CrossMark

ORIGINAL ARTICLE

Select & Apply: understanding how users act upon objects across devices

**Adalberto L. Simeone · Ming Ki Chong ·
Corina Sas · Hans Gellersen**

Abstract As our interactions increasingly cut across diverse devices, we often encounter situations where we find information on one device but wish to use it on another device, for instance, a phone number spotted on a public display but wanted on a mobile. We conceptualize this problem as Select & Apply and contribute two user studies where we presented participants with eight different scenarios involving different device combinations, applications and data types. In the first, we used a think-aloud methodology to gain insights on how users currently accomplish such tasks and how they ideally would like to accomplish them. In the second, we conducted a focus group study to investigate which factors influence their actions. Results indicate shortcomings in present support for Select & Apply and contribute a better understanding of which factors affect cross-device interaction.

1 Introduction

We have access to an ever increasing range of personal devices, such as desktop computers, smart phones and watches, tablets, reader devices, etc. However, each device exists exclusively as an island of interactivity. That is, interaction between applications is possible when they both run on the same device. When applications run on different devices, interaction is restricted and typically needs an intermediary device (e.g., a USB cable, a PC to transfer files between two mobiles) or service (e.g., cloud synchronization, email attachments). With this work, we investigated how users approach these cross-device interaction tasks, which we call Select & Apply tasks. That is, if we want to *select* one object on an application on one device that we wish to *apply* on another, how do we achieve that?

This problem frequently arises because most devices' typical usage scenarios consist of a specific set of activities. For example, we use a smartphone to place calls or navigate us to a destination, but due to the small form factor, we might prefer using a tablet or e-reader device to read a document. We use desktop PCs for more complex activities such as creating a presentation, but part of its content, such as videos and pictures, might have been created by mobile devices. Although research has thoroughly investigated cross-device interaction techniques [2, 10, 24, 26], the prevailing interaction metaphor on consumer devices continues to be based on send and receive. That is, the source device provides functionality to broadcast data to a receiver device either directly, through some form of pairing (e.g., USB cable, Bluetooth), or indirectly through intermediary services (e.g., cloud synchronization, email attachments).

This paper contributes two multi-user studies investigating Select & Apply tasks and a discussion of the feedback we collected. We sought to understand how users approach these tasks with current technology and to analyze features affecting the interaction techniques suggested by our participants. Our work can provide interaction technique designers with a background of Select & Apply tasks and the motives that highlight it as a pressing issue. Further, as results confirmed that there is no technique suitable for all purposes, we discuss how different scenarios and factors affect the suitability of techniques for cross-device interaction tasks.

A. L. Simeone (✉) · M. K. Chong · C. Sas · H. Gellersen
School of Computing and Communications,
Lancaster University, Lancaster LA1 4WA, UK
e-mail: adalberto.simeone@port.ac.uk

A. L. Simeone
School of Creative Technologies, University of Portsmouth,
Portsmouth PO1 2DJ, UK

We designed eight real-world scenarios that involve different combinations of devices, applications and types of data, as a basis for engaging users in analysis of the Select & Apply concept. The first study followed an analytical approach where 20 participants were asked to discuss their experiences in these scenarios. They were asked to describe: (1) the methods they would use to accomplish the tasks according to their current understanding of technology and (2) an ideal technique that would be better suited for each task. The results we obtained from the first study helped us to paint a clearer picture of the status quo in terms of technologies used and to define a set of interaction technique models, from the two perspectives of the Select & Apply concept: selection and application. This study was followed by a second one, consisting of six focus group sessions. Our goals were to investigate in depth which factors motivated users' suggestions, to understand issues they experience and identify desirable properties cross-device interaction techniques should provide.

2 Related work

2.1 Cross-device interaction techniques

Several works have tackled the problem of Select & Apply tasks that extend beyond a single device. In close proximity, the *Pick–and-drop* technique by Rekimoto et al. [24] proposed the idea of a "direct manipulation technique" enabling users to perform copy and paste or drag-and-drop actions across devices through the use of a pen. It used an electronic pen communicating with a "pen server" through a unique ID to enable these features. *HyperDrag* [26] proposed the use of an augmented workspace where a camera-based recognition system allowed users to seamlessly drag data between devices. Hinckley introduced the notion of *Synchronous Gestures* such as bumping two devices together to establish a connection in order to transfer information. This later evolved into "stitching" devices together so that interactions could extend from a source device to the one next to it [11]. In their study, tablets operated by pen devices were used to demonstrate the concept. The time occurrences of pen events were used to disambiguate the interested devices. Similarly, RFID tags were used in [25] to establish a remote connection between a mobile device and a desktop PC. Upon close proximity, a proxy window representing the mobile device is shown in the desktop screen. Users can use this proxy to transfer files.

Other works have studied the use of mobile devices as a medium to perform Select & Apply tasks. In [9] a NFC, capable mobile phone could use *select-and-drop* or *select-and-pick* techniques to trigger the transfer of pictures to

and from a smart picture board augmented by a grid of NFC tags. This was further investigated in *PhoneTouch*, a system allowing mobile devices to interact with a rear-projected tabletop surface [28]. Combining vision-based analysis of the contacts of the phone with the surface, acceleratory and microphone data, the system is able to sense when a user is using her/his phone to interact with the surface. PhoneTouch can be used to pick or drop objects by touching them with the phone. However, the technique only operates in such a described environment.

2.2 Cross-device interaction studies

The works cited previously have presented designs of Select & Apply interaction techniques and evaluated their usability. More recently, researchers have begun to focus on the Select & Apply domain as a whole rather than on single solutions. Scharf et al. [27] presented a taxonomy of cross-device interaction. In their classification, they include all instances of interaction between two devices, such as controlling a presentation on an external device through one's mobile device. Dearman and Pierce reported a study on cross-device interaction practices focusing on academic and industry participants [7], whereas we selected our participants from a non-technical population. Their study suggested future interaction designers to focus on user activities rather than on single devices. Further, they identified the transfer of information as a pressing issue not optimally addressed by existing technology. Marquardt et al. [20] presented a cross-device interaction study informed by the concepts of proxemics of people and device. However, the resulting interaction techniques rely on the presence of an augmented environment and focus on abstract tasks, rather than on real-world applications.

2.3 Understanding users

In this work, we have explicitly involved users in our investigation of the Select & Apply domain. The general consensus in literature on participatory design is that user involvement is likely to lead to a better understanding of system requirements rather than a substantial impact on its effectiveness [14, 15]. Wobbrock et al. [32] inspired a number of studies sought to elicit insights on how to perform specific commands or gestures directly from users, often referred to as "guessability studies". A further study showed that users prefer a user-defined gesture set as opposed to an expert-designed one [23]. In the cross-device interaction domain, various researchers have applied this methodology to elicit gestures to perform commands involving multiple devices. Kray et al. [13] asked participants to freely propose gestures for a list of activities involving combinations of two devices among

smartphones, tabletop and public displays, without concerns for technical limitations. Kurdyukova et al. [16] elicited user-defined gestures involving interaction between two tablets, tablet to tabletop and tablet to wall display. Key distinctions exist between our approach and these previously cited. First, the former (user involvement) is usually employed in the design of a single product or artifact, whereas our perspective focuses on the cross-device interaction domain itself. Second, the latter (guessability studies) have been used to elicit sets of gestures in abstract tasks, whereas we tackle real-world application scenarios. We consider our approach to differentiate from these two approaches. We use the analytical elicitation aspect found in guessability studies in our first study. However, we do not take users' suggestions to be representative of final interaction techniques. Rather, we regard the involvement of users as necessary to provide future designers with background information, current issues, desirable interaction techniques and characteristics affecting the Select & Apply domain (Fig. 1).

3 Select & Apply

Select & Apply tasks are a subset of interaction tasks typically taking place at personal distances [1]. It involves two components: an object to select and an application target. The instantiation of these two components implicitly determines the resulting action. For example, selecting a file in one folder and applying it into another folder will trigger a move action. These tasks are commonly found in our daily interactions. Modern operating systems provide several interaction techniques allowing multiple applications running on the same device to interact with each other. Techniques such as drag-and-drop and contextual menus can be considered to fulfill the Select & Apply

paradigm. However, when the context of use extends to multiple devices, the above is no longer possible. Users need to adapt to the lowest common denominator, as cross-device interaction support is currently limited. Available methods rely on "broadcasting" data from the device where the object of interest is found to the device where is needed. As introduced, this happens either through pairing (e.g., Bluetooth or USB cables) or through intermediary services (e.g., cloud-based synchronization or email attachments). These methods introduce additional steps (connecting a cable, finding the location in the filesystem where the object is, etc.) that hinder the interaction flow. We argue that being able to Select & Apply objects across devices is a desirable interaction capability. However, we believe there is no method suitable for all purposes, due to the wide variety of devices, having different form factors and input capabilities. We designed a set of eight scenarios representative of commonly occurring situations or likely to be in the near future. Our intention was to use these scenarios as starting points for discussions. Each scenario presented our participants with a task requiring them to select data (pictures, documents, event information, etc.) from one device and apply it to another. We considered interactions occurring between PCs/Laptops, smartphones, tablets and public displays. For each, we created two mock-ups: the initial state showing the data within an existing application (e.g., a picture in a mobile gallery app); the final state showing the results of the interaction (e.g., a picture placed in a PowerPoint slide). These are:

3.1 Scenario 1: phone number

Selecting a phone number displayed on a PC screen and applying it to a smartphone dialler (see Fig. 2). The user has bought an item from an online store and wishes to inquire about the order. She finds the customer service
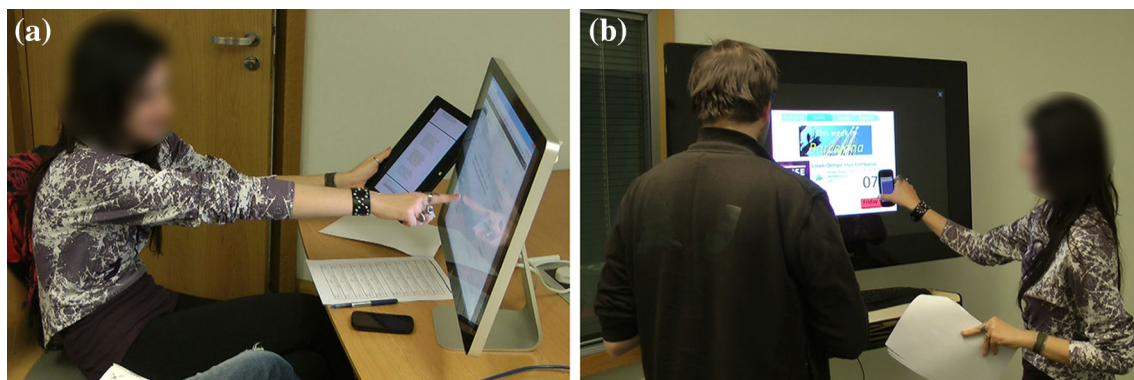

Fig. 1 Video screenshots recorded during the first study—**a** participant simulating a *cross-device drag-and-drop* technique by selecting text from a tablet to and applying it to a text field in a

webpage on the screen; **b** participant using a proximity technique to select information of an event shown on a public display and applying it to her calendar application
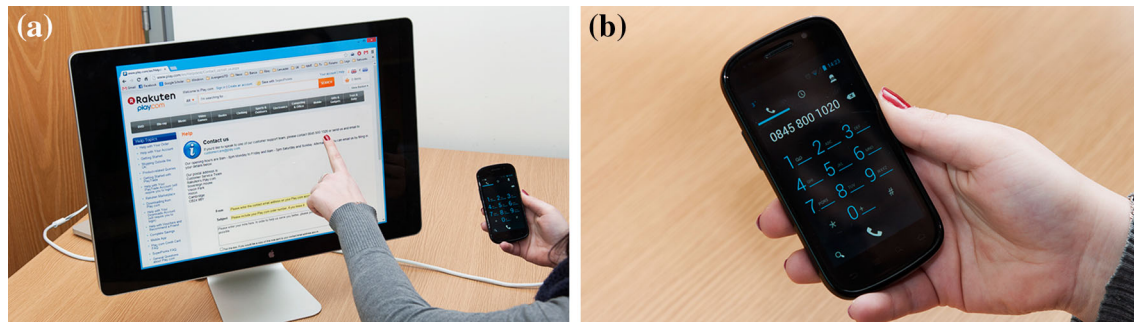
Fig. 2 Scenario 1: phone number—**a** the contact us section of an online store; **b** the customer service number in the dialler application

Fig. 3 Scenario 2: map address—**a** the find us section of a restaurant; **b** a map application showing the path from the user's current location to the restaurant's

number on the "contact us" page of the website. She needs to apply the phone number to the dialler application on her smartphone.

3.2 Scenario 2: map address

Selecting an address displayed on a PC screen and applying it to a smartphone map application (Fig. 3). The user is in a hotel room, looking for a restaurant on her laptop. After browsing a number of potential candidates, the user decides to look up the address of the one where she eventually decides to go to. Having never been there, the user would like to apply the address selected from the webpage to the map application on her smartphone so that being just about to leave the hotel room it may help the user reach their destination.

3.3 Scenario 3: picture

Selecting a picture taken with a smartphone and applying it to a PowerPoint slide (Fig. 4). The user is in an office environment, working on a PowerPoint presentation in which she needs to place a picture that has recently been taken with a smartphone. She is in front of her desktop PC with a smartphone close at hand; she needs to apply the selected picture from smartphone to a specific location in the slide.

3.4 Scenario 4: document files

Selecting a PDF document from a desktop PC and applying it to a viewer application on a tablet (Fig. 5). The user has a PDF of a document that she wishes to read on a tablet or similar reader device as she prefers its more comfortable reading experience. The document is open in a viewer application on her desktop screen side by side with the containing folder.

3.5 Scenario 5: text

Selecting a text paragraph from a book displayed in a tablet and applying it to a desktop blog web application (Fig. 6). The user is reading a book on her tablet when she comes across an interesting paragraph. She would like to use this paragraph inside the body of a new blog post so that she may use it as the incipit of a discussion.

3.6 Scenario 6: passcode

Selecting a PIN code from a text message and applying it to a text field in a web form (Fig. 7). The user is attempting to validate herself on a website. It requires the PIN code in the text to be entered in the appropriate field in the form.

Fig. 4 Scenario 3: picture—**a** gallery application on a smartphone; **b** a slide with the picture applied to it
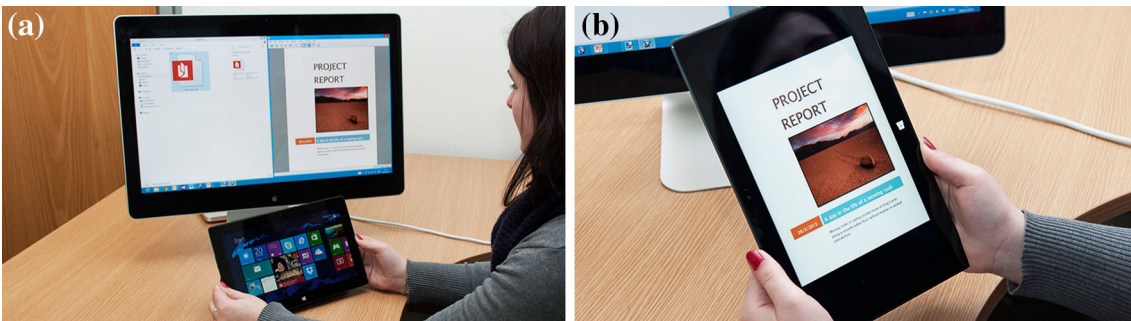

Fig. 5 Scenario 4: document files—**a** PDF document on a desktop PC; **b** the PDF document opened in the tablet
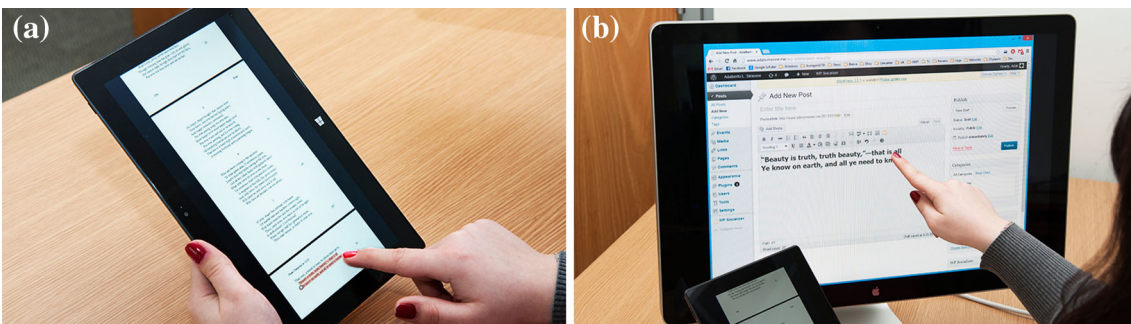

Fig. 6 Scenario 5: text—**a** tablet where a sentence has been highlighted; **b** the paragraph applied in a blogpost
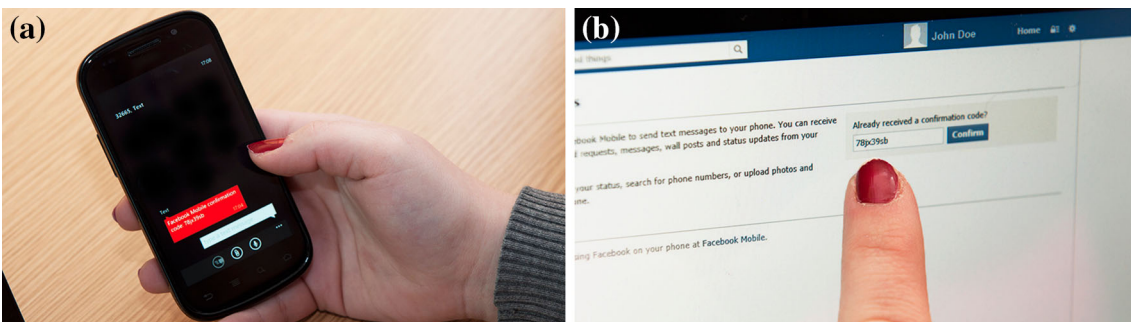

Fig. 7 Scenario 6: passcode—**a** text message showing a PIN code; **b** the PIN code applied to a text field

3.7 Scenario 7: digital ticket

Selecting an electronic ticket purchased from a public terminal and applying it to a digital wallet on a smartphone (Fig. 8). The user is interacting with an automated ticketing machine that sells electronic tickets need a smartphone to be collected. After paying, he needs to collect the ticket and place it inside his digital wallet (e.g., Passbook).

3.8 Scenario 8: event

Selecting event information displayed on a public display and applying it to a calendar application on a smartphone (Fig. 9). The user has just arrived at an airport for a business trip and is currently waiting for the luggage carousel to start. A nearby public display catches his attention so the user begins interacting with it. It provides touristic information, current events, etc. He finds that one of their favorite bands will be playing in this city during the weekend, so he decides to attend. To remember the event, he wants to apply the information (place, date, time, etc.) inside his calendar.

4 First user study: think-aloud interviews

To find out how people Select & Apply information across devices, we conducted a think-aloud study. For each of the eight scenarios, we asked participants to describe how they would approach each task based on their experience with current technology. Successively, we asked them to imagine an ideal interaction technique better suited for the task, regardless of its technological feasibility.

4.1 Apparatus and environment

We used four different devices (see Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9): a 24″ display, which represented a desktop computer (or a laptop); an Android Nexus S smartphone; a Microsoft Surface RT tablet; and a vertically mounted Microsoft PixelSense, which acted as a situated public display (for scenario #7 and #8). We conducted the study in a quiet meeting room. During the study session, only the participant and the experimenters were present. We installed two video cameras to record the participant's actions and conversations, and we also took notes.

4.2 Participants

Twenty paid participants (twelve female), aged between 18 and 29 ($M = 21.7$, SD $= 2.43$), took part in our study. As our aim was to understand the general public, we only recruited people from our university with non-technical background (i.e., no computer science students). Prior to the study, we asked the participants to rate their proficiency
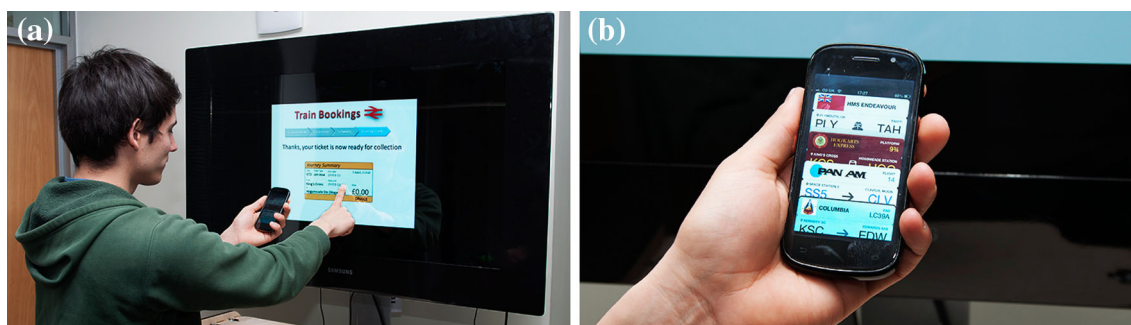

Fig. 8 Scenario 7: digital ticket—**a** mock-up of an electronic ticket ready for collection; **b** the electronic ticket displayed in a digital wallet
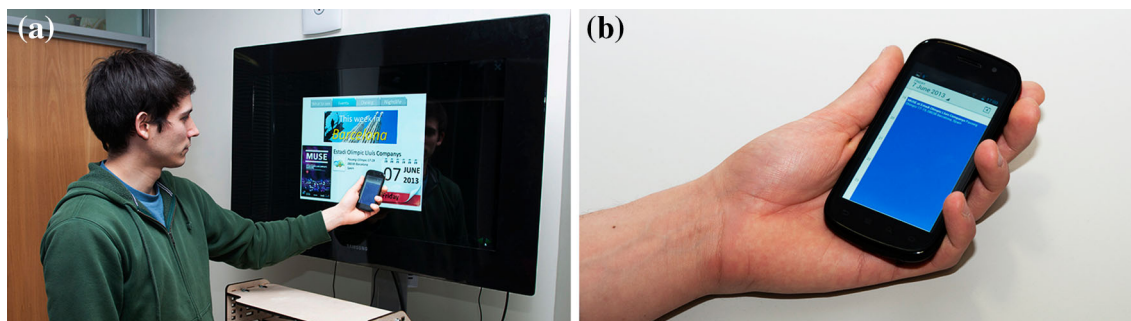

Fig. 9 Scenario 8: event—**a** an event information displayed on a public display; **b** the event information applied to a calendar application

in using the four types of devices on a scale which ranged from 1 (No experience) to 7 (Expert). They gave an average of 5.65 (SD = 1.63) for PCs, 4.75 (SD = 1.33) for smartphones, 4.5 (SD = 1.10) for public displays and 3.56 (SD = 2.35) for tablets.

4.3 Procedure

The experiment followed a semi-scripted think-aloud procedure. Upon arrival, the participants were asked to fill in a short questionnaire about their demographic information and their proficiency in using the four devices. Before the session started, we explained them the purpose of the study and how the study would unfold. The study followed a within-group study design, where every participant experienced all eight scenarios. Their presentation order was counterbalanced.

For each Select & Apply scenario, we first narrated its background setting, using descriptions similar to those reported in the previous section. To help participants relate to the scenario, we used a mock-up UI which portrayed the initial state (prior to the selection) and the final state (after application) of the involved devices. These mock-ups were either screenshots of real applications (Figs. 2, 3, 4, 5, 6, 7) or especially designed (Figs. 8, 9).

We asked the participants how they would accomplish the Select & Apply task based on their understanding of current technology. The investigator handed the participants the devices (which were displaying the mock-up images) relevant to the specific scenario. The devices acted as a thinking aid and helped the participants to conceptualize the scenarios. Thereafter, participants were asked to think beyond the limitations of the methods they currently use. We used this prompt to clarify the details of the proposed technique from the perspective of the two components of the Select & Apply concept. For instance, how the object of the task is selected in the first place and how it is then applied on another device.

4.4 Results

We analyzed the notes we took during the interviews and the video recordings. We divided participants' feedback in two categories of current and proposed techniques. The former represents the solutions currently employed by users to Select & Apply data; the latter represents solutions participants would use if these were implemented and usable. We identified shared traits between solutions and grouped them according to how they approach the two components of a Select & Apply task. This led us to define a list of selection and application categories, which we describe in the following:

4.4.1 Current selection categories

Selection categories represent the means used to identify the object the user wishes to apply to another device and trigger its selection. Participants reported using three main methods:

Direct manipulation refers to selecting of an object through the device's standard input modality (e.g., a click, a tap).

Aiming refers to selecting an object by aligning the device to it (e.g., by targeting the object with the device's viewfinder).

Gaze refers to the act of identifying an object among other elements of a user interface by looking at it. Due to the absence of interaction techniques able to support information such as phone numbers, addresses, event data, "selection" only happens as a mental note. That is, the user identifies the information among other elements through a fixation rather than with an actual interaction command.

4.4.2 Current application categories

Application categories represent the actions used either directly (by the user) or indirectly (by the system) to trigger the actual application of the selected object. The solutions proposed by the participants were grouped into seven categories:

Drag-and-drop refers to using the well-known interaction metaphor to apply an object to its target. The actual action triggered is unambiguously determined by the type of the selected object and the location to where it is applied.

Email refers to sending an email to oneself (or to the intended recipient) with the selected object as an attachment. Email is used as a "carrier" of information, so that it may become available from everywhere the user can access their account.

Synchronization refers to the use of third-party services (e.g., Dropbox, OneDrive, Google Drive) to share the selected object. Similarly to emails, no explicit action is performed by the actual synchronization process, as its purpose is only to make it available from other devices. Further actions are the responsibility of the user.

Bluetooth refers to the use of the Bluetooth protocol to transfer objects between two previously paired devices.

QR detection refers to the process of decoding data stored as a QR-code, through computer vision algorithms. QR codes can typically be decoded into various data types, some of which are able to trigger actions (e.g., loading a web page).

Replication refers to the process of replicating information on another device either through retyping (e.g., a phone number found on a web page) or through redoing the

steps that led to the finding the information in the first place.

Analogous refers to the process of creating an analogous rendition of the information the user is interested into. That is, instead of replicating the exact object on another device as in the previous case, the user creates a new object, analogous to the original. Examples consist in taking a picture of the object, recording a memo or writing down notes.

4.4.3 Proposed application categories

Our analysis of the participants' feedback identified the same three selection categories found for current interaction techniques. The only exception concerns the use of gaze, which now intends the use of eye trackers. Thus, we only report the application categories we have identified:

Cross-device touch drag-and-drop refers to applying an object to another device by means of a cross-device drag-and-drop gesture, initiated by touch. Participants envisioned this technique to be operated when both devices are side by side. This enables users to perform a dragging touch gesture that applies an object selected on one device to an element of the user interface in the destination device (see Fig. 10). A similar technique allowing touch drag-and-drop gestures between desktop PCs and mobile devices has been presented by Simeone et al. [29].

Cross-device direct manipulation refers to being able to operate UIs on different devices as if they were all part of a single logical context and aware of each other. This presumes the existence of a "shared clipboard" between the devices. Thus, objects selected on one device (e.g., by tapping and holding) can be dropped to another device by tapping again on the destination element in the target device's UI (see Fig. 11). A prior work explored the feasibility of a "synchronized clipboard" which allowed copy-

and-paste operations to occur between computers and PDAs belonging to the same network through a client–server architecture [21]. The technique can also be likened to Rekimoto's Pick-and-Drop [24] using touch input instead of pens. A conceptual work explored the use of touch as an input modality [22] to provide a similar interaction metaphor.

Detection represents a method for applying data captured through a camera's video feed. Once the data are selected from another device, the user can apply it to the device he/she is holding by aligning the object through the camera's viewfinder. Users conceptualized this process as automatic. It would be able to distinguish between various data types, triggering the most appropriate action for each (see Fig. 12). DeepShot is an interaction technique that uses a mobile device's camera to capture the state of an application on a desktop screen and migrate it to the mobile device, through the use of computer vision algorithms [6]. Further, the authors also implemented deep posting, a variation of the technique that allows users to push data from a mobile to the desktop.

Remote control refers to using one device to apply an object to another remote device. Similarly, to the "Send To" command in traditional contextual menus, participants envisioned a future version of this metaphor capable of detecting which devices are in close vicinity and of accessing individual applications running on each paired device.

Proximity refers to applying an object either by placing both devices in close proximity to each other or moving the lighter towards the heavier or fixed one. Participants envisioned using the direction the screen faces to determine the action that will be triggered, i.e., facing the screen towards the receiving device to apply an object from the mobile, vice versa to apply an object to the mobile (see Fig. 13). Similar interaction capabilities were presented by
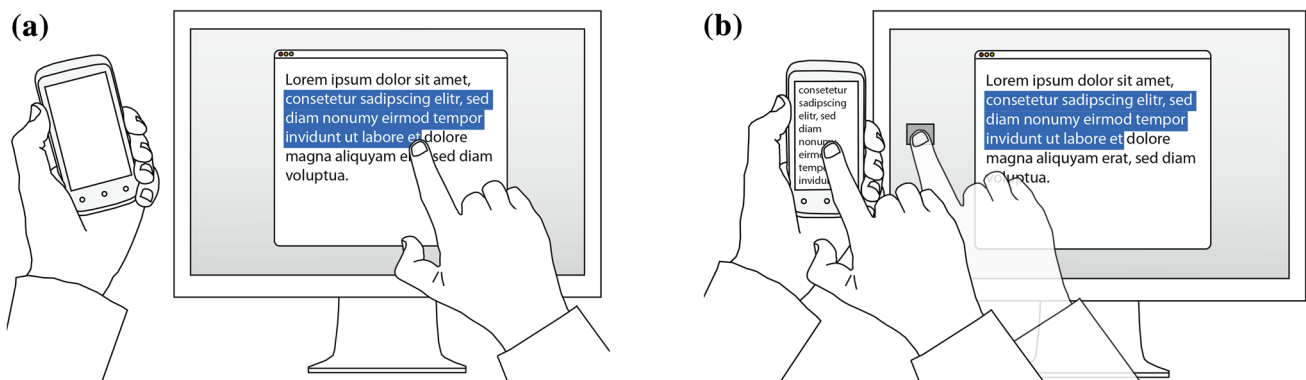

Fig. 10 Cross-device touch drag-and-drop: **a** the user holds the source device next to one of the edges of the target device; he/she then proceeds to select and drag an item towards the edges of the screen and into the target device's screen; upon reaching the intended location on the target device's screen, the user releases their finger to finalize the action (**b**)
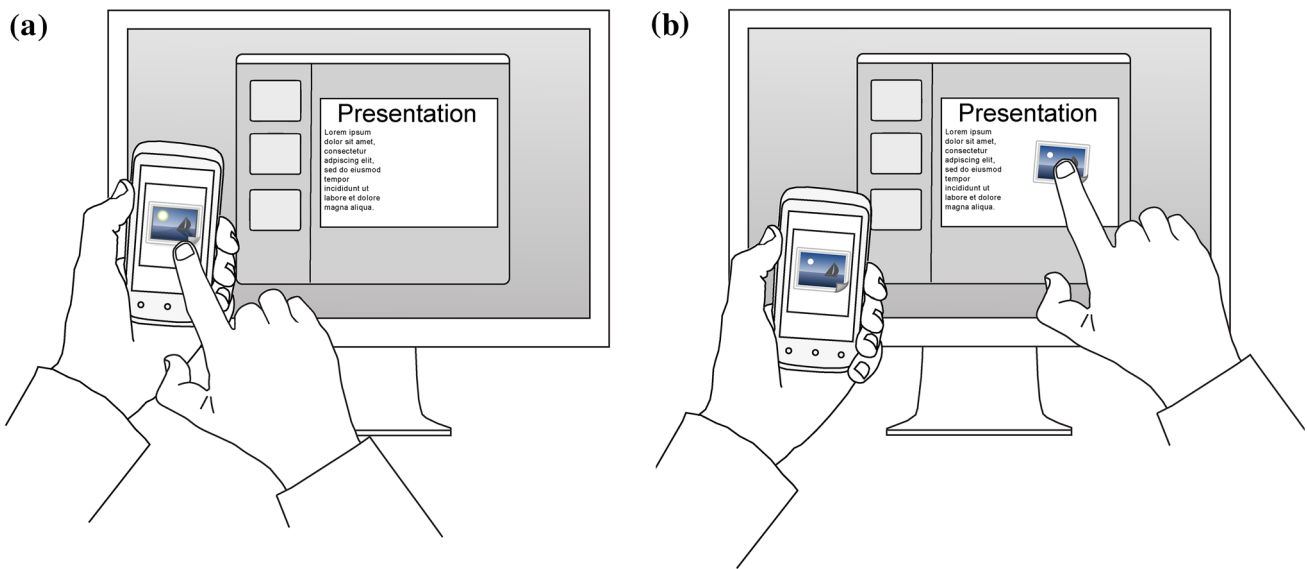
Fig. 11 Cross-device direct manipulation: **a** user taps on the data he/she wishes to transfer from the source device and then taps on the location where it is needed on the target device's screen (**b**)
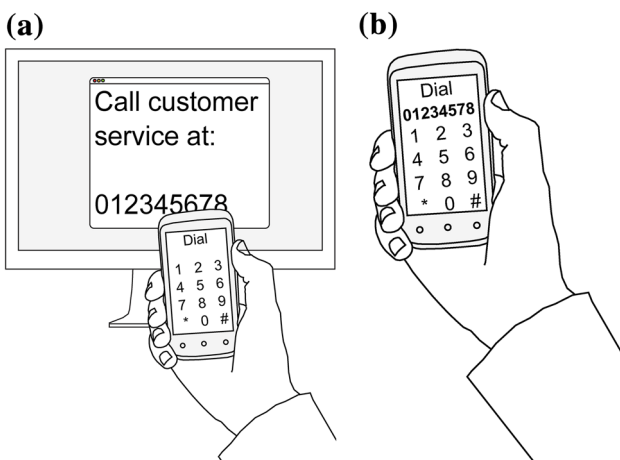

Fig. 12 Detection: a user holds a mobile device over the target device's screen (**a**); the phone recognizes the data focused by its camera and applies it where needed (**b**)

graspable bricks [8] and *PaperWindows* [12]. The former is an example of a tangible tracked prop that can be attached/detached to elements of a projected UI. The latter is an augmented environment that uses paper sheets as tangible input devices. Users can "rub" a sheet on a display to copy its contents. However, the system requires an environment capable of tracking paper sheets so that the UI can be projected over them.

Gesture refers to applying an object through a gesture directed towards the target device. As such, it is entirely performed on or near the source device. The action itself is typically a swipe gesture, although some participants suggested throwing or pointing gestures. Code space provides

similar functionalities allowing users to drag content to and from a shared display to one's mobile [5]. Flicking up or down allows the user to push/pull content.

Gaze refers to the act of applying data to another device by means of gaze. It was suggested by participants concerned about privacy (i.e., when selecting and applying sensitive information). It is intended to work by looking at the location on the target device to apply it. Similar interaction techniques have been recently presented by Turner et al. [31]. These use gaze to acquire an object from an out of reach or large display and place it on the target device (e.g., a tablet or a laptop). Touch (or mouse) hold is used to confirm the selection; releasing triggers its application. However, they were not designed with privacy in mind. Their implementation causes content to become attached to gaze or touch. Our participants' remarked that in this situation, the absence of feedback was desirable.

4.4.4 Frequency of occurrences

We summarized the number of occurrences of selection and application categories in Figs. 14, 15 and 16, respectively. For current techniques, out of 208 suggestions (some participants reported using more than one method in equal measure), 102 (49 %) report that data are conceptually selected through gaze. From Fig. 15, we observe how replication (96 suggestions, 46.2 %) is the most used application technique for extemporary data such as phone numbers, addresses, event info, etc. Creating an analogous of the data (13, 6.3 %), e.g., a photo or a note of information regarding an event, is another option. For media (e.g., pictures and documents), direct manipulation (90,
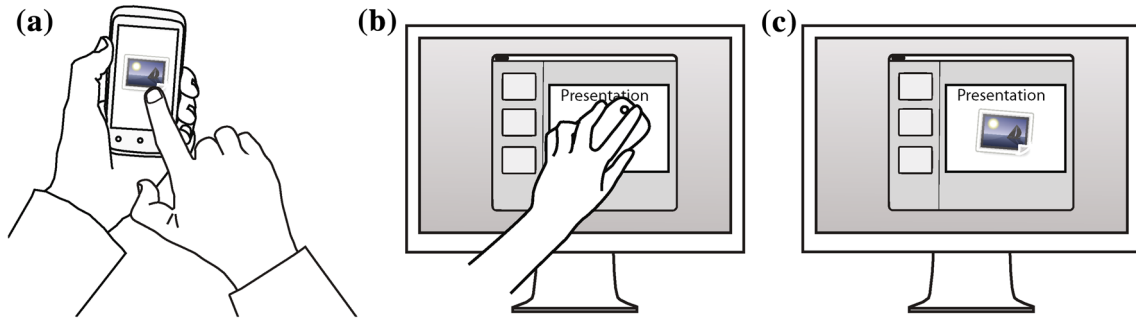
Fig. 13 Proximity: a user selects a picture on her/his mobile device (*left*); then he/she puts the device close to the bigger screen (*middle*) in order to apply it on the desired location (*right*)

Fig. 14 Occurrences of current (**a**) and proposed (**b**) selection categories grouped by the scenario in which they were suggested
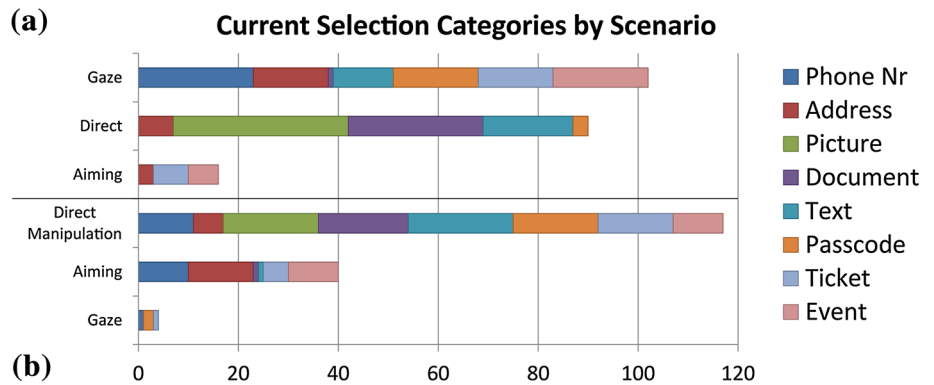

Fig. 15 Occurrences of current application categories grouped by the scenario in which they were suggested
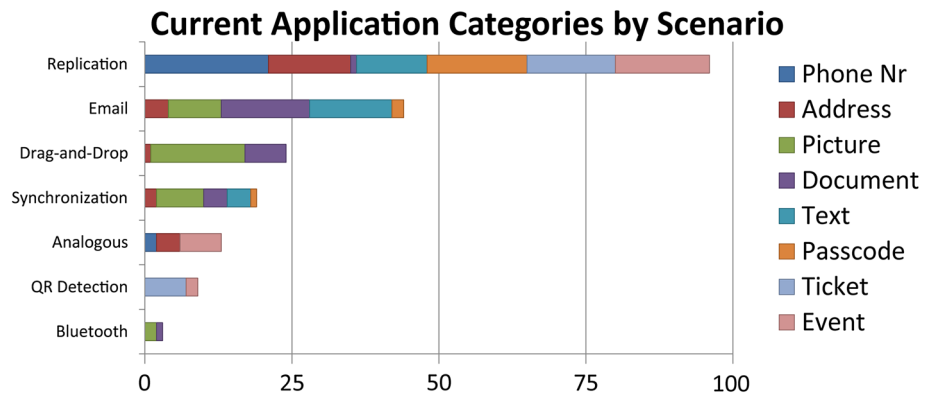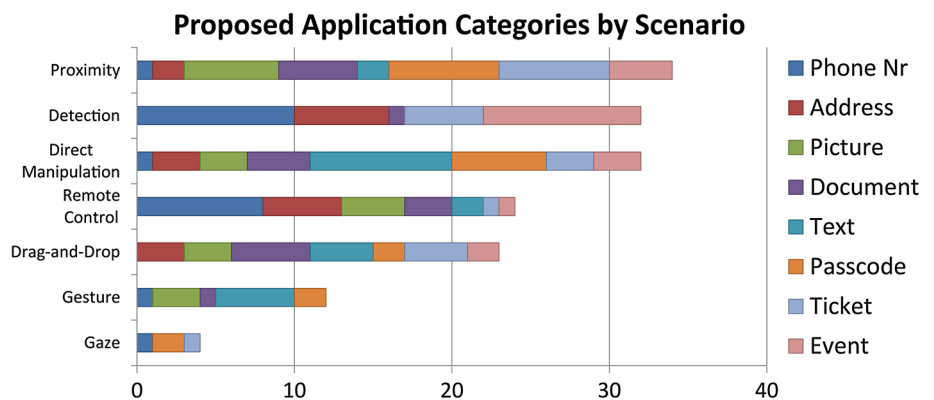

Fig. 16 Occurrences of proposed application categories grouped by the scenario in which they were suggested

43.3 %) is used to select the object (typically through interfaces provided by the operating system), as expected. However, lacking cross-device techniques, users have to resort to different methods. For instance, email attachments (44, 21.2 %), USB pairing (24, 12 %) and cloud-based synchronization (19, 9.1 %). Finally, aiming (9, 4.3 %) is used in those circumstances where a QR-code is provided.

Concerning proposed solutions, Fig. 14 shows that most users wished direct manipulation supported cross-device Select & Apply tasks (116, 72.5 %). Gaze (4, 2.5 %) was suggested as a selection technique supported by tracking devices. Regarding application techniques, using proximity between devices received 34 (21.3 %) suggestions; automatic detection and cross-device 32 (20 %); controlling other devices remotely 24 (15 %); drag-and-drop 22 (13.8 %); performing gestures to apply data to other devices 12 (7.5 %); and using gaze 4 (2.5 %).

5 Second user study: focus group

We wanted to understand in more detail the reasons behind the responses users gave in the previous study: their motivations for the use of one technique instead of another, issues they experience in current practice and interaction properties they find desirable and examples of other Select & Apply scenarios. We organized a second study consisting of focus groups. We expected that being exposed to other participants' experiences and ideas could help foster discussions on the topic.

5.1 Participants, setup and procedure

We interviewed six groups of three participants each from a similar non-technical demographics as in the previous study. Each group was presented with four of the eight scenarios we designed. The order of presentation was counterbalanced, so that each scenario was covered three times. Each session lasted 60 min on average.

These scenarios were again examined from the current and future technology perspective. We used previous knowledge from the first study to draft a set of questions to use as the basis of a semi-scripted interview. Participants were introduced to the study by being asked to think about how they approached each scenario. We then allowed the participants to discuss the topic between themselves, asking for clarifications whenever deemed necessary. Once all avenues were explored, we fell back to the pre-defined set of questions. For instance, regarding current methods: what advantages or issues they experienced with the method they suggested; under which circumstances they could see them using a different approach; if they knew of other apps that provided solutions for the particular scenario; whether they could think of other examples of Select & Apply tasks from their experiences.

Regarding future techniques, we again asked them to think about hypothetical interaction techniques without minding technological limitations. As we did previously, this initial question was used to foster discussions among them, with a pre-defined set of questions to fall back on. For instance, they were asked to act how their proposed solution would work. We used this demonstration to investigate aspects such as the way they held a device and their body posture; which advantages would their approach have over the current status quo and compared to other participants' solutions.

5.2 Results

5.2.1 Choice of method

Participants' answers and their behavior of manipulating data across devices provide insights into their choice of interaction method. The main factors impacting these choices are the relative mobility of the devices (i.e., spatial position and orientation), data characteristics (i.e., containment, if entirely visible) and privacy. These factors impact and constrain how data are selected and applied through the relative movement of the devices. There are two ways in which mobile devices, usually smaller, lighter and private are moved around the larger heavier or public ones: (1) for extending the surface of the larger device's display in both length or height and (2) for extending the depth of larger device. If the former has been previously suggested through stitching [11], the latter finding introduces a category of interaction techniques which involves positioning the mobile device parallel with and in front of the larger device, either in close contact, (i.e., proximity) or at distance, (i.e., detection, direct manipulation).

A useful theoretical lens to explore the choice for these methods is through image schemata. These are representations of specific and repetitive embodied experience of bodily movement through space and manipulation of objects that people develop tacitly from infancy [17, 18]. Image schemata can be classified in different categories relating to space, containment or force and relevant for interpreting our findings are path and container image schemata. Path schemata include a starting and an ending point, contiguous points in between, as well as directionality. Container schemata include an enclosed area delimited through boundaries from the surrounding excluding area, the surface supporting the container and associated actions. Feedback from our second study on the following techniques allowed us to further explore the path and container metaphors by employing the topology of devices in space and the specific types of data they apply to.

Touch drag-and-drop. The mobile device is moved to extend the length or height of the larger device while touching their edges. As a result, the data source and target location are aligned within the plane of the larger display. Movement: the data are applied through the index finger's linear trajectory. Data type: larger, non-sensitive data items such as PDF files scrolling beyond the display of the source device.

Cross-device direct manipulation. The mobile device is moved to extend the width of the larger device while maintaining physical distance between them. As a result, the planes of the two devices become parallel. This tends to occur for larger devices, be them mobile or fixed, in order to reduce the energy expenditure required for manipulating them. For example, mobile devices such as tablets or fixed devices such as public displays require additional physical effort for positioning the former to extend the display of the latter. Movement: the data are applied from the source location to its destination through the index finger's non-linear trajectory. Given that the mobile device is interposed between the data it displays and the larger device, only a non-linear path can be enacted. Data type: larger, non-sensitive data items such as PDF files or paragraphs scrolling beyond the display of the source device, or snippets of private, sensitive visual data, i.e., a PIN code. In the latter case, people seem to prefer to keep the data close to their body, i.e., held inside one's fingertip which acts a virtual clipboard, so that the movement of data from the source location to its destination becomes "embodied".

Detection. This works with camera-based mobile devices moved to extend the depth of the larger device while maintaining physical distance between the two. When the mobile device is the target device, visual alignment is an effective way to select data. Movement: the data are applied through the camera's mobile device, with no need to enact its path from the source location to destination through the finger's trajectory. The only movement required is aligning the data through the mobile device's viewfinder appropriately. Data type: small graphic or textual snippets of non-sensitive data contained within the source display, which do not require scrolling, i.e., phone number, map, e-ticket or event info.

Proximity. The mobile device is moved to extend the depth of the larger device with no physical distance between the two. As they are positioned in parallel touching each other, the Select & Apply interaction could be immediate with no additional physical movement required. Movement: once the mobile device is in the immediate proximity of the larger device, the data are automatically applied, so there is no need to enact its path from the source location to destination through one's finger's trajectory. Data type: small graphic or textual snippets of data contained within the source display which do not require

scrolling. Proximity appears useful for transferring private, sensitive data, i.e., a PIN code, as the body of the mobile device placed on top of the source or destination location within the large public display obstructs its public view. Further, it allows precision in applying the data at the target location. In this case, a smaller mobile device could better serve as pointer to that location. This technique is not used for larger data requiring scrolling, which would be difficult to place precisely within a given target location.

5.2.2 Issues

Participants described several issues they have with current methods. *Replicating* data increase the chance of making mistakes, which can have undesirable consequences: the wrong number might be dialled; a map application might lead us to the wrong place; and our bank account might be locked after failing several attempts to enter the correct PIN. Further, participants noted that websites and applications typically provide a different user experience when used across different devices. Thus, searching for content previously found through a different web layout on another device providing a different web experience is likely to frustrate users. Using methods that were not intended for Select & Apply tasks, e.g., email or synchronization services, make the retrieval of information at a later stage problematic. Indeed, participants mentioned that they do not usually mark an email attachment in any way (e.g., by using a particular subject). Thus, when they need to retrieve that particular information again, it becomes difficult to find an email that is no longer current.

5.2.3 Desirable properties

We identified the aspects shared between techniques suggested by participants of both studies. They reported the *immediacy* of a technique as being the most crucial factor in the decision to adopt a new technique over an established method. Indeed, all proposed techniques provide a direct channel between the selection source and application target, thus avoiding prerequisite or intermediate steps. Another shared aspect is the ability to recognize the *context* in which a Select & Apply task is performed. For instance, detecting if the text applied to another device is just a sentence or an address; if the task involves only personal devices or those owned by others; whether it happens at home or in public. Although existing intra-device Select & Apply techniques do provide some support, current methods used for cross-device tasks are unaware of the content being applied. Further, we observed our participants enact their proposed techniques by using a single hand to interact. All techniques can be operated through a single contact point or gesture (where applicable). Thus, the *simplicity* of

a technique might be an important aspect to factor in the design of new ones.

Privacy was reported as an important concern by the participants we interviewed. They shared the view that as long as the information they send as an attachment or to a synchronization service is not critical, they feel fine using those services. However, when dealing with sensitive information (e.g., business plans, credit card or bank details), participants agreed they would not use them anymore. They suggested that interaction techniques should hide sensible data during the application process, i.e., while applying a PIN code. Participants also indicated that future techniques should be able to operate both in personal environments (i.e., with devices belonging to the same owner) and between devices belonging to different owners. However, users should be able to confirm whether to accept incoming data or opt-out altogether, in order to avoid unsolicited data by third-party sources.

5.2.4 Other scenarios

Participants were asked to describe other scenarios they felt would benefit being supported by Select & Apply techniques. In particular, they described UI migration as a compelling scenario. That is, the ability of migrating workspaces across devices. There are applications that are integrated with cloud-based synchronizations services (i.e., Microsoft Office and OneDrive) providing the ability to access documents on the cloud from different devices. However, participants felt that this does not support extemporary Select & Apply tasks, as it introduces further steps. For instance, the need to select a browser session and apply it to one's smartphone might arise just as the user is about to leave his home or workplace, so that he/she can continue reading once on the bus or tube, or vice versa. Thus, in these situations, the ease and speed with which the application of the information can be completed becomes critical for its adoption.

Other scenarios described by users include the possibility of sharing content such as videos and music to other devices. Current approaches often require a great degree of expertise such as dealing with DRMs, transcoding/encoding formats, etc. Participant #9 expressed the desire of being able to apply not just the video but any metadata associated with it, such as the current position. Similarly, participant #4 indicated sharing music playlists, while participant #16 wished she could be able to share voice recordings easily.

6 Discussion

Our findings advance the understanding of the choices people make for selecting and applying data across devices

through various interaction techniques. These results reiterate the notion that there is no perfect technique for all situations. Instead, we provide insights into which factors might affect the type of technique best suited for the scenarios we designed.

6.1 Physicality

The issues of physicality and efficient effort expenditure strongly impact these choices. We found that mobile devices can be used to extend not only the length of the larger displays, but also their width through both immediate and remote contact. The application of path and container image schemata for further exploring these choices suggest the importance of directionality of transfer from the source to target device and the visibility of the data. Findings suggest that despite its efficiency, detection techniques are less appropriate for sensitive data which people prefer to keep closer to their bodies. They also suggest the value of embodied trajectories for handling sensitive data, i.e., direct manipulation. The containment of the data is another important consideration. Drag-and-drop is preferred for large data whose boundaries exceed the source display and may be difficult to place accurately at target location, especially on smaller device.

This is an interesting outcome, since drag-and-drop is the most costly technique involving the enactment of the physical movement of both the mobile device and data to be applied. If the data are, however, encapsulated in a container, then it could be applied through more efficient techniques such as detection or proximity. Indeed, applying proximity-based techniques on an open PDF file is less appropriate. The icon will work fine, as it acts as a container of data that is easier to place precisely at its target location. For scenarios, where placing the data accurately to its target location matters, using a small mobile device acting as a pointer to the location is particularly useful for this technique. Alternatively, augmenting larger size mobile devices with pointing functionalities could broaden the applicability of proximity-based techniques.

6.2 Differences between current and proposed methods

Current selection methods highlight the fact that participants often need to interact with information that is not represented as a file in the file system, (e.g., documents and pictures). As we have previously explained, phone numbers, addresses and event data are information that frequently arises after browsing the web or is available from both digital (public displays) and non-digital sources (posters). Current support for Select & Apply tasks involving extemporary data does not provide sufficient

advantages for users to stop using replication methods, which come with the issues we have previously described.

Regarding other media, the choice of interaction technique depends on external factors, as highlighted by the second study. For example, how important or private is the information contained in the object might have user prefer using email attachments over synchronization services. If the object is being applied to devices owned by the same user, methods such as an email attachment or drag-and-drop across folders will tend to be used, as opposed to Bluetooth which is more likely to be used between devices having different owners.

Proposed selection methods show that users wish that systems and environments supported interaction capabilities for extemporary data. Further, there is a high degree of fragmentation within each scenario, with different techniques that vary greatly between each other. This highlights specific differences that call for the adoption of techniques able to explicitly address them. For example, participants that would use replication techniques in scenarios #1, #2, #7 and #8 favoured those that did not require serial interaction on both devices such as Detection or Remote. The former appears more suited in situations outside our personal environment, when we happen to find phone numbers or addresses in public. Detection techniques allow us to apply the information without need to touch or interact with its source. On the other hand, techniques based on remote interaction are more suited to personal environments where it is more likely to have access to an interactive display aware of all our devices.

Similarly, the picture sharing scenario showed a high degree of fragmentation for what is arguably one of the most common situations. We believe that the choice is influenced by the amount of control that users perceive to be necessary. Techniques such as direct manipulation, drag-and-drop and proximity allow users to fine-tune the application location. Conversely, techniques based on gestures or remote interaction rely on automatic processes to complete the interaction intent initiated by the user (e.g., extrapolating the target location from the aim of a swiping gesture).

Privacy also impacted the choice of technique in scenario #6. The majority of suggestions favoured techniques do not give feedback about the progress of the task. For example, Select & Apply through gaze or through direct manipulation, was conceptualized as necessitating only of the selection and application location, without any requirements for feedback.

6.3 Implications on the technical feasibility

In *proximity*-based techniques, the main obstacle is identifying the contact point. In "Touch and Interact" [9], a grid of NFC tags was used, while in PhoneTouch [28], a vision-based system infers contact by relating touch events to audio and accelerometer ones. However, all but one participant stated that devices should not touch each other to avoid scratches.

Detection was conceptualized as a tool capable of automatically recognizing data and finding a relevant use for it. Advancements in algorithmic and processing power will be necessary to progress further than the implementation in [6].

Contextual menus are reminiscent of classic desktop interaction paradigm, albeit more advanced. Participants are accustomed to the "Send To" command of contextual menus found in operating systems. An evolution of it would see this command able to recognize nearby devices and also currently running applications. This would allow an intermediary system to establish a direct communication channel between them. Services such as Apple's AirPlay and Microsoft's play to allow users to stream media content on secondary, certified, devices belonging to the same ecosystem, but do not consent users to use it in any other way.

Cross-device direct manipulation requires the existence of a shared clipboard between the two devices (and other devices in the network). From our own observations, we noted that some participants stated that a simple tap is not enough to enable this technique, and it is first needed to invoke a "copy" action by means of a contextual menu. Indeed, a way to disambiguate normal taps from those intended to invoke a cross-device, copy-and-paste action is one of the first technical requirements. A contextual menu would be in line with how users are accustomed to perform copy-and-paste operations both on the desktop and on mobile devices.

Drag-and-drop, similar to how stitching worked, could use the explicit boundary crossing to disambiguate the devices involved in the transfer action. In a personal context, i.e., in a home/work network, the time between exiting the source device's screen and entering the target device's screen would be sufficient for most circumstances such as transferring text paragraphs, phone numbers, contacts, etc. In [29], the technique is implemented by means of either applications that are aware of the presence of other devices or through an intermediary one that allows cross-device communication within sandboxed operating environments.

Finally, *gesture*-based and *remote* interaction techniques are well known in the literature [3, 4, 19, 30]. The major obstacle lies in detecting the implied location aimed by the gesture. If the device itself is used as the pointing device, onboard sensors can be used to determine its aim. Other approaches rely on augmentation of the environment.

7 Conclusion

We often find ourselves in situations where we need to select an object available on one device that we need to apply to another. The first of our two studies revealed that current technological support for these Select & Apply tasks is lacking. Modern operating systems support this class of tasks as long as it happens wholly on one device. When more than one device is involved, we have found that users adopt means that require additional intermediary steps such as email attachments or synchronization-based services. Furthermore, our participants reported that a considerable amount of interactions involve information for which they have to resort to replication methods, due to the lack of technological support.

Our analysis revealed that the overall configuration of devices and their topology, together with associated data characteristics such as containment and privacy are important. In addition, the lens provided by image schemata theory can be particularly useful when designing and developing interaction techniques for acting on data across devices. We believe these findings will help raise awareness of the issues that currently affect Select & Apply tasks and support future designers through the user feedback we have collected and analyzed.

References

1. Ballendat T, Marquardt N, Greenberg S (2010) Proxemic interaction: designing for a proximity and orientation-aware environment. In: ACM international conference on interactive tabletops and surfaces, ITS'10. ACM, New York, NY, USA, pp 121–130
2. Baudisch P, Cutrell E, Robbins D, Czerwinski M, Tandler P, Bederson B, Zierlinger A (2003) Drag-and-pop and drag-and-pick: techniques for accessing remote screen content on touch and pen-operated systems. In: Proceedings of interact'03. IOS Press, pp 57–64
3. Beigl M (1999) Point and click: interaction in smart environments. In: Proceedings of Handheld and Ubiquitous Computing, HUC'99. Springer, London, UK pp 311–313
4. Boring S, Baur D, Butz A, Gustafson S, Baudisch P (2010) Touch projector: mobile interaction through video. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'10. ACM, New York, NY, USA, pp 2287–2296
5. Bragdon A, DeLine R, Hinckley K, Morris MR (2011) Code space: touch + air gesture hybrid interactions for supporting developer meetings. In: Proceedings of interactive tabletops and surfaces, ITS'11. ACM, pp 212–221
6. Chang T-H, Li Y (2011) Deep shot: a framework for migrating tasks across devices using mobile phone cameras. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'11. ACM, pp 2163–2172
7. Dearman D, Pierce JS (2008) It's on my other computer!: computing with multiple devices. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'08. ACM New York, NY, USA, pp 767–776
8. Fitzmaurice GW, Ishii H, Buxton WA (1995) Bricks: laying the foundations for graspable user interfaces. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'95. ACM Press/Addison-Wesley Publishing Co., pp 442–449
9. Hardy R, Rukzio E (2008) Touch and interact: touch-based interaction of mobile phones with displays. In: Proceedings of Mobile HCI'08, pp 245–254
10. Hinckley K (2003) Synchronous gestures for multiple persons and computers. In: Proceedings of the ACM symposium on user interface software and technology, UIST'03. ACM, pp 149–158
11. Hinckley K, Ramos G, Guimbretiere F, Baudisch P, Smith M (2004) Stitching: pen gestures that span multiple displays. In: Proceedings of advanced visual interfaces, AVI'04. ACM, pp 23–31
12. Holman D, Vertegaal R, Altosaar M, Troje N, Johns D (2005) Paper windows: interaction techniques for digital paper. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'05. ACM, New York, NY, USA, pp 591–599
13. Kray C, Nesbitt D, Dawson J, Rohs M (2010) User-defined gestures for connecting mobile phones, public displays, and tabletops. In: Proceedings of mobile HCI'10. ACM, pp 239–248
14. Kujala S (2003) User involvement: a review of the benefits and challenges. Behav Inf Technol 22(1):1–16
15. Kujala S (2008) Effective user involvement in product development by improving the analysis of user needs. Behav Inf Technol 27(6):457–473
16. Kurdyukova E, Redlin M, André E Studying user-defined ipad gestures for interaction in multi-display environment. In: Proceedings of intelligent user interfaces, IUI'12. ACM, pp 93–96
17. Lakoff G, Johnson M (1999) Philosophy in the flesh: the embodied mind and interactive tabletops and surfaces challenge to western thought. Basic books, New York
18. Lakoff G, Turner M (1989) More than cool reason: a field guide to poetic metaphor. University of Chicago Press, Chicago
19. Lucero A, Holopainen J, Jokela T (2011) Pass-them-around: collaborative use of mobile phones for photo sharing. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'11. ACM, pp 1787–1796
20. Marquardt N, Hinckley K, Greenberg S (2012) Cross-device interaction via micro-mobility and f-formations. In: Proceedings of the ACM symposium on user interface software and technology, UIST'12. ACM, pp 13–22
21. Miller RC, Myers BA et al (1999) Synchronizing clipboards of multiple computers. In: Proceedings of the ACM symposium on user interface software and technology, vol 7 of *UIST'99*. pp 65–66
22. Mistry P, Nanayakkara S, Maes P Touch and copy, touch and paste. In: CHI extended Abstract'11. ACM, pp 1095–1098
23. Morris MR, Wobbrock JO, Wilson AD (2010) Understanding users' preferences for surface gestures. In: Proceedings of graphics interface, GI'10. Canadian Information Processing Society, pp 261–268
24. Rekimoto J (1997) Pick-and-drop: a direct manipulation technique for multiple computer environments. In: Proceedings of the ACM symposium on user interface software and technology, UIST'97. ACM, pp 31–39
25. Rekimoto J, Ayatsuka Y, Kohno M, Oba H (2003) Proximal interactions: a direct manipulation technique for wireless networking. In: Proceedings of INTERACT'03. IOS Press, pp 511–518
26. Rekimoto J, Saitoh M (1999) Augmented surfaces: a spatially continuous work space for hybrid computing environments. In:

Proceedings of the SIGCHI conference on human factors in computing systems, CHI'99. ACM, pp 378–385

27. Scharf F, Wolters C, Cassens J, Herczeg M (2013) Cross-device interaction. In: AMBIENT 2013, the 3rd international conference on ambient computing, applications, services and technologies, pp 35–41

28. Schmidt D, Chehimi F, Rukzio E, Gellersen H (2010) Phone-Touch: a technique for direct phone interaction on surfaces. In: Proceedings of the ACM symposium on user interface software and technology, UIST'10. ACM, pp 13–16

29. Simeone AL, Seifert J, Schmidt D, Holleis P, Rukzio E, Gellersen H (2013) A cross-device drag-and-drop technique. In: Proceedings of the 12th international conference on mobile and ubiquitous multimedia, MUM'13. ACM, New York, NY, USA, pp 10:1–10:4

30. Swindells C, Inkpen KM, Dill JC, Tory M (2002) That one there! pointing to establish device identity. In: Proceedings of the ACM symposium on user interface software and technology, UIST'02. ACM, pp 151–160

31. Turner J, Bulling A, Alexander J, Gellersen H (2014) Cross-device gaze-supported point-to-point content transfer. In: Proceedings of eye tracking research and applications symposium, ETRA'14

32. Wobbrock JO, Morris MR, Wilson AD (2009) User-defined gestures for surface computing. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'09. ACM, pp 1083–1092