

At Home with Surface Computing?

David S. Kirk², Shahram Izadi¹, Otmar Hilliges¹, Stuart Taylor¹, Abigail Sellen¹, Richard Banks¹

¹Microsoft Research Cambridge
7 JJ Thomson Ave., Cambridge, UK, CB3 0FB
{shahrami; otmarh, asellen; stuart;
rbanks}@microsoft.com

²Computing Sciences
Newcastle University,
Newcastle, NE1 7RU, UK
david.kirk@newcastle.ac.uk

ABSTRACT

This paper describes a field study of an interactive surface deployed in three family homes. The tabletop technology provides a central place where digital content, such as photos, can be easily archived, managed and viewed. The tabletop affords multi-touch input, allowing digital content to be sorted, triaged and interacted with using one or two-handed interactions. A physics-based simulation adds dynamics to digital content, providing users with rich ways of interacting that borrows from the real-world. The field study is one of the first of a surface computer within a domestic environment. Our goal is to uncover people's interactions, appropriations, perceptions and experiences with such technologies, exploring the potential barriers to use. Given these devices provide such a revolutionary shift in interaction, will people be able to engage with them in everyday life in the ways we intend? In answering this question, we hope to deepen our understanding of the design of such systems for home and consumer domains.

ACM Classification: H5.2 Information interfaces and presentation: Graphical user interfaces.

General terms: Design, Human Factors

Keywords: Interactive surfaces and tabletops, multi-touch, physics-simulation, field-study, home.

INTRODUCTION

Although interactive surfaces and multi-touch are an established area of HCI research – dating back over two decades [3] – it is only recently that these technologies have become the focus of much media attention. The internet is abuzz with talk of new products, and research in this area is growing apace, reflecting this enthusiasm. In many ways this is understandable, given that these technologies change or even revolutionize the ways we can interact with computers. The “natural” style of interaction afforded by the direct multi-touch displays, the form-factor, and their collaborative nature, are the often-cited appeals of such technologies.

It is therefore unsurprising that many companies are now

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

Copyright 2012 ACM 978-1-4503-1015-4/12/05...\$10.00.

beginning to market interactive surfaces, particularly in tablet and phone form-factors. Larger interactive tabletops are becoming commercially available. However, their costs currently are fairly prohibitive for the everyday consumer market. This is likely to change in years to come, as the cost of the hardware falls. Currently however, these technologies have only really emerged in a few niche markets.

As interactive tabletops become more available, we can easily posit a world where people will want to procure such novel technologies for their homes. But how will real users respond to and learn to interact with such technologies, when they are placed in their homes? In this paper we present one of the first deployments of a multi-touch tabletop device in a domestic setting. Our overall goal is to investigate the potential for surface computing in the home. Specifically, in this paper we seek to observe what people's interactions, perceptions and experiences are of such novel computing technologies and interfaces, as a means to further inform the design space. Given that these devices provide such a revolutionary shift in form of interaction, will people be able to interact with them in the ways we intend?



Figure 1: Our interactive surface, a bespoke tabletop, with an embedded multi-touch display and overhead camera for imaging physical objects on the surface (left), with a custom physics-enabled photo application (right).

To explore this question, we have built and evaluated an interactive tabletop, in-the-field with three families, for a month at a time. The vision-based tabletop system, shown in Figure 1, was built from the ground up, in-house. We have developed a media management application for these devices, which provides a central place for users to manage and archive their digital photos (see also Figure 1). Although by no means feature-rich, this application has instead been designed to be rich in the interactions it affords,

to allow us to more closely observe how real families interact with interactive tabletops. The application therefore supports multi-touch input, allowing digital content to be sorted, triaged and interacted with using one or two-handed interactions. A physics-based simulation adds dynamics to digital content, providing users with rich ways of interacting that borrows from the real-world.

Previous work [9] on this system reported on the more sociological aspects of the deployment, describing how the technology integrated, and at many times disrupted, social and organisational practices in the home. In this prior work the technology was very much used as a probe to enrich our understandings of the home, ideas around family and practices around digital media management.

In this paper we describe our findings that pertain more towards the *interactions* that our surface computer offered, in particular around the multi-touch and physics-enabled techniques to manipulate the virtual. We were particularly interested in understanding whether this interaction paradigm shift, away from the desktop, would work outside of the lab. Specifically, with our deployment we wanted to investigate the following questions in detail:

- Can households naturally learn to use multi-touch gestures, and do bimanual or multitouch patterns of use evolve over time?
- How does a physics-based, hands-on interactional model affect the ways such a technology was used? And does such a model actually allow users to develop their own interaction strategies?
- Can we design a system that successfully blends the virtual and physical? In particular how do we maintain the benefits of the virtual when designing an interface that closely resembles the ways in which we manipulate objects in the real world?

While the answers to these questions are specific to the device we have constructed, we believe that exploring these issues will have generalizable implications for informing the design and development of future surface and tabletop systems for home environments. As we conclude, these questions and our observations during the field deployment open up an interesting discussion around the notion of *intuitive* or *natural user interfaces* in general.

RELATED WORK

Lab studies have uncovered many affordances of tabletop systems. Much of this work considers tabletop use in the context of co-located collaboration [14, 16, 17]. These have led to guidelines for the development of tabletops [15, 18].

Although studies and interaction techniques are common in the tabletop literature, real world applications have been less forthcoming. Photo manipulation and organisation has become one popular application area [2, 17]. Given that the field is fairly new, it is understandable that studies of tabletops have mainly been within the confines of laboratories

– to evaluate some of the basic affordances of these systems in more controlled settings. However, real-world deployments are now beginning to emerge, including knowledge workers’ use of large tablet displays [12], use of tabletops within museums [6], and other public and shared spaces [20], including schools [14] and universities [22]. Wigdor et al. [21] presents a longitudinal study of a multi-touch tabletop deployed within an office. Hancock et al. [6] describe the use of physics-enabled tabletops for sandtray therapy for children. Although not a tabletop, CityWall [13] studies the use of a single multi-touch wall within a public space for a week, revealing many findings with regard to the social phenomena surrounding such novel technologies.

We wish to extend this literature by considering the use of tabletops and surface computing in the home. Mazalek et al. have begun to explore this space with their TVViews system [10], a tangible tabletop which allowed multiple users to interact through wireless pucks on a large horizontal display. Rather than considering tangible tabletops, we wish to more closely understand the adoption of direct multi-touch tabletops within the domestic domain, and further also evaluate the more advanced techniques emerging from the tabletop literature, such as physics-based interactions [23]. In doing so, we wish to ground our work in the field, focusing on in-situ rather than lab-based studies. Ultimately if we wish to have such technologies widely adopted in the real world, we need to better understand how people will interact with them as part of their everyday lives.

AN INTERACTIVE TABLETOP FOR THE HOME

For this deployment we purposefully kept the features of the system basic, using the field study itself as a means for eliciting more advanced user requirements for future iterations of our tabletop media management system. Given our dual purpose of also learning more about surface computing in-the-field, we decided to design a system that was *interaction-rich*, even if not *feature-rich*. Therefore we aimed to provide users with hands-on and intuitive multi-touch manipulation of virtual objects, supporting a variety of direct one and two-handed interactions.

We chose to add physical behaviors to virtual objects and use physical gestures wherever possible for interaction. Our aim was to imitate the openness and ease with which we manipulate photos in the real world. Much work has been done recently in the area of physics on tabletops (see [23] for a review), and we felt this was critical to include in the design and study of the system, as it is clearly a relevant technology for surface computing. We were keen to understand whether users would see such features as a novelty, which wore off, or whether they would find them integral to system use.

Physical Design and Underlying Technology

From prior experience of the home deployments of technologies and understanding that the technology would live in peoples’ homes, we decided to design a device that looked more like a bespoke piece of furniture than a piece of technology. We also initially investigated the possibility

of procuring existing off-the-shelf hardware, but given costs of devices and availability, and our aesthetic goals for the device, we settled for developing a solution in-house from the ground-up.

We chose a common Frustrated Total Internal Reflection (FTIR) based technique for multi-touch sensing [5]. The display screen was constructed using an acrylic sheet, covered with a thin sheet of silicon rubber and a layer of drafting film, and finally edge-lit with Infrared (IR) LEDs for touch sensing. A projector was mounted inside the unit in order to back project onto the surface. A camera was similarly positioned inside to image the display surface. These images were processed on an internally stored PC, using custom GPU-based computer vision techniques. This allowed us to process multiple touch contacts at 60Hz, with little latency, giving an acceptable rate of interaction. The projection gave a display size of 47cm x 35.5cm (18½" x 14"), which a single person or two people could comfortably interact around. Due to the anticipated usages and placement of the device, the display was set at waist-level allowing people to stand rather than sit to interact. We were careful to set the height of the surface to support ease of use, without undue strain, when standing.

To fit with the goal of making a device for the home, we were careful in the physical design to try and hide much of the technology from the user. The cameras, projector and connected PC were all encased. A wooden top with an extended space to the side and shelves allowed devices, ornaments and other objects to be kept by the device. A digital camera and dock was placed on one of these shelves to allow people to take and upload photos.

Interacting with the Tabletop Archive

The application running on our multi-touch tabletop is built using a 3D graphics engine (currently Direct3D). We leverage the metaphor of peering inside the tabletop to reveal a 3D virtual world confined by four virtual walls, which correspond to the bounds of the tabletop display. The virtual camera is set so that you are looking at the 3D scene from a birds-eye view (see Figure 1, 2 and 3).

Objects such as photos are controlled using a gaming physics engine (NVIDIA's PhysX). We model multi-touch input in this world using a technique described in [23], which approximates each finger touching the surface to a series of virtual objects that can interact with others in the 3D scene.

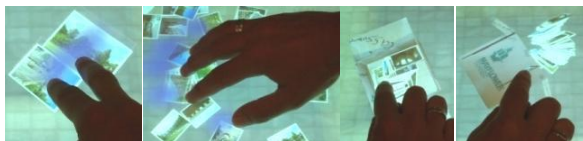


Figure 2: Flicking two photos (left) and gathering multiple objects using multiple touching fingers (middle). Blue shading provides feedback of touch points. Tipping boxes (right).

For users, this means virtual objects begin to move more fluidly and behave more tangibly. Photos and scanned objects have inertia and acceleration, can be flicked from one

side of the surface to the other (see Figure 2 left), can collide with one another and bounce off walls, or be spun and rotated by applying torque. As we model multi-touch input, the user is free to use multiple fingers, and one or two hands to gather multiple objects simultaneously, pile them, separate two piles out using a sweeping gesture and so forth (see Figure 2 middle). This behaviour is unscripted, in the sense that users are free to define their own ways of interacting within the bounds of the physics simulation.

This approach of leveraging physics is clearly inspired by prior work such as Bumptop [1] – although the original Bumptop system only supported interaction with one point of contact. This is one of the core novel aspects of the technology that we wanted to evaluate with households. Of course a literal translation of the physical world sometimes limits the qualities of the digital. Therefore in certain instances we supported scripted gestures, such as a multi-fingered gesture for scaling virtual objects.

Getting Content onto the Tabletop

Content could be automatically uploaded to the table using either a digital camera docking station or other USB memory storage device. The system tracks files that it currently contains in order to prevent multiple downloads of the same picture. Once uploaded, content is automatically arranged in piles within the 3D world, and can be spread out to access individual objects by a flick of the hand.

Inspired by fieldwork around real photos in the home [9] we use the metaphor of *virtual shoeboxes* for containment. This is an easily identifiable real-world analogy, and provides a loose way to partition the media on the tabletop. Boxes can be moved and rotated using similar interactions used for photos. Box lids can be opened and closed using a pinch gesture. Once opened, applying multiple contacts off center, as shown in Figure 2 right, spills box contents. To make the process of getting objects into boxes as straightforward as possible, an open box carries a magnetic force, which essentially attracts any colliding objects into the box.



Figure 3: The photo floor (left) and box basement (right).

Clutter was a concern in early UI versions. As boxes were created less virtual floor space was available for working with digital photos and scans. After iterative testing we decided to separate storage and interaction spaces into two virtual 'floors'. The analogy we use is to have boxes stored in the basement – or at the bottom of the tabletop – from where they can be brought to the top level, to be spilled out, loosely arranged, displayed and subsequently tidied up and stored back in the basement (see Figure 3). To give the user a sense of this two-level arrangement and the ability to

overview the boxes stored in the archive, the “photo floor” is slightly transparent (as shown in Figure 3 left).

So far we have illustrated how users interact with objects already in the system. Interactions such as manipulating photos or boxes can be modeled quite nicely using physics-enabled multi-touch. However, other core features for our application did not easily map to such a literal gesture manipulation paradigm. In order to expose these features to our users we reverted to a collection of virtual buttons along the top edge of the screen. Functionality included the creation of new, empty boxes and also their deletion once no longer needed. Further icons are provided for users to move between the two floors. Retrieving boxes from the lower storage level (or vice versa) was achieved by touching a particular box and the icon to move between floors. The system also provides an easy means for labeling boxes (and objects) with ink – by switching to annotation mode using a virtual button. In this mode users use their fingers to scribble on boxes and draw onto photographs.

We also decided to introduce a set of shortcuts allowing users to rapidly arrange objects an ordered grid in order to get a quick overview over many photos at once. A “slideshow” mode then allows users to rapidly view and cycle through these photos up-close, by repositioning the virtual camera.

FIELD TRIAL

It is a common strategy for evaluations of user interfaces to take place in the lab and to have a formal experimental structure. Our research goals were instead concerned with understanding the practicalities of everyday interaction with multi-touch and physics-enabled surfaces in-the-wild. As such, an observational, in situ analysis of user interaction in the home was more suitable for our purposes.

Accordingly, we conducted an in-depth field study of our tabletop prototype in three different family homes, each deployment lasting for about a month. Given time constraints, we built and deployed three tabletop units in parallel, one for each family. This coupled with the technical complexity of the system – comprising of custom hardware, GPU processing, and a 3D physics-based UI – made for an extremely demanding deployment. Similar systems deployed in the field have only supported a single system deployment [10, 13], but we felt this was important to help generalise our results.

We officially visited each family once per week for several hours at a time, but also took opportunities to capture field data during times when we were called out to fix technical problems. This data consisted of a written diary left with each family and video and audio recordings both of the participants directly interacting with the UI and whilst being interviewed about their use and experiences of the system. This usually involved capturing footage for extended periods, where multiple family members closely interacted with the device to show us the media they had stored.

In terms of spending time in the field with families, we wanted to get a balance between capturing enough data and ensuring the families had time on their own with the system (being careful to ensure that we were not coaching them in how they should be using it). During the field trial all families had at least some members go on short breaks, frequently returning with new content for the tabletop. At the end of the deployment we also collected system logs of device use and copies of all content uploaded and conducted an exit interview.

Participants

Our families were recruited by word of mouth and email advertisement and all came from the local area. We recruited families with the requirement that they did not have children under the age of 6 (as the back of the device was exposed and therefore possibly hazardous to small children), and that they had previous exposure to using digital photography. Participants were rewarded for participation by being able to keep the digital camera and docking station that was provided with the tabletop. Some users had prior experience of using multi-touch technology, such as iPhone use.

Household A: consisted of a recently married couple in their early 30’s, he was a biostatistician and she was an economist at a local university. Both travelled frequently. Both had a post-graduate level of education.

Household B: consisted of a married couple (in their mid-30s), with a six year old son. The father worked in medical research and the mother in biotech. Both had a post-graduate level of education. During the trial for 2 weeks the family had their maternal grandmother staying to look after the little boy during his summer holidays from school.

Household C: consisted of a married couple (late 30s-early 40s) with two teenage daughters (aged 15 and 18). The father worked in software development and the Mother was a housewife. The father had a graduate degree, the mother high school level education. This family had maternal grandparents staying for much of the trial.

Data Analysis

Logged data of system use was collated and summarized after the field trial had ended. The video data of interviews and usage data of the system were then closely analyzed by the research team. Transcriptions of interesting moments of interaction and reflection were transcribed and from this, deeper themes explored.

FINDINGS

General Patterns of Use

All three households chose to have their device situated in their living room. Households B and C kept their systems running continuously, while Household A turned it on only when they intended to use it for uploading materials or displaying photos to friends, family and other guests. This was due mainly to concerns about the possible energy consumption of the unit. Due to the time taken to install and config-

ure the device, participants were given little training from us regarding the device (~30 minutes overview each).

Inevitably, as with most prototype systems, there were occasional technical problems. Due to the fact that the system was vision-based, adverse lighting conditions affected the touch sensing. However, because we rendered feedback in the UI when touch-points were sensed (see Figure 2), households became very adept at detecting when rogue touch points were appearing on the surface, and would devise workarounds. For example, Household B moved the orientation of light sources in the house, and Household C drew curtains at times when the surface was being used under direct sunlight. By exposing the users to some of this technical maintenance we found they were able to devise fairly sophisticated means to work around these limitations on their own.

Despite these problems, we observed substantial use of the tabletop during deployment, as summarized in **Table1**.

	Nr. days deployed	Nr. days used	Nr. photos kept	Nr. boxes kept
Household A	27	13	312	21
Household B	30	21	765	30
Household C	32	18	429	25

Table1: General usage statistics per household. (Note days deployed includes days travelling/vacationing).

In the remainder of this paper we focus on investigating the users’ perceptions of the UI, their uses of multi-touch, physics and 3D, and in general, how they interacted with this novel surface computing technology. We avoid findings pertaining to the media management features of the system (this is instead the focus of [9]).

Ways of the Hands

We were extremely pleased with the speed in which users picked up the gestural vocabulary for interacting, even when training had been kept to a minimum. Manipulation of photos and boxes in particular were rapidly discovered. On the first session most users were able to move, scale, rotate, flick and pile photos. Flicking was particularly evident and users formed strategies for flicking multiple objects into boxes rapidly. Basic interaction with boxes was also straightforward for users, such as moving, opening, closing, rotating and hovering up content.

Throughout the study we observed many one-handed multi-touch interactions, and much bimanual interaction. Users switched effortlessly between using one and two hands. The great degree of bimanual contrasts with some lab studies of photo work on multi-touch tabletops [19].

Users also developed markedly different hand gestures to achieve the same UI actions. This is perhaps best illustrated by Household A. In Figure 4, we show just some of the hand gestures utilised by both husband and wife during a 20 minute period of sustained interaction (observed during the middle of the trial). Here the users switched from using a single finger to move a single small object (a) to grabbing and moving a larger object with multiple fingers of one

hand (b) and then finally moving multiple objects one-at-a-time by using the index fingers of both hands (c). This illustrates how diverse and fairly complex chains of hand gestures could be devised for simply moving virtual objects. We observed how the wife distinguished and switched fluidly between precise selection of a single small target using her index finger, and more coarse selection and movement using multiple fingers of a single hand.

Likewise in the same session, the husband goes on to open a box using one hands’ index and middle finger (f) and then close it using the index fingers of both hands (g), and rotate a photo bimanually (h), whilst the wife rotated a photo using a single finger placed in the corner (i) and finally both rotated using a more typical thumb and forefinger pinch (j). Although this open-ended set of hand gestures may feel complex to master, these were in fact lightweight for users to achieve, and we began to observe these throughout the trial. They felt at ease experimenting with new techniques, and also switching between techniques gracefully.

We often observed sequential interaction with objects as shown in [19], where users interacted with objects one at a time. However, even with these single object interactions, we observed large use of the second hand, which was often used alongside the first to divide the labor between hands [4]. This again contrasts with [19]. That being said we also observed many instances of interaction with multiple objects (although perhaps not as plentiful as sequential use). These were often for more coarse movement of objects, such as gathering up, piling or rummaging through photos.

Here we noted about half the participants would not directly touch on photos, but instead would use more of a collision model, bringing their fingers in on the sides of objects to either push them away or more commonly bring them towards themselves. Although not conclusive, this is interesting as the model does not directly translate to drag-and-drop metaphors where movement is achieved by direct selection, but instead suggests a model that derives more from physical interaction. Returning to Household A, we see later in the session the wife sweeping multiple photos to one side using multiple fingers of one hand, while she searches for a specific photo to show us (d), she then introduces her second hand to speed this process up and again divide up the labor (e).

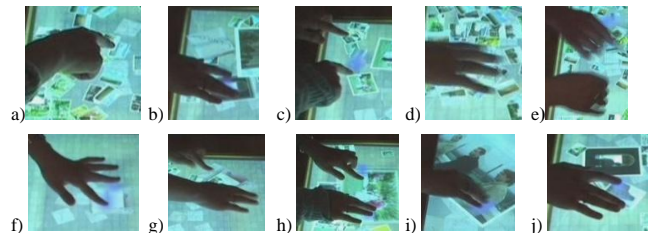


Figure 4: Different hand postures. Users switch fluidly between multiple uni- and bi-manual interactions.

The use of multi-touch and bimanual use is also highlighted in **Table 2**. Here we highlight the total minutes the families interacted with the tabletop on a daily basis (not including

days they were away for travel). We also logged whether during each minute, the user(s) were interacting predominantly with one, two or more fingers (this is shown as a percentage). This gives a *rough* sense of how many fingers were typically used during interactions with the tabletop. Note that we did not count contacts appearing and disappearing rapidly, which did not move, or did not lead to any UI actions, as these may have been falsely triggered by adverse lighting. We also did a simple distance measure between contacts to again *roughly* predict when contacts could not physically belong to same hand, suggesting bi-manual use.

	Total use (min)	Mins/day	1 finger	2 finger	> 2 finger	Bi-manual
Household A	1351	104	37%	44%	19%	31%
Household B	2396	114	42%	47%	11%	22%
Household C	2657	148	29%	38%	33%	45%

Table 2: More detailed usage statistics. Column two shows total interaction time per family, recorded when 1 or more finger was active. Fourth, fifth and sixth estimates the percentage of this time that 1, 2 or more than 2 contacts were sensed. The last column estimates percentage of overall bimanual use.

Overall our logs, video data and interviews demonstrate multi-touch gestures and bimanual use was plentiful. This was interesting given that the size of the display is a little smaller than other tabletops and we had felt occlusion caused by two hands could be an issue. Already through the field deployment we were seeing results that contrast with lab-based studies of tabletop photo work [19].

Curse of Single Touch

The data also highlights single finger use. This in some ways is to be expected, particularly when moving a single object or clicking buttons. However, when we began to unpack other moments *when* and *why* users switched to a single finger, we found a subtle, but yet critical difference between how users interacted with objects in the virtual and the physical. In the real-world, we often use two hands to provide finer motor control (for example imagine trying to control a ball rolling across a table with one finger as opposed to entirety of your hand). In the digital however, more hands or fingers does not necessarily equate to more control. Indeed users often have the opposite way of perceiving fine control, we often observed users reverting to single touch to carry out fine-grained interactions, for example to position an object precisely.

This reverting to a single finger was also the case when physics-based interactions led to unpredictable results in the UI. For example, when boxes filled up they would become more unstable suggesting they may accidentally tip over. Under these conditions using multiple fingers leads to more force to hold the box in place, so in terms of control more fingers are better than one. However, we observed that if the box did show signs of tipping, the users would in fact do the opposite – removing all fingers from the box and using a single finger to carefully move it. This actually made the box less stable, and led to tipping and user frustration.

One consideration here is that when users are touching the virtual box, they are not directly interacting with it. Even if the tabletop gives the impression of directness, there is an intermediary – the computer – so reverting back to the simplest form of interaction, a single finger, allows users to make better sense of the actions of the computer in response to their input.

Beyond the Desktop

In general people appreciated the hands-on nature of the system and the physics-based interactions. It gave a sense of familiarity when interacting digitally, and at the same time distinguished it from standard computing interfaces. Users had a sense that they were interacting with something new and different.

Excerpt 1 – [Household A]

Interviewer 1: have you looked at it like it's a computer?

Anne: No, because we rarely see the first screen (meaning the windows desktop)

This is perhaps best illustrated again through example – the elderly grandfather in Household C walking us through his photo collection on the tabletop. This participant openly admitted he had little if no computer expertise. At first during his interactions he struggled and needed instruction from his daughter to recover his box from the virtual basement. He continually asked “*What button do I press?*” referring to the icons on the top of the screen. However, once the box was on the photo-level, he managed with little instruction to open the box with a pinch gesture, get the contents out of the box, scale and show each photo to us, and rescale them back using bimanual input, before finally flicking each photo back in the box and closing the lid, as highlighted in Figure 5. Whilst interacting, he was fully engaged with us, describing the story behind each photo. It is interesting that the only time he stalled was interacting with the icons at the top of the screen.

When interviewed (with his daughter present), he alluded to what qualities of this interface he found beneficial:

Excerpt 2 - [Household C]

Karen: He (the granddad) doesn't really like computers but as soon as he was here, he was up here [pointing to tabletop]

Karen: I was saying wasn't I mum, that if he'd seen that there was a computer based thing behind it then the block would have come up and he wouldn't have done it

Granddad: Yeah because it's touch init, it's like er fun.

Interviewer 1: do you like the fact that you can do that?

Granddad: Yeah I'm a hands-on kind of person, I've always worked with my hands, everything I do is with me hands, up here [points to head] is a little bit slower than me hands.

The granddad clearly sees a computer as requiring a technical mind, something that requires expertise and a mindset he will never have, but instead this device allows him to engage and express himself with his *hands*. It is interesting that as the daughter points out – there is a PC hidden in the tabletop – but because of the tangible quality of the direct touch and the physics, the granddad feels qualitatively differently about this device. He was confident enough to in-

teract, and manipulate objects based on the hands-on skills he has acquired from the real-world.



Figure 5: The granddad in household C takes his box from the basement and shows us photos using bimanual input to scale the photos and a pinch gesture to close and open boxes.

The physics also added an organic feeling to the UI that people greatly appreciated. As illustrated in the following:

Excerpt 3 - [Household B]

Carol: And the whole physics-I like! It's just that interaction that you don't get. I guess you could put it in a normal computer but it's just a bit more funky, more real.

Excerpt 4 - [Household A]

Victor: The physics of this stuff is really cool. I mean it feels really like it's a real object. A feature that I, now that I'm playing with it, would like to have, is a full hand. Does it do this? (Manages to gather up many objects using his entire hand)

Again the physics gives the interaction a “real” or tangible quality. In excerpt 4, the participant is enticed to play and experiment with the system, and quickly discovers new unanticipated interactions are feasible. In excerpt 3, the participant hints at how the physics adds further distance between this device and a regular PC. She goes on to say:

Excerpt 5 – [Household B]

Carol: The things have a space have a feeling, you are not dealing with files you're dealing with things, I think that has the edge, it is very different it does transform the way you think about these things because you have to think twice to realise that it is a file.

...

Mike: Once you get into it's just a room basically, in the beginning yes you think about it more analytically, but once you get used to it, it just becomes a room, and it becomes a room with real boxes

This further alludes to how the physics and the real-world metaphors, coupled with the tabletop hardware allowed the families to further shift from the idea that they were interacting with a normal computer and desktop metaphors. This seemed critical for all the families, but why? Again Household B provides some answers:

Excerpt 6 – [Household B]

Carol: I think I have watched my pictures more with this thing.
Mike: because you have it always there and available
Interviewer 2: Is that a good thing or a bad thing?
Carol: I think it is good, and if I have to go to the computer to look at pictures I will probably end up doing something related to work.

...

Carol: I do not want a computer in my living room!

Interviewer 1: Why?

Carol: Because the keyboard and cables ... it's ugly.

Here the participants clearly identify the PC with work and the workplace, whereas this device has a more playful or appliance-like quality that seems more suited for the home. As we have seen, the perception of this device as “not a PC” seems due to a combination of factors: its hands-on interaction, the illusion of physics, and its aesthetic qualities. However, as we will go on to discuss, all of this can also create a tension when considering what the device is for, in particular, the tension between work and play.

All Play and No Work?

When interviewed families reported a lot on “playing with the device”. They liked the randomness that the physics gave the process of displaying pictures.

Excerpt 7 - [Household A]

Anne: What I liked about the slide show is that it's not as stiff as the digital photo frame.

Anne: Here you can make it very random and stuff which I like.



Figure 6: Physics used to randomly arrange photos before a slideshow.

Here the non-deterministic nature of the physics added a touch of unpredictability or spontaneity to the interaction. Indeed, we saw users play on this spontaneity further by arranging the photos into an aligned grid (using the relevant UI button), and then proceed to shuffle the photos around slightly before starting a slideshow, as shown in **Figure 6**.

The physics by its very nature added non-determinism to the interaction, which was suitable for playful interactions. However, there were times that the interactions were too literal and almost too real. For example, participants would get frustrated by moments when carefully arranged and sorted boxes would accidentally spill due to adverse collisions. Another example is that participants complained about the need to tidy away photos after displaying them or tidy up the basement when it became cluttered. One couple went to extreme cases, when they wanted to ignore the ensuing clutter in the digital by turning off the display.

Excerpt 8 – [Household B]

Carol: Whenever it was messy yes it was like oh my god not only have I got to clean up my house on a Saturday but now I have to clean up the flippin desktop.

Excerpt 9 – [Household C]

Karen: I'm normally tidying it up after people have used it, but then I have a problem with symmetry.

Other frustrations were caused by the “effort” required by the physics interfaces to do such simple things as tipping over the boxes. Many users struggled to do this gesture efficiently, instead opting to use the “arrange in a grid” button to pull the photos automatically out of the box. Here some of the problem was down to the lack of 3D input on the surface. In the real-world, tipping over a box is a simple

endeavour because we can pick the box up and have all degrees-of-freedom (DOF) available for manipulation. However, with the interactive surface the input is bound to a 2D plane, which must somehow be mapped onto the 3D world – inevitably the DOFs in interacting are lost. The analogy is like trying to interact with small physical boxes on a desk but only being able to push or nudge them with fingertips, instinctively we wish to pick the boxes up and interact with them using all 6DOFs.

This clearly highlights an important design trade-off as to when to create an interaction that borrows from the real-world, and when to instead *offload* to the virtual. Being too literal in the translation of the digital tabletop can lead us to missing some of the benefits and efficiency of the virtual (as also highlighted by [1]). In this case we saw that, while aspects of physics helped create playful interactions, the physical work and unpredictability this entailed added to users' perceptions that this was mainly a device for play rather than one for productivity and efficiency.

It's a Small World!

These issues surrounding the literalness of the UI also play out in the finite space that users had in the virtual 3D world. This coupled with the fact that the space was shared by all family members, children and adult alike meant there was a real "battle" for space. This is perhaps best illustrated in the following vignette:

Excerpt 10 - [Household B]

Mike: It's like he [referring to their 6 year old son] would have his own area with permission only to open boxes to yeah mess about but he cannot change for example where the pictures go in this box or in that box.

Interviewer 1: Would he be able to open your boxes?

Mike: Yeah I guess so but with permission, he cannot modify anything he cannot write on certain pictures for example or certain boxes things like that.

Carol: As long as he has his own space I don't think he'll go and try to mess up ours.

[Pause]

Luke (the son): I will.

Excerpt 11 – [Household B]

Luke: what I found fun was playing around with all of the pictures and letting mum and dad tidy them up.

This vignette demonstrates the contention in the household over the shared space. Because the system was democratic, the son had equal rights to explore the space, open boxes and so forth. He used this space more as a play area. He would upload photos of his toys, make a mess (which was almost encouraged with the physics) and not even need to tidy up. In some ways, this was the only space in the home that he had such freedoms so he made the most of this.

'Hands-off' Computing

The democratic nature of surface computing is often seen as a real benefit of these systems, in particular supporting co-located collaboration more effectively [15, 17, 18]. However, although on the first deployment visit we saw evidence of co-located collaboration, by the end of the trial

it had all but disappeared. Instead, participants got into the practice of interacting on their own to, for example, interacting to arrange photos before calling others over to view the show, rather than directly touch the surface.

This contention in collaboration has begun to be uncovered in the tabletop literature [11, 16, 24], but it really was paramount in the homes we studied. Participants were indeed almost hesitant to use the system at the same time:

Excerpt 12 - [Household B]

Interviewer 2: So when you have used it have you used it together or on your own?

Carol: We mostly use it on our own or otherwise we would start.

Mike: You would start fighting.

Carol: Yeah unless you're in presentation mode.

Mike: In presentation mode it's fine

Part of this could be where the tabletop was placed, or perhaps the form-factor, in particular the limits of the screen size. However, it is interesting that participants comfortably interacted side-by-side during our deployment visit, and were happy to interact bimanually on their own at the tabletop. Instead it seemed that many of the social norms of the real-world were simply ignored in the digital domain. People would be happy taking objects out of each other's hands, competing for screen real-estate, and fighting over objects such as photos and boxes. This clearly was done in a playful way, but participants did mediate access to the device to avoid even these playful conflicts.

DISCUSSION

So to return to the question raised in the title of our paper, were people at home with surface computing? And more specifically did they engage with this novel form of interaction in the ways we intended? When it comes to interacting with such devices, we had assumed that the open nature of interaction supported by a combination of multi-touch and physics would make the system intuitive to use. And indeed this was partly borne out: Both young and old were clearly engaged with the system, and in certain cases with little existing computing expertise.

We found that a number of factors contributed to this. One set of factors had to do with the interaction model, and the use of virtual physics, which led to more open interactions allowing users to more readily experiment and develop their own strategies for interaction. These became extremely elaborate and sophisticated as the trial progressed, allowing users to switch between many different strategies for interaction. We also saw clear evidence of the ways in which households fluidly adopted multi-finger and two-handed methods of interaction, as the situation demanded. Accordingly, we would suggest it is important for designers not to overly prescribe a single interaction model. Bimanual interaction of the kind we observed has not previously been observed in shorter-term studies e.g. [19] and only emerged in our work due to prolonged interaction with the system (> 1 week). Importantly, this *ad hoc* and fluid use of bimanual gesture contrasts with prior work such as [19].

Although our system was designed to be *interaction-rich* rather than *feature-rich* we still had to design a user interface that occasionally had to go beyond the literal to enable certain features. Whenever possible we opted for gestures that at least borrowed from the real world such as two fingers to resize images. However, metaphorical gestures such as the pinch to zoom are not always a feasible choice, especially when exposing functionalities that are *not* possible in the physical such as copy and paste. These functionalities are often abstract concepts therefore we usually revert to non-literal metaphors to expose them to the user (typically menu-entries). If we strive to inform the design of real-world applications based on physics-enabled interactions we need to derive ways to enable all three types of interactions: the literal, the metaphorical and the abstract.

A “Natural” User Interface?

Through our study we have also begun to look more critically at a (somewhat overloaded) term that has become synonymous with surface computing – *natural* – suggesting an interface akin to real-world interaction (also see [8]). However, from our study, we have begun to question what this really means from the perspective of everyday users. One example comes in tipping the box, where the most natural interaction – to simply pick the box up and spill out the contents – was unavailable due to lack of 3D input. This made what at first appears a natural gesture, tipping a box by direct touch, a wholly unnatural one, which was a poor imitation of the real-world. Users opted instead to perform a simple button press to perform this action.

Another example is when users fell back to interacting with a single finger, particularly when the physics-simulation meant virtual objects were behaving unexpectedly. Here even though we tried to provide an illusion of interacting directly with a real object, users clearly saw that their interactions were being mediated and processed – by the computer. To make better sense of the computers actions, they would interact cautiously and limit their interactions to a single finger, and observe how the system reacted. Although almost an implicit action, users were simply limiting input to the computer to better make sense of the output.



Figure 7: Interacting with digital objects. Switching from interacting with a mouse (left), to hybrid bimanual interaction using touch and mouse (middle), to bimanual touch input.

We want to use one final rare and yet remarkable example, which we include here in our discussion, rather than our findings section, to inspire us to think more about how the lines between desktop and real-world metaphors can be blurred. During the last few days of deploying the tabletop, Luke discovered the presence of a mouse plugged into the back of the device. When observing him we saw he was

using both touch and mouse, and had become quite adept at distinguishing what interactions in the UI were optimised for which input. For example, Figure 7 (left) shows how Luke uses the mouse to move around the virtual, something he quotes as being “*hard work*” with touch. He also uses the mouse to interact with the buttons on top of the screen as this aided his reach and accuracy. However, for zooming he understands he needs two touches, and proceeds to touch a single finger down, and use the mouse cursor creating the necessary second touch point Figure 7 (middle). He then almost exclusively switches to manipulating objects – the boxes and photos – by touch Figure 7 (right).

Our field-study uncovered interesting tensions arising from the interplay between real-world and desktop-based metaphors. Despite problems caused by mixing these concepts it is *both* that makes surface computing a compelling experience. Our deployment has highlighted the need to develop a sophisticated understanding of *how* to design user interfaces that that seamlessly blend digital and physical interaction metaphors, optimising interplay between both worlds.

Design Guidelines

We conclude by drawing out our main findings and highlighting their impact on the design of tabletop interfaces in the home.

1) *Flexibility of hand configurations and rapid transitions.* We observed regular single finger, multiple fingers and bimanual hand configurations during the study. Users implicitly transitioned between different hand and finger-based configurations, and we feel that it is important for designers not to overly prescribe a single model. Bimanual interactions of this kind were not observed in shorter term studies e.g. [19] and only emerged due to prolonged interaction with the system (> 1 week). This is a crucial finding that goes against prior work such as [19].

2) *Causes of hand configurations and transitions.* Many different factors impacted user hand configurations. For example, when users perceived that a virtual physics object needed to be stabilized or controlled, they switched to using a single finger (which interestingly is not the way we typically bring a physical object under control). The fidelity of the task also resulted in different hand poses. Users switched from single and multi-finger for finely controlling a single object, to using two hands for coarse multi-object interactions. Another under explored factor was that of the *confidence* the user has in the system. For example, if the system failed to respond in a predictable way, the user typically reduced the number of fingers interacting.

3) *Perceptions of tabletop device and physics-based interactions.* Hand-based interactions and physics allowed users to learn the system in ways they would not have been able to with a regular computer. They saw it as a hybrid virtual/digital world. Beyond a regular PC, but had some reservations due to the fact that it was still ‘technology’.

4) *Battling between virtual and real analogies.* We observed how on many occasions the mental model of the

user was battling between real world analogies and their experiences of WIMP. This can lead to confusion and interesting findings. e.g. users assumed less fingers would give more control for physics-enabled virtual objects (less like the real world), but also trying to open boxes using double click instead of using real world metaphor.

5) *Literal translation of the real world is not always good for tabletop design.* Some of the physics interactions made for hard work for the user, and the virtual space was limited due to the 3D physics. This shows that simply replicating the real in the virtual is not without its downsides, and ultimately these metaphors should be used liberally and not always literally.

6) *Tabletops are not always beneficial for collaboration.* There is a huge body of work on tabletop systems for collaboration. However, in practice we saw little examples of collaboration. People genuinely avoided simultaneous use and would often become agitate each other interacting side by side.

7) *Challenge the term natural user interaction (NUI).* We often think that NUI means touch based input. For children in our study natural meant the mouse and double click. Evidence from the study really demonstrates that ‘natural’ is an overloaded and often confused term.

CONCLUSIONS

We have evaluated a novel multi-touch and physics enabled interactive tabletop in three real family homes, providing a grounded understanding of tabletop UI design in domestic spaces. In doing this we have explored how such devices can be used in everyday homes and with regular users. By capturing a rich data set regarding peoples’ interactions and perceptions of such devices the study has assessed the value of multi-touch, physics and surface computing, thinking deeply about the design of such devices. It has allowed us to see where the emulation of physics fits well, and equally where it may hinder the kinds of things that families want to accomplish. This raises a design challenge of building an interface for a system which supports both work and play, and which at once can incorporate affordances of both the physical and digital world in ways that optimize both.

REFERENCES

1. Agarawala, A. and Balakrishnan, R. 2006. Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. In *Proc. CHI '06* (April 22 - 27, 2006).
2. Apted, T., Kay, J., and Quigley, A. 2006. Tabletop sharing of digital photographs for the elderly. In *Proceedings CHI '06*.
3. Buxton, B., Multi-touch systems that I know and love, <http://www.billbuxton.com/multitouchOverview.html>
4. Guiard, Y., Asymmetric division of labor in human skilled bimanual action: The kinetic chain as a model. *The Journal of Motor Behavior*, 19(4):486–517, 1987.
5. Han, J. Y. 2005. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of UIST '05*.

6. Hancock, M., et al. 2010. Supporting sandtray therapy on an interactive tabletop. In *Proc. CHI '10*.
7. Hornecker, E., “I don’t understand it either, but it is cool” – Visitor Interactions with a Multi-Touch Table in a Museum. *Proc. of IEEE Tabletop 2008*.
8. Jacob, R.J. et al. Reality-based interaction: a framework for post-WIMP interfaces. In *CHI '08*. ACM, New York, NY, 201-210.
9. Kirk, D. et al. 2010. Opening up the family archive. In *Proc. Of CSCW '10*. ACM, New York, NY, 261-270.
10. Mazalek, A., et al. The TVViews Table in the Home, In *Proc. IEEE TABLETOP'07*.
11. Marshall, P, et al (2009) Fighting for control: children's embodied interactions when using physical and digital representations. In *Proc. CHI '09*, 2149-2152
12. Morris, M.R., Brush, A.J., Meyers, B., A Field Study of Knowledge Workers’ Use of Interactive Horizontal Displays. In *Proc. IEEE Tabletop 2008*.
13. Peltonen, P. et al. It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. In *CHI '08* (Florence, Italy, April 05 - 10, 2008).
14. Piper, AM & Hollan, J. D. (2009) Tabletop displays for small group study: affordances of paper and digital materials. In *Proc. of CHI '09*. ACM, 1227-1236
15. Scott, S. D., Grant, K. D., and Mandryk, R. L. 2003. System guidelines for co-located, collaborative work on a tabletop display. In *Proceedings ECSCW '03*.
16. Scott, S. D., et al. Territoriality in collaborative tabletop workspaces. In *Proc. of ACM CSCW '04*.
17. Shen, C., Lesh, N., and Vernier, F. 2003. Personal digital historian: story sharing around the table. *Interactions* 10, 2 (Mar. 2003), 15-22.
18. Shen, C., et al. 2006. Informing the Design of Direct-Touch Tabletops. *IEEE Comput. Graph. Appl.* 26, 5 (Sep. 2006)
19. Terrenghi, L., et al. Affordances for manipulation of physical versus digital media on interactive surfaces. In *Proceedings of CHI '07*. ACM, New York, NY, 1157-1166.
20. Van de Mortel, D., Ju, H., Apartgame: a multi-user tabletop game platform for intensive public use. In *IUI '07*: pages 49–52, New York, NY, USA, 2007.
21. Wigdor, D. et al., Living with a Tabletop: Analysis and Observations of Long Term Office Use of a Multi-Touch Table. In *Proc. IEEE Tabletop*. 2007.
22. Wigdor, D. et al (2009) WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. In *Proc. CHI '09*, 1237-1246
23. Wilson, A.D. et al (2008) Bringing Physics to the Surface. In *Proc. of UIST '09*. ACM, 67-76
24. Zhang, X. and Takatsuka, M. Put That There NOW: Group Dynamics of Tabletop Interaction under Time Pressure. In *TABLETOP 2007*. pp. 37—43.