Affordances for Manipulation of Physical versus Digital Media on Interactive Surfaces

Lucia Terrenghi[1], David Kirk[2], Abigail Sellen[3], Shahram Izadi[3]

[1]LMU University of Munich
Amalienstrasse 17, D-80333
Munich
lucia.terrenghi@ifi.lmu.de

[2]Mixed Reality Lab
University of Nottingham,
UK, NG8 1BB
dsk@cs.nott.ac.uk

[3]Microsoft Research
7 JJ Thomson Ave.
Cambridge, UK CB3 0FB
{asellen; shahrami}@microsoft.com

ABSTRACT

This work presents the results of a comparative study in which we investigate the ways manipulation of physical versus digital media are fundamentally different from one another. Participants carried out both a puzzle task and a photo sorting task in two different modes: in a physical 3-dimensional space and on a multi-touch, interactive tabletop in which the digital items resembled their physical counterparts in terms of appearance and behavior. By observing the interaction behaviors of 12 participants, we explore the main differences and discuss what this means for designing interactive surfaces which use aspects of the physical world as a design resource.

Author Keywords

Affordances, Digital, Physical, Manipulation, Interactive Surfaces, Tabletop, Interface Design.

ACM Classification Keywords

H5.2. User Interfaces

INTRODUCTION

One of the most recent trends emanating from the ubiquitous computing paradigm is the rise in prominence of "surface computing" [27, 42]. By this we mean the use of new technologies and techniques to create interactive surfaces in the spaces we inhabit (such as walls, floors and ceilings) and using the artefacts that permeate our everyday lives (such as tables, mirrors, whiteboards and doors [36]). Partly this is driven by new technologies and techniques. Here we can point to new display technologies that afford both input and output at the same point of interaction [8, 28], or advanced computer vision techniques in combination with projection onto surfaces that make it possible to recognize real objects, hand gestures and body movements, for example [29, 42].

In tandem with these technological advances, more attention to the development of interfaces for wall and tabletop displays has driven a number of new and compelling applications in this area (for a review see [7]). Most make heavy use of physical metaphors as the basis for interaction, the increased size of display surfaces making it possible to represent virtual objects in a life-size way. Most also recognize the power of both bimanual (and even multi-touch) input to support gesture in order to perform integrated natural actions on digital objects [20, 25, 30, 43]. And many combine the manipulation of physical artefacts with digital interaction, as seen in many of the "tangible computing" ideas that have emerged over the years [12, 15, 18, 40, 41]. While the seeds of these ideas were sown many years ago, they are now becoming more widespread and diverse as the ideas and technologies mature.

Given the emerging popularity of surface computing and the new interaction paradigms they make use of, it is a good time to examine more deeply what specific aspects of the physical world and physical interaction are being drawn upon as a resource in their design (whether this be consciously or not). For example, to elaborate on some we have already mentioned and to name a few more, we can point to:

- The use of the physical metaphor in the way objects and actions on those objects are graphically represented. For instance, the desktop metaphor can be interpreted in various ways [see 1, 3, 9, 35].
- The use of spatially distributed input (such as bimanual input) to interact with virtual objects. Early work on this is exemplified by Bier et al.'s ToolGlass technique [4]. More recently, many tabletop applications make use of multi-touch input to manipulate objects [20, 43].
- Continuity of action in input (as distinct from discrete actions or gestures). For a good discussion of why this is important, see Buxton's paper on "chunking and phrasing" [5].
- A direct mapping between input and output so that an action produces feedback at the point where the input is sensed, as typically seen in pen-based interfaces such as [1, 8].
- A 3D space of manipulation making new kinds of actions and feedback from those actions possible: this

is also the basis of much of the tangible computing work [12, 18, 41].

- Rich multimodal feedback, not limited to visual and audio feedback, such as is possible in the physical world. Hinckley's work on using physical "props" for input is a good example of this work [15].

Although existing work builds on many of these different aspects, there has been less theoretical research that looks at these issues more deeply in the context of today's new and emerging interfaces. In other words, we almost take for granted that mimicking aspects of the physical world is the best way forward in designing ubiquitous computing interfaces. But, what aspects of the physical world should we be concerned with in the interface design of digital media? How do different aspects affect people's mental models and behavior in interaction? To what extent can emulating the physical world result in a behavior similar to that exhibited in the physical world (given that this is desirable)? We believe that looking more systematically at the relationship between different aspects of physicality and interactional patterns can help guide design decisions about how and to what extent we apply aspects of physical interaction to digital interface design.

In the following, we consider how aspects of physicality have been addressed in related work. We then describe a study in which we compare the manipulation of physical objects with those on a digital tabletop, which emulates analagous physical tasks in a number of important dimensions. We report on how interaction is different in the two contexts, and discuss how these results can inform interface design for future interactive surfaces.

RELATED WORK

The use of metaphors for user interface design has been largely discussed in the literature [e.g. 6, 10, 22], its most familiar example being the graphical user interface based on the "desktop metaphor" [35]. In the desktop metaphor, many elements of the interface are modeled on artefacts (e.g. wastebasket, folders, buttons) and behaviors (e.g. direct manipulation [17] and paper-based interaction [9]) from the physical world. Ark et al. [3] investigate the effects of 3 dimensional graphical representations for the desktop metaphor. Recently a more "physical" and realistic representation of the desktop environment has been suggested in [1] to afford pen-based interaction on tablet PCs, building on previous work on pile metaphors [24].

As computing moves beyond the desktop and becomes more integrated in our physical environment, the work on tangible user interfaces (or TUIs) has provided different ways of integrating physicality in the interaction with digital media. Beginning with early work by Fitzmaurice, Ishii, Buxton, and others [11, 18], there have been many instantiations and variations of the TUI paradigm [e.g. 15, 40]. Fishkin [11] provides a useful taxonomy for the analysis of tangible interfaces based on the dimensions of "metaphor" and "embodiment".

More general research on the affordances of physical artefacts has examined how these have been used in product and interaction design [26], and also how they support communication, organization and collaboration [33, 37]. Such approaches provide an understanding of how people deal with physical artefacts to guide interaction design.

With the emergence of multi-touch and multi-user interactive displays [8, 28, 29, 42], similar approaches have looked at the ways in which people work and organize themselves around surfaces [19, 21, 32]. These investigations have focused on how people place artefacts on surfaces, exchange and re-orient documents, communicate with each other, and so forth.

A different way to look at the interaction design issues raised by interactive surfaces is to observe how people interact with digital media while using applications for multi-touch large displays [30, 34]. This work inspires the design of multi-fingered [20, 28, 38, 43], and multi-user [25, 30] vocabularies of gestures. Many of the applications in this area focus on the manipulation of digital images on multi-touch displays which mimic physical interaction to different extents [2, 16, 39].

Overall, the work discussed above has inspired interface and interaction design in essentially three different ways: i) by providing insights into how people manipulate physical artefacts and act in the physical world; ii) by suggesting some examples of how to integrate aspects of physical manipulation in the design of interaction techniques for digital media; and iii) by assessing how people interact with digital media in some applications for interactive surfaces which use physical metaphors.

However, from our perspective, some of the benefits of integrating aspects of physical interaction in the design of digital media have been taken for granted and not systematically discussed. Uderkoffler and Ishii [40] assert that the "proposition of giving additional meaning and animate life to ordinary inert objects is a cognitively powerful intriguing one". Scott et al. [31] state: "Understanding the natural interaction practices that people use during tabletop collaboration with traditional media (e.g. pen and paper) can help to address these issues. Interfaces that are modeled on these practices will have the additional advantage of supporting the interaction skills people have developed over years of collaborating at traditional tables". Even though we generally share this point of view, we feel there are still many open questions: For different kinds of interactional experiences in the digital world, what specific aspects of the physical world are mimicked? What are the consequences of this for people's behaviors and expectations about a given system? What are the fundamental differences between interaction in a 2-dimensional world and a 3-dimensional world? This requires understanding not only people's expectations and mental models about digital versus physical media, but also an understanding of the associated affordances for interaction in these different situations.

METHODOLOGY

Design of the Study

Wanting to understand at a deeper level how interactions with digital and physical objects might differ, we designed an experimental comparison of interactions in the two modalities for the same tasks. To ground this understanding we felt it important that this should be performed during the completion of common tasks in which the digital artefacts had been mapped as closely as possible (in terms of appearance and interactive nature) to the physical objects they represented. In this case, we designed digital tasks on an interactive tabletop deliberately modeled on the physical tasks which shared the following features: They used a physical metaphor, presenting the objects in the digital world in the same way (same physical size and high resolution) as their physical counterparts. Input was bimanual and multi-touch, with a direct mapping between input and output. Gestural actions similar to the physical world for moving objects on the surface were used: thus, only rotation and translation of items was possible in the digital mode, and multiple items could be simultaneously manipulated, so as to mimic the manipulation of paper-based objects on a plane.

The key differences between the physical and digital conditions were the lack of the ability to manipulate objects in 3-dimensional space in the digital condition, and the corresponding lack of multimodal feedback (through touch, for example) that accompanied this. This design enabled us to focus our analysis on how these differences affected the nature of interaction in comparison to the physical task.

We used a design in which participants engaged in both a puzzle and a photo sorting task. The two tasks were chosen to explore a diversity of potential interactions with artefacts but, at the same time, each being relatively common tasks that might be performed in the future with an interactive tabletop interface. To further increase the validity of the study, the photos used for the sorting task were provided by the participants themselves (providing their most recent unsorted photos).

The tasks chosen facilitated two forms of analysis. First, it allows for basic measurements of performance in order to map out broad differences in physical and digital interactions. This includes time to complete the task as well as observations about the form of interaction in the digital versus physical tasks (for example, one versus two-handed interaction) which can be categorically coded and statistically compared. Second, a deeper, potentially more informative qualitative analysis of behaviours at the interface could also be derived from the video record, this latter analysis allowing us to interpret our observations in grounded instances of interaction with digital and physical artefacts. This combination of both qualitative and quantitative analysis is consistent with the exploratory approach we adopted in this study.

Technical Set-up

For the study we used the prototype of a projection- and camera-based interactive tabletop. Such a system provides a projection area of 62 x 43 centimeters (at a resolution of 1024 by 768 pixels), capable of sensing when multiple fingers or hands are placed on the surface. The physical dimensions of the table are 77 x 92 x 69 centimeters. The puzzle and photo sorting applications allow a user to move and rotate an object (a puzzle piece or a photo) in 2-dimensional space on screen. To translate an object on the screen, the user specifies one or more contact points (these can be hand, finger or multiple fingers/hands) on the object, and moves these in a particular direction. To rotate an object, the user specifies one pivot contact point on the object that determines the centre of the rotation and another relative contact point on the object to specify the angle of the rotation (i.e., similar to the "two-point rotation and translation mechanism" as described in [14]. The choice of these techniques was deliberate in order that the digital objects in our study could be translated and rotated in a fashion similar to their paper-based physical counterparts, using one or two hands, and multiple fingers,.

Participants and Procedure

Our study participants were 12 adult volunteers (6 female, 6 male), from both technical and non-technical backgrounds, all right-handed, all with normal or corrected vision, and all of whom had little or no prior experience of direct manipulation tabletop interfaces.

Prior to the study, participants provided 80 of their most recent digital photographs (in digital format). These were randomly split into two groups for the photo sorting task (one to be printed and one to be accessed digitally). At the beginning of the session, each participant was given an explanation of the nature of the tasks in which they would be engaged, and were then introduced to the interactive tabletop surface.

In the first stage, each participant completed two 25 pieces puzzles: one digital and one physical (the puzzle pieces were previously disarrayed on the tabletop by the examiner, they were approximately 2 inches square and were matched as closely as possible for size in the digital version). A picture of the completed puzzle was attached to the wall in front of the tabletop for reference. Prior to the digital trial, a demonstration of the interactive tabletop was provided and participants were given 5 minutes to practice, interacting with and manipulating digital shapes. During the physical task, the interactive surface was covered over with a black board and the puzzle was assembled on top. Participants were told that they must complete the puzzle as quickly and as accurately as possible, but no time limit was given. The order in which digital and physical trials occurred was counterbalanced across participants (along with the picture used for each puzzle).

After both trials of the puzzle task had been performed, participants completed a questionnaire asking which modality they had preferred and exploring their frustrations/difficulties with each method using a series of Likert scales.

The second stage of the study involved a sorting task in which the participants were given their photos and asked to sort them into 3 groups: those photos they would probably discard, those photos they would like to keep but not share and those photos they would like to keep *and* share with others. This task was performed in two trials, one with digital photos and one with physical photos (40 in each group). Order of trials was counterbalanced, and again the surface of the digital tabletop was covered for use in the physical photo condition. The digital photos were sized to be as similar as possible to the physical photos.

After the second trial, a final questionnaire was administered, asking about participants' satisfaction and frustrations when completing the sorting task in the two modalities. All trials were video recorded for subsequent analysis.

RESULTS

We base the majority of our results on a descriptive video analysis, supplemented by a statistical analysis of both the videos and the answers to the questionnaires.

Quantitative Results

The first stage of the analysis compared the relative amounts of time spent on each task (puzzle and photo sorting) in each of the two interaction modalities (digital and physical) (see Figure 1).

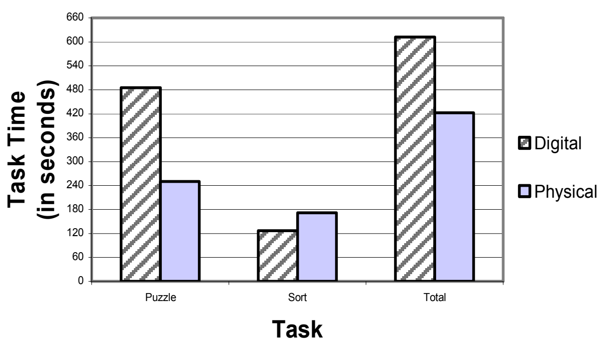

Figure 1. Time to complete each task for digital and physical conditions, for both tasks, and in total.

The puzzle task completion time and the total of the two tasks was longer in the digital than in the physical condition (within-subjects test) ($t(11) = 3.72$, $p< 0.01$ and $t(11) = 2.95$, $p\leq 0.01$, respectively), but was not significantly different for the sorting task. It is perhaps understandable that the digital puzzle task took longer. This may have been due to the difficulty of manipulating the smaller pieces on the interactive surface. Certainly when asked, 11/12 participants felt the digital puzzle had taken longer, 8/12 had enjoyed doing the physical puzzle more than the digital puzzle and 11/12 found the physical puzzle easier than the digital one. When asked to rate their experience (on a Likert scale of 1=very relaxed to 5=very frustrated) participants rated their experience of the digital puzzle as more frustrating than the physical puzzle (average score for physical puzzle 1.67, for digital puzzle 3.08). When asked about the sorting task, participants were equally split over which method was easier, but they were confident that physical photo sorting was more enjoyable (10/12 participants). Again, for the sorting task there was more frustration when using the digital tabletop than the physical one (2.33 versus 1.25 respectively) although overall levels of frustration were lower for the digital sorting than for the digital puzzle task.

Such statistics can only shed a limited amount of light on the differences between the physical and digital tasks. To understand further what constituted these differences, the method of interaction with the task artefacts was observed and coded. One aspect which immediately attracted our attention was the degree to which one versus two hands was used, and the nature of these interactions. We noticed that one-handed interaction was characterized by periods spent using only one hand whilst the other arm was used to support the weight of the body over the table (perhaps similar in principle to our common use of the mouse in GUI interactions). For bimanual interaction, both hands were active in the space, either being used conjunctively in largely symmetric actions (i.e., the index fingers of two hands being used collaboratively to move a single artefact) or being used for differing elements of task action in asymmetric actions (such as one hand moving a piece a large distance whilst the other hand makes fine adjustments at destination). For a more detailed discussion of the differences between symmetric and asymmetric bimanual interaction, see [13]. From video records of both the puzzle and sorting tasks a log was made of the time spent in one-handed and bimanual interaction (see Figure 2).

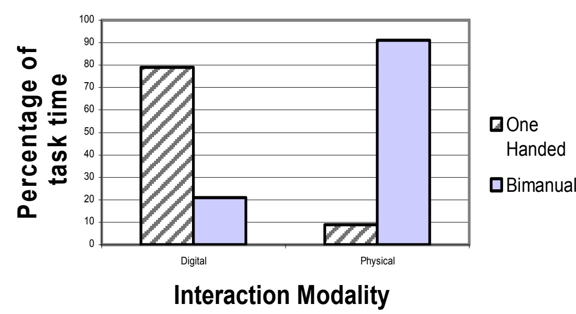

Figure 2. Percentage of task time (both tasks) engaged in either bimanual or one-handed interaction.

Some consistent patterns were observed for both tasks. In both of the digital tasks, there was a predominance of one-handed interaction despite the fact that two-handed interaction was possible, and its use was shown in training. In the equivalent physical tasks, bimanual interaction was much more prevalent. This difference is statistically significant ($t(11) = 8.49$, $p<0.001$). In terms of numbers of participants, in the digital tasks, 9 of 12 participants used one-handed interaction more than bimanual interaction. In the physical tasks, all 12 of the participants used bimanual interaction more than one-handed, and of these, 7 out of 12 of them used *only* bimanual interaction.

Furthermore, there were important differences in the nature of two-handed interaction in the two conditions: In the digital tasks, we have already remarked that 9/12 participants relied mainly on one-handed interaction, but of the 3 who used bimanual manipulation more, 2 of them used proportionally more symmetric actions than asymmetric. In fact, for all 12 participants, *symmetric* bimanual action was more prevalent in the digital condition than *asymmetric* action. By contrast, in only *one* instance in the physical tasks did a participant engage in a bimanual symmetric action – the bimanual use of hands was otherwise almost entirely asymmetric in nature. In summary, the nature of manipulation in the digital and physical tasks was qualitatively very different: although asymmetric bimanual interaction was possible in the digital tasks, participants adopted very different methods of manipulation. We will elaborate on this in the next section.

Qualitative Observations

The video analysis provided rich material for observing salient differences in the way each participant engaged in both physical and digital interactions, and highlighted differing strategies for task completion between participants. Below, we report our observations on the basis of how the tasks unfolded over time. Despite the fact that the puzzle and sorting tasks are obviously different, they have several common aspects, both giving rise to spatial-temporal patterns of interaction which we have segmented as follows:

- General posture and patterns of manipulation
- Getting an overview of the task
- Focusing on a single item
- Comparing multiple items
- Holding items in “stand-by”
- Creating spatial structures.

General posture and patterns of manipulation

First of all we observed a remarkable difference in the postures adopted by the participants in the physical versus the digital tasks. In the physical modality, participants put their forearms on the table from the very beginning of the interaction and generally used both hands to interact with the items in 3 dimensions as well as 2 dimensions (e.g. sliding the puzzle tiles on the board).

Figure 3. General posture adopted in the physical tasks (left) and digital (right).

In the digital modality, all 12 participants (all right-handed) started both tasks by resting their left arm, and often their left elbows as well, on the side of the tabletop and moving the items with the right hand only (as shown in Figure 3).

We also noticed a difference in the relationship between the dominant (right) and non-dominant (left) hands in the two domains. According to Guiard’s description of asymmetric bimanual interaction [13], in the physical realm the motion of the dominant hand occurs relative to the motion of the non-dominant hand, the non-dominant hand acting as the frame of reference for the dominant one. Furthermore, the motion of the dominant hand tends to happen later in the course of bimanual action, and to be finer-grained than that of the non-dominant hand.

Such spatial-temporal patterns were partly confirmed by our analysis of the physical interaction. However, we noticed some differences between the physical puzzle and the physical sorting tasks in this regard:

At the beginning of the physical puzzle, the dominant and non-dominant hands were alternatively used depending on the location, rotation and predisposed position of the puzzle tiles disarrayed on the board. Thus, there was little dependency of the dominant hand on the non-dominant hand, the actions of the two hands being more dependent on the location of the artefacts (i.e., the left hand picked up pieces lying on the left side of the board, and the right hand on the right side). As the puzzle neared completion, participants tended to use their dominant hand to sequentially pick up the missing pieces from the board and place them in the gaps of the puzzle picture. The non-dominant hand was used to secure the tiles already arrayed, creating a physical constraint against which the pieces had to fit (similar to Guiard’s “frame of reference” [13]).

In the physical sorting task, the photos were initially piled in a stack instead of being disarrayed on a 2-dimensional plane. Eight participants held the pile with one hand in the air (4 subjects did so with the right hand and 4 with the left hand): with a slight movement of the thumb of the holding hand they shifted the photo on the top so as to serve the other hand, which then picked it up and placed it on the board. The interaction was therefore asymmetric, with one hand passing the artefact to the other one.

The remaining 4 participants left the pile on the table during the sorting task (see Figure 4, left hand image, for example). Thus, they did not need one hand for holding the pile and passing the photo to the other hand for selecting and placing. Instead, in this case, one hand (left or right) would select the photo from the pile which was rested on the table, both hands would hold the photo for a while, and then the left hand would move the photo to the left side, and the right hand would move the photo to the right or to the top/centre area of the board depending on the spatial layout of the sorted piles on the plane. Only one participant among the ones who kept the pile on the table performed the task with just the right hand and kept the left arm rested on the border of the table for the whole duration of the task.

In the physical puzzle task, rotation of the pieces was mostly allocated to the dominant hand and happened more in the 2-dimensional than in the 3-dimensional space, so as to exploit the reference and proximity to other tiles on the board. In the physical sorting task, the rotation of photos only happened in the air, before they were placed onto a pile on the tabletop: thus, the friction of the photo on the board or on other photos of sorted piles was avoided.

Figure 4. Two-handedness.

In the digital modality, interaction in 2-dimensional space resulted in a different distribution of actions between hands. The range of actions we observed in the physical tasks (e.g. holding a puzzle tile with one hand while the other hand hits the edges to align the pieces; holding a pile in the air with one hand and placing the photo with the other hand) were not possible.

Even though the *moving*, *placing* and *rotating* actions were mostly right-handed actions in the digital modality, some exceptions occurred. Sometimes the left hand was used for for moving items to a pile on the left side (i.e., similar to the physical interaction shown in Figure 4, left image.). Sometimes both hands were used to slowly translate an object from one location to another or to rotate it by symmetrically using two hands, usually with only the index fingers (for example see Figure 4, right image). Coarser and faster symmetric movements of both hands (and multiple fingers) were sometimes used to simultaneously move multiple objects.

To move one or more items from one location to another in the digital setting, participants dragged the items "passing over" other digital ones displayed on the interactive table. In the physical interaction however, participants picked up the pieces from the tabletop and dropped them in a new location; they rarely slid them along the tabletop, and only for small distances, as they would have bumped into other objects lying on the board. This affordance of the *digital* medium seemed very strong, all participants observably at times moving digital items whilst intersecting others.

Focusing on a single item

Before deciding where to place a puzzle piece or a photo, participants first focused on each item in isolation. Here we saw a clear spatial-temporal pattern emerge: In the physical setting, participants often brought each item closer to their eyes, thus facilitating focusing while spatially isolating the item from the rest of the array of items (see Figure 5, left image). They then placed the item on the tabletop in relationship to existing structures (i.e., puzzle or piles of photos).

Figure 5. Focusing on an item.

In the digital case, a pattern also emerges: The 2-dimensional space of the digital setting forced participants to move their whole body closer to the item to get a more focused view. In the sorting task, for example, they dragged the picture from the unsorted digital pile to a blank region first (often towards their body) to visually isolate the item, focused on the picture while keeping their fingers on it, and then dragged it to a virtual pile (see Figure 5, right image).

Comparing items

In the puzzle and sorting tasks, comparison of items was seen in assessing the relationship between items (both spatial and semantic), often prior to making a decision about those items.

The physical setting affords bringing multiple items closer to each other both in the air and on the table surface. In the physical puzzle, comparison happened much more on the tabletop, as proximity on the plane was essential for assessing whether two pieces could fit together. In the physical sorting task, comparison happened both by holding items off the surface as well as manipulating them on the table surface. In the former case, multiple items were compared in isolation from the rest of the visual landscape before being placed in different categories on the tabletop. The left and right hand could alternatively bring one photo or another one closer to the user's eyes (see Figure 6, left image). More than 2 photos were sometimes kept close to each other in the air using multiple fingers (see Figure 6, right image). Or they were held close to each other and at the same time close to the visual landscape (although with a different visual angle) when the participant moved their forearms toward the table. Items were also compared on the tabletop by placing them close to each other on the board. In this case it was interesting to notice that they were touched most of the time, despite the fact there was no obvious need to hold them (see Figure 7, left image).

Figure 6. Comparing items in 3-dimensional space.

Figure 7. Comparing items on the physical (left) and digital (right) tabletops.

In the digital tasks, the 2-dimensional space of the interactive table meant that arranging multiple pictures close to each other happened within the plane. The isolation of multiple items from the rest of the visual landscape was constrained by the borders of the real estate of the display. Thus, when the display was already cluttered, participants needed to overlap the pictures, which meant it was sometimes difficult to isolate them visually (see Figure 6). Also in this case, fingers were often kept on the pictures during comparison.

Getting an overview of the task

In both tasks, participants could be seen engaged in actions which helped them gain an overview of the number of items that needed to be dealt with in each task.

In the puzzle task, the pieces were equally disarrayed on the 2-dimensional plane in the two modalities, thus, the subjects could simply gain an overview of how many pieces needed to be arranged by looking at the displays. In the photo sorting task however, participants were provided with a pile of pictures so only the top photo was visible. The two modalities afforded different strategies for gaining an overview of quantity and content.

In the physical setting participants often lifted the pile in the air,, hitting its edges perpendicularly on the board or tapping the sides with one hand: the weight and physical thickness of the pile conveying an approximation of quantity. For some participants, visual feedback appeared to be sufficient to convey approximate quantity, this information was suggested by the physical depth of the pile**.** In these cases, as we saw earlier for 4 of the 12 participants, the pile was placed on the table and the photos were sequentially picked up and placed elsewhere with one hand.

For some participants, previewing the *content* of the pile appeared to be important: for 3 of the participants, the affordances of the tabletop were more extensively exploited to spread the photos out. In this case, they did not create an ordered spatial structure from the beginning of the task**,** going sequentially through the pile. Rather, they displayed the photos on the table first, partly coping with the geometric limitations of the board by keeping some photos in their hands (Figure 8, left image). In this way they could simultaneously view and visually compare multiple photos before proceeding with the clustering task. This type of interaction seemed more exploratory than goal-driven.

Figure 8. Getting an overview of the content.

The unsorted pile in the interactive table did not afford an estimation of quantity in the same way because of its 2-dimensional appearance. Most of the participants (9) coped with such an issue by sequentially dragging one picture after the other one out of the pile and creating spatial clusters progressively, until the end of the pile was reached. Other participants developed some alternative techniques for gaining a preview of the size and content of the pile. In Figure 8 (right image), for example, the participant first "unfolded" the whole pile with symmetric, coarse-grained and rapid movements of both hands. Another subject first unfolded the whole pile by sequentially dragging one picture after the other one from the unsorted pile to another location, and started creating spatial structures afterwards.

Holding items in "stand-by"

During the decision making process regarding the placement of the single puzzle tiles/photos in the different grid cells/categories, we saw that participants tended to keep some items in a "stand-by" state. Thus, they postponed taking a decision about the grouping or placement of an item until later in the task. In the physical tasks, different strategies were recognizable. In Figure 9 (left image), for example, the participant held the same puzzle tile in her left hand for some time, while she moved another item on the table with the right hand. She was not really looking at the piece in the left hand, although this provided a physical reminder of an item still to be placed. In the sorting task, some participants happened to keep several photos between their fingers, while picking up another one from the unsorted pile. The photos in the holding hand were therefore not sorted right away, but rather in a second stage. In other cases, the participants placed one photo in a blank region of the board, often at the periphery and away from existing piles, thus creating a visual cue.

Figure 9. Holding items in "stand-by" mode.

The interactive table supported this stand-by mode in 2 dimensions only. Thus, participants dragged an item to a blank region of the display to create a visual cue, to place or sort the item later on. There were also some examples (e.g. Figure 9, right image) in which one hand was kept as kind of place-holder on the items for some time, while looking at other items (similarly to how the left hand is used in the physical modality in Figure 9, left image).

Creating spatial structures

In both the puzzle and sorting tasks, participants were asked to spatially structure their pictures/photos.

To complete the puzzle, it was necessary to build temporary structures as pieces (and larger parts of puzzle) were aligned. Differences in how this was achieved were observed between the modalities. In the physical puzzle, multiple pieces that were already aligned could be shifted en masse because of the physical constraints that the tiles created for one another. The use of two hands facilitated such an action. Digitally, it was harder to simultaneously drag multiple pieces whilst preserving alignment.

In the sorting task, there was less constraint on the resulting spatial structure (constituent piles), however, observation revealed a strong predominance of spatial order in the way people organized their sorted piles. Here we found that 9 of the participants tended to arrange the "keep and share", "keep", and "discard" piles in a horizontal line across the tabletop, 6 of them in that order from left to right, and 3 in the opposite order. There was little difference in this between physical and digital modalities, and a remarkable correspondence was noticed between the final spatial layout in the physical and digital settings for almost every participant. Indeed, even those who in the physical setting did not arrange the piles in a row, but rather in a virtual triangle (2 of them) used a similar spatial organization on the interactive table (see Figure 10). Another participant created just two piles (only keep and share, and discard) in both settings, again laid out in approximately the same way.

Figure10. Using the same spatial layout in both conditions.

SUMMARY AND IMPLICATIONS

Our attempt to simulate as much as possible many aspects of the physical world in tabletop interaction (including for example multi-touch input, use of gestures, adherence to a physical metaphor, sizes of the artefacts, and so on) allowed us to elicit and reflect on the fundamental differences between interacting with tangible objects in 3-dimensional space as compared to digital objects in 2-dimensional space.

We observed that, despite the different spatial affordances of the physical and digital modalities, there were fundamental elements of both tasks that they had in common. These include, for example, the need to get an overview of content and quantity, the need to compare objects, focus on particular ones, hold some objects distinct from others, and keep some in a "stand-by" mode. However, the means and strategies by which these subtasks were accomplished were intrinsically different across conditions.

First, one of the most striking findings was that although the digital tabletop interaction was designed to support the kind of bimanual interaction used in the physical world, we in fact saw predominantly one-handed interaction. Further, any bimanual interaction we did observe in the digital domain was largely symmetric in nature, which is quite different from the kind of asymmetric bimanual interaction typical of physical manipulation. In fact, interaction in the digital realm appeared to some extent almost "mouse-like" in terms of the posture participants adopted, and in the way they chose to deal with digital objects.

Second, the tasks in the physical realm highlighted the many ways in which two hands work together both in 3-dimensional space and with tactile objects in terms of the non-dominant hand providing a frame of reference for the actions of the dominant hand. A good example here is the non-dominant hand holding a pile of photos, while the dominant hand selects and places the photos. Such interactions were much rarer in the digital case, suggesting a lack of tangibility and 3-dimensional space undermines the natural allocation of hands to these asymmetric roles.

Third, the presence or absence of the physicality of individual objects affected the way in which users manipulated the artefacts. The thickness of the physical puzzle tiles, for example, provided physical constraints against which other pieces could be laid. On the other hand, the lack of thickness of the digital puzzle tiles, suggested that one single piece could simply be dragged from a location to another one across the display passing over other pieces. Physicality also allowed for implicit assessment of the quantity of objects such as photos in a pile through touch. Such assessments in the digital world required other more effortful strategies and actions, such as we saw when participants needed to spread out piles to visually judge the quantity and content they were about to deal with.

Finally, the use of 3 dimensions in physical space supported a diverse range of strategies people could use to focus, select, and keep some objects separate from others (such as in stand-by mode, or in ad-hoc categories). The third dimension also meant that participants could either bring objects close to the body or the body toward objects, offering a greater range of flexibility for dealing and manipulating the artefacts in the tasks.

In terms of design, this implies that the simple mimicking of physical space through graphical representation, multi-touch input, and the like may not be sufficient to encourage interaction which is really like the physical world. Rather, it suggests that the actions and strategies for accomplishing the key elements of tasks across physical and digital modalities (such as focusing, comparing, and so on.) may in fact be quite different when some but not all aspects of the physical world are emulated. Design solutions must therefore take account of this fact and think about how different parts of a task might be best supported. The point is here not that we necessarily have to mimic physical properties, but rather we have to recognize what those physical affordances achieve for people when working with tangible objects, and ask how we can employ perhaps different methods to attain those same ends digitally.

A case in point is the finding that digital interaction may not naturally engender the kind of bimanual interaction we see in the physical world (even if it supports it). This suggests that in order to confer the benefits of bimanual interaction [23], one approach is to design specific tools and techniques which more explicitly *require* asymmetric bimanual interaction, e.g. "ToolGlass-type" interfaces [4, 39]. As another example, to compensate for the lack of physical constraints in the digital realm, "magnetic snapping" between pieces and grouping gestures are some possible solutions. Likewise, scaling possibilities, "elastic regions" with rubber-band borders, and interactive visualizations such as zooming and fish-eye views, are some possible strategies for the visual design of interactive digital media for focus-and-context purposes. Furthermore, graphical visualizations that suggest depth, as is proposed in the pile metaphor interface [24], or embodying a kind of virtual resistance into the interface in order to mimic the friction of a physical pile on top of a surface, can support the estimation of quantity. Semi-transparent overlays displaying the number of items in the pile or small thumbnails of the pictures are other examples of how design could address such issues.

It is important to be aware, though, that the emulation of the physical can only go so far in shaping interaction in the digital realm. Ultimately the designer must face a decision about the extent to go down this path. This makes sense to the extent that this improves the quality of the interaction, and confers on the digital interface new and compelling interaction techniques. However, a better design decision may sometimes be to preserve those physical aspects of interaction through tangible or hybrid physical-digital user interfaces. As designers and researchers we need to recognize and comprehend those limits. One important step in doing so is a deeper understanding of the relationship between affordances and interaction.

CONCLUSION

We have reported on a largely exploratory study in which we wanted to compare physical and digital manipulation on a tabletop in such a way that the effect of actually being able to "get to grips" with objects on interaction could be evaluated. While there were advantages in terms of people's level of satisfaction and task time in the physical case, in many ways this was less informative than an in-depth and detailed examination of the nature of interaction in the two cases. As we have seen, physical metaphors and methods of input may appear to encourage manipulation in a physical way, but in the digital realm it is essentially quite different.

It is important that we understand these differences, especially as more and more artefacts in our everyday life assume a digital instantiation (e.g. photos and documents). These changes, together with technological advances in interactive displays, call for the design of novel ways of manipulating, sharing and integrating those artefacts with other existing ones. We need to think more deeply about how we can use physical affordances as a design resource while at the same time exploiting the new and exciting possibilities of digital media. We view this study as one step towards that goal.

ACKNOWLEDGMENTS

Thanks to all the people who participated in the study and gave freely of their time and effort.

REFERENCES

1. Agarawala, A., Balakrishnan, R. Keepin' it Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen. In *Proc. of CHI 2006,* 1283-1292.
2. Apted, T., Kay, J., Quigley, A. Tabletop Sharing of Digital Photographs for the Elderly. In *Proc. of CHI 2006,* 781-790.
3. Ark, W., Dryer, D.C., Selker, T., Zhai, S. Representation Matters: The Effect of 3 Dimensional Objects and Spatial Metaphor in a Graphical User Interface. In *Proc. of HCI on People and Computers XIII*, 1998, 209-219.
4. Bier, E. A., Stone, M., Pier, K., Buxton , W., DeRose, T. Toolglass and Magic Lenses: the See-through Interface. In *Proc. of SIGGRAPH 1993*, 73-80.
5. Buxton, W. Chunking and Phrasing and the Design of Human-Computer Dialogues. In *Proc. of the IFIP World Computer Congress*, Dublin, Ireland, 1986, 475-480.
6. Carroll, J.M., Thomas, J.C. Metaphors and The Cognitive Representation of Computing Systems. IEEE *Transactions on Systems, Man and Cybernetics*, 12 (2), 1982, 107-116.
7. Czerwinski, M., Robertson, G.G., Meyers, B., Smith, G., Robbins, D., Tan, D. Large Display Research Overview. In *Proc. of CHI 2006*, 69-74.
8. Dietz, P., Leigh, D. DiamondTouch: A Multi-User Touch Technology. In *Proc. of UIST 2001*, 219-226.
9. Dragicevic, P. Combining Crossing-Based and Paper-based Interaction Paradigms for Dragging and Dropping Between Overlapping Windows. In *Proc. of UIST 2004,* 193-196.
10. Erickson, T. Working with Interface Metaphors. In *The Art of Human-Computer Interface Design*, Ed. by B. Laurel, Addison-Wesley, 1990.
11. Fishkin, K. P. A Taxonomy for and Analysis of Tangible Interfaces. *Journal of Personal and Ubiquitous Computing*, 8 (5), September 2004, 347-358.
12. Fitzmaurice, G. W., Ishii, H., and Buxton, W. A. Bricks: Laying the Foundations for Graspable User Interfaces. In *Proc. of SIGCHI 1995,* 442-449.

13. Guiard, Y. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. In *Motor Behavior*, 19 (4), 1987, 486-517.
14. Hancock, M., Carpendale, S., Vernier, F., Wigdor, D. Shen, C. Rotation and Translation Mechanisms for Tabletop Interaction. IEEE Workshop on TABLETOP 2006, 79-88.
15. Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. Passive real-world interface props for neurosurgical visualization. In *Proc. of SIGCHI 1994,* 452-458.
16. Hinrichs, U., Carpendale, S., Scott, S., Pattison, E. Interface Currents: Supporting Fluent Collaboration on Tabletop Displays. In *Proc. of the 5th Symposium on Smart Graphics*, 2005, 185-197.
17. Hutchins, E., Hollan, J., Norman, D. Direct Manipulation Interfaces. In D. A. Norman & S. W. Draper (Eds.) *User Centered System Design: New Perspectives in Human-Computer Interaction.* Lawrence Erlbaum Associates: Hillsdale, NJ, 1986.
18. Ishii, H. and Ullmer, B. Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms. In *Proc. of SIGCHI* 1997, 234-241.
19. Kirsh, D. The Intelligent Use of Space. In *Artificial Intelligence*, 73, 1995, 31-68.
20. Koike, H., Xinlei, C., Nakanishi, Y., Oka, K., Sato, Y. Two-handed Drawing on Augmented Desk. In *Extended Abstracts of CHI 2002,* 760-761.
21. Kruger, R., Carpendale, M.S.T., Scott, S.D., Greenberg, S.. Roles of Orientation in Tabletop Collaboration: Comprehension, Coordination and Communication. In *Journal of Computer Supported Collaborative Work*, 13(5-6), 1994, 501-537.
22. Laurel, B. Interface as Mimesis. In DA Norman & SW Draper (eds.), *User Centered Systems Design*, Hillsdale, NJ: Lawrence Earlbaum Assoc, 1986, 67-85.
23. Leganchuk, A., Zhai, S., Buxton, W. Manual and Cognitive Benefits of Two-Handed Input: An Experimental Study. In *Transactions on Computer-Human Interaction*, Vol5 (4), 1998, 326-359.
24. Mander, R., Salomon, G., & Wong, Y. A "Pile" Metaphor for Supporting Casual Organization of Information. In *Proc. of CHI 1992*, 627-634.
25. Morris, M. R., Huang, A., Paepcke, A., and Winograd, T. Cooperative Gestures: Multi-user Gestural Interactions for Co-located Groupware. In *Proc. of SIGCHI 2006*, 1201-1210.
26. Norman, D.A. The Psychology of Everyday Things. Basic Books, New York, 1998.
27. Rekimoto, J., Saitoh, M. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In *Proc. of CHI 1999*, 378-385.
28. Rekimoto, J. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In *Proc. of CHI 2001*, 113-120.
29. Ringel, M., Berg, H., Jin, Y., Winograd, T. Barehands: Implement-Free Interaction with a Wall-Mounted Display. In *Extended Abstracts of CHI 2001*, 367-368.
30. Ringel, M., Ryall, K., Shen, C., Forlines, C., Vernier, F. Release, Relocate, Reorient, Resize: Fluid Techniques for Document Sharing on Multi-User Interactive Tables. In *Proc. of CHI'04*, 1441-1444.
31. Scott, S.D., Grant, K., D., Mandryk, R. System Guidelines for Co-located, Collaborative Work on Tabletop Displays. In *Proc. of ECSCW 2003*, 159-178.
32. Scott, S.D., Carpendale, M.S.T, Inkpen, K.M. Territoriality in Collaborative Tabletop Workspaces. In *Proc. of CSCW 2004*, 294-303.
33. Sellen, A., Harper, R. The Myth of the Paperless Office. MIT Press, Cambridge, MA, 2003.
34. Shen, C., Vernier, F.D., Forlines, C., Ringel, M. DiamondSpin: An Extensible Toolkit for Around-the-Table Interaction. In *Proc. of CHI 2004,* 167-174.
35. Smith, D., Irby, C., Kimbal, R., Verplank, B., Harslem, E. 1982. *Designing the Star User Interface*. Byte, 7/4, 242-282.
36. Streitz, N.A., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P.,Seitz, P., Steinmetz, R. i-LAND: An Interactive Landscape for Creativity and Innovation. In *Proc. of CHI 1999*, 120-127.
37. Tang, J.C. Findings from Observational Studies of Collaborative Work. In I*nternational Journal of Man-Machine Studies*, 34, 1991, 143-160.
38. Terrenghi, L. Design of Affordances for Direct Manipulation of Digital Information. In *Proc. of Smart Graphics Symposium 2005*, 198-205.
39. Terrenghi, L., Fritsche, T., Butz, A.: The EnLighTable: Design of Affordances to Support Collaborative Creativity. In *Proc. of Smart Graphics Symposium 2006*, 206-217.
40. Underkoffler, J., and Ishii, H. Urp: A Luminous-Tangible Workbench for Urban Planning and Design. In Proc. *of CHI 1999*, 386-393.
41. Wellner, P. The Digitaldesk Calculator: Tangible Manipulation on a Desktop Display. In *Proc. of SIGGRAPH 1991*, 107-115.
42. Wilson, A. PlayAnywhere: a Compact Interactive Tabletop Projection-vision System. In *Proc. UIST 2005*, 83-92.
43. Wu, M., Balakrishnan, R. Multi-finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In *Proc. of UIST 2003*, 193-202.