The Screen Is Yours—Comparing Handheld Pairing Techniques for Public Displays

Matthias Baldauf[1], Markus Salo[2], Stefan Suette[1], and Peter Fröhlich[1]

[1] FTW Telecommunications Research Center Vienna, Austria
[2] University of Jyväskylä, Finland

Abstract. Whereas mobile devices have been heavily investigated as remote controls for distant displays, research on the fundamental first step, the pairing with the display, is scarce. In a comparative user study with 31 participants we evaluated five potential pairing techniques concerning their performance and acceptance for connecting to a public display and gained insights into the general requirements for pairing techniques in this special usage context. Besides four established mobile interaction techniques (touching an NFC tag, capturing a QR code, scanning, and manual input), our study considered a recent appropriate pairing technique called display pointing, which allows passers-by to connect to available displays just by pointing their smartphone at the installation, for the first time. Our results show that display pointing is superior to traditional alternatives to a large extent. Yet, its quick response times can result in a perceived lack of control if not handled appropriately. Further, we learnt that long distance techniques are generally preferred while the type of device gesture as well as security concerns are of less relevance.

Keywords: mobile interaction, pairing, pointing, visual recognition.

1 Introduction

While digital signage technologies have been increasingly deployed in urban environments over the last few years, we currently observe a shift toward interactive content on public displays. Such emerging installations allow passers-by to fetch coupons, play games, or take part in contests and polls. Besides gestural and touch-based interactions which require special hardware to be deployed, smartphones are promising ubiquitous and versatile remote controls for such interactive applications. Respective previous research focused on the investigation of various mobile interaction styles [1], only little attention has been attached to the crucial first step in the interplay of mobile devices and public displays, the pairing process. Related work has mainly studied post-pairing interaction and innovative application concept for smartphones and public displays assuming an already established wireless connection and neglecting this obstacle for usage. However, creating a low threshold of use and thus allowing for highly serendipitous interactions might be one of the critical success factors for upcoming interactive applications [1].

J.C. Augusto et al. (Eds.): AmI 2013, LNCS 8309, pp. 32–47, 2013.

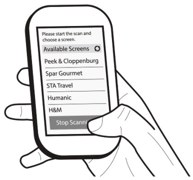
In this paper, we present a user study comparing potential pairing techniques. The most recent one, in this paper called “display pointing”, allows for connecting the mobile device with a public display through a simple pointing gesture. Interested passers-by can just briefly target the distant display through the mobile camera viewfinder to immediately establish a connection to it (Figure 1). We compare display pointing with the following more established physical mobile interaction techniques [2,3] for the very specific use case of pairing a mobile device with a public display, deliberately independent of the subsequent actual interaction style:

- *Touching.* Bringing a mobile device very close (5cm or less) to a smart label near the public display to read the pairing information without contact (Figure 1b).
- *Short-range pointing.* Identifying an object by capturing a corresponding unique visual marker attached to the public display through the camera (Figure 1c).
- *Scanning.* Selecting the desired display from a list by scanning the environment for available public displays based on proximity (Figure 1d).
- *User-mediated pairing.* Manually entering a code on the mobile device to specify the public display to be paired (Figure 1e).

(a) Display pointing

(b) Touching

(c) Short-range pointing



(d) Scanning

(e) User-mediated

Fig. 1. We compared the recent display pointing approach (a) and four more established pairing techniques (b-e)

Our comparative study aims at investigating the characteristics of display pointing and evaluating the overall applicability of these pairing techniques in the context of public display interaction. Further, we are interested in the general requirements of such mobile pairing techniques for the given scenario. Do any of these arouse special security concerns? What role does the distance to the public screen play? And have explicit smartphone gestures such as pointing become socially accepted in public spaces?

2 Related Work

In the following, we summarize previous related research in the field of mobile interaction with real-world objects, so-called physical mobile interactions, and respective results from user studies in different contexts.

2.1 Mobile Real-World Interaction

Today's feature-rich smartphones provide several technological means to act as mediators to real-world surroundings and nearby smart objects and thus realize abovementioned interaction styles [4]. One popular approach are geospatial calculations based on built-in positioning technologies such as GPS. While the plain device location can be used to realize simple scanning features for typical location-based services, it can be combined with orientation information gathered from a compass to realize long-range pointing styles [4].

Another way to realize both short- and long-range pointing gestures is the visual detection through the built-in camera of a mobile device and suitable computer vision methods. While early prototypes utilized custom hardware such as laser pointers [5] or obtrusive visual codes [6] acting as physical hyperlinks attached to real-world objects (e.g. popular QR codes), more recent research enabled the markerless recognition of objects on mass-market devices in real-time exploiting natural image features and comparing snapshots with a set of target photos [7]. A popular example based on a remote image database is Google Goggles, a mobile application for identifying landmarks by pointing with a smartphone.

A third technological enabler for mobile interaction are wireless communication technologies such as NFC, Wifi, and Bluetooth. NFC (near field communication) is a short-range technology and can be used to realize touching gestures to fetch data from or write data to unpowered chips, so-called tags (e.g. [8,9]). In contrast, Bluetooth and Wifi are the typical cornerstones for medium-range scanning techniques for detecting nearby smart objects.

2.2 Evaluating Physical Mobile Interaction

Rukzio et al. [2,3] were amongst the first researchers who compared different interaction techniques to derive guidelines on what technique to apply for what scenario. According to their studies, scanning and user-mediated object interaction are the preferred techniques for tourist guide applications while in a museum context where visitors are close to objects touching is favorable. For smart (home) environments which are probably most related to the public display scenario explored in this paper, the distance is pivotal: again, for close objects touching is preferred, for object further away yet visible it is pointing (implemented by laser pointer attached to mobile device), for not visible object (in other room) scanning. Similar results are reported by Välkkynen et al. [10]. At that time, camera-based pointing with visual codes suffered from high execution times requiring up to a few seconds for decoding and informing the user about the success of the action. A more recent study comparing product identification techniques [11] shows that NFC is still preferable in terms of speed, however, very close to capturing visual codes. Yet, both techniques outperform manual code entry. Other studies proved advantages of NFC touches and visual codes over Bluetooth scanning approaches for dedicated use cases [12,13]. Comparing NFC and 2D codes directly in the context of smart posters unveiled the easier learnability of 2D code capturing, whereas trained participants preferred NFC [14]. However, related studies unveiled security concerns regarding touching gestures in public places [9].

In contrast to such previous studies, we investigate these techniques as pairing modes for public displays and thus as enablers for subsequent interaction. Further we are interested in the specific requirements in the context of public displays and the impact of contemporary technology such as improved QR code readers and NFC sensors.

2.3 Screen Pairing Techniques

Previous research on mobile interactions with public screens focused on innovative interaction styles to use smartphones as remote controls. However, work dedicated to the required pairing step is scarce. When applying visual markers (e.g. [15,16]) or custom displays equipped with a grid of NFC tags (e.g. [17]), these smart labels can be utilized for both determining the screen area a mobile device was targeted at and encoding the address of the hosting computer to establish a data connection it. For markerless setups, Bluetooth scanning approaches have been proposed to determine available public displays and select the desired one based on descriptive names. Whereas this approach probably takes more time, it is also feasible for environments with multiple displays deployed.

Purely location-based approaches which automatically connect to a nearby screen [18] might be suitable for home environments, however, fail for urban surroundings densely equipped with public displays. Pointing-like techniques for visual markerless screen recognition are considered in recent work on smart lens-based interaction with displays [19], however, only as part of the actual interaction, not as dedicated pairing techniques.

So far, no studies have been conducted to compare pairing techniques for enabling mobile interactions with public displays in depth. The experiment presented in this paper tries to shed light on this neglected aspect and studies pairing techniques independent from any task-related interaction with the remote display.

2.4 Research Hypotheses

Based on the abovementioned literature research and own experiences from a preliminary qualitative study [20], we formulated our expectations in the following research hypotheses.

H1. Due to its intuitive nature and common "photographing" gesture, display pointing is easier to learn than alternative techniques.
H2. Display pointing outperforms the alternative in terms of several key factors in physical mobile interactions: ease, mental and physical load, speed, innovativeness, fun and accuracy.
H3. Display pointing arouses less security concerns than for example NFC due to its long-distance contactless pairing approach.
H4. According to previous research in smart home environments, long-distance techniques for pairing with remote display will be preferred in the context of public screens.
H5. Due to the public context, participants prefer pairing techniques based on typical smartphone interactions such as manual input over explicit device gestures.

3 Method

To address our research questions and answer the hypotheses, we implemented a functional research prototype featuring five pairing techniques and conducted a comparative lab study.

3.1 Participants

32 participants (13 females, 19 males) were recruited from our institute's user database and through social media announcements took part in our experiment. The data of one male participant had to be excluded due to insufficiently provided data in the final interview phase. The participants' age ranged between 23 and 66 years (mean=36.8, median=31.5). We carefully selected the participating users to resemble the structure of the entire local smartphone users and thus gain representative results: the test group consisted of 15 participants (5 females, 10 males) younger than 29 years, 10 participants (5, 5) aged between 30 and 49, and 6 persons (2, 4) over 50 years. As remuneration for taking part in our study we offered vouchers for supermarkets and online shops.

Each participant owned on average 1.3 mobile devices. On five point Likert scales between strongly agree (1) and strongly disagree (5), the test persons stated a mean of 1.6 for the statement I like to experiment with new information technologies and of 2.1 for I often seek out information about new technology products. Further, they answered both I regularly use many kinds of mobile applications and I think I am skillful at using technology and mobile applications with 2.1 on average. All participants had seen visual codes before (20 had used one in practice), 23 participants had heard about NFC (3 had used it), and 27 participants knew Bluetooth scanning (25 had used it).

Fig. 2. The test setup contained two large screens with exemplary interactive applications (left). One of them was a public poll showing a topical question and three possible answers after the successful pairing (right).

3.2 Setup

In our test room, we arranged two flat screen TVs with screen diagonals of 47 and 55 inches, placed on bar tables at a typical height of public displays. The two screens were facing the same direction to resemble the scenario of adjacent windows in a shopping street (Figure 2, left). These two screens were no separate systems but were hosted by one desktop computer through a multi-monitor graphics card in order to facilitate data logging. Another smaller screen was attached for the test manager to change the system configuration via a simple console, e.g. remotely specifying the connection technique for the current user. For the two large screens we chose two typical applications: one was designed as an official information terminal of the city showing a poll for citizen participation after connecting to it (Figure 2, right), the other one acted as an advertisement screen of a travel agency offering a quiz to win a journey. In front of the screens, we attached custom boxes containing an NFC tag and a QR code to the tables. Labels with the screens' names and three-digit numeric identifiers were placed at the bottom right corner of the display as suggested by Rashid et al. [21].

As mobile device, we used a Samsung Nexus S smartphone running Android 4.0.3. This device meets all the required hardware requirements including a suitable camera and an NFC sensor.

3.3 Mobile Prototype

Our mobile application communicated with the remote system over Wifi and featured all five pairing modes as depicted in Figure 1:

- *Display Pointing.* We realized the display pointing technique utilizing *Qualcomm's Vuforia*, an advanced augmented reality toolkit for mobile devices. Our mobile user interface showed the camera viewfinder in full-screen. Instead of typical graphical overlays, we only used the toolkit's image recognition capabilities to visually identify the background images shown on the two screens. The toolkit provides an efficient markerless image recognition based on natural image features (i.e. no obtrusive visual markers such as black/white patterns are required) and works highly robust also from varying distances and angles. In the mobile application, we associated each background image used in the study with a screen identifier (i.e. a fictive IP address). While this simplified approach with predefined images is feasible for a lab study, a more sophisticated back-end system is required to enable display pointing on a large scale. We presented a respective framework continuously informing nearby mobile devices about the currently visible display content in previous work [22].

- *Touching an NFC tag.* The NFC approach was implemented using the standard Android framework and appropriate method calls. Our mobile user interface just showed the message "Please touch the NFC tag". As soon as the smartphone was brought into close proximity of about three centimeters or less to one of the tags, the screen identifier was read.
- *Capturing a QR code.* To efficiently decode the QR codes, we made use of the popular ZXing software library. On top of the full-screen camera viewfinder we placed a rectangular border as target area like in typical barcode scanner applications. Our solution worked in continuous camera mode and was very responsive, i.e. explicitly pushing a button to photograph the marker was not required.
- *Scanning.* For study purposes, we created a view resembling the well-known Bluetooth scanning process (Figure 1d). The scan could be started and stopped with a push of a button. Then, a scrollable list of 12 faked display names (including the two available in the test room) was built up in random order with random delays between 0.5 and 1.5 seconds for each list item resulting in a typical Bluetooth scanning experience. As soon as a display name was visible, it could be selected by touching it.
- *Manual input.* Finally, we implemented a traditional numerical on-screen keyboard featuring a backspace and a confirm button for entering numeric identifiers.

When a screen identifier was correctly recognized by the smartphone (whether visually from the display content, read from an NFC tag, decoded from a QR code, selected from the scanning list or entered manually), it was sent to the desktop computer over Wifi and the successful connection to the screen was indicated by a short vibration. Further, the corresponding screen changed its content and showed a quiz question or a poll, respectively, with multiple answer buttons (Figure 2, right). The questions were selected from a set of prepared questions to avoid routine and keep the participants interested. To select an asnwer, the mobile application featured two interaction modes:

- *Touchpad.* In analogy to publicly available mobile applications like *Logitech's Touch Mouse*, the participants could relatively move the remote enlarged mouse cursor by strokes on the mobile display and simulate mouse clicks by pressing a button.
- *Mini video.* This interaction technique showed a cloned scaled-down version of the distant display's content on the mobile device for direct touch interaction similar to available remote control software such as *Teamviewer*.

3.4 Procedure

During the study, each participant used each pairing technique in a training and a test phase to connect to the prepared screens. However, each participant was only confronted with one randomly assigned interaction technique. To collect the participants' experiences and final resumes, we prepared a questionnaire enquiring demographic data and including questions to estimate the participant's technology affinity and experience. We decided to follow an established method for comparing technologically heterogeneous interaction techniques (cf. [2,3]) by engaging participants in interactions and afterwards asking them for the subjectively perceived performance. As an example for the heterogeneity of the compared techniques and the arising difficulties for objective assessment, the time for scanning can be easily measured as the time between starting the scanning and selecting an item from the result list, whereas the precise start time of capturing a QR code cannot be determined.

For evaluating each pairing technique in practice, each participant started with a training phase and subsequently went through the test phase. The order of the pairing techniques was systematically varied to avoid any learning effects. When the connection to a screen was successfully established, the participant answered the question related to the given interaction technique. Then the screen showed a short "Thank you" message and both the display and the mobile device were set back to the original pairing mode.

Training phase. During the training phase, we asked the participants to use only one screen and connect to it using the given technique for several times until they felt comfortable with it. Afterwards, we wanted them to answer how easy it was to learn this technique on a five-point Likert scale, ranging from "very easy" to "very hard" and to give reasons for their decision.

Test phase. Familiar with the pairing technique, the participants were asked to use the technique again, but now to move over to the other screen after each screen interaction, imagining that they would be using it within an everyday situation on a shopping street. As a constraint for the test phase, we asked the participants to switch between four and eight times between the screens. Having completed the test phase, the participants were asked to state on five point Likert scales how easy they considered the pairing process through the given technique and rate it in terms of mental and physical effort, speed, innovativeness, fun and accuracy and give short explanations for their rating. These aspects have been previously identified as key factors in physical mobile interactions [2,3]. A further rating scale was related to whether subjects would have any security concerns when using their personal mobile device for this technique at the aforementioned public place.

Final Interview. Having tested all five pairing techniques, we asked the participants to rate the techniques according their overall preference. Finally, we wanted to know whether the participants in general preferred short-distance (NFC, QR code) or long-distance (display pointing, manual input, scanning) techniques and explicit gestures (display pointing, NFC, QR code) or traditional on-screen interaction (manual input, scanning) at a public place.

4 Results

Overall, our participants conducted 1,478 successful pairing operations: 569 in training mode and 909 in test mode. In the following, we report on results from the user inquiry directly after usage, as well as comparative preference rankings and comments from the final interview. Error bars in the figures indicate 95% confidence intervals. For inferential statistical analysis, we ran GLM repeated measures tests with SPSS to identify main effects of the registration techniques. In case of violated sphericity assumptions, the values reported are using the Greenhouse-Geisser degrees of freedom adjustment. When main effects were significant, post hoc pair wise comparisons with Bonferroni corrections were performed. To test potential biases by the sample composition or test design, we also investigated interaction effects of registration technique vs. several independent factors such as age groups, gender, technology affinity or the used interaction technique after registration, again using SPSS GLM analyses. For none of these independent factors, we identified a significant interaction effect with registration technique.

We found significant main effects for learnability ($F_{4,65.58}$=8.38, $p<0.001$), ease of connecting ($F_{4,59.43}$=6.88, $p<0.01$), speed ($F_{4,79.59}$=10.52, $p<0.001$), mental ($F_{4,116}$=8.90, $p<0.001$) and physical demand ($F_{4,87.69}$=5.64, $p<0.01$), fun ($F_{4,83.33}$=32.06, $p<0.001$), innovativeness ($F_{4,79.17}$=17.45, $p<0.001$), accuracy ($F_{4,116}$=11.87, $p<0.001$), overall preference ($F_{4,116}$=6.96, $p<0.001$), and preparedness to use in public ($F_{4,116}$=4.57, $p<0.01$). However, we did not find significant main effects for mean rating scores on security concerns (rating scores were around 2.9, the average SD was 1.35).

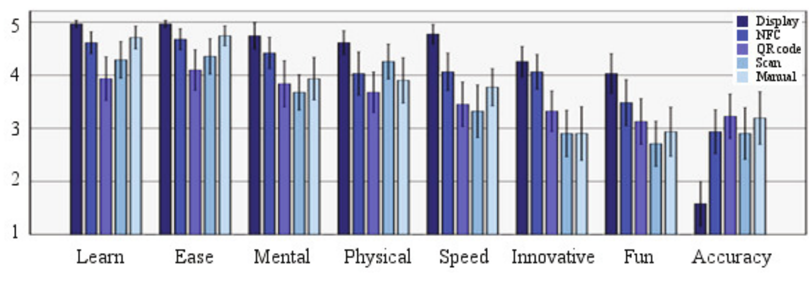


Fig. 3. Rating results for the five pairing techniques

Learnability. In general, participants' ratings after the training phase for each technique were rather positive (Figure 3), with mean scores of 4 and higher. When analyzing pairwise differences for learnability, we found that display pointing received an almost maximum mean rating of 4.97 (SD=0.18), followed by manual input (M=4.71, SD=0.59), NFC (M=4.61, SD=5.58), scanning (M=4.29, SD=0.94) and QR code (M=3.94, SD=1.12). We found significant pairwise differences between display pointing and NFC (p=0.28), display pointing and scanning ($p<0.001$), display pointing and QR code ($p=0.004$), as well as manual and QR code ($p=0.015$).

Ease of pairing. Also the ratings on how easy it was in overall to connect to the screen were quite positive for all techniques. A nearly perfect mean score was received for display pointing (M=4.97, SD=0.18). This was again followed by manual input (M=4.74, SD=0.51), NFC (M=4.68, SD=0.54), scanning (M=4.35, SD=0.92) and QR code (M=4.10, SD=1.04). Significant differences were found between display pointing and scanning ($p=0.012$), display pointing and QR code ($p=0.001$), as well as between manual input and QR code ($p=0.023$).

Mental demand. The rating scores for mental demand (1=high demand, 5=low demand) were best for display pointing (M=4.74, SD=0.68), second best for NFC (M=4.42, SD=0.81), followed by QR code (M=3.84, SD=1.19), manual input (M=3.94, SD=1.09), and scanning (M=3.68, SD=0.91). Significant pairwise differences were found between display pointing and QR code ($p=0.002$), display pointing and manual input ($p=0.004$), display pointing and scanning ($p<0.001$), as well as NFC and scanning ($p=0.015$).

Physical demand. Display pointing was rated best (M=4.61, SD=0.62) and QR code was rated worst (M=3.68, SD=1.05); only this pairwise difference was significant ($p=0.001$).

Speed. Display pointing (M=4.77, SD=0.48) was qualified significantly faster than the other techniques NFC (M=4.06, SD=0.96, $p=0.009$), manual input (M=3.77, SD=0.95, $p<0.001$), QR code (M=3.45, SD=1.15, $p<0.001$), and scanning (3.32, SD=1.35, $p<0.001$). No other pairwise differences were found.

Innovativeness. Figure 4 indicates that display pointing (M=4.26, SD=0.77) and NFC (M=4.06, SD=0.89) were perceived more innovative than the other techniques. We did not find a significant pairwise difference between them, but each of them had significant pairwise differences to the other techniques, namely to QR code (M=3.32, SD=1.05, both $p<0.01$), scanning (M=2.90, SD=1.19, both $p<0.01$), and manual input (M=2.90, SD=1.38, both $p<0.01$).

Fun. Display pointing provided most fun (M=4.65, SD=0.84), significantly more than the other techniques NFC (M=3.33, SD=1.16), QR code (M=2.87, SD=1.26), manual input (M=2.00, SD=0.86), and scanning (M=2.10, SD=1.04), all differences were highly significant ($p<0.001$). Further significant differences were NFC vs scanning ($p=0.006$) and NFC vs manual input ($p<0.001$).

Accuracy. Overall, the participants' accuracy ratings were comparatively low, with scores of 3.20. Display pointing was qualified as least accurate (M=1.58, SD=1.15), significantly less accurate than scanning (M=2.90, SD=1.33), NFC (M=2.94, SD=1.12), QR code (M=3.00, SD=1.15), and manual input (M=3.19, SD=1.33), all differences highly significant ($p<0.001$). There were no further significant pairwise differences.

Overall preference. Figure 4 shows that display pointing received the highest mean ranking score, i.e. it was significantly more often preferred than the other techniques (M=4.06, SD=1.34), NFC (M=3.00, SD=1.32), manual input (M=2.84 SD=1.10), QR code (M=2.81, SD=1.38), and scanning (M=2.29, SD=1.40), all pairwise comparisons significant ($p<0.01$). Figure 5 (left) shows that when participants were explicitly asked about the distance to a public screen while registering to it, 87.1% preferred long distance methods (such as display pointing, manual input or scanning) to short distance methods (such as NFC or QR code, 9.7%), only 3.2% were undecided. Confronted with the question whether they would like to apply a gesture (such as with display pointing or touching an NFC tag) vs a typical smartphone interaction (by manual input or scanning), the distribution of answers was balanced between these two alternatives (45.2% vs. 38.7%), and 16.1% participants were undecided (Figure 5, right).

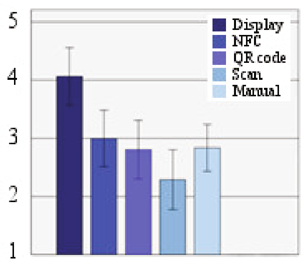


Fig. 4. Mean ranking scores for 'Overall Preference'

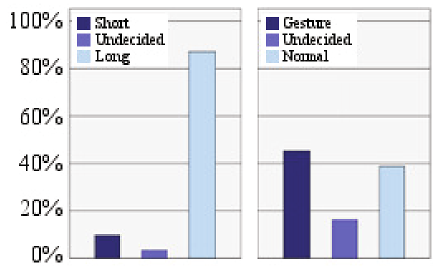


Fig. 5. Percentage of preference for distance (left) and gesture (right)

4.1 Observations and Statements

While most of the participants did not have any remarkable problems with NFC, scanning and manual input, some test persons struggled with capturing the QR codes. Even though the QR code scanner worked very well and responsive, targeting the code was difficult for them. In contrast, several people were obviously surprised by the speed of display pointing while some of them, especially elderly, described display pointing even as "too fast" and "too sensitive".

However, we observed that the subjective speed of a mobile connection technique in the context of public displays is perceived differently. While most of the participants experienced NFC, QR code and display pointing as fast techniques due to their short response times, others surprisingly considered scanning and manual input as faster alternatives. These techniques allow passers-by to start establishing the connection even while approaching the screen (given the screen identifier is known or well readable also from distance, respectively) and not to walk to a smart label, thus shorten the distance to be travelled and resulting in an earlier completed connection attempt.

During observation, we especially directed our attention to the moment after a successful connection attempt when the participants moved from the connection phase to the interaction phase and tried to identify issues in the combination of different connection and interaction techniques. As potential obstacle several participants mentioned that switching device orientation is impractical, e.g. when combining manual input in portrait mode (Figure 1e) with interaction in landscape mode such as mini video. Surprisingly, many of these participants used display pointing and the QR code reader in portrait mode, even though both also worked in landscape mode.

Further, some combinations resulted in a complete departure from the large display by some participants. In the previous example combining manual input with mini video, those were entirely focused on the mobile device throughout both pairing and interaction. Some even did not notice changes on the large screen at all and considered the action happening only in the mobile application. Typical statements in such cases included that the public screen actually would not be needed.

5 Discussion and Conclusions

In this section, we refer back to our research hypotheses and discuss the results of our study in view of our previous assumptions.

5.1 H1. Learnability

The hypothesis on the learnability of display pointing is confirmed. Display pointing received outstanding learnability scores and was considered significantly better than NFC, QR code and scanning and also better, yet not significantly, than the very trivial manual input method in terms of learnability. Whereas scanning consists of several interaction steps, QR code needs accurate pointing and NFC requires a special hand posture to position the NFC sensor of the device over the tag. Display pointing, by contrast, consists of one intuitive gesture and is very robust to varying distances and angles when based on state-of-the-art recognition algorithms. Since the recognition of parts of the remote display is already sufficient for establishing the connection, no precise pointing and targeting is required making this technique also tolerant against hand or arm trembling and thus suitable for both young and elderly users with varying experiences in smartphone usage and different physical conditions.

5.2 H2. Perceived Performance

The study results support this hypothesis in major parts, however, not entirely. Concerning the occurring mental load, NFC and display pointing are perceived best. Scanning and manual input require reading, mistakes in capturing a QR code easily leads to frustration. This is also true for the physical load, where display pointing is rated best with a significantly lower load than the QR code alternative. Furthermore, display pointing significantly outperforms all alternatives in terms of speed and fun. This finding is also supported by observations within the training and test phases, where participants positively noted the game-like targeting of display pointing. However, in contrast to capturing QR codes, display pointing was always successful and led to a sense of achievement for the participants. In our context of enabling mobile interactions with public displays, we especially regard the criteria ease of connecting, speed and fun as key factors and thus consider display pointing as very suitable: pairing techniques should be fast to allow users to quickly start doing the actual task, should be comfortable and easy to use to lower the threshold of starting an interaction and should be fun to animate passers-by to participate.

However, the major drawback of display pointing turned out to be the perceived accuracy: Participants were not satisfied with the technique's accuracy, they rated display pointing significantly worse than all other techniques. The participants' statements and our observations explain this drawback: for several (especially elderly) participants, the tested implementation of display pointing worked too fast and sensitively what resulted in a perceived lack of control and the fear of selecting a nearby screen by mistake. As appropriate countermeasures we suggest the integration of optional confirmation features to avoid such wrong selections. As soon as the remote display is detected, the mobile application should communicate the successful recognition and provide a simple approval mechanism. This might either be a short dwell time indicated by a progress bar or an additional button for explicit pairing. We consider the second approach as more promising since the mobile device can be brought down into a normal position after the pointing gesture.

5.3 H3. Security Concerns

The results of the study do not confirm our hypothesis on security. Surprisingly, we could not find any significant differences with regard to security: the participants considered remote connection techniques as secure as the alternatives in the evaluated context. With regard to the results of related previous studies, we interpret this as a consequence of the emerging ubiquity of mobile services based on QR codes and NFC tags and thus a familiarization and growing trust in these techniques. Further, we consider the type of interactive screen application to be crucial: With the quiz game and the poll we chose two very typical examples for our study, namely none security-critical applications. In contrast to mobile payment processes or similar, connecting to a public display to play a game or take part in an informal survey seems not to give rise to any security concerns.

5.4 H4. Distance

Our hypothesis on the preference of techniques allowing for connections from distance is fully confirmed. Whereas short-range techniques such as NFC were proven to work well for security-relevant interactions in the public, e.g. for purchasing tickets [8], more than 80% of our participants clearly stated that they appreciate long-distance techniques for connecting to a public display over pairing methods requiring immediate vicinity to the screen. As reasons for their decision, the test persons mentioned not to be forced to walk to the display as well as the applicability for multi-user applications at busy locations where interested passers-by could simply connect from distance instead of breaking through the crowd or queuing up in front of a corresponding smart label. Another major reason was the opportunity to literally operate from a position in the background and not to be immediately recognizable as the one responsible for the actions on the public display.

This clear preference for long-distance techniques in the context of mobile interactions with public displays is a strong point for scanning, manual input (if the screen identifier is known) as well as display pointing. Due to high-quality cameras and zoom features and dependent on the size of the remote display, this technique can also be applied from larger distance.

5.5 H5. Gesture

Our hypothesis on the preference of "normal" smartphone interactions over explicit device gestures is not confirmed. Slightly more participants preferred an explicit gesture for setting up the connection than a normal smartphone interaction such as on-screen typing. We ascribe this outcome to the positive experiences with the tested gestural techniques (especially touching the NFC tag and display pointing) and primarily their quick response times with regard to scanning and manual input. This result indicates that directional device gestures like touching and pointing are accepted when they are perceived useful. Further, this result may be a sign of changes in technology usage and acceptance over time: While in related early studies participants tended to avoid explicit smartphone gestures in public space or mentioned feeling awkward or embarrassed (cf. [5,14]), such gestures have become more mainstream in the meanwhile and thus socially acceptable. This is further confirmed by the technique ranking concerning the overall preference: the gestural display pointing method is significantly preferred to the alternatives.

6 Limitations

To gain systematic and comparable results for this first comparative study on screen pairing, we decided to conduct our study under controlled lab settings, accompanied by a test assistant to maximize data richness. We took several measures to alleviate the potential limitation of contextual validity by arranging a setup with two screens according to a typical scenario of a shopping street. Further, the test assistant emphasized the public context of the study several times when going through the questionnaire and also used photos showing public displays at busy locations. We assume that these points help to make clear the public context and were confirmed by several participants who mentioned the realistic setup and atmosphere of their own accord. In the light of the promising findings obtained, a field study to verify and contextualize these is definitely justified. This should especially be beneficial for the exploration of further benefits and functional restrictions of display pointing, e.g. depending on the impact of crowded places, varying distances and different light conditions.

7 Outlook

We presented a user study evaluating different pairing techniques for enabling serendipitous mobile interactions with public displays. Surprisingly, capturing

QR codes turned out to be difficult for a considerable number of users and lowers the learnability. NFC, by contrast, received good ratings in general, however, suffered from the general preference for distant techniques in the context of public displays. The distant techniques scanning and manual input are easy to learn and to use, but are perceived as less innovative and fun. Display pointing was proved to be good combination of a directional technique supporting distant pairing while being efficient and fun. Yet, mentioned improvements must be considered to address the perceived lack of control and avoid wrong selections.

Besides a replication study in the field, further work should investigate how to communicate available interactive features in the case of sophisticated markerless setups, i.e. how to optimally inform passers-by about the pairing opportunity. Further interesting research questions include the impact of multi-user applications on the preferred pairing technique. While we consider display pointing a suitable technique due to its support for pairing from distance and several participants of our study implicitly considered this aspect in their ratings, a dedicated study with groups of participants might clarify respective issues.

Acknowledgments. This work has been carried out within the projects b-Part and PRIAMUS financed by FFG and A1. Further funding was received from Nokia Foundation, the Research and Training Foundation of TeliaSonera Finland Oyj, and HPY Research Foundation. FTW Forschungszentrum Telekommunikation Wien GmbH is funded within the program COMET by BMVIT, BMWA, and the City of Vienna. The COMET program is managed by the FFG.

References

1. Ballagas, R., Borchers, J., Rohs, M., Sheridan, J.G.: The smart phone: A ubiquitous input device. IEEE Pervasive Computing 5(1), 70–77 (2006)
2. Rukzio, E., Leichtenstern, K., Callaghan, V., Holleis, P., Schmidt, A., Chin, J.: An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 87–104. Springer, Heidelberg (2006)
3. Rukzio, E., Broll, G., Leichtenstern, K., Schmidt, A.: Mobile interaction with the real world: An evaluation and comparison of physical mobile interaction techniques. In: Schiele, B., Dey, A.K., Gellersen, H., de Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., Buchmann, A.P. (eds.) AmI 2007. LNCS, vol. 4794, pp. 1–18. Springer, Heidelberg (2007)
4. Fröhlich, P., Oulasvirta, A., Baldauf, M., Nurminen, A.: On the move, wirelessly connected to the world. Commun. ACM 54(1), 132–138 (2011)
5. Mäkelä, K., Belt, S., Greenblatt, D., Häkkilä, J.: Mobile interaction with visual and rfid tags: a field study on user perceptions. In: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, pp. 991–994. ACM (2007)
6. Rohs, M.: Real-world interaction with camera phones. In: Murakami, H., Nakashima, H., Tokuda, H., Yasumura, M. (eds.) UCS 2004. LNCS, vol. 3598, pp. 74–89. Springer, Heidelberg (2005)
7. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D.: Real-time detection and tracking for augmented reality on mobile phones. IEEE Transactions on Visualization and Computer Graphics 16(3), 355–368 (2010)

8. Broll, G., Rukzio, E., Paolucci, M., Wagner, M., Schmidt, A., Hussmann, H.: Perci: Pervasive service interaction with the internet of things. IEEE Internet Computing 13(6), 74–81 (2009)
9. Riekki, J., Salminen, T., Alakarppa, I.: Requesting pervasive services by touching rfid tags. IEEE Pervasive Computing 5(1), 40–46 (2006)
10. Välkkynen, P., Niemelä, M., Tuomisto, T.: Evaluating touching and pointing with a mobile terminal for physical browsing. In: Proc. of the 4th Nordic Conf. on Human-Computer Interaction: Changing Roles, pp. 28–37. ACM (2006)
11. von Reischach, F., Michahelles, F., Guinard, D., Adelmann, R., Fleisch, E., Schmidt, A.: An evaluation of product identification techniques for mobile phones. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 804–816. Springer, Heidelberg (2009)
12. Scott, D., Sharp, R., Madhavapeddy, A., Upton, E.: Using visual tags to bypass bluetooth device discovery. SIGMOBILE Mob. Comput. Commun. Rev. 9(1), 41–53 (2005)
13. Seewoonauth, K., Rukzio, E., Hardy, R., Holleis, P.: Touch & connect and touch & select: interacting with a computer by touching it with a mobile phone. In: Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services, pp. 36:1–36:9. ACM (2009)
14. O'Neill, E., Thompson, P., Garzonis, S., Warr, A.: Reach out and touch: Using NFC and 2D barcodes for service discovery and interaction with mobile devices. In: LaMarca, A., Langheinrich, M., Truong, K.N. (eds.) Pervasive 2007. LNCS, vol. 4480, pp. 19–36. Springer, Heidelberg (2007)
15. Ballagas, R., Rohs, M., Sheridan, J.G.: Sweep and point and shoot: phonecam-based interactions for large public displays. In: Ext. Abstracts on Human Factors in Computing Systems, pp. 1200–1203 (2005)
16. Madhavapeddy, A., Scott, D., Sharp, R., Upton, E.: Using camera-phones to enhance human-computer interaction. In: Sixth Int. Conf. on Ubiquitous Computing (Adj. Proc.: Demos) (2004)
17. Broll, G., Reithmeier, W., Holleis, P., Wagner, M.: Design and evaluation of techniques for mobile interaction with dynamic NFC-displays. In: Proc. of the Fifth Int. Conf. on Tangible, Embedded, and Embodied Interaction, pp. 205–212. ACM (2011)
18. Dachselt, R., Buchholz, R.: Natural throw and tilt interaction between mobile phones and distant displays. In: Ext. Abstracts on Human Factors in Computing Systems, pp. 3253–3258. ACM (2009)
19. Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P.: Touch projector: mobile interaction through video. In: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, pp. 2287–2296. ACM (2010)
20. Baldauf, M., Salo, M., Suette, S., Fröhlich, P.: Display pointing - a qualitative study on a recent screen pairing technique for smartphones. In: Proc. of the 27th Int. BCS Human Computer Interaction Conference. ACM (2013)
21. Rashid, U., Terrenghi, L., Quigley, A.: Labeling large displays for interaction with mobile devices: recognition of symbols for pairing techniques. In: Proc. of the Int. Conf. on Advanced Visual Interfaces, p. 417. ACM (2010)
22. Baldauf, M., Fröhlich, P., Lasinger, K.: A scalable framework for markerless camera-based smartphone interaction with large public displays. In: Proc. of the Int. Symposium on Pervasive Displays, pp. 4:1–4:5. ACM (2012)