activation gestures outperformed in both qualities, hedonic and pragmatic. The state of the avatar did not impact the user experience as the user experience was found to be optimum with the moving and fixed state of the avatar. On the contrary, activation gestures influenced the user experience. All conditions with activation gestures received higher user experience questionnaire scores as compared to conditions without activation gestures.

A square layout with activation gestures only appears most likely to be promoted. Because activation gestures were perceived as helpful in making selections, the conditions with activation gestures were found to be more promoted by participants. Huge differences were found in the NPS score between the presence and absence of activation gestures. This shows that the design elements that assist in making and confirming selections are likely to be more promoted and appreciated by users. Activation gestures are one such example. This is in line with previous work that compared point-and-dwell to grab-and-pull, to find that the former is more suitable for selections whereas the latter is more suitable for navigation [38]. A possible explanation of this result in our study could also be attributed to legacy bias; that is, users are more accustomed to separating the tasks of pointing and activating, and may thus prefer activation gestures compared to selection by dwelling at targets.

Use of two-handed interaction facilitates the touchless interaction with the display. Interaction with displays is not just about a few selections but sometimes it can also be as complex as navigating a map. In complex scenarios, the use of two-handed interaction is favored by users. The use of two-handed interaction not only makes the interaction easier but also facilitates it by decreasing the time required to perform the task. As mentioned earlier, it is important to allow users to also use one hand especially if the display is deployed in a public space, as users of public displays may have an occupied hand.

6 Conclusion

To make displays more interactive, numerous efforts have been made such as introducing avatars to represent the user, finding new ways for triggering interactive items, and identifying general rules for optimal visual layout arrangements. While such innovations are impressive, they have a risk of impacting the user experience and usability of the system. Without optimum user experience and usability, users are less likely to use the system. For this reason, we investigated two possible layouts, and the use of avatars and activation gestures, and evaluated their impact on usability and user experience in visual interfaces for interactive displays. We found that the square layout is preferred more than the middle layout, showing higher usability and better user experience. We also found that the presence of a moving avatar had a positive impact on the perceived usability of the interface. As far as the usage of an activation gesture is concerned, we found that using one has a positive impact on both perceived usability, the user experience, and general appreciation. The results and findings of our study are summed up as key takeaways which could be a quick and useful reference list for the design of future interactive displays.

Acknowledgements

This work was partly supported by the EPSRC (EP/V008870/1), and the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which is also funded by the EPSRC (EP/S035362/1).

References

1. Ackad, C., Clayphan, A., Tomitsch, M., Kay, J.: An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. p. 1227–1238. UbiComp '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2750858.2807532, https://doi.org/10.1145/2750858.2807532
2. Ackad, C., Tomitsch, M., Kay, J.: Skeletons and silhouettes: Comparing user representations at a gesture-based large display. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. p. 2343–2347. CHI '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858427, https://doi.org/10.1145/2858036.2858427
3. Alroobaea, R., Mayhew, P.J.: How many participants are really enough for usability studies? In: 2014 Science and Information Conference. pp. 48–56. IEEE (2014)
4. Alt, F., Schneegaß, S., Schmidt, A., Müller, J., Memarovic, N.: How to evaluate public displays. In: Proceedings of the 2012 International Symposium on Pervasive Displays. PerDis '12, Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2307798.2307815, https://doi.org/10.1145/2307798.2307815

5. Alt, F., Shirazi, A.S., Kubitza, T., Schmidt, A.: Interaction techniques for creating and exchanging content with public displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1709–1718 (2013)
6. Brignull, H., Rogers, Y.: Enticing people to interact with large public displays in public spaces. In: INTERACT'03-9th IFIP TC13 Int. Conf. on Human-Computer Interaction. vol. 3, pp. 17–24 (2003)
7. Brooke, J., et al.: Sus-a quick and dirty usability scale. Usability evaluation in industry **189**(194), 4–7 (1996)
8. Dalton, N.S., Collins, E., Marshall, P.: Display blindness? looking again at the visibility of situated displays using eye-tracking. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. p. 3889–3898. CHI '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2702123.2702150, `https://doi.org/10.1145/2702123.2702150`
9. Davies, N., Clinch, S., Alt, F.: Pervasive displays: understanding the future of digital signage. Synthesis Lectures on Mobile and Pervasive Computing **8**(1), 1–128 (2014)
10. Francik, E.: Five, ten, or twenty-five – how many test participants? (2015), `https://www.humanfactors.com/newsletters/how_many_test_participants.asp`, retrieved May 09, 2023
11. Gentile, V., Khamis, M., Milazzo, F., Sorce, S., Malizia, A., Alt, F.: Predicting mid-air gestural interaction with public displays based on audience behaviour. International Journal of Human-Computer Studies **144**, 102497 (2020). https://doi.org/https://doi.org/10.1016/j.ijhcs.2020.102497, `http://www.sciencedirect.com/science/article/pii/S1071581920300999`
12. Gentile, V., Khamis, M., Sorce, S., Alt, F.: They are looking at me! understanding how audience presence impacts on public display users. In: Proceedings of the 6th ACM International Symposium on Pervasive Displays. pp. 1–7 (2017)
13. Gentile, V., Sorce, S., Malizia, A., Milazzo, F., Gentile, A.: Investigating how user avatar in touchless interfaces affects perceived cognitive load and two-handed interactions. In: Proceedings of the 6th ACM International Symposium on Pervasive Displays. pp. 1–7 (2017)
14. Gentile, V., Sorce, S., Malizia, A., Pirrello, D., Gentile, A.: Touchless interfaces for public displays: can we deliver interface designers from introducing artificial push button gestures? In: Proceedings of the International Working Conference on Advanced Visual Interfaces. pp. 40–43 (2016)
15. Hart, S.G.: Nasa-task load index (nasa-tlx); 20 years later. In: Proceedings of the human factors and ergonomics society annual meeting. vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA (2006)
16. Houben, S., Weichel, C.: Overcoming interaction blindness through curiosity objects. In: CHI '13 Extended Abstracts on Human Factors in Computing Systems. p. 1539–1544. CHI EA '13, Association for Computing Machinery, New York, NY, USA (2013). https://doi.org/10.1145/2468356.2468631, `https://doi.org/10.1145/2468356.2468631`
17. Ju, W., Sirkin, D.: Animate objects: How physical motion encourages public interaction. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) Persuasive Technology. pp. 40–51. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
18. Khamis, M., Becker, C., Bulling, A., Alt, F.: Which one is me? identifying oneself on public displays. In: Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems. CHI '18, ACM, New York,

NY, USA (2018). https://doi.org/10.1145/3152832.3157813, https://doi.org/10.1145/3152832.3157813

19. Khamis, M., Bulling, A., Alt, F.: Tackling Challenges of Interactive Public Displays using Gaze. In: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM (2015). https://doi.org/10.1145/2800835.2807951
20. Khamis, M., Klimczak, A., Reiss, M., Alt, F., Bulling, A.: Eyescout: Active eye tracking for position and movement independent gaze interaction with large public displays. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software & Technology. UIST '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3126594.3126630, https://doi.org/10.1145/3126594.3126630
21. Khamis, M., Trotter, L., Mäkelä, V., von Zezschwitz, E., Le, J., Bulling, A., Alt, F.: Cueauth: Comparing touch, mid-air gestures, and gaze for cue-based authentication on situated displays. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **2**(4) (Dec 2018). https://doi.org/10.1145/3287052, https://doi.org/10.1145/3287052
22. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Symposium of the Austrian HCI and usability engineering group. pp. 63–76. Springer (2008)
23. Mäkelä, V., Khamis, M., Mecke, L., James, J., Turunen, M., Alt, F.: Pocket transfers: Interaction techniques for transferring content from situated displays to mobile devices. In: Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems. CHI '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3173574.3173709, http://dx.doi.org/10.1145/3173574.3173709
24. Mäkelä, V., Radiah, R., Alsherif, S., Khamis, M., Xiao, C., Borchert, L., Schmidt, A., Alt, F.: Virtual field studies: Conducting studies on public displays in virtual reality. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–15. CHI '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376796, https://doi.org/10.1145/3313831.3376796
25. Memarovic, N., Clinch, S., Alt, F.: Understanding display blindness in future display deployments. In: Proceedings of the 4th International Symposium on Pervasive Displays. p. 7–14. PerDis '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2757710.2757719, https://doi.org/10.1145/2757710.2757719
26. Microsoft: Microsoft human interface guidelines (2014), https://download.microsoft.com/download/6/7/6/676611b4-1982-47a4-a42e-4cf84e1095a8/kinecthig.2.0.pdf, retrieved January 31, 2023
27. Müller, J., Alt, F., Michelis, D., Schmidt, A.: Requirements and design space for interactive public displays. In: Proceedings of the 18th ACM International Conference on Multimedia. p. 1285–1294. MM '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1873951.1874203, https://doi.org/10.1145/1873951.1874203
28. Müller, J., Bailly, G., Bossuyt, T., Hillgren, N.: Mirrortouch: combining touch and mid-air gestures for public displays. In: Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services. pp. 319–328 (2014)

29. Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F.: Looking glass: a field study on noticing interactivity of a shop window. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 297–306 (2012)
30. Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., Krüger, A.: Display blindness: The effect of expectations on attention towards digital signage. In: Pervasive Computing: 7th International Conference, Pervasive 2009, Nara, Japan, May 11-14, 2009. Proceedings 7. pp. 1–8. Springer (2009)
31. Ng, P.C., She, J., Jeon, K.E., Baldauf, M.: When smart devices interact with pervasive screens: A survey. ACM Trans. Multimedia Comput. Commun. Appl. **13**(4) (aug 2017). https://doi.org/10.1145/3115933, `https://doi.org/10.1145/3115933`
32. Nielsen, J., Landauer, T.K.: A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems. pp. 206–213 (1993)
33. Sauro, J.: If you could only ask one question, use this one. (2010), `https://measuringu.com/single-question/`, retrieved January 31, 2023
34. Vom Lehn, D., Heath, C.: Displacing the object: mobile technologies and interpretive resources. Archives & Museum Informatics **2** (2003)
35. Walter, R., Bailly, G., Müller, J.: Strikeapose: revealing mid-air gestures on public displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 841–850 (2013)
36. Walter, R., Bailly, G., Valkanova, N., Müller, J.: Cuenesics: using mid-air gestures to select items on interactive public displays. In: Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services. pp. 299–308 (2014)
37. Wikipedia, t.f.e.: Net promoter score (2022), `https://en.wikipedia.org/wiki/Net_promoter_score`, retrieved January 08, 2023
38. Yoo, S., Parker, C., Kay, J., Tomitsch, M.: To dwell or not to dwell: An evaluation of mid-air gestures for large information displays. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. p. 187–191. OzCHI '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2838739.2838819, `https://doi.org/10.1145/2838739.2838819`