LMU

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Ph.D. Thesis, October 2007

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

von

Lucia Terrenghi

Designing Hybrid Interactions through an Understanding of the Affordances of Physical and Digital Technologies

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilans-Universität München

von

Lucia Terrenghi

Oktober 2007

Berichterstatter:

Prof. Dr. Andreas Butz
(Ludwig-Maximilians-Universität München, Deutschland)

Dr. Abigail Sellen
(Microsoft Research Cambridge, Großbritannien)

Prof. Dr. Bill Buxton
(Microsoft Research Redmond, USA, und University of Toronto, Kanada)

Tag des Rigorosums: 17 Dezember 2007

Acknowledgments

A relevant part of this thesis argues that the social and physical contexts affect one's subjective perception of the surrounding environment. Similar to that, I feel like I am at the end of a long journey, on which I have visited several places and met a lot of people who in different ways have made this trip possible. I have also had a rewarding and diverse set of experiences, which have affected my personal way of seeing the world. Like on a trip, whilst you learn about others and the places you visit, you also learn a lot about yourself. The people I associate with this trip are portrayed in a kind of virtual photo album, which is alive in my memory and emotions, and which will accompany me in my future travels.

I deeply thank my advisor Andreas Butz, who made my expedition possible and gave me a kind of return ticket to go and explore an exotic land called Human-Computer Interaction. He encouraged me, tracked my movements, and took care that I would not get lost or get into danger, whilst trusting me with the choices I have taken along the path. A warm thanks goes to Abigail Sellen for her very generous help: She has given me advice for the route to take, motivated me to reach further destinations, interpreted my thoughts and taught me how to express them in the language of academic research. And like on any important trip, I also met my "guru", who has inspired and enlightened my path: I thank Bill Buxton, whose passion for research, and most of all for life in general, has challenged me and transmitted me the sake for exploration. I feel extremely fortunate to have benefited from the diversity of my advisors' perspectives: They have stimulated my curiosity, and made my expedition rich of discoveries.

Along the way, I have experienced different situations: like exciting markets, whose colors, faces, scents and noises stimulated me and my creativity. I thank the smiles and faces I met in those inspiring places, those people who shared their time, thoughts, and their perspectives on my path: Abigail Durrant, Gerhard Fischer, George Fitzmaurice, Richard Harper, Shahram Izadi, David Kirk, Reinhard Oppermann, Dynal Patel, Thomas Pederson, Yvonne Rogers, Chia Shen, Ben Shneiderman, Albrecht Schmidt, Marcus

Specht, Alex Taylor, Stuart Taylor, and Ken Wood. Many of these people have become good friends and our acquaintance has influenced my path in diverse ways.

I have had the fortune of being hosted in a fruitful environment and constructive working atmosphere, with competent and supportive colleagues who accompanied me for much of my trip: I thank Heinrich Hussmann and my companions of the Media Informatics group Michael Aron, Richard Atterer, Dominikus Baur, Gregor Broll, Thorsten Büring, Alexander De Luca, Heiko Drewes, Rainer Fink, Paul Holleis, Andreas Pleuß, Enrico Rukzio, Sara Streng, Sybille Thomsen, Arnd Vitzthum, Raphael Wimmer, and especially Otmar Hilliges and Sebastian Boring, whose knowledgeable collaboration was very precious for the realization and elaboration of many of the ideas explored in this thesis. Likewise, I sincerely thank the students I worked with for their valuable creative contribution: Manuela Altendorfer, Mark Breisinger, Torsten Fritsche, Sebastian Kraiker, Sebastian Krämer, David Kim, Matthias Marm, Armin Prosch, and Hendrik Richter. And many thanks go to Amy Ko, for helping me with the English language.

During my trip, there have also been steep mountains to climb. Without the great psychological and emotional support of my friends I wouldn't have overcome the difficulties of the ascent. I thank the ones who pushed me further, sheltered me, and shared my accomplishments: Silvia Aramayo, Michela Bertero, Jacques Caillaud, Valentina De Tomasi, Ulrike Herb, Laurent Larivière, Magda Marchesotti, Sabrina Marozin, Sara Mazzetto, Anna Pou, Maddalena Prini, Silvia Raabe, Dagmar Schimanski, Mónica Solórzano, Cristina Tilotta, Carla Valle, Anne Wirth.

A special thanks goes to Timo Ropinski, who was by my side for a large part of this trip.

And, finally, I thank those very unique and special persons who did not come along, but always gave me a place to go back to, which I could call "home": I thank my brothers Francesco, Marco, Nicola, and my parents. Most of all I thank my mum, for being always there, whatever happened.

Thanks to the numerous people who contributed to my life and work during the past years. If their names have not found their way into the acknowledgements, their trace is somewhere hidden but vivid in the diary of my trip.

Abstract

Two recent technological advances have extended the diversity of domains and social contexts of Human-Computer Interaction: the embedding of computing capabilities into physical hand-held objects, and the emergence of large interactive surfaces, such as tabletops and wall boards. Both interactive surfaces and small computational devices usually allow for direct and space-multiplex input, i.e., for the spatial coincidence of physical action and digital output, in multiple points simultaneously. Such a powerful combination opens novel opportunities for the design of what are considered as *hybrid interactions* in this work.

This thesis explores the affordances of physical interaction as resources for interface design of such hybrid interactions. The hybrid systems that are elaborated in this work are envisioned to support specific social and physical contexts, such as collaborative cooking in a domestic kitchen, or collaborative creativity in a design process. In particular, different aspects of physicality characteristic of those specific domains are explored, with the aim of promoting skill transfer across domains.

First, different approaches to the design of space-multiplex, function-specific interfaces are considered and investigated. Such design approaches build on related work on Graspable User Interfaces and extend the design space to direct touch interfaces such as touch-sensitive surfaces, in different sizes and orientations (i.e., tablets, interactive tabletops, and walls).

These approaches are instantiated in the design of several experience prototypes: These are evaluated in different settings to assess the contextual implications of integrating aspects of physicality in the design of the interface. Such implications are observed both at the pragmatic level of interaction (i.e., patterns of users' behaviors on first contact with the interface), as well as on user' subjective response. The results indicate that the context of interaction affects the perception of the affordances of the system, and that some qualities of physicality such as the 3D space of manipulation and relative haptic feedback can affect the feeling of engagement and control. Building on these findings, two controlled studies are conducted to observe more systematically

the implications of integrating some of the qualities of physical interaction into the design of hybrid ones. The results indicate that, despite the fact that several aspects of physical interaction are mimicked in the interface, the interaction with digital media is quite different, which suggests the existence of mental models and expectations resulting from previous experience with the WIMP paradigm on the desktop PC.

Kurzzusammenfassung

Zwei aktuelle technologische Entwicklungen haben dazu geführt, dass interaktive Computersysteme in völlig neuen sozialen und physikalischen Situationen eingesetzt werden können. Diese Technologien sind in Alltagsobjekte eingebettete Computer und interaktive Oberflächen. Beide erlauben komplexe Eingaben (z.B. mit mehreren Fingern und Händen) und zwar direkt, also am gleichen Ort wie die zugehörigen Ausgaben (space multiplexing), was neue Möglichkeiten für die Gestaltung hybrider interaktiver Systeme schafft. Die vorliegende Arbeit untersucht den Designraum dieser Systeme.

Den Ausgangspunkt für die Gestaltung hybrider Benutzerschnittstellen bildet dabei das Konzept der "affordances" physikalischer Interaktionsobjekte. Die hier entwickelten Beispielsysteme unterstützen verschiedene soziale und physikalische Situationen, wie z.B. das gemeinsame Kochen in einer häuslichen Küche, oder die gemeinsame Ideenfindung in einem Designprozess. Dabei werden Metaphern eingesetzt, um verschiedene physikalische Aspekte der jeweiligen Interaktionssituation zu vermitteln. Hierdurch kann der Nutzer bereits erlernte Fähigkeiten auf die neuen Systeme übertragen.

Die Arbeit untersucht verschiedene Methoden, solche direkte und anwendungsspezifische Benutzerschnittstellen zu konzipieren. Dabei baut sie auf verwandten Arbeiten im Bereich anfassbarer Schnittstellen (graspable user interfaces) auf und erweitert den Designraum auf direkt berührbare (direct touch) Schnittstellen unter Verwendung interaktiver Oberflächen in verschiedenen Größen und Orientierungen, wie z.B. TabletPCs, interaktiver Tische und Wände.

Die verschiedenen Methoden wurden dann dazu eingesetzt, eine Reihe von Prototypsystemen zu bauen, die jeweils bestimmte Benutzungserfahrungen vermitteln (experience prototypes). Dabei wurde immer untersucht, welchen Einfluss die Integration physikalischer Aspekte in der jeweiligen Situation, sowohl auf pragmatischer als auch auf emotionaler Ebene, hat. Eine wesentliche Beobachtung ist, dass die Anwendungssituation maßgeblichen Einfluss auf die Wahrnehmung der jeweiligen "affordances" hat, und dass einige physikalische Qualitäten, wie z.B. haptisches Feedback und echte Drei-

dimensionalität insbesondere die emotionale Reaktion der Benutzer beeinflussen und dazu führen, dass diese sich stärker an der Interaktion beteiligt fühlen. Auf Basis dieser Beobachtungen wurden zwei kontrollierte Benutzerstudien durchgeführt, die den Einfluss physikalischer Interaktionselemente in hybriden Schnittstellen systematisch untersuchen.

Die Ergebnisse dieser Studien belegen, dass es nicht ausreicht, physikalische Aspekte in hybriden Schnittstellen nachzubilden, sondern dass die Interaktion mit digitalen Medien oft durch bereits vorhandene mentale Modelle aus der PC-Welt beeinflusst wird, die demnach für den Entwurf hybrider Benutzerschnittstellen genauso wichtig sind.

Contents

List of Figures

1 Introduction

This thesis explores and analyzes the integration of aspects of physical interaction in the design of hybrid interactive systems.

This chapter introduces the reader to the motivations, perspectives, and approaches which constitute the foundations of this dissertation. It then provides the context in which its contribution is to be considered and anticipates its structure and content.

1.1 Motivation and Design Space

The ubiquitous embedding of digital technologies in everyday environments and activities increases the complexity, variety, and occurrence of human interactions with digital media. Furthermore, advances in display and input technologies bring digital information and interaction possibilities to the very artifacts of our physical space, such as tables and walls. For input and navigation into the digital space, we obviously need physical handles in the analogue one, be they tangible (e.g., a mouse), or not (e.g., speech). In this sense, one can consider every kind of interaction with digital media as "hybrid" in nature, since it involves a physical as well as a digital component.

Within such a broad class of physical-digital interactions, this thesis focuses on interactions characterized by direct input. This can be effected using either a physical transducer, such as a stylus or some other physical device, or with fingers, by direct touch.

The handles we most commonly use are indirect input devices such as mouse and keyboard, which enable the manipulation of Graphical User Interfaces (GUIs) in the WIMP (Windows, Icons, Menu, Pointer) paradigm. But such a paradigm becomes inappropriate for coping with the heterogeneous, distributed, mobile, and multi-user scenarios of ubiquitous computing, and poorly reflects the richness of our physical interactions.

In the desktop PC (Personal Computer) environment, for example, the appearance of GUIs for widgets relies on visual cues and office related metaphors (see Fig. 1.1) in order to suggest affordances for mouse and keyboard interaction (e.g., 3D effects for clicking buttons, white fields for text entry, ripples on the moving part of scrollbars for dragging), as well as a conceptual model of the system (e.g., files and folders for hierarchical organization). When digital information is displayed for a different interaction style (e.g., gesture-based interaction) and is embedded in different domains and physical artifacts (e.g., interactive tabletops, wall displays), novel handles, affordances, and paradigms need to be designed for users to "get a grip" and "grasp" the conceptual model in the diverse contexts of hybrid interaction.

Figure 1.1: The problem statement: How can design afford manipulation and the mental model thereof when digital and physical realities interweave?

Previous work on Graspable User Interfaces (cf. (Fitzmaurice, 1996) and Chapter 3 for more details on this topic) had already grounded the basis for context-specific input paradigms adopting physical objects as transducers. As described by Fitzmaurice (1996) "A graspable function consists of a specialized physical input device which is bound to a virtual function and can serve as a functional manipulator." Thanks to the persistent association between a physical object and its function, graspable UIs reduce the number of phases of interaction, as illustrated in Figure 1.2, b. Indeed, while the mouse needs to be alternatively associated with different functions in different moments in time (i.e., it is a generic, time-multiplex input device), graspable UIs are specialized tools embodying a certain function, which has its physical representation in the space (i.e., they are space-multiplex input devices, cf. Section 2.4 in Chapter 2 for more details on this topic).

This dissertation investigates how the characteristic of "directness" of direct touch interfaces (such as the ones provided for example by interactive surfaces, cf. Section 3.1 in Chapter 3) can allow for a similar mapping between the acquisition of the interface (i.e., the handle for manipulation in the physical world) and the logical device (see Fig. 1.2, c). Furthermore, it explores how, if the shapes of the graphical UIs of direct touch interfaces suggest their functions, domain-specific tools can be designed that afford space-multiplex input as graspable UIs do.

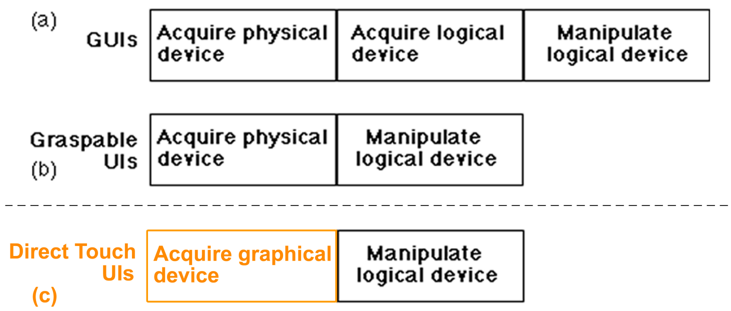


Figure 1.2: The upper part of the figure illustrates the reduction of the number of interaction phases as represented and articulated by Fitzmaurice (1996). The lower part (c) shows an extension of the same concept, explored in this thesis.

Hence, to create hybrid UIs that suggest their functions, this thesis investigates and explores the affordances of physical interactions as resources for the design of domain-specific tools, be they graspable or graphical representations on touch-sensitive surfaces. The goal is to inform the design of interactive systems so as to draw on such affordances, and finally to reduce the cost of users' transition from novice to experts.

1.1.1 The Cost of Transition

As the level of complexity rises, we, as users of technology, run the risk of being increasingly trapped in problem solving tasks at the operational level of interaction, rather than being supported by technology to leverage our skills and acquire new ones, unless our interfaces evolve in response to the increment of complexity. Put differently, from an ecological perspective, our tools need to support our adaptation to the environment and evolutionary process. This requires good design of what Hutchins (1995) defines *notational devices*, i.e., tools that are able to represent a given problem in a different way so as to facilitate its solution.

The adoption and skilled use of a tool inevitably implies costs, e.g., the costs in terms of time for learning and experience. Similar to the idea of technology transfer, if we are able to transfer a skill from one domain to another one, we can potentially reduce costs and bring innovation in the target domain. In other words, if we, as designers, are able to design tools that reflect existing human skills, we have a good chance of reducing the cost of acquisition of skill in the target domain. This is in line with Buxton's metaphor of the "three mirrors" as human-centered criteria for assessing design (Buxton, 1994). According to this thesis, the more a design reflects humans' existing physical, cognitive and social capabilities, the better the design supports such a skill transfer.

Coming from a design background, the most natural approach is to tackle the problem from the physical angle, i.e., from the pragmatic level of the input interface (as discussed in Chapter 2, Section 2.2). This deals with gestures, spatial, as well as device issues, and is the first level of contact between the user and the system: In other words, it is the handle for hybrid interaction with the digital space. Such a level has an impact on the whole experience of interaction and it is in the scope of this thesis to leverage an understanding of the properties - and the implications - of interactive systems at this level.

To this end, this work aims at understanding what it is about physicality, in terms of multi-sensorial as well as cognitive and emotional aspects, that affects the quality of hybrid experiences. In order to address those questions, the thesis will focus on:

- Identifying the affordances (physical, cognitive, functional, sensorial, and social affordances, cf. Chapter 2, Section 2.7) of physical media in a systematic way, so as to consider which aspects thereof can be integrated in the design of hybrid interactions;
- Understanding how such an integration in a specific hybrid context of use can be beneficial for the design of augmented, meaningful experiences, that go beyond the ones that are possible in the purely physical realm;
- Given that affordances are goal and context-dependent, understanding how the domain, the physical context, and the social context can affect the perception of such affordances.

A first step in this direction is to start unpacking some of the main qualities of physical interaction.

1.1.2 The Affordances of the Physical World for the Design of Interactive Systems

To understand how different aspects of the physical world can be integrated in the design of interactive systems it is necessary, first of all, to thoroughly examine what specific qualities of physical interaction could be drawn upon as design resources (whether this be consciously or not). For example:

- The use of the physical metaphor in the way objects and actions on those objects are represented. For instance, the desktop metaphor can be interpreted in various ways, as in (Smith et al., 1982), (Dragicevic, 2004), and (Agarawala and Balakrishnan, 2006).
- A direct mapping between input and output, so that an action produces feedback at the point where the input is sensed. This is typically the case for direct touch interactive surfaces, such as the Diamond Touch (Dietz and Leigh, 2001), Smartskin (Rekimoto, 2002), and TouchLight (Wilson, 2004).
- Continuity of action in input (as distinct from discrete actions or gestures), similarly to analogue interaction. This is the tenet of several marking interaction techniques, e.g., (Kurtenbach and Buxton, 1991b). For a discussion of the benefits of continuity see Buxton's work on chunking and phrasing (Buxton, 1995).
- 3D space of manipulation, enabling a high degree of freedom. In these cases, the movement of interactive objects is not necessarily bound to a surface. Hickley's passive props (Hinckley et al., 1994), for example, are designed for navigation of visual information in 3D.
- Physical constraints, that are provided, for instance, by the alignment and physical contact of material objects, or by the geometry of the devices. The ConnecTable (Tandler, 2001) and DataTiles (Rekimoto et al., 2001) are two examples illustrating this concept.
- Multimodal feedback, such as it is possible in the physical world. Haptic and especially proprioceptive[1] feedback, for example, are the basis

[1] *Proprioception* is the sense of position and movement of the limbs and the sense of muscular tension. The term is often used as alternative to the term *kinesthesis*. Propri-

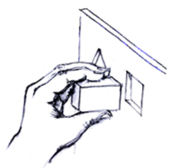
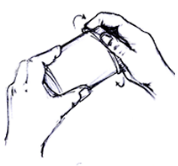
	**Metaphorical representation:** The shape of interactive objects symbolically evokes physical objects to suggest their functions and the way in which they can be manipulated.
	**Directness:** The point of physical input action is the same as the digital output one.
	**Continuity of action:** The actions required to interact are continuous, i.e., they do not need to be chunked in discrete sub-actions.
	**3D Space of manipulation:** Interactive objects can be manipulated in 3 spatial dimensions, i.e., not necessarily on a 2D display only.
	**Physical constraints:** Geometrical constraints, gravity law, or friction, for example, are integrated in the design to suggest how the interactive object can be manipulated.
	**Multimodal feedback:** The interaction with the object provides multimodal feedback, e.g., haptic (be that passive or active), audio, etc.
	**Two-handed cooperative work:** Objects can be interacted with two hands, asymmetrically.

Figure 1.3: Affordances of physical interaction integrated in the design of hybrid interactive systems of much of the work on tangible computing (Fitzmaurice et al., 1995), (Ishii and Ullmer, 1997). Likewise, physical tools can be used for sound generation and control as in (Jordà et al., 2007).

- Two-handed cooperative work to interact with virtual objects, thus exploiting spatially distributed input. Early work on this is exemplified by Bier et al.'s ToolGlass technique (1993). Other examples often occur in the design of interactive objects, such as Toolstone (Rekimoto and Sciammarella, 2000) and Tuister (Butz et al., 2004).

For an overview of these qualities see Figure 1.3. Starting from an analysis of how existing related work has built on such different aspects (cf. Chapter 3), the work presented in this thesis explores how to draw upon them in the design of hybrid interactions.

1.1.3 The Design of Hybrid Interactions

The term *hybrid* has been used on different occasions in the HCI (Human-Computer Interaction) literature to indicate the coupling of heterogenous interaction properties. Feiner and Shamash (1991) developed the concept of "hybrid user interfaces" to combine the form factors and resolution capabilities of different display technologies, i.e., conventional PC displays and 3D virtual environments perceivable through head-mounted displays. Rekimoto and Saitoh's work (1999) on "hybrid surfaces" shifts the focus from virtual to augmented reality. From the authors' perspective, the space can be considered as a display continuum, in which people can interact with both digital and physical displays/surfaces, the latters being augmented through front projection. Fitzmaurice's (1996) work on graspable user interfaces (cf. Chapter 3, Section 3.2) directly couples physical objects to the manipulation of virtual information, thus "hybrid objects" were conceived: These are meant to merge the physical and virtual affordances for the manipulation of digital media.

In this thesis, the term "hybrid" is used to indicate interactions that directly couple physical (analogue) and virtual (digital) worlds, and that integrate some of the aspects identified in the previous paragraph. The designs presented in this dissertation (see Fig. 1.4 for some examples) alternatively integrate several of those qualities, and three of those in particular:

oception is one element of haptic feedback: another one is *touch*, or *somesthesis*. This consists of somatic sensibilities aroused by stimulation of bodily tissues such as the skin. For more discussion on these topics cf. Section 4.3.

Figure 1.4: Some examples of the designs of hybrid interactions presented in this thesis. From left to right, the Mug Metaphor Interface, the Learning Cube, the Living Coobook, the EnLighTable.

- **Metaphorical Representation**: The visual representations (and/or physical shapes) of the user interfaces designed in this thesis, be they graphical or graspable, metaphorically evoke the ones of physical artifacts. These are referenced in different ways:
  - for their manipulation vocabulary in the physical world, thus providing cognitive affordances for the manipulation of digital media in 2D (e.g., the Mug Metaphor Interface) and/or physical affordances for manipulation in 3D (e.g., the Learning Cube), cf. Chapter 4;
  - for their role/function in the specific domain the interface is designed for (e.g., the Living Cookbook, cf. Chapter 5, or the EnLighTable project, cf. Chapter 6).
- **Directness**: This work focuses on interactions which are not based on a pointer whose control is remotely operated by the user. Rather, like in the analogue world, the type of interactions which were designed implies that users directly touch the display of information. This can be done with fingers as well as with other physical transducers, such as a pen or a tool, and implies the contact of two materials, i.e., a part of our body (e.g., our hand or fingers) or an extension thereof (e.g., a pen) together with the display surface. In this sense, interfaces for interactions with a mouse or a light pen on a WACOM tablet, for example, are not in the main scope of this work, as they imply a remote interaction with the display surface.
- **Continuity**: Whilst one can consider the mouse-click as a discrete interaction (i.e., it is a binary interaction, either we *do* click something or not), sliding, rotating, and pushing are actions which expand along continuous dimensions (e.g., the length of a sliding path, the angle of the arc of a curve, the depth of pressure, and the speed of rotation).

> Sliding the mouse on a 2D surface is a continuous action, but is not a direct one, as we do not directly operate on the information display of our focus. On the other hand, typing the characters of a soft keyboard on a touch-sensitive display is a direct, but not continuous interaction.

In the type of hybrid interactions addressed in this work, one can think of surfaces and 3D objects as interfaces, and hands as controls. In this sense, the main differences between hands and mice as operating tools, for example, need to be taken into account. A first simple difference is that while the ratio between the pointer and the display sizes remain constant in a mouse-based interaction (i.e., the pointer area displayed on a screen scales proportionally to the screen size), in a hands-based interaction the ratio varies as a function of human metrics. Furthermore, hands allow for multiple simultaneous input points (i.e., space-multiplex input, cf. Chapter 2, Section 2.4): Considering how we manipulate physical objects, we can easily notice hands' asymmetric cooperative work (for a good explanation see Guiard's (1987) theory of hands' kinematic chain). For instance, we usually hold a jar with the non-dominant hand and open the lid by rotating it with the dominant one. Additionally, the fact that there is no spatial distance between physical input and digital output also implies other types of issues, such as occlusion, preciseness of input, and visual angles.

Building on these considerations, one can then start distinguishing some main dimensions to define and position the design space in relation to other types of hybrid interaction paradigms, as described below.

1.1.4 Defining the Design Space

The previous paragraph has anticipated one of the dimensions, directness, that one can consider in order to characterize different types of interactions based on the type of physical (spatial) relationship between the user and the interface at the pragmatic level.

Another dimension is the persistence of the association of a transducer with a virtual function (i.e., space-multiplex vs. time-multiplex input, cf. Chapter 2, Section 2.4). Space-multiplex interfaces can provide handles which are specific for the task at hand. In these cases, the transducer, be it graspable (e.g., Ullmer and Ishii's (1997) models and lenses in the MetaDesk interface) or graphical (e.g., Butler and St. Amant's (2004) HabilisDraw), can have a shape and/or perform its function consistently with its use and manipulation vocabulary in the physical space. One can then talk of *semantic continuity* of the transducer. On the other hand, a physical cube like in the case of the Bricks project (Fitzmaurice et al., 1995), for example, can alternatively be associated with different functions depending on the context. In this case, its interaction vocabulary is *diverse*, and the binding between physical shape and virtual function is looser. Additionally, a transducer can be malleable, thus implying that the user can change its shape, as for example in the cases of Piper et al.'s (2002) Illuminating Clay and Balakrishnan et al.'s (1999a) ShapeTape.

Based on these dimensions, one can then define a taxonomy of interaction paradigms in order to delimit and collocate the design space (see Fig. 1.5).

			direct	indirect
space-multiplex	physical shape	semantic continuity	- MetaDesk (Ullmer and Ishii, 1997) - HabilisDraw (Butler and St. Amant, 2004) - Designers' Outpost (Klemmer et al., 2001)	- Props (Hinckley et al., 1994)
		diverse	- Bricks (Fitzmaurice et al., 1995) - Flatland (Mynatt et al., 1999) - DataTiles (Rekimoto et al., 2001)	- Toolstone (Rekimoto et al., 2000) - Navigation Blocks (Camarata et al., 2002)
		malleable	- Illuminating clay (Piper et al., 2002)	- ShapeTape (Balakrishnan et al., 1999)
time-multiplex	physical shape	diverse	- light pen on tablet PC	- mouse - Wacom graphic tablet - touch pad

Figure 1.5: A taxonomy of hybrid interaction paradigms. The design space is highlighted in orange.

Wanting to explore the interaction with computing technologies in specific domains and social contexts beyond the multi-purpose PC environment, the designs elaborated in this thesis focus on function-specific, space-multiplex interfaces, which integrate different affordances of physical manipulation to suggest affordances for hybrid ones. In this respect, the tendency is towards appliances showing a metaphorical representation of the transducer, be it graphical or graspable, which holds a semantic continuity of its manipulation vocabulary in the physical world. The area which is explored by design is highlighted in the schema of Fig. 1.5. Most of the interaction paradigms inserted in such a taxonomy are considered in Chapter 3 in further detail, thus clarifying the boundaries of the design space.

1.2 Context

This thesis aims at contributing to the field of Human-Computer Interaction (cf. Section 1.2.2). The work has developed within the FLUIDUM research project [2], at the Ludwig-Maximilians-Universität München, in Germany. The goal of the project is to study interaction techniques and metaphors for differently scaled ubiquitous computing scenarios within everyday life environments. The setting in which a large portion of this work is implemented is an interactive room, which is instrumented with large interactive displays, both vertical and horizontal, a steerable projector, as well as several other mobile displays in different formats. Such a set-up has been used to instantiate and experience some of the designs for interaction techniques that are presented in this thesis.

The following sections introduce the main research field and analytical perspective of this dissertation in order to position and contextualize its contribution.

1.2.1 An Agenda for Ubiquitous Computing

The term *Ubiquitous Computing* originated at the Xerox PARC Research Lab in the early '90s, as Mark Weiser was leading the Computer Science Lab. The term labeled Weiser's (1991) vision and a research program for a new era of computing technologies: The multi-purpose and centralized set-up of the Personal Computer was to be augmented by a distributed architecture of connected computing devices, which are specialized and embedded in the environment and activities of everyday life.

Two main aspects of this vision are particularly relevant in this context: invisibility and diversity of display sizes and functionalities. According to Weiser, indeed, computing capabilities will become so spread and embedded

[2]http://www.fluidum.org. The project is funded by the Deutsche Forschungsgemeinschaft (German Research Foundation).

in the environment as to actually "disappear", thus letting the user concentrate on the task and activity rather than on the tool in use (Weiser and Brown, 1995). In this sense, the disruption caused by the clumsiness of the multi-purpose PC, which like a Swiss army knife overwhelms us with its many functionalities, would be overcome by specialized, ad-hoc designed devices: These would disappear in the same way a pen, for example, does not steal our attention in the action of writing.

As a consequence, ubiquitous computers were to come in different sizes which are suited to particular tasks. Weiser classifies them into *tabs*, *pads* and *boards*, and metaphorically relates them to paper formats, i.e., "inch-scale machines that approximate active Post-it notes, foot-scale ones that behave something like a sheet of paper (or a book or a magazine), and yard-scale displays that are the equivalent of a blackboard or bulletin board". According to their formats, they are more or less suitable to mediate different activities, types of information, as well as social contexts.

These thoughts have motivated further investigation in different directions. The distribution of displays and devices has stimulated, for example, the design of scenarios for multi-user interaction, thus affecting the field of CSCW (Computer Supported Cooperative Work). The goal of unobtrusiveness of computing technologies and invisibility has inspired some of the work on context-modeling and context-adaptive systems. Thus, a number of different agendas have evolved within the field if HCI in response to the ubiquitous interweaving of digital technologies in the physical environment.

Starting from Weiser and Brown's vision of "calm technology" (Weiser and Brown, 1995), the Ubicomp (UBIquitous COMPuting) agenda has first focused on invisible technology and context-aware systems. From that perspective, automation has been considered as one of the most promising features of ubiquitous technology, pushing the design of "smart objects" and "smart environments" which could adapt to users' implicit intentions and even anticipate them. In this case, the need of users' physical actions for input and interactions is minimized, and most of the control is delegated to the system. On the other hand, the difficulty of modeling and predicting the users' dynamic and often "irrational" context in such a way that reliable inferences can be drawn has highlighted the limitations of such an approach.

Other critical approaches, e.g (Gaver et al., 2003), (Shneiderman, 2003), (Rogers, 2006), and (Terrenghi, 2006a), have suggested alternative perspectives from which to look at the relationship between users and interactive environments. In these cases, the migration of usage scenarios of digital technologies and applications from the traditional office domain to a variety of other ones - such as the home, the school, and the city - has been considered for its potential to engage users in novel hybrid experiences and interactive contexts whose value does not necessarily lie in efficiency or automation of activities. From these perspectives, technology can be seen as a tool for reflecting upon, interpreting and interacting with the environment, rather than the other way around. In other words, users play a more proactive role in the interaction.

The agenda pursued in this thesis is in line with these ideals. Its claim is that the design of such interactive technological tools needs to comprehend the expressiveness of the physical world as design resource. Thus, this work takes into consideration a diversity of physical and social contexts, as well as a diversity of domains (e.g., a domestic kitchen, collaborative brainstorming, and graphic design) in the design of hybrid experiences. When designing for such domains, the focus is on the exploration and analysis of how the design choices that are taken at a pragmatic level of the interface are perceived and can have an impact on users' subjective experience of hybrid interaction (e.g., exploration of the interface, communication, cognitive and evocative associations). The observation and assessment of such subjective experiences can in turn raise an understanding of human values and needs in those specific domains, thus informing the design of hybrid interactive systems which strive for diversity and engaging interactions. From a design perspective, one can then begin to tease out aspects of Information Technologies (IT) which solicit and encourage self-expression and creativity on the part of the users.

1.2.2 HCI from a Design Perspective

This thesis aims at contributing to the Ubicomp agenda from a design perspective, within the field of HCI. The relationship between HCI and design can be explained by distinguishing between *field of research* and *discipline*. The ACM Curricula for Human-Computer Interaction (Hewett et al., 1997) defines HCI in the large as an interdisciplinary field of research, emerging as a specialty concern within several disciplines, each with different emphasis. The field arose from the evolution of the relationship between computer and behavioral sciences. In the '80s, in particular, the introduction of Graphical User Interfaces and of direct manipulation (cf. Chapter 2, Section 2.6) pushed this field forward in the investigation and design of the computing technologies with which we are familiar today, e.g. the Personal Computer.

Along with the distribution of Information Technology in everyday life activities, targeting different user groups and enabling novel scenarios of computer mediated social interactions, HCI has progressively embraced other disciplines, such as sociology, anthropology, as well as design. The contribution of the design discipline and of design thinking to the field has been encouraged in different venues and publications, e.g. (Winograd, 1996), (Shneiderman et al., 2002), (Zimmerman et al., 2004), and (Zimmerman et al., 2007), but the issues of how to effectively formalize and integrate design approaches and methods in the HCI field remains a topic of discussion still today. Despite several attempts, no universal agreement has been reached for the definition of design research (see (Laurel, 2003) for a review).

In order to appreciate the specific contribution of the design discipline to the field of HCI, it becomes appropriate to identify, as a first step, the main aspects that distinguish design from other disciplines.

As by Simon (1969), design is based on "making", since it contributes to a body of knowledge about artificial objects and phenomena designed to meet certain desired goals. Shaping and making a concept evident (i.e., representing a design vision) is essential to designers for introspective as well as for interpersonal communication, in alternative and iterative phases of problem setting and problem solving in the design process. One could then say that designers "create to communicate", while other disciplines "communicate to create", articulating knowledge in a more explicit and prescriptive form.

In order to create, designers are trained to sketch, bricolage, model, prototype, and present so as to externalize and communicate a design concept, i.e., a message. Sketching, in particular, has been recognized as the archetypal design activity, characterizing designers' way of thinking (Arnheim, 1993), (Fallman, 2003), and (Buxton, 2007b), and shaping their cognitive process (Gedenryd, 1998), (Suwa and Tversky, 2002), (Goldschmidt, 1991). In Figure 1.6, Buxton's (2007b) sketch illustrates Goldshmidt's (1991) conceptualization of sketching as a dialogue between the mind and the sketch. By creating a sketch, we externalize our ideas (as we see them in our minds): By reading a sketch, we can visually reason on this idea and acquire new knowledge through the interpretation of the sketch. Designers are trained in both creating and reading sketches, which shape their minds in a distinctive way. Such a practice is not part of the educational program of other disciplines, but it is probably right in the different analytical (reading) approach to sketches that lies the main difference between design and engineering mind-sets. The first one is mostly based on association, thus bringing to generative thinking; the second one is more logic, thus leading to a reduction of alternatives. One of the most interesting points, here, is the fact that a sketch can be read by different minds, thus offering a diversity of interpretations. In this sense, as discussed in (Buxton, 2007b), there is a social value in the openness and ambiguity afforded by the lack of refinement and by the evocative style which are characteristic of sketches.

Because of the dynamic nature of the hybrid interactions that are considered in this thesis, sketches which are merely based on pen and paper are often a starting point, but not sufficient to convey more articulated de-

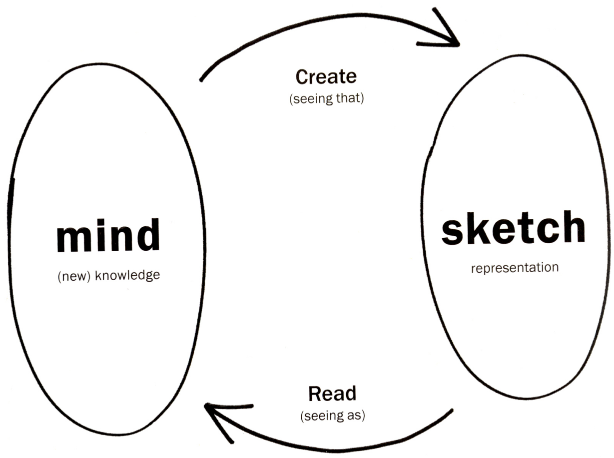


Figure 1.6: The conversational relationship between the mind and the sketch as represented by Buxton (2007b) and described by Goldschmidt (1991).

sign concepts. That means that they have limitations for externalizing and representing visions of interactions which unfold across time and space, and for sharing such visions with others. Thus, this thesis presents sketches of interactions and experience prototypes realized in different ways and with different levels of refinement, for different audiences and domains.

With respect to the design contribution to HCI, one can then conceive of design research as a discipline dealing with the creation of artifacts (sketches, experience prototypes, and probes) which can raise the understanding of a certain problem and can facilitate the communication among design stakeholders in iterative phases of problem setting and problem solving. In this sense, design research also promises to contribute to the identification of users' benefits and identification of requirements through the creation of artifacts which are open for users' expression of needs and preferences. Such creative and iterative activities, which distinguish the design practice, are reflected in the approach adopted in this dissertation.

1.3 Approach

The approach that follows is explorative as well as empirical and can be described as consisting of three main activities: 1) Sketching interactions; 2) Reading sketches; 3) Comparing and assessing. These activities are described in the following paragrph.

Figure 1.7: Examples of transitions across different levels of refinement of an idea for photo browsing on an interactive tabletop: a) Exploratory sketches with pen and paper; b) A tool made of cardboard representing a tangible UI: By placing it on a PC screen and taking a sequence of photos of different screen-shots, a storyboard of the interaction was created; c) The prototype implemented on an interactive tabletop.

1.3.1 Sketching Interactions, Prototyping Experiences

The sketching of interactions and the creation, reiteration and analysis of experience prototypes instantiating interaction design concepts has been the basis for the work presented in this thesis and the communication thereof. In this sense, referring again to Buxton's (2007b) approach, this work presents "sketches of experiences", which were refined to different levels (see Fig. 1.7) so as to share design concepts, reflect upon, and learn from them.

Starting from an identification of the different aspects of physical interaction (cf. Fig. 1.3), and building on an analysis of how related work has integrated some of those in the design of hybrid ones, the design work here presented draws upon the consideration of physical artifacts, spaces, and interactions in a number of ways. First, analogue technologies are considered in order to extrapolate some of the qualities thereof that characterize and affect human interactions and experiences; secondly, different aspects of physicality are metaphorically referenced or integrated in the design of experience prototypes for different types of domains (e.g., kitchen, design, and domestic environments) and different social contexts of use (e.g., simultaneous vs. asynchronous shared interactions). In doing so, those designs metaphorically reference physical artifacts at different levels of the interface design, from pragmatic to conceptual (cf. Chapter 2, Section 2.2), and exploit different aspects thereof as metaphorical sources: e.g., their manipulation vocabulary in the physical world, the way they are used in a social context, and the way they mediate interpersonal communication for utilitarian as well as decorative purposes (e.g., picture frames or mirrors on the home mantelpiece).

In such an approach, qualities of physical technologies are considered in the light of their potential for being coupled or augmented with digital ones so as to create meaningful hybrid experiences, which could extend humans' communication, creativity, and self-expression possibilities. These should go beyond what is possible in the purely physical reality, but still build on an understanding of humans' needs and values, and consciously exploit humans' mental models of how things (physical and digital) work for them.

By evaluating those sketches, such values and mental models can be elicited or further understood, as discussed below.

1.3.2 Reading Sketches

Each design was evaluated in order to "read" the potential benefits of those interaction sketches and experience prototypes, as well as to learn from them. In this sense, those designs acted as probes and test-beds, i.e., research tools for validation as well as elicitation of design issues to be considered, and for leveraging an understanding of users' expectations and mental models of hybrid interaction.

The methods that were used for assessment were diverse, mostly qualitative and applied in different settings due to the different technologies supporting their implementations. Whilst smaller devices such as tablets or graspable UIs were tested in situ (e.g., the Learning Cube, the Living Cookbook, and the Time-Mill Mirror), large and multi-display appliances (e.g., the Mug Metaphor Interface, the EnLighTable and Brainstorm) had to be tested in the lab because of the heftiness of the display hardware. Also in these latter settings, the trials were designed to be as plausible as possible, either in terms of target users, or in terms of tasks. This is because, as suggested by Hutchins (1995), the physical and social contexts of interaction affect users' mental models of how the tools are supposed to be operated, their functionalities, as well as the users' whole experience of use.

Therefore, when evaluating those designs, the focus was set both on the user interface at the pragmatic level and on the user experience provided by such prototypes. It was then taken into consideration whether or not (and how) users appropriated the manipulation vocabulary of the interface, and how this influenced their subjective assessment of the experience of use, thus trying to understand the link between interaction patterns and perception of experience.

1.3.3 Comparing and Assessing

The evaluation of the experience prototypes and their consideration in relation to other existing work suggested the identification of some critical issues for the design of hybrid interactions: These are mostly concerned with aspects of manipulation in 2D vs. 3D, and different aspects of tangibility.

To assess the relevance of those aspects and their design implications, two empirical studies in controlled experimental settings were conducted. In such comparative studies the critical variable was represented by the integration - or not - of some of the qualities of physical interaction (cf. Fig. 1.3) in the design of hybrid interfaces for direct input. The results indicate that despite the fact that several aspects of physical interaction are mimicked in the interface, the interaction with digital media is actually quite different. In this respect, the findings suggest that people's previous experiences with the WIMP paradigm on the desktop PC affect their interaction behaviors and their expectations of functionalities and location of computing technologies, within the ecology of everyday life activities and socio-physical contexts. This fact provokes a reflection on users' interaction attitudes, expectations and associated values for leisure, social technologies.

1.4 Contribution

The main contribution of this thesis lies in deepening an understanding of how the physical component of hybrid interactions has an impact on users' mental models and on the whole user experience. Thus, its goal is to inform the design of hybrid interactions by shedding light on how we can draw upon the affordances of the physical world as resources for the design of interactive, hybrid systems which are envisioned to support humans' social and creative capabilities beyond the office domain.

With its approach, this thesis seeks to contribute to the field of HCI in several additional ways. By addressing physical and social contexts beyond the office environment and work activity, this research provides an opportunity for investigating and reflecting on human values and needs in a diversity of domains. Such an investigation can in turn inform the design of hybrid experiences which engage users and strive for the augmentation of humans' self-expression and creative possibilities.

Furthermore, through the design of experience prototypes, this work aims at suggesting a variety of design solutions integrating elements of physical interaction in different ways. Hence, it provides both designers and engineers, as well as users, with an opportunity for "reading" the sketches it presents, thus stimulating multiple, complementary interpretations. In this respect the value of the design process is acknowledged: This lies in the visibility of its transitions/iterations (e.g., Fig. 1.7). Such transitions are a fertile terrain for a democratic involvement of design stakeholders.

Finally, the generation of comparable alternatives and their assessment in empirical studies suggest an approach for analyzing the effects of some specific qualities of physical interaction on users' interaction behaviors and mental models. In this sense, this thesis seeks to contribute to the theoretical knowledge of HCI and provokes a reflection upon the values that could/should be addressed in emerging scenarios of ubiquitous computing.

1.5 Thesis Outline

Although the sequentiality of a manuscript can hardly reproduce the iterative nature of the learning process articulated in this thesis, its structure tries to guide the reader through the reasoning, explorations, design choices, and reflections which were taken in such a process.

First, in **Chapter 2** the theoretical foundations underlying the thesis analytical perspective are defined and discussed. These basic concepts are fundamental as background for the review of the related work (cf. **Chapter 3**) because they underscore the reasons why those specific examples were considered. Such a review also distinguishes which elements of physical interaction are integrated in the design of hybrid systems of different scales (i.e., interactive surfaces, objects, and environments) so as to present and critically analyze a spectrum of alternative solutions.

Drawing upon these considerations, **Chapter 4** presents the exploration of design concepts through sketches of interactions: These metaphorically build on the manipulation vocabulary and conceptual models of physical artifacts for creating manipulation affordances of digital media in 2D (i.e., on interactive surfaces, such as the Mug Metaphor Interface, (Terrenghi, 2005)), or in 3D (i.e., with interactive objects, such as the Learning Cube appliance, (Terrenghi et al., 2006c)).

These alternative design approaches are further investigated and assessed through the design of two hybrid artifacts for social engagement in the home, i.e., the Living Cookbook (Terrenghi et al., 2007a) and the Time-Mill Mirror, cf. **Chapter 5**. The evaluation of their experience prototypes in situ, as kinds of probes, elicits methodological and design implications. These are concerned both with the representation of the interface at the pragmatic level, as well as with people's expectations, values, and benefits for engaging home technology.

Chapter 6 presents further instantiations of design concepts metaphorically integrating aspects of physical interaction for enhancing the manipulation of digital media in 2D. In this case, the focus shifts to large interactive surfaces and to their physical affordances for shareability: These are considered for their potential for being augmented in order to support and foster collaborative creativity in hybrid environments of interaction. Thus, the experience prototypes of the EnLighTable appliance (Terrenghi et al., 2006a) for collaborative photo editing, and of the Brainstorm appliance (Hilliges et al., 2007) for collaborative problem solving are presented. Their evaluation contributes to leveraging an understanding of how interactive surfaces and environments can affect collaborative creative processes, and raises some considerations on users' expectations of transducers' interaction vocabulary in hybrid interactive systems for creativity support.

Chapter 7 reflects on the issues raised by the evaluation of the interaction sketches and experience prototypes, in relation to other existing literature. Based on these considerations, two controlled comparative studies are presented which assess the implications of integrating some specific qualities of physical interaction into hybrid ones. Such implications are discussed both in terms of different affordances (Terrenghi et al., 2007b), as well as in consideration of how those differences impact mental models, interaction behaviors and subjective perception of experience.

Finally, **Chapter 8** summarizes the work presented in the thesis and discusses the lessons learned in terms of methodology and design. Hence, it articulates the thesis contribution and its implications for future work.

2

Underlying Concepts

This chapter introduces the theoretical background and analytical perspective of this thesis. Basic concepts of interface design are defined and discussed. As such, experts in this field can proceed to the next chapter.

2.1 Conceptual vs. Mental Models

The definition of *conceptual* and *mental models* helps in this context to describe a designer's role in shaping a system image, i.e., an interface, in relation to the target users of the system. Norman (1983) distinguishes between the two models: "Conceptual models are devised as tools for the understanding or teaching of physical systems. Mental models are what people really have in their heads and what guides their use of things". The concept was further discussed later on by Norman (1988) and in much of the HCI literature to explain the relationship between designers, system, and users (see Fig. 2.1). Designers develop a conceptual model of how a system should work, and try to convey such a model through the system image in order for it to appear understandable and coherent to the user, i.e., to suggest a certain mental model. Users perceive the system image and develop their mental models of how the object works and is interactable. A comprehensive definition, which takes the system appearance into account, is provided in by Preece et al. (2001): "A conceptual model is a description of the proposed system in terms of a set of integrated ideas and concepts about what it should do, behave and look like, that will be understandable by the users in the manner intended [by the designer]".

This thesis explores how the representation of the system beyond its mere visual appearance has an impact on the creation of users' mental models. The aim, as anticipated in the introduction, is to better understand the implications of the pragmatic level of the system on the conceptual one: Such levels are defined in the next paragraph.

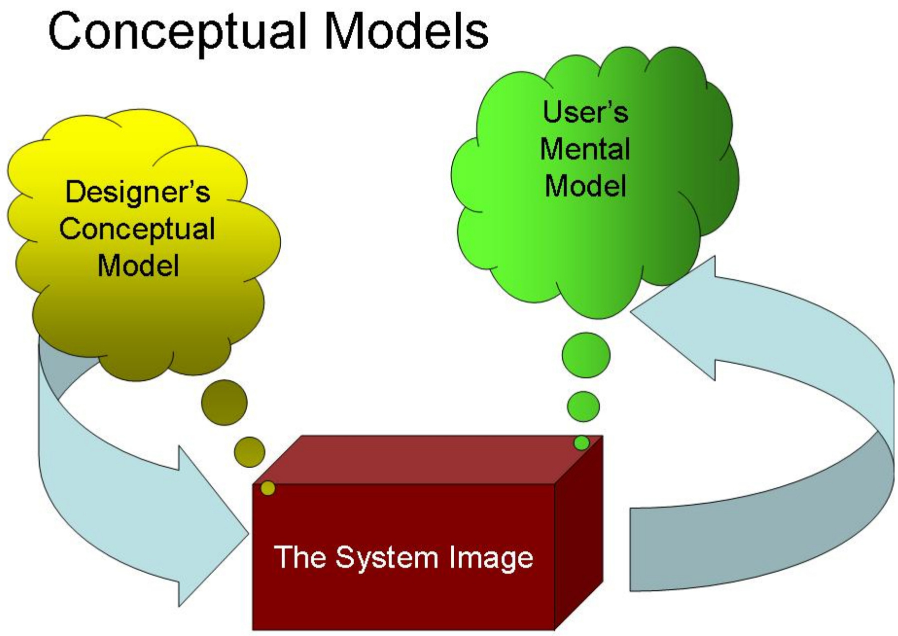


Figure 2.1: Conceptual Model: The system image reflects the designer's conceptual model and is supposed to suggest a mental model to the user. Image from (Norman and Draper, 1986).

2.2 Pragmatic vs. Lexical Level of Interface Structures

The designer of a system needs to consider and create different levels of an interface structure for users to develop a mental model which is coherent with the conceptual one. This section defines such levels so as to support the comprehension of the relationships between them and their implications on the design space of user interfaces.

Borrowing from linguistics models, Foley and Van Dam (1982) first proposed a layered structure for the analysis and definition of the design space of user interfaces for interactive systems. In their top-down model, they distinguish:

- the conceptual layer, which describes the main concepts of the interactive system as it is seen by the user;
- the semantic layer, which defines the functionality of the system, sequences of user actions and system responses;
- the syntactic layer, which defines interaction tokens (words) and how to use them to create semantics;
- the lexical layer, which describes the structure of these tokens.

This model provides the possibility of comparing different systems at different levels. As discussed by Buxton (1983), though, the level of detail of the lexical level is too coarse to describe the properties of the tokens in relation, for example, to the physical space, to other tokens, and to the user. This fact, in turn, does not allow for an elicitation and analysis of the implications of those properties on the overall interaction with the system. To this end, Buxton (1983) suggests a distinction between lexical and pragmatic levels:

- lexical level: issues having to do with the spelling of the tokens (e.g., the alphabet they use);
- pragmatic level: issues of gesture, space, and device.

In accordance to Buxton's and others' work in this area (e.g., Fitzmaurice's (1996) and Hinckley's (1997)) it is a tenet of this thesis that the physical component of the interface, which is incorporated in the pragmatic level, contributes to conveying a mental model of the system: Thus, it has a relevant impact on the other layers, and ultimately on the overall conceptual one and on the subjective experience of use. To this end, some of the main qualities of physical interaction were unpacked in this work (cf. Chapter 1, Fig. 1.3) in order to analyze and compare different systems at the pragmatic level (cf. Chapter 3), as well as to make informed choices in the design of the experience prototypes of hybrid systems which are presented in this thesis.

2.3 Epistemic vs. Pragmatic Actions

If we are to consider physical interactions as resources for the design of interactive systems, we need an understanding of physical actions. Kirsch and Maglio (1994) distinguish *epistemic* vs. *pragmatic* actions: "Epistemic actions are physical actions people take to uncover information that is hidden or difficult to compute mentally". As such, they are meant more to simplify people's internal problem solving processes, rather than to bring people physically closer to an external goal (the latter being a pragmatic action).

Epistemic actions enable cognitive offload through externalization of mental computations. Some examples for it are the use of fingers while counting, or the arrangement of ingredients and tools on the kitchen counter for cooking preparation.

Pragmatic actions, on the other hand, are meant as physical actions that serve to achieve the goal of transforming physical states: For example, turning a door handle to open a door and walk through it to enter a room. The benefits of epistemic actions are not to be considered in terms of motor effort, but rather of cognitive one. The authors discuss and identify such benefits in more detail in terms of:

- space complexity, i.e., reduction of the memory involved in mental computation;
- time complexity, i.e., reduction of the number of steps in mental computation;
- unreliability, i.e., reduction of the probability of error of mental computation.

The cognitive advantages of epistemic actions are reinforced as we consider the sensorimotor theory of perception (O'Regan and Noe, 2001). The main account of such a theory is that perception does not happen in the brain, seen as a black box, but rather it is something humans do as explorative, motor activity. For any stimulus which can be perceived, there is a set of motor actions which will produce sensory changes regarding this stimulus. In TVSS (Tactile-Visual Sensory Substitution), for example, one human sense (tactile) is used to receive information normally received by another human sense (visual). This implies that the more an interface facilitates epistemic action, the more our senses can synergetically process information, thus potentially leading us to a more cost-effective and reliable construction of a mental model (in terms of cognitive effort).

Our interface to digital information has traditionally been embodied by mouse, keyboard, and computer screen. In turn, our experience with digital information has mostly been shaped by the WIMP interaction paradigm of the desktop PC, which relies mainly on our visual capabilities and less so on our auditory capabilities. Furthermore, the continuous way in which we manipulate and explore physical artifacts in the analogue world does not find an equivalent counterpart in the interaction with digital environments.

Here, our manipulation vocabulary is usually reduced to the degrees of freedom provided by a mouse. These facts end up limiting our possibilities for exploration through physical activities and,therefore, our epistemic actions and the sensorimotor perception of the object we interact with.

This thesis explores different ways of supporting epistemic actions through the design of interfaces that extend the possibilities of physical actions normally afforded by mouse and keyboard at the pragmatic level. One way to do so is by taking advantage of some of the advances in touch-sensitive surfaces and context-sensitive technologies (for example, cf. Chapter 4): These, indeed, provide novel possibilities to design interfaces for space-multiplex input and, therefore, a richer manipulation vocabulary, as discussed below.

2.4 Space-Multiplex vs. Time-Multiplex Input

The distinction between *space-multiplex* vs. *time-multiplex* input is instrumental in this context to recognize some of the differences between the ways in which we normally manipulate physical vs. digital media. The mouse, for example, is a time-multiplex input device, as it alternatively defines its function through the selection of different modes, which enable us to manipulate digital media. Our interactions with mundane physical artifacts, on the other hand, are mostly space-multiplex, as we can use different parts of our body (e.g., in bimanual interaction, such as playing the piano, or writing on a piece of paper) to interact with different parts of the object, simultaneously. In this way, space-multiplex input provides opportunities for parallel execution of operational tasks (e.g., two-handed interaction), thus potentially improving performance.

For a clear distinction between space-multiplex vs. time-multiplex input devices one can refer to Fitzmaurice's (1996) definition: "With space-multiplex input, each function to be controlled has a dedicated transducer, each occupying its own space. [...] A space-multiplex input style affords the capability to take advantage of the shape, size and position of the multiple physical controllers to increase functionality and decrease complexity. [...] In contrast, time-multiplexing input uses one device to control different functions at different points in time. Hence, the device is being repeatedly attached and unattached to the various logical functions of the GUI."

Such a distinction is important to recognize and exploit some of the advantages brought along by emerging technologies, such as multi-touch interactive surfaces, as well as context-sensitive sensors that can be embedded in interactive physical objects (cf. Chapter 3 for more detail). These provide novel opportunities for the design of manipulation vocabularies that are more articulated than the one afforded by the WIMP paradigm, and that can better reflect humans' manipulation skills with physical objects. Furthermore, the use of space for physical manipulation affords epistemic actions (Kirsh, 1995): The physical representation of a variety of arrangements and temporal structures contributes to the externalization of our cognitive computations of alternative arrangements. Additionally, it enables space-multiplex feedback and, therefore, a richer sensorimotor perception of the object (e.g., proprioceptive haptic feedback).

It is necessary to bear in mind, though, that increasing the number of spatial locations where input is simultaneously possible leads to a trade-off between physical manipulation and cognitive computation costs. Indeed, as discussed by Fitzmaurice (1996), "the amount of effort and attention needed to manipulate the physical objects must be less than the internal cognitive computational effort to make it attractive to use". In other words, the operational cost of interaction must be decreased as much as possible by the interface to enable users to perform bigger chunks of actions at a functional level, as is discussed below in further detail.

2.5 Skill Acquisition and Skill Transfer

As anticipated in the introduction, this thesis investigates how the transfer of some aspects of physical interactions (e.g., the space-multiplex input capabilities discussed above) in the design of hybrid ones can contribute to suggest a mental model coherent with the system conceptual one, so as to result in a lower cost of skill acquisition.

Decreasing the cost of skill acquisition, i.e., the cost of transition from novice to expert, implies a reduction of the cognitive resources which a user needs to invest in the *means* (i.e., operational problem solving) rather than in the *content* (i.e., functional problem solving) to accomplish the task (Buxton, 2007a). Given that humans' cognitive resources are limited, an expert (i.e., a skilled user) typically has a low operational cost, so that s/he can invest more resources in the functional solution of the problem. The way in which an expert approaches a problem can indeed benefit of the automatization of some actions at the operational level. This, in turn, implies a different segmentation of the task into sub-tasks, i.e., a different granularity of the clustering of the actions to perform in order to solve the problem (Buxton, 1995). Typically, an expert can perform bigger chunks of actions to accomplish the task, which can have an effect, for example, on the time of task performance.

Now, if we, as designers, are able to conceive of a representation of a conceptual model of the system that at the pragmatic level reflects the way in which an expert would approach the problem at an operational level, we are likely going to maximize compatibility (i.e., the way in which a system responds to users' expectations (Buxton, 2007a)) and, in turn, users' performance. In other words, we can facilitate the dialogue between a user and the system by creating affordances which are coherent and consistent with users' set of skills, thus in turn shaping users' mental model of the system.

As suggested by Buxton (1994), the motor, cognitive, as well as social skills which we have acquired in the experience with physical reality provide us with a reach material for skill transfer in the design of hybrid interactions. To this end, we clearly need to investigate how chunks of actions can be represented and supported by the interface, at the pragmatic level. The use of interaction metaphors which go beyond the design of visual cues for mouse and keyboard input constitutes an interesting approach in this respect.

2.6 Interaction Metaphors

Metaphors were already used in written languages in early writings, as in Sumerian epics. Throughout history, the concept of metaphors has been developed in different fields, especially in philosophy, literature, cognitive linguistics and interaction design, with mutual influences among the fields. In cognitive theories, metaphors are presented in the form of "A is B", where B is said to be the source (or *vehicle*) and A is the target (or *tenor*). Generally, the source metaphor is determined based on the common knowledge in the real world and the target is the complex entity (often an abstract concept) that should be represented. In Black's interaction theory of metaphor (Black, 1972), a metaphor is not simply a process of transferring properties from the source to the target, but a complex interaction between them, in which our knowledge of the target is equally across the target and the source. Thus, the properties that are highlighted by the comparison are determined by the interaction of the tenor and the vehicle.

Such a theory has had most influence on the following cognitive linguistics theories. Lakoff and Johnson's (1983) work *Metaphors We Live By* is the most popular example in this area. Their main account is that a large part of the human conceptual system is metaphorical itself. Thus, metaphors are not just constructs of a language, but they rather constitute some fundamental tools for human reasoning and understanding. Metaphors underly our reasoning structure and can shape our perceptions and actions by creating relationships between what we know about our physical and social experience and other topics. In this sense, the authors argue against the idea of a priori objective similarities: Metaphors do not just *point out* existing similarities, but rather *create* them.

From these ideas, both metaphorical as well as anti-metaphorical approaches to interaction design have generated. On the one hand, metaphors have been recognized to play an important role in user interface design as descriptions and representations of conceptual models, e.g. (Carroll and Thomas, 1982), (Madsen, 1994), (Laurel, 1986), (Marcus, 1998), and even as essential tools for design thinking (Erickson, 1995). From these views, they can serve as means for making an unknown complexity into an understandable format as they provide an intuition of how things work, transferring the world knowledge.

On the other hand, the lack of a priori, objective similarities implies interpretation, which is subjective. This has motivated the skepticism of other approaches towards the value of metaphors: Users' cultural differences, their different experience and levels of skill, as well as metaphors' scalability and coherency, are some of the main concerns of these accounts, e.g. (Richardson, 1993) and (Nelson, 1990). According to Halasz and Moran (1982), for example, the use of metaphors is limiting and inappropriate to teach new users how to interact with computing systems. These, indeed, have a syntactical complexity which cannot be efficiently learned and understood by analogy. Rather, the authors propose the use of abstract conceptual models for teaching users how to reason about the system.

Another thesis considering the relationship between metaphors and learning is the one presented by Cooper and Reimann (2003). From the authors' perspective, most elements of intuitive Graphical User Interfaces are actually *idiomatic*. Idiomatic interfaces are based on the learning of simple, non-metaphorical and behavioral idioms (i.e., principles) to accomplish goals and tasks. Intuition, on the other hand, is a mental comparison between a new experience and things we have already learned. For Cooper and Reimann, "windows, title bars, close boxes, [...] are things we learn idiomatically rather than intuit metaphorically".

The point here is to recognize that a large part of idioms are actually what, in linguistic terms, one would call *dead metaphors*. These are metaphors which have been so much embedded in a culture or a language, that their sense of a transferred image has become unnoticed. A curious observation is that several dead metaphors are actually embodying a physical action as a source of transition (em-bodiment and under-standing are metaphors themselves). In this sense, it becomes artificial to some extent to draw a hard line between metaphorical intuition and idiomatic learning. One could argue, indeed, that the learning activity is in large part metaphorical itself, as we relate new concepts to other knowledge we already have (which is often implicit knowledge, that comes from our physical experience) in order to reason and create meaning in a semantics, i.e., in a language. The difference, rather, relies in the granularity of the analysis, i.e., whether we consider the chunks (the parts that constitute the metaphors) or the idioms (the phrases consisting of chunks). In other words, it makes a difference whether we just consider metaphors which directly emulate the source entity into the target domain, vs. lateral metaphors which transfer an aspect (e.g., a physical gesture) from a domain to another one to convey new meaning.

In this sense, despite their limitations, we can consider the PC desktop metaphor and the direct manipulation paradigm as successful examples of transferring knowledge and gestures (e.g., the tension in dragging an object from a location to another one) from the physical world into the digital one, thus representing the complex command-line structure of computing systems into a more accessible format. The next section introduces these concepts.

2.6.1 The Desktop Metaphor

The computing domain has been characterized by the nearly universal acceptance of the desktop metaphor for decades. In 1981, the Xerox Star workstation set the stage for the first generation of Graphical User Interfaces (Smith et al., 1982), establishing a metaphor which simulates a desktop on a bit-mapped screen and is operable with mouse and keyboard. The Star also set several important HCI design principles, such as *seeing and pointing* vs. *remembering and typing*, and *what you see is what you get*. The Apple Macintosh brought this new style of interaction into the public's attention in 1984 (Williams, 1984), creating a new trend in the PC industry which was further widespread through the large diffusion of Microsoft Windows.

Although this was a fundamental contribution to the enhancement of human-computer interaction, the limited vocabulary of the pragmatic level has somewhat restricted the semantics of interaction with digital media. The emerging scenarios of ubiquitous computing drive the design of alternative computing tools and novel usage paradigms encompassing multi-user and multi-display environments, as well as multiple input/output modalities. In this respect, one can expect that the concept of direct manipulation - which has been the basis of the desktop metaphor and the WIMP paradigm - will evolve, as novel technological possibilities for direct input interfaces become available. The principles of direct manipulation are explained below.

2.6.2 Direct Manipulation

In the Personal Computer environment, *direct manipulation* describes the activity of manipulating objects and navigating through virtual spaces by exploiting users' knowledge of how they do this in the physical world (Shneiderman, 1987). The three main principles of direct manipulation are:

- continuous representation of the objects and actions of interest;
- physical actions or presses of labeled buttons instead of complex syntax;
- rapid incremental and reversible operations whose effect on the object of interest is immediately visible.

Direct manipulation is the basis for the dominant WIMP paradigm, with which we manage different applications. According to the activities they support, applications rely on different metaphors. In the Microsoft Office software package, for instance, visual and auditory icons mimic the objects of a real physical office. In software programs for graphic design, icons resemble brushes and pencils. While the metaphor varies according to the application domain, the general paradigm does not change as the appearance of widgets for desktop GUIs remains consistent. Graphic elements are mapped to objects of the real world and those in turn provide affordances for mouse and keyboard interaction, e.g. 3D effects for clicking buttons, white fields for text entry, and ripples on the moving part of scrollbars for dragging (cf. Chapter 1, Fig. 1.1).

Despite talking about direct manipulation, in the Desktop environment we mostly need indirect input devices, such as mice, track pads or joy-sticks, to interact with the system. As interactive environments become more complex, encountering a variety of displays, both physical and digital, as well as a diversity of input and output modalities, novel affordances for the manipulation of information need to be designed and encoded.

2.7 Affordances

Ecological approaches focus on perception and action based on human attributes to investigate the interaction between an agent and the environment in which s/he is situated. In this context, *affordances* were first described by Gibson (1979) as a property of the relationship between an actor and a physical artifact in the world, reflecting possible action on those artifacts. Over time, this concept has extended to the field of product design first, with

Norman's book "The Psychology of Everyday Things" (1988), and to interaction design later on (Gaver, 1991), to indicate properties of an interface which suggest its use.

In moving from physical to digital artifacts in the design space, Norman (1999) specifies the concept of affordances distinguishing between *real affordances* and *perceived affordances*. The first ones are properties of an object which enable an action (e.g., a computer screen can be touched, looked at, pointed at, and one could possibly click on every pixel on the screen), while the second ones refer to cues (mostly visual cues) of an object which suggest its use (e.g., the already mentioned 3D effect of a graphic suggests that clicking with the mouse on that area of the screen triggers a certain effect).

Perceived affordances are often relying on cultural conventions, or *stereotypes*, which we have learned as being part of a specific culture and/or domain (see Fig. 2.2 for an example). The establishment of such conventions in hybrid interaction (e.g., dragging downwards the ridge of a scrollbar implies the content of the window pane to move upwards) is one of the main challenges for interface designers. Indeed, the creation of a conceptual model can rely much less on physical constraints and "real" affordances, and rather needs to work much more on the design of "perceived affordances", which require people to learn. The use of stereotypes as metaphorical sources for the design of perceived affordances can be effective only if such stereotypes are coherent with users' language and existing mental models, i.e., if they are compatible with users' expectations. Otherwise, they generate interferences in the users' analogical reasoning.

a) b) c)

Figure 2.2: Different dining arrangements rely on different stereotypes. These spatial arrangements of artifacts can suggest how to sit and eat only to the members of the culture in which the stereotype/convention is embedded. Respectively: a) A table arrangement for eighty place settings from the 18th Century; b) The Lazy Susan of a sushi bar; c) One plate on the floor for sharing couscous among several hand-eaters.

Hartson (2003) builds on Norman's distinction and further explores the topic of affordances by providing a good classification for interaction design, based on a review of existing definitions. He distinguishes between *physical*, *functional*, *cognitive*, and *sensorial* affordances for interaction design. Physical affordances are design features which enable users to execute a physical action in the interface: These correspond in a certain way to Norman's real affordances. Functional affordances extend the concept to purposeful actions: I.e., functional affordances are design features which support users to achieve a certain goal when acting in the interface. Cognitive affordances are design features which "tell" users something, mostly about a physical action they can do with the interface, and usually through visual cues and conventions. Additionally, sensory affordances are defined as "design features that help users to sense something (especially cognitive affordances and physical affordances)". Auditory icons, such as the sound of crumpling paper which occurs when a digital document is dragged to the icon of a wastepaper basket, for example, can be considered as sensory affordances, which redundantly support the metaphoric conceptual model and its cognitive affordances. On top of this classification, Gaver (1996) explores an ecological approach to social interaction using the concept of affordances to describe properties of the environment that affect how people socially interact. *Social* affordances have then been defined by Kreijns and Kirschner (2001) as properties of collaborative environments which serve as facilitators of the social context. They enable and enhance social interaction between a member and the group.

These different aspects of affordances are summarized and illustrated in Figure 2.3. In Fig. 2.3, a, a physical context is represented, consisting of a table and two chairs. These artifacts afford the placement of objects on their horizontal planes, as well as sitting: These are real/physical affordances (Fig. 2.3, b). When the character ("John") gets hungry (i.e., has a motivation for interaction), he spatially displays artifacts in such a way to functionally support his goal: cutlery at the sides, plate in front, bottle and glass at a reachable distance, and chairs arranged accordingly (thus creating functional affordances, Fig. 2.3, c). The way in which he displays artifacts also relies on cultural conventions/stereotypes (e.g., cutlery and napkin at the right side of the plate). Furthermore, such a spatial arrangement indicates to a second person at which side of the table to sit, if s/he wants to eat (cognitive affordance). When the food is ready, John tries to call the attention of his partner ("Lisa") by hitting his glass with his knife. The smell of the food, as well as the sound produced by John, will be for Lisa multi-sensorial cues to understand that the food is ready (sensorial affordances, Fig. 2.3, d). The way in which John has arranged the artifacts will suggest Lisa where to sit and will facilitate the social context (social affordances, Fig. 2.3, e). As seen from the example, cognitive affordances can as well be social, and be reinforced by multi-sensorial ones. This thesis alternatively refers to different

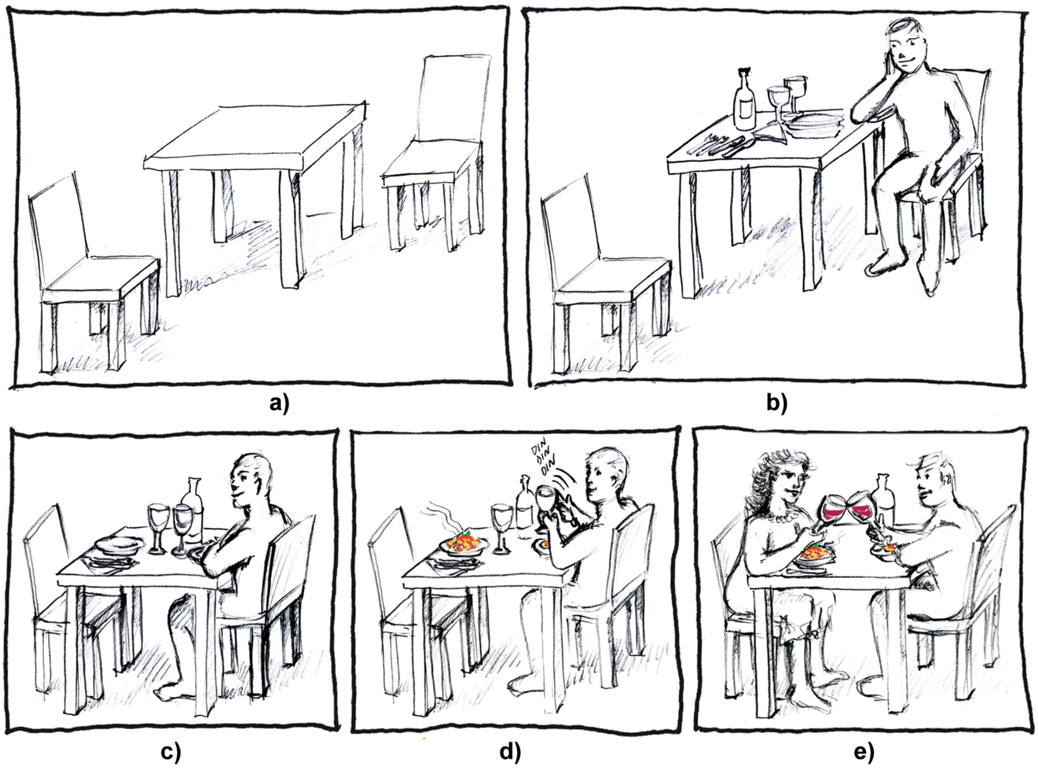


Figure 2.3: Different types of affordances: a) A physical context; b) Real/physical affordances; c) Functional and cognitive affordances; d) Multi-sensorial affordances; e) Social affordances.

types of affordances, without specifying the type it is referring to all the time: The distinction is important, though, for sharing with the reader the vocabulary which is used in this dissertation.

This survey of the HCI literature on affordances has also shown that the use of such a concept in the community has moved quite far from the original Gibson's definition (1979). This dealt with a purely physical relationship between an agent and the environment: "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill". There is another aspect of Gibson's definition that is important to consciously consider, that is: "[an affordance] implies the complementary of the animal and the environment. [...] It is equally a fact of the environment and a fact of behavior. [...] An affordance points both ways, to the environment and to the observer". In this respect, even though in the HCI community we refer to the *design* of affordances, and have found a way around the original concept by distinguishing *real* vs. *perceived* affordances, we need to be aware that affordances are kind of phenomena which emerge from the interaction between an agent and an environment/system, thus are not independent properties of either entities. This implies that what we actually design are opportunities for affordances to emerge and take place, and to this end we need to gather an understanding of both the agent and the environment, in relationship with one another.

2.8 Conclusion and Discussion

This chapter has reviewed different concepts of interaction design and discussed some of the approaches to the representation of conceptual models and interfaces for interactive systems.

Working with GUIs in the desktop metaphor, interface designers have concentrated almost uniquely on the design of cognitive affordances, by representing visual cues that suggest certain mouse actions. Staying with Norman's distinction between real/physical and perceived affordances (Norman, 1999), the cognitive, functional, sensorial as well as social affordances can be considered as perceived ones. It is in the scope of this thesis to investigate how the integration of real/physical affordances and constraints in the design of interactive systems can contribute to effectively represent the system image for the user. Its hypothesis is that by taking spatial issues, epistemic action, and haptic feedback into account, one can design affordances which potentially better reflect and further enhance users' existing skills - be their motor, cognitive, or social skills - and is compatible with users' existing mental models (cf. Section 2.1., and (Buxton, 2007a)).

A main aspect of physical affordances is that the physical attributes of the thing to be acted upon are compatible with those of the actor. Thus, representing affordances for the emerging multi-display, multi-user, multi-touch environments means to consider ergonomic aspects, such as dominant hands, hands size, users' height, reachability, territoriality, and so forth. The traditional usability guidelines for visual displays[1] will most likely need to be revised in order to address such aspects, and one can expect that interface design will need to merge graphic and product design competencies.

The next chapter considers how existing work in the field of interface design has coped with those aspects in various ways, when computing technologies are embedded in the spaces and artifacts beyond the desktop PC environments.

[1]ISO 9241-11:1998. Ergonomic requirements for office work with visual display terminals /(VDT) - Part 11: Guidance on usability

3 Related Work

This chapter provides an overview of how qualities of physical interaction (cf. Chapter 1, Fig. 1.3) have been used as a design resources in different interaction paradigms for ubiquitous computing; and of how, at the same time, computing technologies have been integrated into physical artifacts and spaces such as tables, walls, objects, and rooms. The survey is organized on the basis of the device form factor in which these paradigms are embedded.

3.1 Interactive Surfaces

This section considers the design of interfaces for the delimited real estate of interactive surfaces, i.e., surfaces which enable the direct spatial mapping between input and output. In this sense, it focuses on interfaces which do not rely on a pointer controlled by a remote input device, such as a mouse, a track-ball or a touch-pad. Nevertherless, it is worth mentioning in this context Sutherland's work on the Sketchpad (Sutherland, 1988) as a predecessor of the stroke-based interaction which often characterizes interactive surfaces. In Sutherland's work, the idea of using a touch pad for the recognition of strokes and pen gestures as input modality already aimed at supporting a more analogue and continuous style of interaction with digital information.

Sutherland's early attempt towards more fluid interaction techniques is further explored as pen-based interaction becomes possible directly on the screen, rather than on a remote surface. Thanks to the technological advances which started in the early '70s with the invention of the electronic touch interface by Samuel Hurst, touch screen surfaces have become gradually available, in different formats and relying on a wide variety of technologies (e.g., resistive, capacitive, camera based, infrared, and ultrasonic). The different underlying technologies have enabled different sorts of "touch" (see also Buxton's overview[1]): single vs. multiple input points, finger vs. pen, one-handed vs. two-handed, and single user vs. multi-user interactions. These aspects obviously have an influence on the design of the interface, both in terms of graphical appearance as well as interactive behavior. A first and early discussion of the design implications of touch screen interaction is presented by Nakatani and Rohrlich (1983). In that context, the advantages and trade-offs of touching interactive graphical UIs are discussed in comparison to interacting with hard, mechanical, and more specific machines.

The interaction with such interactive surfaces is often based on gestures, as it shall be discussed later on in further detail. Most gesture interactions using a stylus are qualified as *marking interactions*, that means "interactions where the pointing devices leave an ink trail on the display similar to writing with a pen" (Kurtenbach and Buxton, 1991b). Other types of gestures are the ones in 3D space which are based on the orientation, placement, speed, and direction of movement of the device itself as input modality. This type of interaction obviously occurs only in the case of smaller, portable devices, such as Personal Digital Assistants (PDAs).

This survey covers the solutions for interface and interaction design which, over time, have been developed in the field of interactive surfaces.

3.1.1 Portable Displays

Portable displays such as PDAs and tablet PCs are typically held with one hand and controlled with the other one. This implies that keyboard-based interaction can only happen with one hand, which obviously hinders the efficiency and comfort of text entry in comparison to the interaction with a desktop keyboard. Such considerations, together with size limitations and portability, are some of the factors which have motivated the use of direct touch and stroke-based interactions for portable interactive displays.

A first approach to text entry on touch screens is to display a soft keyboard: In this case, users need to select one character at a time by tapping it on the screen. Another approach is the one adopted in the Graffiti system used by Palm PDAs or in the Unistroke alphabet proposed by by Goldberg and Richardson (1993) (see Fig. 3.1, a). These techniques require users to learn how to draw characters and allows them to enter one character at a time. In this case, characters can be drawn on a specific part of the screen only, aiming to minimize the required screen space for text entry and to allow for a more analogue interaction at the same time.

[1]http://www.billbuxton.com/multitouchOverview.html

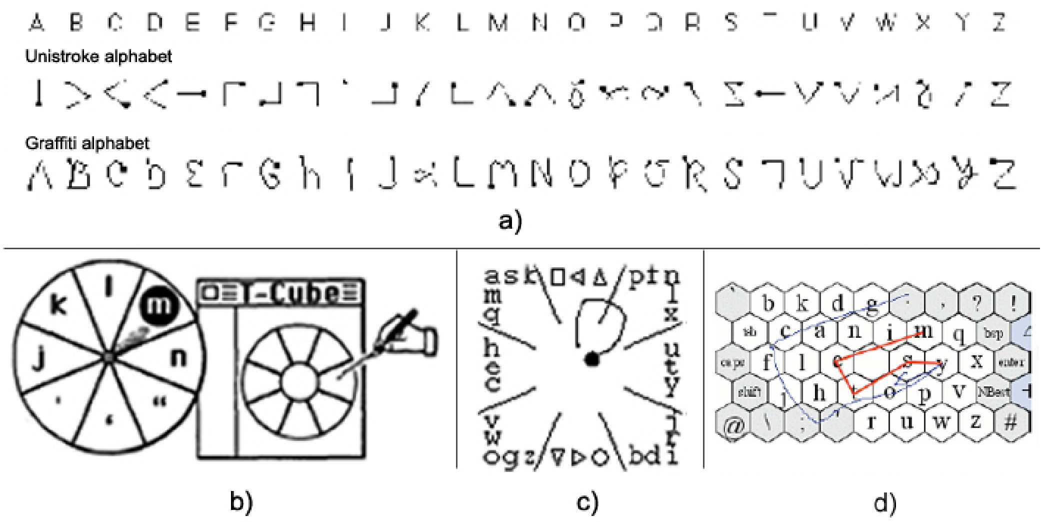


Figure 3.1: Different techniques for entering text with strokes on a display: a) The Unistroke and Graffiti alphabets; b) T-Cube (Venolia and Neiberg, 1994); c) Quickwriting (Perlin, 1998); d) SHARK (Zhai and Kristensson, 2003).

A solution in between, which aims at affording continuity without forcing for the user to learn a certain set of characters, is the one of marking menus for the selection of characters as well as commands (i.e., objects and verbs). Marking menus are extensively treated in (Kurtenbach and Buxton, 1991a) and (Kurtenbach and Buxton, 1991b), and build on earlier work on pie menus by Callahan et al. (1988). T-Cube (Venolia and Neiberg, 1994) and Quickwriting (Perlin, 1998) are but two examples of text entry techniques for mobile applications which are inspired by marking menus (see Fig. 3.1, b and c). In both cases, the characters can be selected from a radial array. The idea is that the muscular memory can support users to learn the gesture corresponding to an entire word, rather than to a single character. This is a good example of how a gesture can support different granularity of chunks (cf. the concept of chungking and phrasing in Chapter 2, Section 2.5).

SHARK (Shorthand Aided Rapid Keyboarding, by Zhai and Kristiansson (2003), in Fig. 3.1, d) is another project related to this work. This technique turns the stylus-based typing on the soft keyboard of a tablet PCs into shorthand fluid gesturing. Users can learn how to enter a word by memorizing a pattern of gestures on the characters of an ATOMIK keyboard (Alphabetically Tuned and Optimized Mobile Interface Keyboard). The technique has lately been extended by the same authors for entering short commands such as “copy” and “paste”, thus avoiding the invocation of pop-down menus (Kristensson and Zhai, 2007).

Another, more natural approach to text entry is to recognize normal handwriting on the device screen, such as Jot in the Microsoft PocketPC 6 operating system. The drawback of such an approach is that handwriting has a limited speed, i.e., it is not an efficient technique for entering longer texts. Furthermore, it is not always recognized, i.e., it is not always possible to correctly transform digital ink produced with stylus entry into machine readable text. Since the input of words can invoke commands and operations, their correct recognition is even more important than for characters. Indeed, when a gesture/word is misrecognized, it will cause an unintended operation to be performed and users may have difficulties determining what happened and, therefore, correcting the error (Long et al., 1998).

While handwriting recognition is still challenging to some extent, crossing gestures are technically easier to recognize, especially when supported by user interfaces unambiguously designed for this kind of interaction. The main principle of crossing-based user interfaces (Accot and Zhai, 2002) is that in order to trigger an action, the cursor needs to be moved beyond the boundaries of a targeted graphical object (i.e., a 1 dimensional bar), instead of pointing and selecting a 2 dimensional target area (see Fig. 3.2, a). This enables a continuous (e.g, drawing a line) rather than a discrete (e.g., pointing and clicking) style of interaction.Such analogue and continuous in-

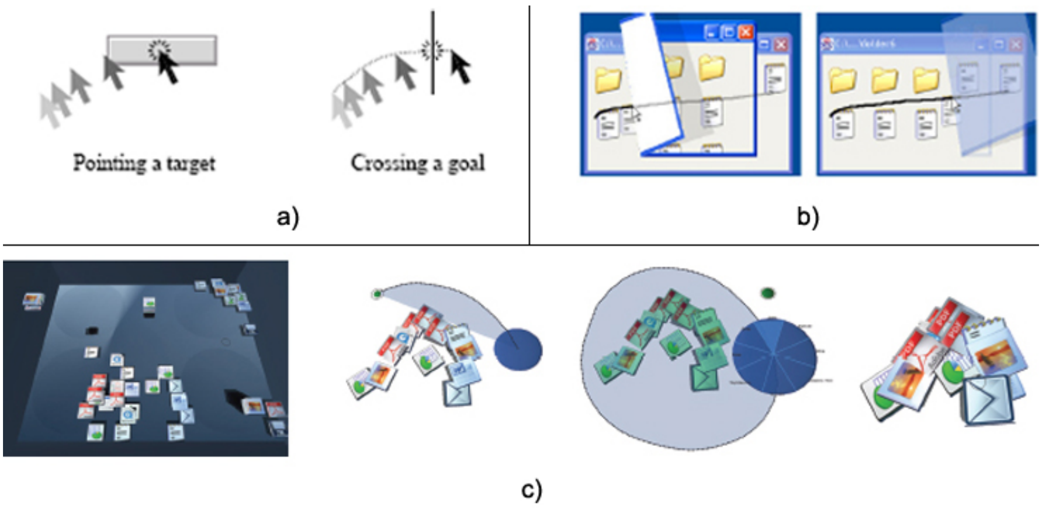


Figure 3.2: Crossing-based interfaces: a) The basic concept as explained in (Accot and Zhai, 2002); b) Folding windows with one mouse gesture as proposed in (Dragicevic, 2004); c) A screenshot of the Bumptop interface (Agarawala and Balakrishnan, 2006) and a sequence illustrating a crossing gesture for grouping physical-like icons.

teraction technique has been combined with more physical-like visualizations of graphical UIs so as to metaphorically hint the type of gestures to perform. An example in the WIMP paradigm is the fold-and-drop technique by Dragicevic (2004) (see Fig. 3.2, b), which in turn builds on Beaudouin-Lafon's (2001) peeling-back techniques .

These last examples still rely on mouse-based interaction, but the continuous characteristic of crossing-based interfaces is obviously of particular interest for the design of direct pen-based interactions (e.g., the Crossy drawing application (Apitz and Guimbretière, 2004) and the BumpTop UI for a tablet PC (Agarawala and Balakrishnan, 2006)). In the BumpTop user interface, a more physical and realistic representation of the desktop environment has been combined with crossing gestures for interaction (see Fig. 3.2, c). Such a visualization builds on previous work on pile metaphors (Mander et al., 1992). Physical behaviors are simulated so that icons can "be dragged and tossed around with the feel of realistic characteristics such as friction and mass" (Agarawala and Balakrishnan, 2006).

Another approach to direct manipulation in the desktop metaphor environment is the File Browser for two-handed interaction proposed by Ka-Ping Yee (2004). In this work, a touch screen display is combined with a tablet PC in order to enable pen-based as well as touch-based interaction, with respectively the dominant and non-dominant hands. The dominant hand, holding the pen, is used, for example, for drag and drop tasks, while the non-dominant hand is used for browsing through the windows in order to select the target folder, thus creating the reference frame (see Fig. 3.3, a). Similar to that, Matsushita et al.'s Dual Touch system for PDAs (Matsushita et al., 2000) is based on the tapping of graphical UIs with a thumb and on stroking with a pen. The thumb is used, for example, for mode selection, as anchor for rotation, as well as control of pop-down menus, while the dominant hand performs more fine-grained tasks with the pen (see Fig. 3.3, b).

a) b)

Figure 3.3: Examples of two-handed interaction for portable devices using both a finger and a pen for input: a) On a tablet PC (Yee, 2004); b) On a PDA (Matsushita et al., 2000).

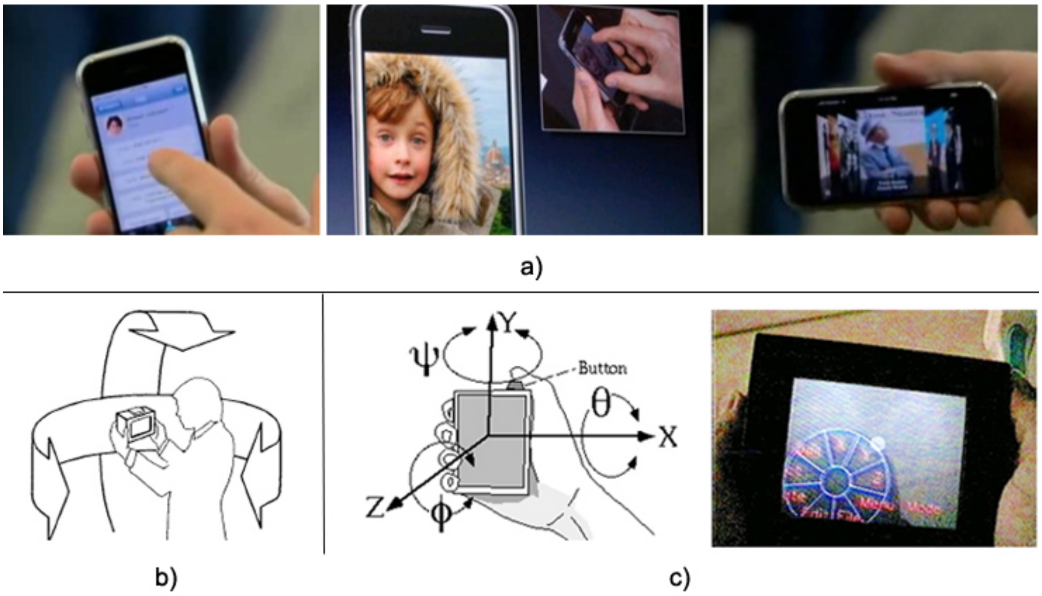


Figure 3.4: Direct touch and 3D gestures for handheld interaction: a) The Apple IPhone; b) The Camaleon concept (Fitzmaurice et al., 1993); c) The 3D gestures for interaction proposed in (Rekimoto, 1996).

Lately, the Apple IPhone[2] (see Fig. 3.4, a) has enabled a variety of tasks, such as text entry, scrolling, and resizing with a purely direct touch interface. Users can interact with one as well as two fingers (e.g., for scaling), and browse through e-mails, pictures, and music files by simply stroking the screen with their fingers, in some ways similar to the metaphorical gestures proposed by Pirhonen et al. (2002). When the device is tilted horizontally, the visual display reorients automatically. Tilting and other 3D gestures had been already used for a semantics of commands in other related work as well, for example (Fitzmaurice et al., 1993) (see Fig. 3.4, b), (Rekimoto, 1996) (see Fig. 3.4, c), and (Harrison et al., 1998). This survey won't get into the detail of 3D gestures interaction with portable devices such as PDAs and mobile phones because this is not really in the focus of this thesis. Still, it is note worthy, from a design perspective, that whilst direct touch interfaces imply constant visual attention in order to correctly tap the desired target, 3D gestures interfaces often enable eyes free interaction, as they can rely on human kinesthetic perception. This aspect is further discussed in this thesis with respect to the use of physical objects as transducers of interaction (cf. Chapter 4, Section 4.3).

[2]http://www.apple.com/iphone/

3.1.2 Wall Displays

The decreasing cost of display technologies has progressively made large, high resolution, interactive surfaces more available in a variety of domains, e.g. meeting rooms, classrooms, and control rooms. The affordances determined by the scale of conventional physical whiteboards, such as shared visualization, collaborative sketching of ideas, awareness, and presentation, have been augmented in several projects by the affordances of digital technology, such as data storage, history tracking, scaling, reproducibility, automation, and motion, just to name some.

The Tivoli project (Pedersen et al., 1993) is one of the first examples of systems for meeting support using a digitally augmented whiteboard. In the Tivoli interface, running on a LiveBoard, three electronic pens can be recognized at a time, thus enabling multi-user simultaneous interaction. Users can scribble and manipulate such scribbles as objects (e.g., they can copy, paste, and delete) or as real strokes (e.g., they can wipe parts of a scribble as if they were using an eraser). In order to switch from the scribbling mode to the gesture command mode (e.g., drawing a closed line for the selection of multiple items, as in Fig. 3.5, a), a button is pressed on the pen. The graphic layout of the Tivoli interface displays buttons in a fixed location, similarly to the desktop environment, where commands are operated with a pointer. On a large screen, this implies that the user needs first to reach the icon (by physically moving or stretching her arm) of the tool/mode she wants to use (e.g., a certain stroke line thickness), and perform the action afterwards.

In the Flatland project (Mynatt et al., 1999), such an issue is addressed with the use of pie menues (see Fig. 3.5, b) appearing in proximity of the input point. In this way, the user can invoke locally a set of functions by pressing a pen button and tapping the pen on the display, thus swapping between different interaction vocabularies of the same transducer, i.e. the pen. Whilst the Tivoli system is geared on supporting temporally isolated instantiations of a specific meeting, Flatland is meant for supporting continuous, individual office work. To this end, the tracking of the history of personal interaction has a lower level of granularity, till the stroke creation: In this way, the user can navigate back and forwards through her scribbling interactions, as it is shown in Figure 3.5, b.

FlowMenus (Guimbretière and Winograd, 2000) provide a solution similar to the one adopted in Flatland for invoking and interacting with a local menu on a wall screen. In this case, rather than a simple pie menu, a kind of multi-hierarchy marking menu (see Fig. 3.5, c) can be invoked right at the spot where the user is currently working. Such a menu enables localized, gesture-based fluid interactions for direct manipulation. The FlowMenus system

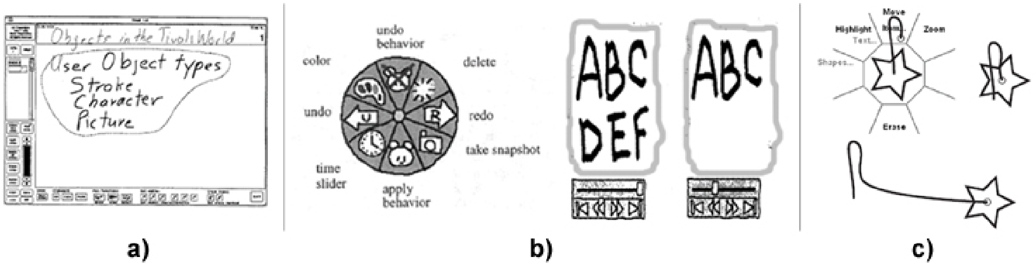


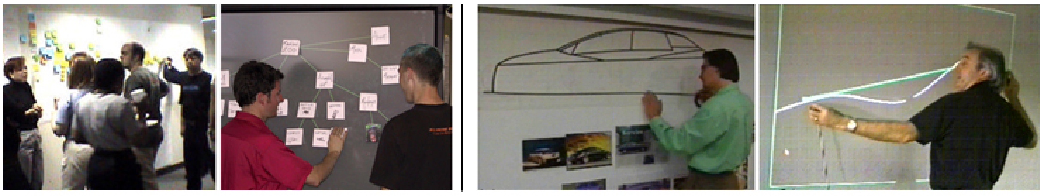
Figure 3.5: Pen-based gestures for large vertical displays: a) Grouping scribbles on the Tivoli interface (Pedersen et al., 1993); b) Left, the pie menu. Right, the history tracking control on the Flatland interface (Mynatt et al., 1999); c) The FlowMenu, which enables selection and movement with a single gesture (Guimbretière and Winograd, 2000).

has been extended for text entry with handwriting recognition in the Typed Drag and Drop technique, which is used in a brainstorming application as explained by Guimbertière et al. (2001).

The use of large interactive displays for supporting creativity and collaboration has been researched in different domains, which are often related to design. The Designers' Outpost system (Klemmer et al., 2001) is intended to support websites information architects to arrange their ideas, to structure them, and to record changes (see Fig. 3.6, a). By dynamically tracking the location and spatial relationships of paper Post-its on a SmartBoard[3], the system enables the designers to scribble their ideas and to naturally structure them by moving their paper notes on the board. In this way, the affordances of paper (e.g., immediacy and tangibility) are integrated and augmented in such a physical/digital system. Indeed, when users physically annotate their ideas by scribbling, this results in a tag in the digital database so that they can easily keep track of the project evolution and use automatic searching.

The Digital Tape Tool system (Balakrishnan et al., 1999b) also builds on a traditional paper-based design technique (i.e., curve-drawing with physical tape for car design) and augments it with the possibilities of digital media (e.g., storage, easy reproducibility, easy editing, and 3D effects) in a system for two-handed interaction (see Fig. 3.6, b). Finally, the Portfolio Wall (Buxton et al., 2000) is another example for the use of large displays in automotive design. The main goal, here, is to create a sort of shared visual server which displays the work going on in distributed design offices so as to establish mutual awareness among designers and project managers.

[3]http://www2.smarttech.com/st/en-US/Products/SMART+Boards/



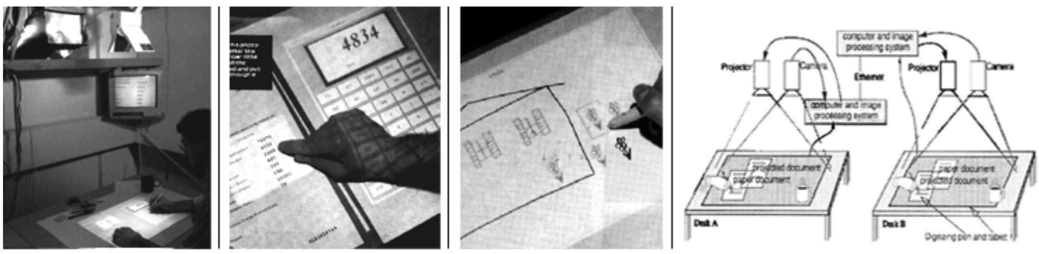
a) b)

Figure 3.6: Interaction techniques which are inspired by design practices in the physical world and are intended to support design work on large vertical displays: a) The Designers' Outpost (Klemmer et al., 2001); b) The Digital Tape Tool (Balakrishnan et al., 1999b).

3.1.3 Tabletops

Depending on their format and location, physical tables (e.g., coffee tables, desktops, meeting tables, and dining tables) typically support a wide variety of mundane activities: These can be individual as well as collaborative, and be related to work as well as leisure. Such activities have been augmented by digital technologies in a variety of ways and projects in the history of HCI.

Wellner's Digital Desk project (Wellner, 1993) explores the combination of paper-based documents together with digital functionalities. A normal table affords the support of physical objects, such as paper sheets, on its surface: These are captured by a camera (which also tracks users' interaction such as pointing with a finger, as shown in Fig. 3.7, b and c) and are augmented by a projector mounted on top of the table (see Fig. 3.7, a and d). Thus, digital functionalities such as search, copy, and paste become possible. The same concept underlies a calculator application (Wellner, 1991), a drawing application (Wellner, 1993), and a system to support remote collaboration (Freeman, 1996) (see Fig. 3.7, respectively b, c, and d).



a) b) c) d)

Figure 3.7: Wellner's Digital Desk (Wellner, 1993): a) The set-up; b) The Digital Calculator; c) The drawing application; d) The Double Digital Desk set-up for remote collaboration (Freeman, 1996).

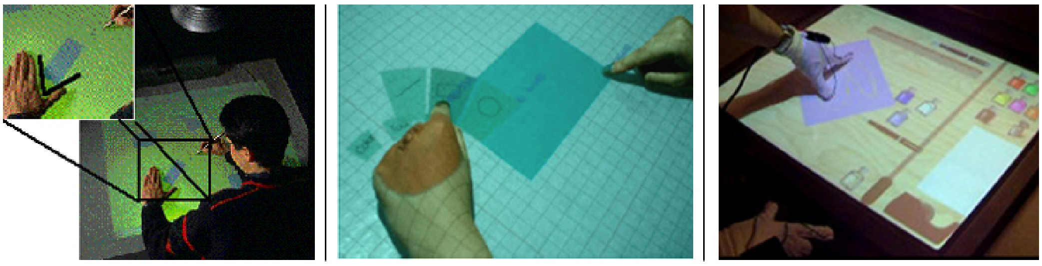


Figure 3.8: Interfaces for drawing on digital desks: a) One version of the Active Desk based on hands shape recognition (Buxton, 1997); b) Bimanual cooperative work in the drawing interface by (Koike et al., 2002); c) The HabilisDraw DT application (Butler and St. Amant, 2004) for two-handed interaction on the DiamondTouch.

Building on the physical affordances of tables, drawing and collaboration are indeed some of the main activities on which related research has focused so far. Fitzmaurice et al.'s Bricks application for the Active Desk (Fitzmaurice et al., 1995), for example, adopts physical handles for the manipulation of graphical shapes. In this work, the authors introduce the concept of Graspable User Interfaces, which is discussed in further detail in the next section. Another version of the Active Desk tracks the angle of openness between the thumb and the index fingers of the non-dominant hand (see Fig. 3.8, a), thus indicating whether the user is gripping a virtual tool (Buxton, 1997). Similarly, the distance between thumb and index is mapped to a grabbing gesture of graphical shapes by Koike et al. (2002). In such a system, the two hands work cooperatively: The non-dominant hand is used for displaying and manipulating a pie menu, while the dominant one is used for selection and drawing (see Fig. 3.8, b).

Two-handed cooperative work in a drawing application is also suggested by (Butler and St. Amant, 2004) in the HabilisDraw DT project. In this case, the system builds on the MERL DiamondTouch device (Dietz and Leigh, 2001) and a ceiling-mounted beamer for front projection. The DiamondTouch uses capacitive coupling to register a unique input from multiple users, recognizing the chair where the user is sitting. In each chair, indeed, a receiver is embedded which sends a signal identifying the parts of the table that are touched by each user. This fact, on the other hand, limits the user's input to either a single point of contact or to a bounding box around the contact point with the surface. For this reason, in the HabilisDraw DT system the authors designed a pair of gloves so as to provide two unambiguous input points per hand, i.e., at the thumb and forefinger of each hand (see Fig. 3.8, c). In this way, the user can have four input points, two per hand, and perform direct manipulation tasks such as grabbing, moving, scaling, and drawing with 2D tools, e.g., rulers, cutters, ink bottles, and tape. The interface is designed to metaphorically resemble a drafting table (see Fig. 3.8, c).

The DiamondTouch system has been used for a variety of projects investigating interaction techniques for tabletops. Wu et al. (Wu and Balakrishnan, 2003), (Wu et al., 2006), explore multi-fingers and whole hands gestures for interaction (see Fig. 3.9, a and b). Another approach to the recognition of hands' and fingers' shapes, as well as their location, is proposed in the Smartskin system (Rekimoto, 2002). This system uses capacitive sensing as underlying technology. The sensors, in this case, are embedded in the table surface itself, thus making possible the tracking of the distance between the hand and the surface too (see Fig. 3.9, c, top). This feature adds another dimension (i.e., depth in the 3D space) to the semantics of possible gestures.

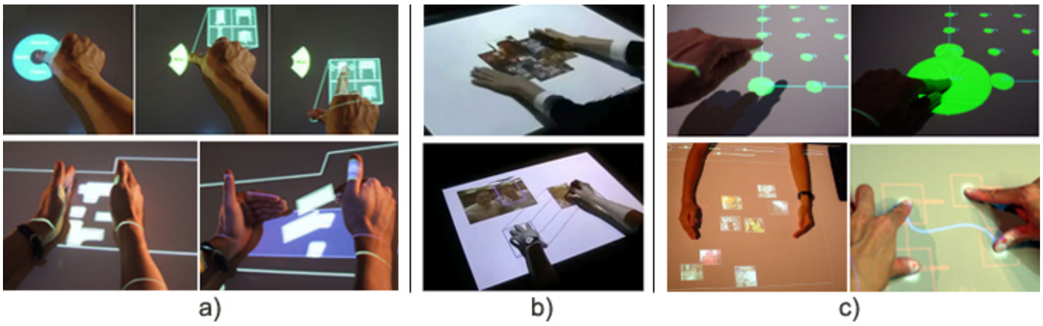


Figure 3.9: Examples of multi-finger and two-handed gestures for tabletops: a) A gesture vocabulary for interacting with a room planning application as described in (Wu and Balakrishnan, 2003); b) Two-handed gestures for bimanual cooperative work as described in (Wu et al., 2006); c) The Smartskin system, enabling the recognition of multiple fingers, hands shapes, as well as the distance between the hand and the surface (Rekimoto, 2002).

As already anticipated, collaboration is a main focus of the research on tabletops. (Scott et al., 2003), (Scott et al., 2004), (Kruger et al., 2003), (Inkpen et al., 2005), (Shen et al., 2006) have highlighted some of the main issues and challenges for the design of systems supporting co-located collaboration on interactive tables: Orientation of content, reachability, territoriality, occlusion, and allocation policies are some of those issues.

The problem of orientation is addressed by Shen at al. (2002) in the Personal Digital Historian project. In this work, the authors designed a circular interface for the sharing of a digital photo collection, building on

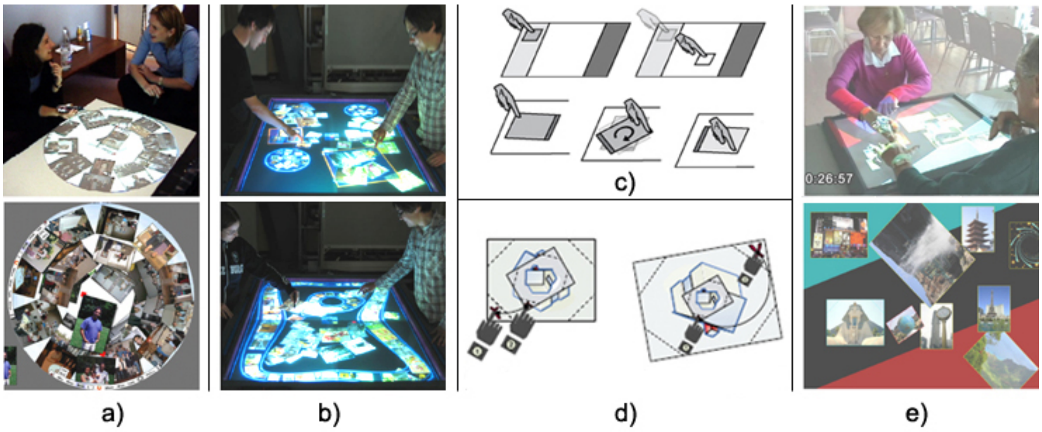


Figure 3.10: Examples of interface design addressing orientation and territoriality issues for supporting collaboration on tabletops: a) The Photo Digital Historian (Shen et al., 2002); b) On top, Storage Bins (Scott et al., 2005). Below, Interface Currents (Hinrichs et al., 2005); c) Reorientation and translation gestures for collaborative interaction in a shared area of the display, as suggested in (Ringel et al., 2004); d) The reducing/rotate gesture (left) and enlarging/rotate one in the SharePic interface (Apted et al., 2006); d) The layout design of the SharePic application for elderly users (Apted et al., 2006).

the DiamondSpin toolkit (Shen et al., 2004) for orientation. The radial coordinate system and its radial symmetry allow for no privileged point of view and are intended to support casual communication and collaboration. By interacting with a pen directly on the photo items, users can scale and explicitly reorient the content towards one-another, thus facilitating users' content-related communication and story telling (see Fig. 3.10, a). In this sense, the design provides social affordances for media exchange.

Other examples of user interfaces addressing orientation issues are Scott et al.'s (2005) Storage Bins and Heinrich et al.'s (2005) Interface Currents, shown in Fig. 3.10, b. These systems rely on the DVit technology by Smart Technologies[4] - which can track two single input points simultaneously - and propose solutions to both orientation and territoriality issues for co-located, collaborative photo browsing. In the Storage Bins interaction metaphor, temporary working areas can be casually created by drawing a closed line on the table: These areas can represent both personal and shared spaces of interaction. The Interface Currents interface combines such a system with a Lazy Susan metaphor. The photos move automatically along the perimeter of the table and orient towards its sides. Users can control the speed of the

[4]http://www.smarttech.com/dvit

flow and create shared bins in the center of the real estate, where photos are rotating too.

Apted et al.'s (2006) SharePic system for photo sharing similarly explores territoriality issues by dividing the table real estate in personal and shared areas (see. Fig. 3.10, e), thus relating also to Ringel at al.'s (2004) work on allocation policies (see. Fig. 3.10, c). By analogy to the bins metaphor, in SharePic black holes can be used in the shared area, in this case for deleting pictures. Whilst the Storage Bin and Interface Currents systems rely on Rotation and Translation mechanisms (RNT by Kruger et al. (2005)), where rotation and translation occur in a single action with single input point, the SharePic system (which is implemented on the DiamondTouch technology) combines rotation and resizing in a unique gesture with a single input point (see Fig. 3.10, d). The aim is indeed to privilege zooming functionalities because the system is meant for elderly users, who might have a poor sight.

Other work investigates interaction techniques for supporting co-located collaboration through collaborative gestures (Morris et al., 2006) and multi-modal interaction (Tse et al., 2006), as shown in Figure 3.11, a and b.

Figure 3.11: Gestures and multimodal input to support collaboration: a) The combination of human-human gestures with human-display gestures for collaborative interaction (Morris et al., 2006); b) The combination of gesture and speech input for collaborative interaction (Tse et al., 2006).

Finally, the recent advances in multi-touch sensing technologies (e.g., Jazzmutant[5]) and camera-based systems (e.g., TouchLight (Wilson, 2004), PlayAnywhere (Wilson, 2005), and Han's sensing technique based on frustrated total internal reflection (Han, 2005)) enable the tracking of multiple fingers and hands simultaneously (see Fig. 3.12). These technologies promise to support the design of a heterogeneous semantics of fluid gestures for a rich

[5]http://www.jazzmutant.com

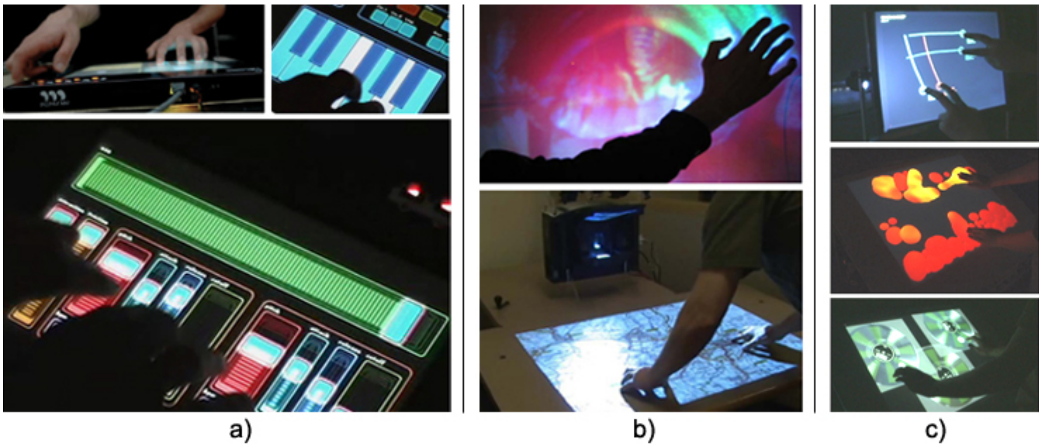


Figure 3.12: Multi-touch sensing techniques: a) The Lemur system for the direct touch control of digital sound on the Jazzmutant interactive surface. b) Wilson's work on Touch Light (2004), on top, and Playanywhere (2005), below. c) Han's demonstration of the interaction possibilities enabled by a sensing technique based on frustrated total internal reflection (Han, 2005).

manipulation vocabulary in a variety of mundane contexts of social interactions, as recently promoted by the new Microsoft Surface tabletop[6].

3.2 Interactive Objects

As computing moves beyond the desktop and becomes more integrated in our physical space, the integration of physical objects as mediators of human-information interaction has been explored. In some cases the mundane artifacts of our everyday life have been augmented with a semantic meaning in the digital world. Examples for this are the MediaCup (Gellersen et al., 1999) and Passage (Konomi et al., 1999), where sensors are embedded either in the artifact itself (e.g., in MediaCup, see Fig. 3.13, a) or in the related infrastructure (e.g., Passage, see Fig. 3.13, b) to enable the overlay of a digital semantics onto the physical one. In these cases, although the physical artifacts are coupled to a digital meaning, their appearance is not altered.

A more systematic design of physical artifacts for the manipulation of digital information is the focus of other approaches, such as the ones of Graspable and Tangible User Interfaces. This body of research explores the integration of physical artifacts in the human-computer interaction in order to control, organize, and manipulate digital information, thus establishing

[6] http://www.microsoft.com/surface

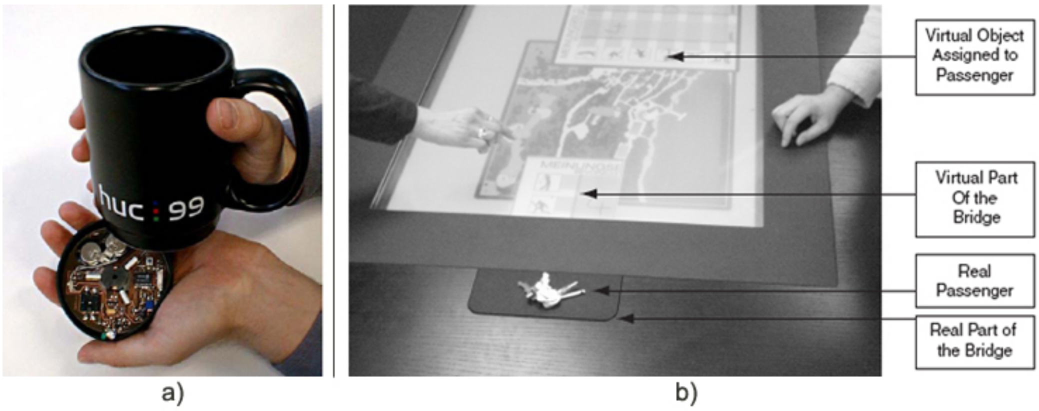


Figure 3.13: Digitally augmented mundane artifacts. a) In the MediaCup sensors are hidden in the bottom of the artifact so as to track where the cup is, and whether it is hot or cold (Gellersen et al., 1999). b) In the Passage system any object and be augmented of a digital meaning when it is placed on the Bridge system (Konomi et al., 1999).

a hybrid interaction dialogue. The main goal is to merge and exploit the possibilities of virtual media as well as the affordances of physical artifacts, such as their rich manipulation vocabulary, the mental models of everyday objects, and spatial reasoning.

Beginning with early work by Fitzmaurice et al. (1995) and Ishii and Ullmer (1997), there have been several instantiations and variations of the TUI paradigm in different ways, to different extensions, in different projects. This fact has in turn motivated further work to the end of systemizing the design space of tangible user interfaces and tangible interaction, e.g., (Ullmer and Ishii, 2000), (Holmquist et al., 1999), (Fishkin, 2004), (Hornecker and Buur, 2006). Fishkin (2004), for example, provides a useful taxonomy for the analysis of tangible interfaces based on the dimensions of *metaphor* and *embodiment*. The following analysis of the related work is structured in correspondence to the conceptual model underlying the integration of physical objects in the interaction. This does not aim to add just one more framework or taxonomy to the existing work, but rather to support an analysis of the affordances of physical objects for the design of interaction. The reader needs to bear in mind that the borders between the distinct categories are not sharp, but they facilitate recognition of the different ways in which physical objects contribute to shape and convey the conceptual model for hybrid interaction.

3.2.1 Objects as Handles for Navigation

This category comprehends those systems which exploit the proprioceptive feedback and manipulation vocabulary afforded by 3D physical objects to facilitate spatial reasoning and navigation.

Hinckley et al. (1994), for example, adopt passive props in the neurosurgical domain to support spatial navigation. By moving in 3D a physical ball in which sensors are embedded, a surgeon can correspondingly change the visualization of a brain on a computer monitor and manipulate the ball/brain by using other two physical props as tools (see Fig. 3.14, a). The system supports two-handed, modeless interaction.

Cao and Balakrishnan's (2003) VisionWand interface combines remote pointing and navigation, together with touch screen selection. The system consists of a SmartBoard (which enables touch screen selection by tapping the display with the wand) together with two cameras which can track the rotation, alignment and distance of the wand from the screen in 3D (see Fig. 3.14, b, top). This set-up allows for a high degree of freedom of the wand tool, which in turn supports the design of a rich manipulation vocabulary for the navigation of data. By pointing with either the edges of the stick towards the screen, changing the distance and orientation of the stick with respect to the wall, and by performing different types of gestures, users can select, move, and scale objects, control the data visualization by zooming and scrolling, as well as invoke and navigate pie menus (see Fig. 3.14, b).

With Toolstone, Rekimoto and Sciammarella (2000) explore the use of tangibles to support two-handed interaction through the design of an input device for the non-dominant hand. Toolstone is a cordless prism in which rotation, tilting, flipping, and the contact face on a touch pad can be sensed. This enables a user to manipulate position and rotation of a virtual camera in a 3D scene by rotating and dragging the Toolstone with the non-dominant hand, as well as to change the camera viewing angle by dragging its projection with a stylus with the dominant hand (see Fig. 3.14, c)

Whilst these examples mainly combine the manipulation of physical artifacts in 3D with the navigation of visual information, other work focuses on the association between physical manipulation and navigation of information spaces and semantic structures.

The AlgoBlock system (Suzuki and Kato, 1993) is based on a set of physical blocks that can be connected to each other to write a program in the Logo programming language. Each block has a designated atomic function, corresponding to a single command. Blocks can be linked together to compose a more complex syntax. The outcomes of the assembled structure, accessible

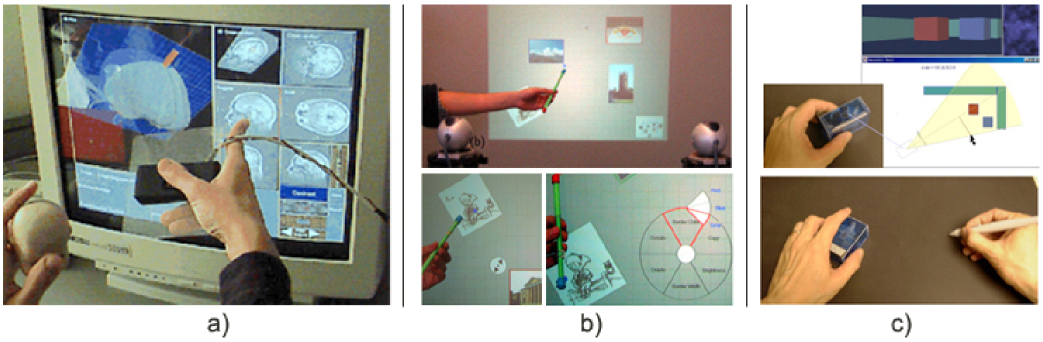
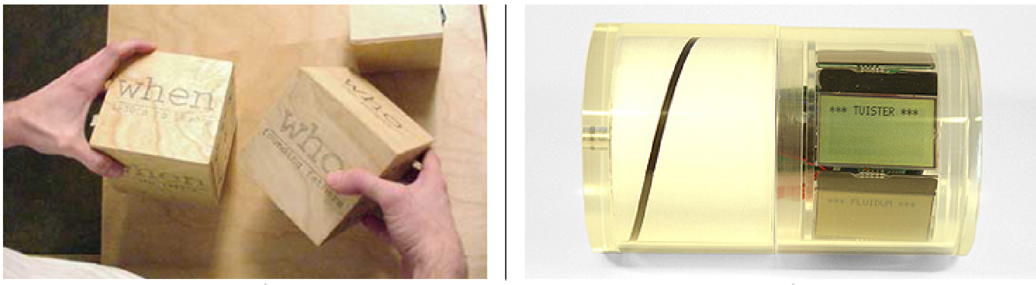


Figure 3.14: Objects as handles for visual navigation. a) Props for navigation of neurosurgical images (Hinckley et al., 1994). b) Top, the VisionWand set-up, consisting of a SmartBoard and two cameras; Bottom, navigation gestures with the VisionWand (Cao and Balakrishnan, 2003). c) Changing the viewing angle of a virtual camera with the Toolstone interface (Rekimoto and Sciammarella, 2000).

by multiple users and especially conceived for young pupils, is displayed on a PC screen.

Camarata et al. (2002) associate the manipulation of Navigational Blocks with the exploration of the history of Seattle's Pioneer Square. The interface consists of four blocks that represent some categories of the history of the square, such as "who" (the people), "what" (the events), "where" (the locations), and "when" (1850's-1890's). Each face of each block represents a subcategory, e.g. the six sides of the *who* block represent the founding fathers, the women, the merchants, the miners, the native Americans, and all of the people associated with the history. Users can explore a virtual gallery, displayed on the monitor of a tourist booth, by creating compound database queries. This is possible by aligning multiple blocks (e.g., the *when* and *who* blocks, as shown in Fig. 3.15, a). The system can track the placing, the sliding, and the combination of Blocks in the active space thanks to a gravity-fed six-sided sensor inside each block, which wirelessly communicates with an onboard microcomputer.

Tuister (Butz et al., 2004) is a tangible device for two-handed manipulation that supports the navigation of hierarchical structures, such as multi-layer menus. Input and output appear in the same device, consisting of two cylinders of approximately the same size and aligned on the same axis. In one of the cylinders, called "head", six displays are arranged. While the non-dominant hand holds the other cylinder, the "handle", the dominant hand can rotate the head till the item which is desired from the list appears on one of the displays (see Fig. 3.15, b). The level of hierarchy in the multi-menus structure is managed with the non-dominant hand. When the



a) b)

Figure 3.15: Examples of objects as handles for semantic navigation: a) The Navigational Blocks (Camarata et al., 2002); b) Tuister (Butz et al., 2004).

handle is rotated clockwise, towards the user's body, the navigation moves down along the hierarchical path, and vice versa when the handle is rotated anti-clockwise.

3.2.2 Objects as Containers

In Holmquist et al.'s (1999) conceptual analysis of tangible user interfaces, *containers* are defined as "generic objects which can be associated with any type of digital information". Within this class of devices, this section focuses on those physical objects which, differently from digitally augmented mundane artifacts such as the MediaCup and the Passage systems previously considered, are designed ad-hoc for the purpose of moving digital information from one context to another one, and/or to exploit the manipulation vocabulary of the object to trigger different data.

A first example in this direction is Durrell Bishop's conceptual design for the Marble answering machine, as described by Crampton Smith (1995). In this project, voice messages are represented by physical marbles. These can "fall" into a bin when a new message is received and can be moved into a slot on the machine to play the message, or on a slot aside the phone to dial the calling number (see Fig. 3.16, a).

In a similar manner, Mediablocks (Ullmer and Ishii, 1999) are physical wooden blocks embedded with digital ID tags, which turn them into containers for online media (see Fig. 3.16, b). In this way, digital content, such as a URL path, can be copied from one environment and be pasted onto another one, such as a display, or can be printed on a printer.

Similar to that, PhotoCube (Want et al., 1999) is a kind of physical holder of digital documents. In this case, different faces - accommodating a disk-sized RFID tag each - correspond to different people's homepages. A photo of the person is pasted on the block face: When the photo is put in contact

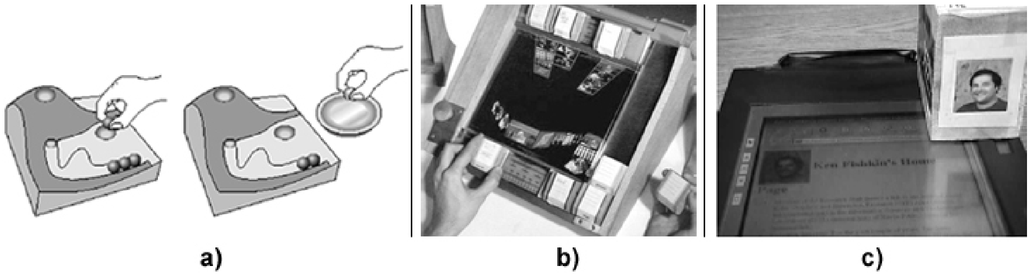


Figure 3.16: Examples of objects as containers: a) The Marble answering machine (Crampton Smith, 1995); b) The Mediablocks interface (Ullmer and Ishii, 1999); c) The PhotoCube (Want et al., 1999).

with the computer, the URL of the person's homepage is triggered on the screen (see Fig. 3.16, c).

3.2.3 Objects as Tools

By analogy to the physical world, *tools* are here considered as those interfaces which are used for editing and transforming an object, and in this specific context for editing a *digital* object. As interactive surfaces have become progressively available, they have enabled the spatial merging of the physical input tool with the digital output in the same area in which the tool is applied. This is analogous to what actually happens in the physical world.

The examples referenced in this section conceive of the use of physical tools for editing digital information which is spatially mapped directly to the physical object. The resulting changes can affect the overall visual or audio system display (thus using tools as handles for manipulation), as well as a portion of it, in proximity of the point in which the tool is applied.

Fitzmaurice et al.'s (1995) mocked-up prototype of the GraspDraw application on the Active Desk is one of the first examples illustrating the manipulation of graphical shapes with physical handles (see Fig. 3.17, a). As sensing technologies have become more mature, such a concept has been applied in other contexts. The Urban Planner Workbench (URP) is an application which aims at supporting urban planning (Underkoffler and Ishii, 1999). It relies on a vision technique for tracking the position and orientation of physical objects on a unique pattern of colored dots on a tabletop (see Fig. 3.17, b). Users can reorient and move the physical models of buildings, so as to visualize casted shadows with a front projection system. Furthermore, they can use tools for changing system status, such as the hour on a clock tool, so as to arbitrarily change the light condition and preview its shad-

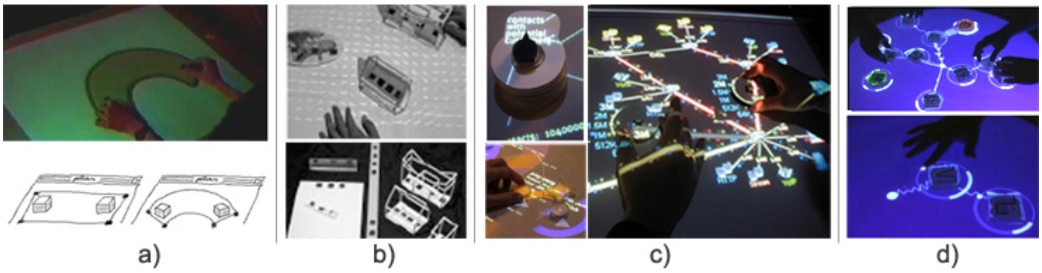


Figure 3.17: Examples of objects as handles for manipulation. a) Bricks for changing graphic shapes on the Active Desk (Fitzmaurice et al., 1995). b) Tools for changing buildings position, light effects and measures on the URP system (Underkoffler and Ishii, 1999). c) Pucks and knobs for changing system as well as local conditions with the Sensetable (Patten et al., 2001). c) Pucks for the creation and modification of electronic sound on the reacTable (Jordà et al., 2007).

owing effects. Additionally, they can use measuring tools, which track and dynamically display the distance between two buildings.

Patten et al.'s (2001) Sensetable tracks the position of wireless pucks electromagnetically. The movement of the pucks on the touch-sensitive surface (consisting of two Wacom[7] tablets placed next to each-other) dynamically edits the graphical representation which is front projected on the tabletop. Like handles for manipulation, the physical tools can be bound to the graphical representation to scratch it, move it, and edit it. Furthermore, a dial is mounted in each puck so that users can as well change the state of the puck, thus applying local changes. This concept has been integrated in a diversity of domains by the authors, such as chemistry and system dynamics, business supply chain management, urban planning, interactive visual art, as well as performance and composition of electronic music. This last domain has been explored in the Audiopad project (Patten et al., 2002). This relies on an infrastructure similar to the Sensetable, although it integrates a matrix of antenna elements which track the positions of electronically tagged objects on the tabletop. As the user interacts with the tools, s/he can physically manipulate the sound as well as the visual display (see Fig. 3.17, c). Each object represents either a musical track or a microphone, and they have different shapes to afford different functionalities.

Similar to that, Jordà et al. (2007) create a semantics of objects/tools for the reacTable. The objects, consisting of plastic pucks in different shapes, have a marker attached on the side which is in contact with the table surface. The marker is tracked with a camera-based system mounted under the table together with a beamer for back projection of the visual display. Each

[7] http://www.wacom.com/productinfo/

puck represents a modular synthesizer component with a specific function for the generation, modification, or control of electronic sound. By changing the number, position, spatial relationship, and local rotation of the pucks, an unlimited number of users can casually interact with the system and dynamically edit the sound output. An interesting aspect of this system is that fingers can also be used for interacting and locally altering the function of single pucks. This is made possible by simple, small paper markers that users can stick on their finger tips, thus augmenting the interaction vocabulary of bare fingers.

DataTiles (Rekimoto et al., 2001) and TViews (Mazalek et al., 2006) are other examples in which the use of tools has an effect on the local area in which the tool is applied, rather than on the overall system display. DataTiles are physical transparent tiles which a user can slide on a grid of rails over a screen. Similar to interactive filters, these tiles allow for different visualizations and manipulations of the information displayed underneath. Users interact with a pen directly on the tile and across multiple tiles (see Fig. 3.18, a).

a) b)

Figure 3.18: Examples of objects as tools for local modification: a) Interacting with DataTiles for editing the information displayed underneath, or aside a tile (Rekimoto et al., 2001); b) Using physical pucks for moving photos on TViews (Mazalek et al., 2006).

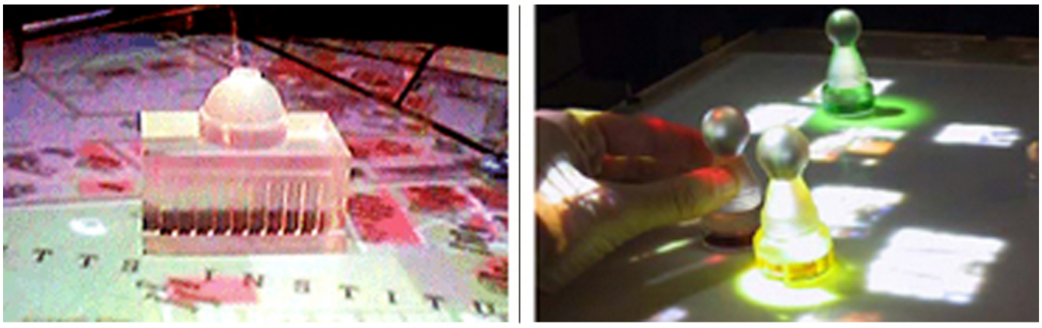
TViews (Mazalek et al., 2006) is a general platform for tangible interaction with digital media on a tabletop for everyday life social environments, e.g. domestic living rooms. In this case, physical objects are the only input device and the sensing technology relies on both ultrasonic and infrared sensors. In PhotoSorting, one of the applications implemented on TViews, physical pucks are used as kind of magnets to “attach” photos and move them around on the table surface (see Fig. 3.18, b).

3.2.4 Objects as Symbolic Embodiment

This section considers those cases in which physical objects are designed and used as 3D metaphorical representation of an entity. To this purpose, their shapes resemble the features of the entity they stand for.

In Ullmer and Ishii's (1997) MetaDesk, a small plastic model of the MIT Great Dome can be rotated and translated for reorienting and sliding the map of the MIT campus, which is displayed on an interactive tabletop (see Fig. 3.19, a).

In Tangible Viewpoints (Mazalek et al., 2002) physical pawls represent the characters of a story in an application for collaborative interactive narratives. Users can activate the viewpoints (i.e., video clips, sound, and pictures) of different characters as they place and move the pawls on the tabletop (see Fig. 3.19, b).



a) b)

Figure 3.19: Examples of objects as symbolic embodiment: a) The MIT Great Dome on the MetaDesk (Ullmer and Ishii, 1997); b) The pawls representing characters in the Tangible Viewpoints application (Mazalek et al., 2002).

3.3 Interactive Multi-Display Environments

The vision of ubiquitous computing has initially been instantiated in several research labs through the instrumentation of closed rooms with different kinds of interactive surfaces and objects, displays, as well as devices in different sizes, orientations, and able of different degrees of mobility: iRoom at the Stanford University (Johanson et al., 2002) and I-Land (Streitz et al., 1999) at Fraunhofer IPSI (see Fig. 3.20), in Germany, are some examples.

In the context of such interactive rooms, the focus has often been put on the design of interaction techniques for moving and editing information across multiple displays. With Pick-and-Drop, Rekimoto (1997) tackles this

Figure 3.20: Roomware, deriving from the I-Land project (Streitz et al., 1999), is a multi-display environment incorporating different devices. From the left: ConnecTable, Dynawall, CommChair, and InteracTable.

problem by translating the conceptual model of "drag and drop" we are familiar with in the desktop PC to a multi-display environment. With the use of a stylus with a unique ID, a user can pick up a data (represented by an icon as in the desktop GUI) from a display and drop it onto another one by tapping on its surface. From the implementation point of view, the data is simply transferred through the network, but, from the user interface point of view, this technique resembles the way in which we pick a physical object to move it to another location (see Fig. 3.21), thus using the pen as a prosthesis of our fingers.

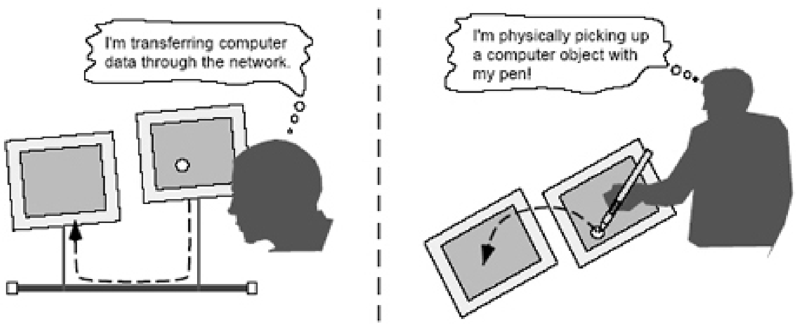


Figure 3.21: The conceptual difference between remote copy and Pick and Drop (Rekimoto, 1997).

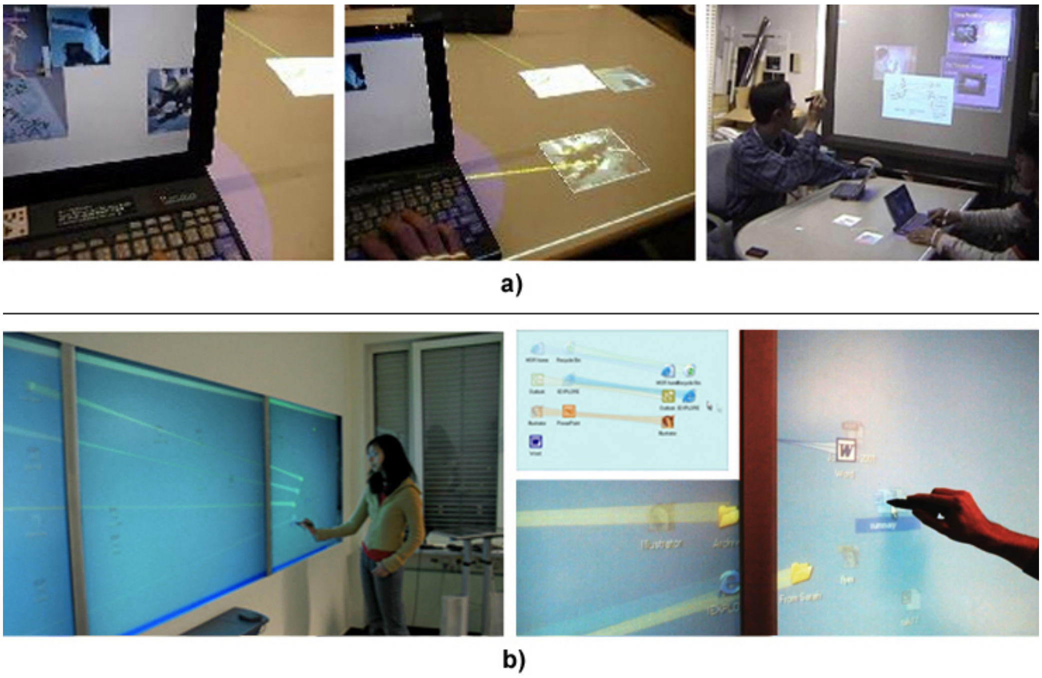


Figure 3.22: Techniques for drag and drop across multiple displays: a) The Hyperdragging system (Rekimoto and Saitoh, 1999), enabling the movement of data across physical and digital surfaces; b) The Drag-and-Pop technique (Baudisch et al., 2003), enabling the user to choose among different targets displayed in a proxy near the input point.

Hyperdragging (Rekimoto and Saitoh, 1999) builds on a similar concept, but it extends it in the way that digital information (represented by an icon or open in a window) can be continuously dragged from one surface to another one without requiring the user to tap onto the target surface. When the user drags the icon of the data file towards the physical borders of the screen real estate with a pointer, the object migrates to the next available display. This can be a physical surface, e.g. a table on which the file is projected, as well as an interactive whiteboard. Other users can than interact with the data. The interaction in this case is remote, and the "extension" of the link between input device and output is represented as a line (see Fig. 3.22, a). The authors name this "anchored pointer" to provide visual feedback for the relationship between the user and the object of interaction.

Pick-and-Drop and Hyperdragging have inspired a number of projects which focus on the translation of desktop GUIs across screens, giving novel physical behaviors to cursors and icons. Drag-and-Pop (Baudisch et al., 2003) (see Fig. 3.22, b), RADAR (Nacenta et al., 2005), and Vacuum (Bezerianos and Balakrishnan, 2005) are some related examples in the area of large displays. In these cases, the user's physical action of dragging an icon in a small display region is kind of "exaggerated" so as to make the icon reach other targets - in the same or a different display - which could not be easily reached by the user with her arm otherwise.

Another approach builds on the proximity between devices and displays to create a semantics for the transfer of data. In the ConnecTable project (Tandler, 2001), for example, bringing together two mobile table appliances enables two users to create a shared display. On such a combined display the working area can be a single and shared one, or a doubled one for symmetric orientation and interaction (see Fig. 3.23, a). In a similar way, Hinckely et al.'s (2004) work on Bumping (see Fig. 3.23, b) and Stitching (see Fig. 3.23, c) techniques aims at supporting displays synchronization, and it enables the variation of users' relative orientation.

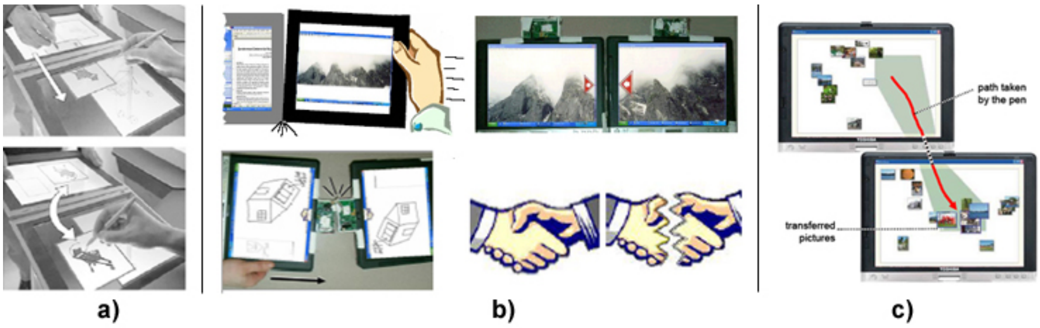


Figure 3.23: Different techniques for synchronizing mobile displays. From the left: a) ConnecTable (Tandler, 2001); b) Bumping, and c) Stitching (Hinckley et al., 2004).

Rather than displaying the same content, different screens can also be used for displaying different visualizations of the same data. Baudisch and Good's (2001) work on focus plus context visualizations exploits the different properties of different screens (i.e., different size and resolution of TFT vs. projection-based displays) for providing detailed view and overview in a desktop set-up. Building on this work, a mobile version of such a concept is proposed in the Ubiquitous Graphics project (Sanneblad and Holmquist, 2006). The interface relies on the concept of Magic Lenses (Bier et al., 1993). A tablet PC is used for a more detailed visualization of the data appearing on a large projection-based display (see Fig. 3.24, b). In a similar way, and on a larger scale, Boring et al. (2007) make a complete wall interactive, integrating different kinds of display technologies and resolutions. The wall hosts three back-projected displays: the central one (in focus) is a commercial SmartBoard with high precision optical tracking. When the cement wall is illuminated by a ceiling-mounted steerable projector, it becomes a display itself. Four webcams are mounted at the corners of the wall to provide lower precision tracking so that input is enabled in two points on the entire wall simultaneously (see Fig. 3.24, c). Thus, different resolutions are possible in two ways, both in the output and input directions.

a) b) c)

Figure 3.24: Instantiations of the focus plus context visualization concept in different scales. From the left: a) On an augmented desktop (Baudisch et al., 2001); b) On a projected area (Sanneblad and Holmquist, 2006); c) On an entire wall (Boring et al., 2007).

3.4 Summary and Discussion

This chapter has presented an overview of the various ways in which different qualities of physical interaction have been integrated in the design of interactive systems of different scales. The identification of those qualities allows for the mapping of such features to the projects that were considered, as illustrated in Figures 3.25, 3.26, 3.27, and 3.28. This responds to the aim of highlighting the very aspects of physical interaction that, consciously or not, were integrated in the design of such systems and whether it is possible to recognize patterns among projects.

This analysis was based on the scale of interactive systems as underlying criterium. When looking at Figure 3.25, one can observe that most of the work on interaction techniques for multi-display environments builds on physical constraints (i.e., the geometrical boundaries of the display real estate in most cases) in order to afford cognitive as well as physical affordances for interaction. That means, in other words, that the physical boundaries of the screens are given a semantic meaning in the design of the interaction conceptual model. On the other hand, the work on interactive surfaces (see

Fig. 3.25, 3.26, and 3.27) mostly combines directness (which was one of the characteristic features of the examples considered in the survey) together with continuity. This suggests that as soon as the pointer is not necessary anymore, the chunking (Buxton, 1995) of an action into "pointing", "clicking", "dragging", and "releasing" can be reduced into a continuous phrase, such as "touching", "dragging" (i.e., stroking), and "releasing". Similarly, such a continuity is afforded by the use of physical objects for interaction, as highlighted in Fig. 3.27 and 3.28. This is in line with Fitzmaurice' s (1996) analysis of the phases of interaction with classical GUIs (i.e., 1: "acquire physical device"; 2: "acquire logical device"; and 3: "manipulate logical device") in comparison with the phases of interaction with graspable interfaces (i.e., 1: "acquire physical device"; and 2: "manipulate logical device"). This aspect, which was anticipated in Chapter 1, cf. Fig. 1.2, tackles the meaning of direct input and is one of the main issues addressed in this thesis.

Furthermore, one can notice that the use of interactive physical objects often affords two-handed interaction as well, likely due to the natural space-multiplex way in which we interact with physical artifacts (cf. Chapter 2, Section 2.4). Finally, interactive objects are often used in combination with a 2D surface, typically a tabletop, for the visual output. This can be noticed when looking at the rather limited number of cases in which the manipulation of an object in 3D triggers an interaction with digital media independently from a surface of reference. This aspect is obviously motivated by the physical law of gravity, which affords the placement of objects on a horizontal plane.

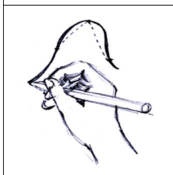
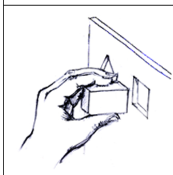
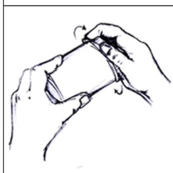
On top of these considerations, this survey underscores different approaches to the design of interfaces for ubiquitous computing. A large portion of the work on interactive multi-display environments explores ways for translating the traditional desktop GUI across different displays. On the other hand, much of the work on large interactive displays and tangible user interfaces moves away from the WIMP traditional UIs towards a more domain-specific design, i.e., towards appliances. One can speculate that both approaches will evolve in the future, and one won't probably replace the other. Desktop GUIs will likely continue to exist and support certain types of contexts, while appliances will arise to support other more specific ones.

This thesis investigates how ubiquitous computing technologies can be embedded and appropriated into various domains of everyday life beyond the office (e.g., cooking, design, learning, creativity): Thus, the design work it presents is more appliance oriented. In the next chapter preliminary considerations and explorations of how some of the qualities of physical manipulation can be integrated in the design of hybrid interactions are discussed.

	Pick-and-Drop	Hyperdragging	Drag-and-Pop	ConnecTable	Hinckley et al. Bumping	Hinckley et al, Stitching	Baudisch and Good, F+C	Ubi. Graphics	Boring et al., Wall Display	Graffiti & Unistroke	T-Cube, Quickwrt., SHARK	Dragicevic, Fold-Drop	BumpTop
metaphor	□	□	□	□	□	□	□	□	□	□	□	■	■
directness	■	□	■	□	□	■	□	■	■	■	■	□	■
continuity	□	■	■	□	□	■	□	□	■	■	■	■	■
3D space	□	□	□	■	■	■	□	□	□	□	□	□	□
contstraints	■	■	■	■	■	■	■	■	■	□	□	□	■
multimedia	□	□	□	■	■	■	□	□	□	□	□	□	□
two-hands	□	□	□	□	□	□	□	□	■	□	□	□	□

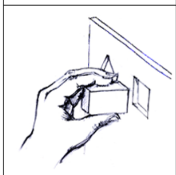
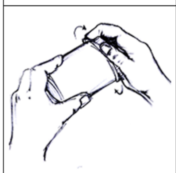
← multi-display environments → | ← surfaces

Figure 3.25: Physical qualities in the design of interactive systems (1)

		Ka-Ping Yee File Browser	Dual Touch	IPhone	3D Gestures for Handhelds	Tivoli	Flatland	FlowMenu	Deisgners' Outpost	Digital Tape Tool	Digital Desk	Koike et al. drawing. app.	HabilisDraw DT	Smartskin
metaphor		□	□	□	□	□	□	□	■	□	□	□	■	□
directness		■	■	■	□	■	■	■	■	■	■	■	■	■
continuity		□	■	■	■	■	■	■	■	■	■	■	■	■
3D space		□	□	■	■	□	□	□	■	□	■	□	□	■
contstraints		□	□	□	□	□	□	□	□	□	□	□	□	□
multimedia		□	□	□	□	□	□	□	■	□	□	□	□	□
two-hands		■	■	□	□	□	□	□	□	■	□	■	■	■

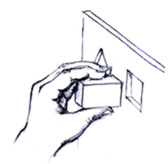
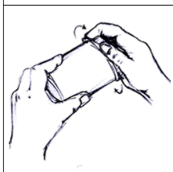
surfaces

Figure 3.26: Physical qualities in the design of interactive systems (2)

		Wu et al. Multi-fingers Gest.	Personal Digital Historian	Storage Bins & I. Currents	SharePic	Collaborative gestures	Tse et al. Speech and Gest.	Lemur, TouchLight, Han '05.	Hinckley's props	VisionWand	Toolstone	Algoblock	Navigational Blocks	Tuister
metaphor		□	□	■	□	□	□	□	□	□	□	□	□	□
directness		■	■	■	■	■	■	■	□	□	■	□	□	■
continuity		■	■	■	■	■	■	■	■	■	■	□	□	■
3D space		□	□	□	□	■	□	□	■	■	□	□	■	■
contstraints		□	□	□	□	□	□	□	■	□	□	■	■	□
multimedia		□	□	□	□	■	■	■	■	■	■	■	■	■
two-hands		■	□	□	□	■	□	■	■	□	■	■	■	■

surfaces ←→ objects

Figure 3.27: Physical qualities in the design of interactive systems (3)

		Marble answering machine	Mediablocks	Photocube	Bricks on GraspDraw	URP	Sensetable	Audiopad	reacTable	DataTiles	TViews	MetaDesk	Tangible Viewpoints
metaphor		■	□	□	□	□	□	□	□	□	□	■	■
directness		■	■	■	■	■	■	■	■	■	■	■	■
continuity		□	□	□	■	■	■	■	■	■	■	■	■
3D space		■	■	■	□	□	□	□	□	□	□	□	□
contstraints		■	■	■	□	□	□	□	□	■	□	□	□
multimedia		■	■	■	■	■	■	■	■	■	■	■	■
two-hands		■	■	■	■	■	■	■	■	□	□	□	□

objects →

Figure 3.28: Physical qualities in the design of interactive systems (4)

4

Exploring Affordances for 2D and 3D Manipulation through Sketches

Drawing upon the analysis of the related work for the design of interactive systems presented in Chapter 3, this chapter introduces two preliminary instantiations of design concepts through a "sketching approach". It discusses the meaning and value of such an explorative approach: Then, it reflects upon some of the possibilities and implications of integrating aspects of physical interaction in the design of interactive systems for manipulation of digital media in 2D and 3D spaces.

4.1 The Meaning of Sketching

The analysis of how interface design has addressed interaction paradigms for interactive surfaces and objects (cf. Chapter 3) has suggested a number of different issues and possible solutions, and has highlighted how aspects of physical interaction can be embedded into hybrid interactive systems: This, in turn, has created the ground for an exploration of design alternatives.

In order to externalize, explore, elaborate, as well as share alternative solutions, the natural design attitude to problem statement is to start out and "sketch" possibilities, i.e., "sketch interactions" in this particular case. As anticipated in the introduction, sketching is a fundamental activity of design thinking, as it allows designers to visualize, read, and interpret their own ideas so as to further reflect and build on them (see also (Buxton, 2007b)). Designers can learn both from the process of sketching as well as from its products, i.e., from their own sketches. In the process, designers think whilst creating, discover construction issues as well as representational ones, and shape their abstract ideas during the action itself. The products of the sketching activity are incomplete and undefined but through this incompleteness and ambiguity designers can read opportunities and alternatives for further elaboration towards problem solving.

Furthermore, sketching is the natural way in which designers share ideas with others. They provide a shared reference and a tangible common ground for reflection and elaboration. According to their target (e.g., other designers, project stakeholders, potential users), sketches can have different levels of refinement. From a sketch, designers as well as others can gain an intuition of the weaknesses, potentials, as well as strengths of an idea, which can be validated in further iterations.

The two examples presented in this chapter can be considered as *sketches of interactions* in the sense suggested by Buxton in (Buxton, 2007b): Indeed, they respond to the aim of externalizing and sharing an idea, but they go beyond pen and paper in order to represent and simulate the dynamism of interaction. Since they are at a level of refinement that enables others to actively experience the interaction with the artifact to a certain extent, they are sometimes referred to as *sketches of experiences*. Their value is not in the products themselves, which are incomplete and at an early stage of development, but rather in their capability to raise further questions, discussions, as well as collaboration.

The following sections introduce the rationale behind these sketches and their meaning for the process of this investigation.

4.2 Reflecting on 2D Surfaces

The sketches presented in this section explore the integration of elements of physical interaction in the design of affordances for direct touch interfaces on interactive surfaces.

Chapter 3 reviewed a number of research projects addressing the manipulation of digital information on interactive surfaces. Several of these examples adopt a stylus or an interactive object as transducers. These enable the design of manipulation vocabularies that are based on the gestures afforded by the transducers themselves (e.g., crossing-based and marking gestures for pen input (Accot and Zhai, 2002); tilting and rotating gestures for input with tangible tokens (Rekimoto and Sciammarella, 2000)). Other work exploits the richness of hands and multi-finger manipulation in the physical world to design a set of gestures for interacting with digital information (e.g., (Wu et al., 2006)). Such a vocabulary of gestures mostly needs to be learned and remembered by the user in order to interact with the system.

In this context, emerges the question: How can we display digital information on a surface in such a way that it suggests users what gestures to perform, thus minimizing their learning effort? An hypothesis is that if we represent digital, abstract information in such a way that it metaphorically represents a physical artifact, we can design cognitive affordances for users' gestures by relying on the physical affordances and manipulation vocabulary of the metaphorical source (cf. Chapter 2, Section 2.7, for a distinction of physical vs. cognitive affordances).

While designing cognitive, visual affordances for digital information in 2D, two main design aspects need to be addressed. On the one hand, visual cues of the displayed information can suggest its pliancy, i.e., its characteristic to be interactive (Cooper and Reimann, 2003). For example, frames can define semantic areas of interaction on the screen; 3D effects, colors, and shadings can bring the information in fore- or background and suggest information status (e.g., active/inactive). These visual cues provide what Pirolli and Card (1999) name *information scent*, which can enhance perceptual as well as cognitive viewers' processing of information: I.e., they can accelerate users' visual "scanning" (Nielsen, 1999) of the information display.

On the other hand, the metaphoric link to real world objects and to their affordances in the physical world can provide rich material for the design of affordances for the manipulation of digital information. The virtual representation of a lever doorhandle, for instance, can be mapped to the natural gesture that we make when we push down a real physical lever. The representation of a steering wheel can be mapped to the turning gesture we perform while driving. In these cases, the mapping relies on the analogy between the elements representing digital information and the affordances of their referents in the physical space. In this respect, as designers, we need to think thoroughly about a manipulation vocabulary (and the representation thereof) that is suitable for the type of transducer used for input: e.g., multi-finger or single input point, such as in the case of pen input.

To demonstrate this concept, and building on the considerations above, these ideas were explored through the design (or sketch, rather) of a metaphorical interface for multi-finger hand gestures on a wall display. The next paragraph illustrates the design of such an interface.

4.2.1 The Mug Metaphor Interface

The underlying idea of this sketch of interaction was to explore how one can map the affordances of real world objects to gestures, relying on the

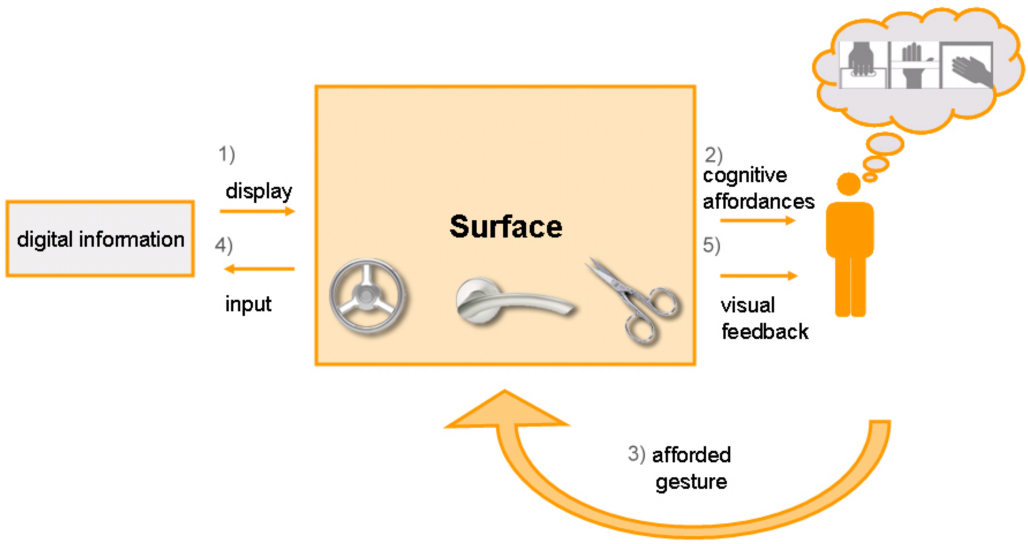


Figure 4.1: The figure illustrates the idea of metaphorically mapping the physical affordances of real world objects to cognitive affordances for gesture-based manipulation of digital information on interactive surfaces.

conceptual model in which such real objects are embedded. Referring to the qualities of physical interaction identified in Chapter 1 and represented in Fig. 1.3, the Mug Metaphor Interface explores the integration of the following aspects:

- Metaphorical representation;
- Directness;
- Continuity of action;
- Two-handed cooperative work.

The Mug Metaphor Interface relies on the physical affordances provided by a real mug, and metaphorically represents it as a container of information. When manipulating a real mug we know we can move it around by holding its handle and incline it to pour its content (see Fig. 4.2, b and c). Empty mugs are expected to be lighter than full ones (e.g., contain less data); steaming mugs are expected to be hot (e.g., contain recent data). Additionally, a mug is a mundane object which we use in different environments, e.g., in the office, in a living room, in a kitchen: Thus, the metaphor is not strictly related to the office domain, and can be valid in others as well.

In such a concept, mugs and units of information - the latter represented as kind of drops - can be manipulated across the display. Pie menus appear in correspondence of the hands, thus “following” the user while moving across the display, rather than being operable just in a fixed location on the screen. This responds to the need of freedom of movement of the user, and to enable two-hands cooperative interaction. The dominant hand, e.g. the right one, is devoted to the manipulation and navigation of information. A pie menu displaying containers of information is displayed in correspondence of the right hand (see Fig. 4.2, a). The non-dominant hand (e.g., the left one) works as command invocation, managing a menu of resources (e.g., drain, displays, and printers). Such a menu appears when the non-dominant hand double-taps the interactive surface. The circular menu can be scrolled with a movement of the finger on a holed gear, which makes the circle segments rotate (see Fig. 4.2, d). The dominant hand moves units of information to the preferred resource (e.g., to cancel an information unit as in Fig. 4.2, e).

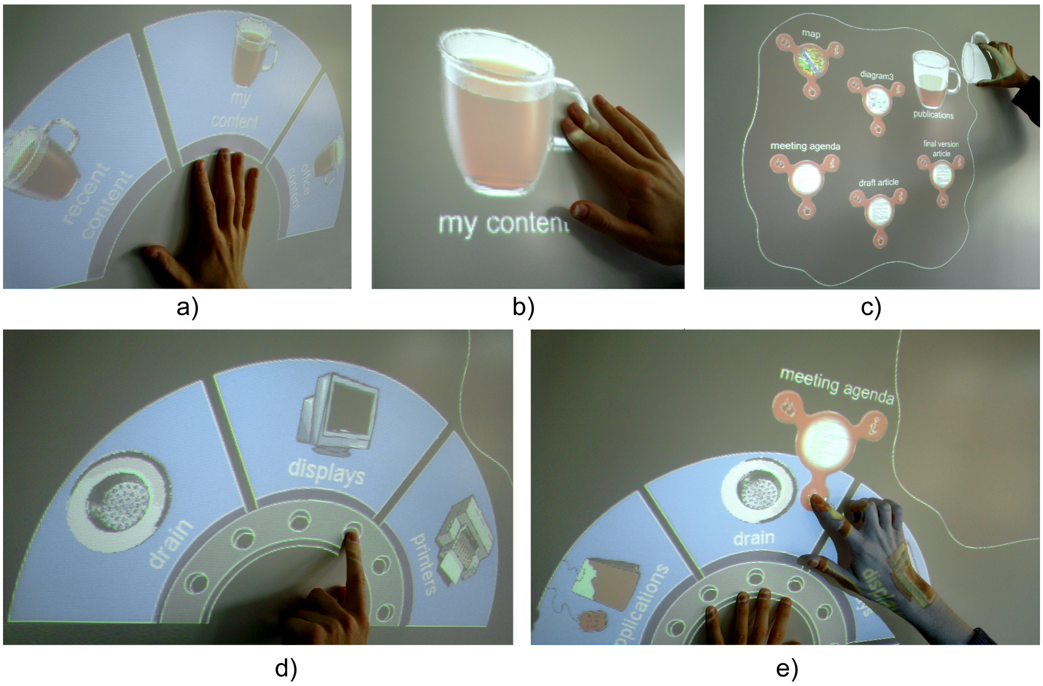


Figure 4.2: The Mug Metaphor Interface: The figure illustrates different stages of the mock-up realized in Flash.

4.2.2 Observations and Implications

The mock-up of the Mug Metaphor Interface was realized in Flash according to a storyboard of possible manipulation tasks, such as for example “open”, "close", "move", and "delete". Such an interface was informally evaluated in our lab, both with colleagues and with people who are not familiar with this research and with the technology it adopts. We first demonstrated the set of tasks we had implemented in the Flash interface on a SmartBoard and mimicked the gestures we wanted to suggest (e.g., rotating the mug for opening; dragging the mug handle for moving; tapping the screen surface with a hand in a certain place and time sequence for invoking the pie menu) so as to show that different gestures were possible for different tasks. We then asked individual users to perform a certain set of tasks such as "explore the content of the mug", "move the mug", "cancel the information unit", "scroll the menu", and so forth, in such a sequence that could be supported by the mocked-up storyboard.

The participants easily interacted with the interface and could accomplish the tasks. It was curious to observe, though, that despite the mimicking demo in the introduction, they seldom used multiple fingers for interacting with the surface. In order to open or close the content of the cup, for example, some users were just quickly tapping on the body of the cup itself with their index and middle fingers (i.e., similarly to a mouse click gesture) instead of continuously rotating it with the whole hand.

In a Flash mock-up it is not possible to implement, obviously, the use of multiple fingers as a constraint: Thus, the graphical output behaves according to a single input point, such as tapping. The implementation of other constraints in the software could have possibly influenced users' interaction behaviors in a different way. This simple mock-up, though, was useful to explore potentials and limitations of the design ideas, and suggested that some issues about users' expectation of digital behaviors in hybrid interactions should be investigated further and more systematically (cf. Chapter 7). Additionally, it inspired the idea that if one was able to implement or simulate haptic feedback to the action afforded by a visual cue in the display, this fact could possibly affect users' behaviors and learning process: E.g., the force to be applied on a cup in order to incline it could be coupled by a haptic resistance; full cups could be harder to move than empty ones. Such a conceptual model could be reinforced by audio cues (e.g., the sound of splashing water when the content of the cup is poured) so as to provide sensorial affordances (cf. Chapter 2, Section 2.7).

In this sense, the multimodality of interaction with digital information can potentially provide more analogue interfaces. Muscular memory (rather than purely cognitive one) could then be exploited by users to structure frequent, patterned tasks into stylized or abbreviated gestures (such as is the tenet of crossing-based gestures, cf. Chapter 3, Section 3.1.1), thus providing us, as designers, with the possibility to build a gesture vocabulary that exploits humans' sensorimotor perception and haptic memory.

All in all, this first exploration made clear that the potential benefits of haptic feedback needed to be further unpacked, experimented, and assessed: Below, another interaction sketch is described, which explores ways of taking advantage of physical affordances and the haptic feedback they provide.

4.3 Reflecting on 3D Volumes

Referring to different aspects of haptic feedback (cf. Section 1.1.3, footnote 1), one can distinguish between *touch* (or somesthesis) and *proprioception* (or kinesthesis): The first one consists of somatic sensibilities aroused by stimulation of bodily tissues such as the skin, and provides *passive haptic feedback*. The second one is the sense of position and movement of the limbs and the sense of muscular tension, which provides *active haptic feedback*. The related literature on input devices usually refers to active haptic feedback and is still very controversial in terms of the benefits and implications thereof (Buxton, 2007a). Such controversies are especially about what kind of input devices (e.g., isometric, isotonic, or elastic) can better support different types of controls (i.e., rate and/or position) according to different criteria (e.g., learnability, muscular memory, accuracy, and feeling of control). For a thorough and critical analysis of the related literature on this topic see (Buxton, 2007a) and (Zhai and Milgram, 1993). Although it is beyond the scope of this thesis to enter such a discussion, it is worth reflecting on some of the implications of haptic feedback on affordances.

What is relevant in this context is that, despite several issues remaining unsolved in the understanding of haptic feedback, it seems that proprioceptive feedback from the control device is a facilitator of control actions. Now, when considering the feedback afforded by direct touch interactive surfaces, one can actually argue that they also afford haptic feedback in terms of somesthesis, due to the skin sensations emerging from the contact and movement of our limbs in relation to the surface of the display. The type of haptic feedback afforded by graspable 3D objects, though, is quite different in terms of motor control, as they normally return a higher force-feedback. Furthermore, 3D graspable objects imply the involvement of more muscle spindles, which are currently considered the major source of proprioception (Buxton, 2007a). Proprioceptive feedback, in turn, is considered as the main source of the feel of control (Zhai and Milgram, 1993).

These considerations are functional to a reflection on the affordances that we can design for the manipulation of digital media. As computing technologies really become smaller and embeddable, we can start integrating both input sensors and output displays into physical objects, environments, as well as tasks. This creates a number of new challenges and opportunities for designing hybrid tools that afford haptic proprioceptive feedback. The manipulation vocabulary designed for such tools can metaphorically refer to the one of 3D objects and provide cognitive as well as physical affordances for interaction. In everyday life, physical controls have indeed a manipulation vocabulary which derives either from physical constraints (e.g., gravity and viscosity forces) or from stereotypes (e.g., when steering the car wheel clockwise the car moves right and vice-versa). If we metaphorically refer to those manipulation vocabularies in the design of interactive objects, we can then provide both cognitive as well as physical affordances for users to build a mental model of manipulation in hybrid interactions (see Fig. 4.3).

To this end, a first step is to start thinking in terms of exploration and manipulation of physical objects in the 3D space, so as to consider the haptic feedback this can provide. In experimental psychology the word *haptic* refers indeed to "the ability to experience through active exploration, typically with our hands, as when palpating an object to gauge its shape and material properties" (Robles-De-La-Torre, 2006). In this respect, the WIMP paradigm has limited to some extent the way in which we experience digital information: Mostly, our visual perception is engaged, and less so our other senses. On the other hand, there is evidence (O'Regan and Noe, 2001) that we make sense of the world with all our senses and body.

Furthermore, physical activity helps to build representational mapping (Rieser et al., 1994). Kirsch (1995), in particular, reports on the value and explorative nature of epistemic actions (cf. Chapter 2, Section 2.3): These make possible the externalization and the trying out of different alternatives in temporary spatial structures so as to offload the cognitive effort of mentally imagining the unique right solution.

In the interaction with the desktop PC, our possibilities for epistemic action, physical exploration, and sensorimotor perception are very limited, mainly because of the time-multiplex nature of the input modality (cf. Chapter 2, Section 2.4). This fact, in turn, poorly supports our spatial reasoning. To put it differently, the desktop PC is a poor notational tool for representing spatial tasks because the way of interacting with the object does not support exploration and sensorimotor perception, which we use in the experience with physical tools. On the other hand, computing technologies offer other features, such as sensing of events (e.g., gestures), automatic response (e.g., update of the information display), and reversibility of actions, which go beyond what our physical tools can normally do. This raises the question: How can we take advantage of the possibilities of digital technologies so as to create hybrid tools that support our exploration and, in turn, spatial rea-

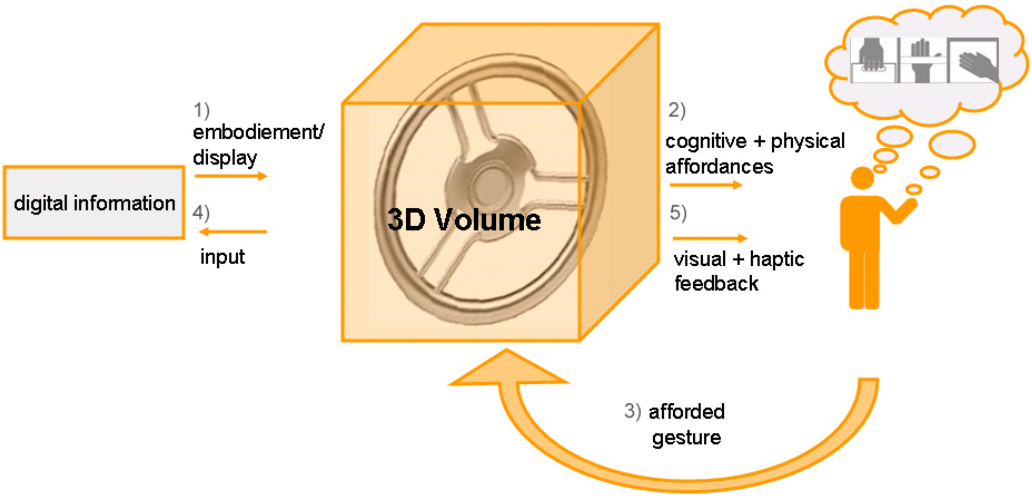


Figure 4.3: The figure illustrates the idea of metaphorically mapping the physical affordances of real world objects to cognitive and physical affordances for gesture-based manipulation of digital information with interactive objects.

soning? By breaking the conventions about how a computer looks, how it can be "grasped", and about the place in which computing resides, we can possibly create tools that support our reasoning and better reflect the way we naturally explore and learn in the physical space. To demonstrate this idea, the Learning Cube (Terrenghi et al., 2006c) was designed and prototyped.

4.3.1 The Learning Cube

As discussed by Gutierrez (1996) "everyday life provides plenty of interactions between plane and space, and most of them imply the dissemination of some kind of spatial information by means of plane data (drawings, schemas, pictures, figures, etc.). But text books are still plane". The same consideration can be done for much of our computing technologies. With the Learning Cube appliance the enhancement of spatial reasoning to teach children space geometry is investigated by exploiting qualities of both digital and physical technologies.

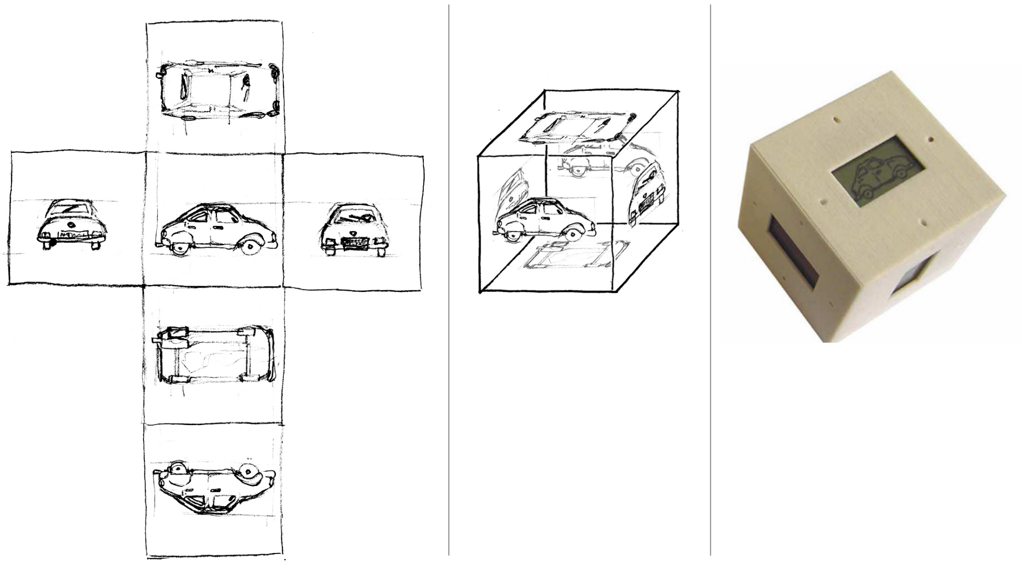
As discussed by Druin and Inkpen (2001), children activities are mostly disconnected from the desktop PC environment. Children, especially young ones, play and move around in the real world, manipulate different objects, talk loud, and like to explore. The playing activity is as well a way of exploring and, in turn, of learning. This is in line with constructivist learning theories (Piaget, 1972), whose tenet is that children learn while actively being engaged in explorative and problem solving activities which are embedded in the physical experience. Rieser (1994) has shown that physical movement can enhance categorization and recall in tasks of perspective taking and spatial imagery. Furthermore, recent neuro-scientific research suggests that some kinds of visual-spatial transformations (e.g. mental rotation tasks, object recognition, and imagery) are interconnected with motor processes and possibly driven by our motor system (O'Regan and Noe, 2001).

Besides the static, individual nature of the interaction provided by the desktop PC, this presents additional constraints for children: As reported by Smets (1994), children encounter difficulties when interacting with control devices (e.g., the mouse) and a detached 2D screen because action and perception are spatially separated.

The design of tangible user interfaces has addressed some of these issues and the learning activity in a number of examples (e.g., (Resnick et al., 1998), (Price and Rogers, 2004)). In the Learning Cube appliance here presented, the physical affordances of a cube are mapped to the conceptual model of orthogonal representation, thus leveraging such physical affordances to cognitive ones. This appliance for learning spatial geometry builds, indeed, on the design of a semantic link between physical manipulation/control, digital output, and abstract concept, thus providing a redundant learning interface. Referring to the qualities of physical interaction identified in Chapter 1 and represented in Fig. 1.3, in this case the integration of the following aspects is explored:

- Metaphorical representation;
- Directness;
- Continuity of action;
- 3D Space of manipulation;
- Physical constraints;
- Multimodal feedback;
- Two-handed cooperative work.

Orthogonal views are very common as representations of 3D objects in technical drawing education: The object is supposed to be in a cube and projected orthogonally on the six faces of the cube. Accordingly, orthogonal representation is often illustrated through the idea of an unfolded box, whose faces are displayed on a plane (see Fig. 4.4, a and b). The possibility to embed



a) b) c)

Figure 4.4: a) The usual 2D representation of orthogonal views. b) The box as conceptual model for teaching spatial geometry and technical drawing. c) The Learning Cube appliance for learning spatial geometry.

computing capabilities into playful tangible objects can possibly support children to explore, move, and analyze mental images of 3D objects generated from the information brought by a plane drawing.

The spatial geometry learning appliance was designed in collaboration with the Embedded Interaction group at the University of Munich, who developed and implemented the general Display Cube platform (Kranz et al., 2005). The Display Cube is a digitally augmented physical cube: It is enriched with a small display on each of the six faces and a speaker inside (see Fig. 4.5, a). The displays and the speaker are controlled by an embedded hardware platform. This is a low-power micro-controller board with wireless communication capabilities and several built-in sensors and actuators. With this set-up it is possible to track the movement of the Display Cube in 3D. The acceleration sensors are calibrated for delivering meaningful information about the extent and the direction of the movement of the Learning Cube. So it is possible to determine which face of the cube is on top (parallel to the floor plane) and whether the cube is shaken.

Such a basic platform offers the function of a multiple choice test system. In the spatial geometry appliance, the faces of an object are represented orthogonally on the six faces of the cube (see Fig. 4.4, c) and one of them does not match the representation. The goal is to identify the wrong orthogonal representation. The interaction operated by the user is to turn the cube to the side with the right answer (i.e., the wrong orthogonal view) on the top, and then to shake the cube (see Fig. 4.5, b). If a shake is detected by the embedded sensors, it is checked whether the correct answer is printed on the display currently on top of the cube. If so, the user gets positive feedback and the next test is displayed on the cube. If the answer is incorrect - that is, if any other face than the one with the correct answer is facing up - the answer is considered as false. Negative feedback is then given to the user showing that this answer is incorrect and the loop is then started again.

a) b)

Figure 4.5: a) The Digital Cube disassembled; b) Children interacting with the learning appliance.

4.3.2 Observations

The appliance was tested in informal experience trials with children in the age from 7 to 12. Whilst such trials do not provide generalizable results about the effectiveness of the learning experience (whose assessment would imply longer time of analysis and more controlled experiments) a number of observations could be drawn.

The 3D features of the device supported exploration through physical manipulation. E.g., children would turn the object around in the air and on the table, try different gestures for picking it up, shaking, and putting it down again. Furthermore, the feeling of physical control on the digital display seemed to engage children in looking for the right solution so as to change the displayed object. In this sense, the playfulness of the appliance seemed to trigger a motivation effect. Additionally, the 3D volume of the device makes it seeable from different perspectives: This appears not only to support exploration, but the sociability of the task too (see Fig. 4.5, b, for example). In other words, the artifact provides social affordances for a collaborative solution of the task. Children played together, helping each-other, arguing, and showing solutions. Whilst this can be partially due to the novelty of the appliance, one can also speculate that the "graspability" and "multi faceted" features of the device can inherently afford social protocols which strongly divert from the way in which computers have assisted our learning paths so far (which are normally individual, due to the PC interface for single user).

4.4 Open Discussion through Open Sketches

The two designs presented in this chapter have suggested a few insights, thus stimulating further investigation. To summarize:

- The metaphorical representation of affordances for manipulation in 2D was easily understood, but interaction was not always correctly executed in terms of gestures, i.e. at the pragmatic level: Some users seemed to use preferably one single finger rather than multi-touch, and discrete actions, rather than continuous ones.

- The 3D manipulation vocabulary of the Learning Cube afforded exploration and social collaboration among children, and it seemed that a sense of physical control on the digital output had an engaging and motivating effect.

There are other considerations on these two sketches that go beyond the designs themselves and regard the style of representation of an interaction design concept. The Mug Metaphor Interface was initially sketched with a marker on a whiteboard in the lab to explain the idea and share it with other members of the research team. Those sketches provoked discussion and brainstorming among a small group of colleagues. To gain further feedback from other people outside the team, we developed the Flash mock-up in the graphical style previously shown in Fig. 4.2 and made a movie simulating the different interaction tasks. This was a very suitable tool for "presenting" the idea, but less so, to a certain extent, for "brainstorming" on it with others. The style of representation was such that when some people viewed the video they thought this was a fully working technology. As suggested by Buxton and presented in (Buxton, 2007b), we then developed another movie identical to the previous one, but with a more sketchy representation style (see Fig. 4.6). When other people saw this movie they took for granted that this was a mock-up, an externalization of an idea to be further developed.

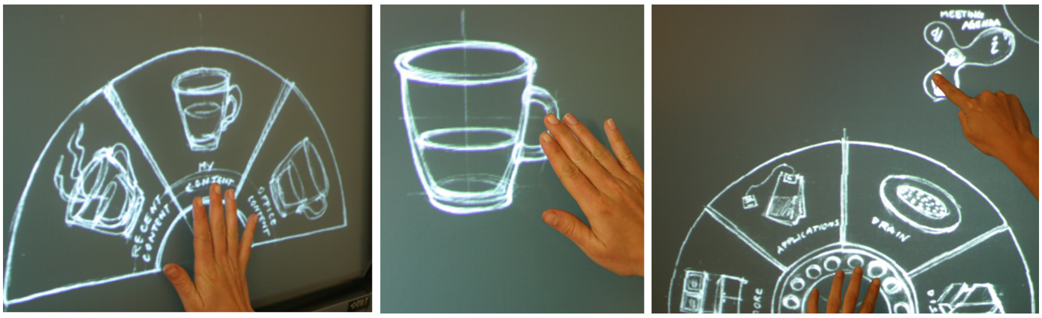


Figure 4.6: Sketches of the Mug Metaphor Interface.

On the other hand, the level of refinement of the Learning Cube was such as to allow a real hands-on experience. Despite the fact that the look of the device was rather primitive and the appliance was not completely reliable, the main functionalities were actually implemented, the prototype relying on an existing platform (i.e., the Display Cube (Kranz et al., 2005)). In this particular case, it would have been hard to try out interaction design ideas with children if the sketches had been "too sketchy". Put differently, it was necessary to represent the design ideas at a level of refinement sufficiently credible and engaging. This fact implies that, as designers, we need to thoroughly think about the targets of our sketches and what we want to gain from them (e.g., brainstorming, feedback, assessment). There is indeed a trade-off between the level of refinement of a sketch and its ability to generate collaborative brainstorming. On the other hand, the sketch of an interaction needs to convey a level of credibility sufficient to let others imagine its functioning, especially when the targets of those sketches are potential users. These considerations might lead to different styles of representations to share ideas with different stakeholders, at different stages of their elaboration.

The work presented in this thesis builds on these considerations and has gone through different iterations, depending on the feedback and level of assessment that it was sought to gain from others. In the following chapters (cf. Chapter 5 and Chapter 6), different examples of what can be considered "experience prototypes" are presented. In these cases, the goal was to achieve a stage of refinement of the design concepts that was sufficient to let other people experience them, in a context that was as real or plausible as possible, and compatible with hardware constraints.

Thus, the explorative design process moves from sketches to experience prototypes: These are used as kind of "probes" (Gaver et al., 1999) to create a common ground with their potential users (cf. Chapter 5), or as test-beds to observe interaction and communication patterns (cf. Chapter 6).

In this way, it was sought to gain an understanding of the implications of integrating some of the aspects of physical interaction in their design, as well as the implications of introducing digital interactions in those specific physical and social contexts.

5

Design of Hybrid Artifacts for Social Engagement in the Home

This chapter presents two case studies. Both take into account the reflections discussed in the previous chapter on alternative approaches to the design of affordances for the manipulation of digital media. These approaches were implemented into the interface design of experience prototypes of hybrid artifacts for the home.

Building on a consideration of how some physical artifacts (e.g., a cookbook and a mirror) mediate interpersonal communication and embody home organization as well as social relationships, the goal is here to extend and enhance those aspects. To this end, this chapter suggests design opportunities to augment physical artifacts as well as mundane practices by taking advantage of some of the qualities of digital, ubiquitous computing technologies.

5.1 The Living Cookbook

As anticipated in Chapter 1, the work presented in this thesis builds on the consideration and acknowledgement of different affordances of physical artifacts (i.e., physical, functional, cognitive, sensorial, and social affordances, cf. Chapter 2, Section 2.7) and explores design opportunities of exploiting some of the qualities of digital technologies in order to augment those artifacts and the engagement, communication, and self-expression they mediate. Within this scope, the design of the Living Cookbook builds on the idea of augmenting a physical cookbook by providing:

- family generated content, thus reinforcing a sense of community;
- a sense of presence, extended across space and time;

- an opportunity for self-expression and passing on of family practices and values;
- an engaging experience, which can motivate people to hands-on, contextualized learning (as well as teaching) of cooking.

This section further discusses the motivation of the project, its development and iterations, the validation of the design goals in context, and the lessons that were learned from these activities.

5.1.1 Motivation

The main idea motivating the design of the Living Cookbook appliance is to make people's cooking experiences recordable and shareable across time and space constraints, in such a format that can foster a sense of presence. Instead of simply exchanging written instructions, with the Living Cookbook people can capture the whole cooking process with annotated audio and video and make it available for others so that they can asynchronously reproduce the dish. In this sense, the appliance is similar to a family photo album, composed of recorded and shareable family "kitchen stories" (Terrenghi et al., 2007a). The high level goal is indeed to support social bindings so as to preserve cultural and social roots on the one hand, and stimulate cultural and generational fertilization on the other.

The design approach draws upon the consideration of the mundane practices of the home and of the artifacts that support these practices. Families grow and evolve as communities of practice (Wenger, 1998), which as such rely on their members, rituals, as well as artifacts. The instruction and apprentice of home practices within the family walls happens in a large part through storytelling, performance, observation, and practical routines. Adult family members "play" the model for younger generations in the way they manage domestic activities such as cleaning, tiding up, as well as cooking. When family members are remote from each-other, e.g. a parent is away for work or a child moves out, instructions are mediated via different channels: for example, by text for exchanging recipes, by instruction notes next to home appliances to illustrate how to operate them, and by telephone calls for synchronous communication of instructions.

The focus of the Living Cookbook appliance is on domestic learning of the cooking practice. By exploiting the possibilities offered by digital technologies for capturing, archiving and displaying multimedia instructions "on demand", its design aims at supporting the collaborative practice of the cooking activity.

By enabling parents to record their "special pasta" or "unique roast beef" for their children - customizing each recipe with personal tips and tricks - one can expect that very personal experiences can be created and communicated. People often call their parents and friends to ask "What was the recipe of that dish?"; "What does the sauce have to look like?"; "How thick should it be?". Much of the communication around cooking, such as the exchange of recipes, and especially within families or close social networks, is one-to-one and supported by different media (e.g., paper, telephone, and e-mail), and is a way of tightening social bindings. The emotional quality of content created by family members or intimate friends is indeed very different in comparison to the cooking sessions published in books or broadcast on television programmes, which are produced for a larger audience. Although TV programmes provide a multimodal presentation of food preparation and take advantage of the popularity of acknowledged chefs, they cannot be consulted on demand as a paper cookbook, nor they can be personalized. Furthermore, TV cooking shows are often watched in spaces and time slots which are detached from the actual cooking activity: They are rather watched in contexts which depend on the location of the TV display and on the TV schedule.

The goal of the Living Cookbook appliance is to provide an alternative way for people to personalize their cooking experience and, as a consequence, their communication. To this end, its design draws upon some of the qualities of computing technologies, and in particular:

- storage and retrieval of multimedia content on demand: This allows for a personalization of the creation and consumption of content;
- multimodality: This contributes to augmenting the sense of presence;
- embeddable technology: This provides the possibility of creating and consuming content in context.

Thus, this project explores ways in which technology can support time-spending and engagement in the home which are motivated by the added value of accomplishment, self-expression, and social exchange. In this sense such an approach to ubiquitous computing technologies for the home diverges from the ones focusing on optimization of efficiency, invisibility, automation and time-saving (cf. Chapter 1, Section 1.2.1). One could actually argue that the way in which multimedia technologies have supported time-spending in the home (e.g., home entertainment such as TV or video games consoles) has confined users to a rather passive, and not very creative role.

It is a claim of this thesis that ubiquitous computing technologies also offer a good potential for supporting and motivating creativity and learning, which has not been thoroughly exploited by the Ubicomp agenda so far. Put differently, these technologies can also provide rich resources for the design of novel learning experiences, which can be situated in the very space and social context of the home.

In this sense, cooking is here considered as an engaging, collaborative, learning experience, rather than as a mere working activity. Below the evolution of technologies for the kitchen is treated in more depth, to characterize the design domain.

5.1.2 Technology in the Kitchen

The kitchen has traditionally been populated by a diverse range of artifacts supporting food preparation. Such a domain provides a rich collection of specific as well as open tools, which often embody and represent the material culture of a society and of a generation. In western kitchens, most of the electric appliances have focused on lessening users' mechanical effort, or on maximizing the efficiency of certain tasks, or even their complete automation (for a good review see (Bell and Kaye, 2002)). Digital display technologies, on the other hand, have had a very limited application so far. Mostly, they have been used as interfaces to control electric appliances, as alternatives to - or in combination with - physical dials and buttons. More recently, some companies producing domestic appliances have addressed the potential of digital displays and internet technology for augmenting the kitchen environment, bringing information and entertainment in the kitchen. The GR-D267DTU Internet Refrigerator by LG[1] contains a server which controls the communication to the other connected appliances. On the display different functionalities are embedded: It is possible to watch TV, listen to music, or surf the internet. A microphone and a camera are built-in, thus enabling multimedia communication.

In academic research, some ways to augment the kitchen environment have been investigated as well. At MIT, a smart kitchen space named "La Cantina", was set up (Bell and Kaye, 2002): Here, displays are embedded in the space for different augmentation purposes, mostly dealing with artificial intelligence. One of the proposed scenarios is instantiated in the Counter-Active project (Ju et al., 2001). This is an interactive cookbook, projected down onto the kitchen counter; the cook touches the countertop to navigate through the recipe or to glean greater details. Recipes incorporate pictures, audio, and video.

[1] LG GR-D267DTU freezer: http://www.lginternetfamily.co.uk/fridge.asp

Similar to the CounterActive project, the Living Cookbook aims at augmenting the cooking experience and the traditional cookbook by delivering and displaying multimedia content. Here the focus, though, is on augmentation by social and family relationships and real life experiences, rather than on multimodality per se. Indeed, one aspect to consider in the design of kitchen appliances is that cooking is often social, and involves several rituals and symbolic aspects. Some people enjoy cooking together; several people enjoy cooking for others as a sign of care; friends and relatives often exchange recipes, which assume a cultural and communicative value. In this sense, cooking seems to offer a great potential for communication and the enforcement of the social bindings of a community of practice.

The design of the Living Cookbook explores how communication and display technologies available today can offer the possibility to support the communication and sociability of cooking, and bring new aspects to its social and engaging character.

5.1.3 Design

The Living Cookbook appliance consists of an application running on a tablet PC: This has a touch-sensitive surface enabling users to interact either with a pen or directly with a finger. On the tablet PC a digital cookbook is displayed (see Fig. 5.1). On the same interface people can either author a new recipe and add it to their personal virtual book, or consult the book and learn someone else's recipe. In the authoring/teaching mode, the video of the cooking session is captured by a camera. In the learning mode this video is played back and the learner can cook along. When recording, the cook can indicate phases of activity and inactivity in the user interface. When playing back, the device projects the recorded video of activities, pauses during times of inactivity, and the cook can speed up or slow down the playback of the recorded session by advancing to the next section or pausing to catch up.

The design of the GUI relies in several ways on the considerations on metaphorical representation for manipulation of digital media in 2D, which were discussed in Chapter 4, Section 4.2. In this respect, physical artifacts are here metaphorically referenced both for their manipulation vocabulary in physical interactions, as well as for their role in the specific domain.

In the Living Cookbook, the metaphor of a traditional cookbook is used to represent the conceptual model and suggest the types of actions to perform. The book, indeed, can metaphorically draw upon some of the affordances of paper, where people can both write and read, as well as flip pages. Furthermore, in the learning domain the book is often associated with the archiving of knowledge and is the medium to transfer it. Users can go back to their personal digital cookbook, see the recipes they authored as well as the ones authored by others, thus enhancing a sense of ownership and of shared memory. These analogies are adopted to display both the authoring (i.e., teaching) and rendering (i.e. learning) environments, thus using a consistent interface for both modes of interaction.

Besides the use of the book, the screen design also makes use of metaphors. Several widgets refer to artifacts of a normal kitchen and semantically relate to different functions: The dial (see Fig. 5.1, a) embodies the cookbook selection; portions can be specified by dragging plates on a table (see Fig. 5.1, b); and video control is operated on an egg-shaped widget (see Fig. 5.1, c). On the dial people can choose among a set of cooks/buddies, and among courses. This combined selection triggers the appearance of the cover of the book displaying the picture of the selected cook/home inhabitant, and of the desired course. By tapping on the cover, users can open the book, which displays a list of recipes. Pages can be flipped back and forward by tapping on a plied corner in respectively the left and right bottom corners of the book: Such page corners are smoothly animated to capture users' attentions and suggest their pliancey. A single tap triggers an animation simulating the flipping of a single page.

The general idea behind the design of those metaphorical widgets is to minimize text input, which is tiring on a soft keyboard, especially when the display is vertically mounted (e.g., on a cupboard). Thus, wherever possible, the interface affords direct manipulation, such as tapping and dragging. Instead of entering the text for ingredients' names and quantities, for example, cooking authors select them from categories represented by pictures (e.g. Fig. 5.1, d). This, in turn, responds to the goal of reducing the display of text lists that require the reading and scanning of the screen, which could be awkward in the meantime of cooking. Furthermore, it responds to the goal of representing an interface which would appear simple and familiar for an extended audience of users, who do not necessarily use desktop PCs on a regular basis.

To validate the design goals and relative design choices, the development of the project went through different iterations, and different technical set-ups were adopted for the implementation of the appliance. First, the kitchen of the lab in Munich was instrumented with two cameras and a projector. In a following iteration, a portable version of the appliance was developed, which made possible to test it in real domestic kitchens, as well as in a focus group with pupils. Below, such different set-ups are described, with which different user tests were conducted, adopting different methods.

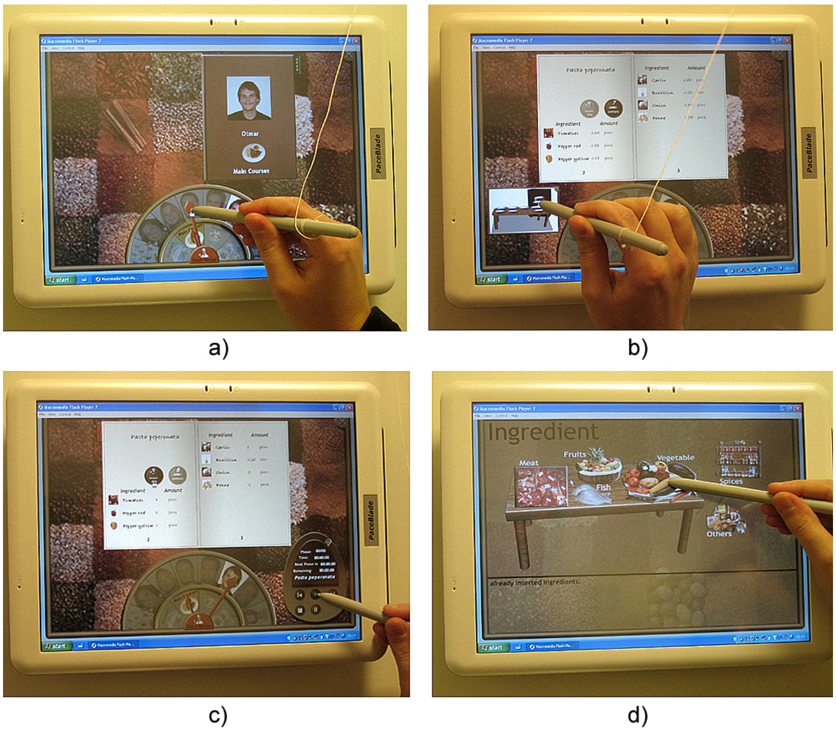


Figure 5.1: a) Selection of cook and course on the dial; b) Selection of number of portions; c) Video Control; d) Selection of ingredients category.

5.1.4 Evaluation in Different Settings

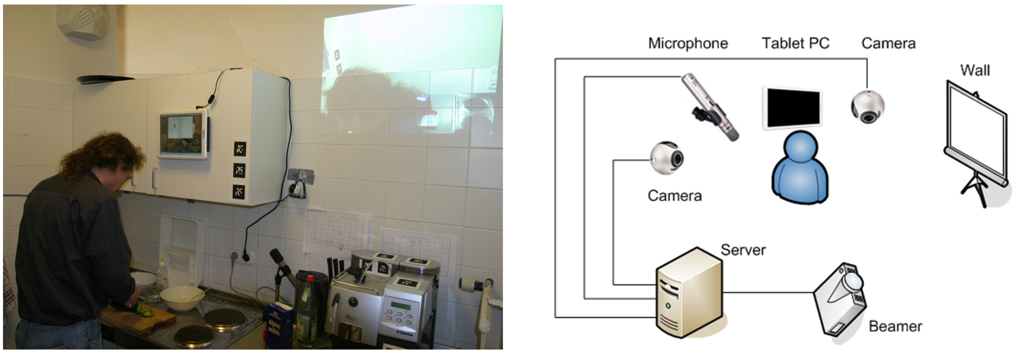
The main goals in evaluating the appliance were:

- observing how people interact with the interface at the pragmatic level;
- assessing the subjective experience of using it as a home appliance within a family;
- assessing its potential for domestic learning.

Addressing these topics required for adopting different techniques and for involving different groups of participants. The methods which were used are clearly qualitative: Such an approach is rather common for exploring idiosyncratic domains such as the home, where generalizable arguments are hard to make even on quantitative data. The next sections report on the different settings and the techniques which were adopted for evaluation.

In the Kitchen of a Research Lab

In the first iteration, the appliance was implemented and tested in the kitchen of the lab (see Fig. 5.2, a). One PC was used as a server and a tablet PC was mounted on the cupboard above the stove. A beamer, connected to the server, was used for displaying the video on a wall beside the cupboard, as illustrated in the schema of Fig. 5.2, b. A camera was connected to the server too, for recording the cooking session.



a) b)

Figure 5.2: a) A participants is interacting with the Living Cookbook in the kitchen of the lab; b) Schema of the hardware of the Living Cookbook.

In a preliminary phase, some experience trials were conducted in the lab together with team members and colleagues, who were invited to instruct or learn in a cooking session. Such preliminary trials suggested that two cameras rather than a single one were desirable for capturing both the person and the details of food preparation. Thus, the server was connected to two cameras for capturing audio and video, and a split screen was designed for the playback interface. Furthermore, those tests elicited the need to reduce as much as possible pen-based interaction because users are busy with handling kitchen tools and ingredients (Terrenghi, 2006b). This suggested the introduction of a speech recognition component for the control of the video: Such a component was based on the Sphinx system, developed at the Carnegie Mellon University[2], and was customized to process certain commands only, such as “record”, “play”, “pause”, and “stop”. The speech input was conceived of as alternative input modality to control the video recording and playback. The limited quality of performance of the sound hardware embedded in the tablet PC motivated the decision of implementing the speech

[2]Sphinx, Speech Recognition System, Carnegie Mellon University, http://www.speech.cs.cmu.edu

recognition module on the server PC, thus connecting the server PC to an external microphone.

Such a set-up was tested in 6 cognitive walkthroughs with participants external to the research team. At this stage, the goal was to validate the usability of the interface at the pragmatic level, as well as to gather a more general feedback about people's experience of interacting with the appliance.

Four of the participants were usability experts, thus providing a heuristic evaluation. The other 2 participants (both psychologists professionally dealing with topics of Computer Supported Collaborative Learning) were members of the families in which the portable Living Cookbook was tested later on (see next paragraph). Each participant was introduced to the appliance and then instructed to accomplish either the authoring or the learning of a recipe task, alternatively. While interacting with the interface, they were invited to use kitchen utensils and reproduce cooking tasks, like cooking water or cutting something, so as to simulate as much as possible a complete cooking session in a domestic kitchen. During the test, they were asked to talk aloud. At the end of the cognitive walkthrough, they were asked to answer a questionnaire inquiring about their general satisfaction in using the appliance: what they found useful and what was instead irritating; their impressions about the speech input modality; suggestions for improvements, and whether they could expect using it in their own kitchens.

All the participants were able to accomplish all tasks and reported a positive feedback, especially regarding the graphical user interface of the appliance. Some interactions occurred, though, which caught the researchers' attention. Despite the fact that every user understood that s/he was supposed to select a name and a course in the dial for triggering a corresponding cookbook, the way in which people did so were diverse. Instead of dragging the edge of the pointer on the desired sector (see Fig. 5.1, a) in the way the interface was implemented for, most of the participants directly pointed the stylus on the sector they wanted to select, and were expecting the lever to automatically move there. After some time or with some help from the researcher, they were then able to point on the lever and drag it over the desired picture with a continuous gesture. When doing that, some of the participants commented: "Of course, that's clear, I don't know why I didn't think about that immediately". Another person mentioned "We are so much used to click with the mouse that I would expect to click here as well, instead of moving the pointer. May be a housewife or someone who does not use the computer on a daily basis would do it differently". These observations motivated the implementation of the automatic movement of the levers of the dial to the selected sector in the next iteration. Consistently, it was decided that the pages would flip automatically after a single tap on the plied corner of the page, rather than requiring the user to drag the page from one side of the book to the other one (which would actually be the case when manipulating a physical book).

Such interaction patterns suggest that people potentially develop different mental models, which might derive from previous interaction in the WIMP paradigm. The observations of the relationship (and mismatch in a certain sense) between analogue actions in the physical world and discrete ones in hybrid interaction contexts are somehow similar to the ones emerging from the trials of the Mug Metaphor Interface (cf. Chapter 4, Section 4.2.2): These are further discussed in Section 5.1.5.

The other interaction tasks, like inserting the name of a new recipe, inserting ingredients, and specifying the number of portions, proceeded smoothly. Sometimes, though, users were curious to go back in the process to check the effect of the actions that they had performed: Several testers seemed particularly concerned about the correctness of the step-by-step interaction process. From a design point of view, this suggests that feedback, reversibility, and error tolerance are very relevant requirements for such an interface that is operated in the meantime of other tasks, such as cooking.

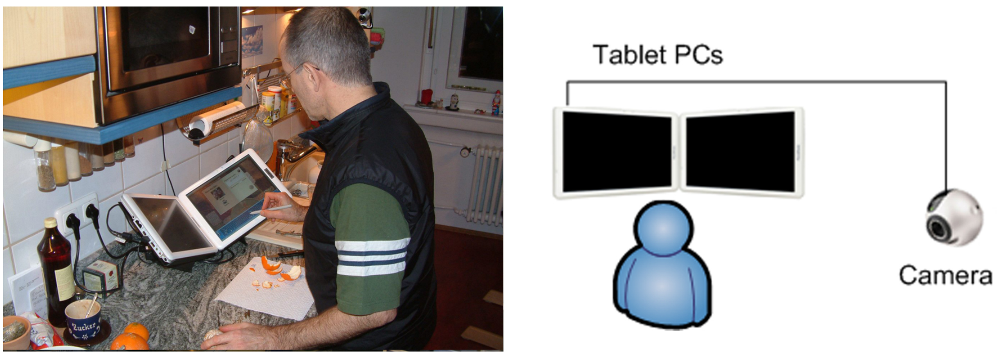
The speech command did not perform in a satisfactory way: The system recognized repetitively the instructions spoken by the researcher, thus revealing that including such a feature in a real domestic kitchen, where more people are present, would hardly work.

Regarding participants' expectations of use in their personal kitchens, mixed answers were reported. Most of the participants saw a possible use in the playback/learning mode, rather than the teaching one, and could imagine the "utility" of cooking along video instructions, which provides a better feedback. A main concern, still, was due to the hardware, both in terms of the space and the cables it requires, as well as its audio performance in playback mode, which is rather poor.

Overall, despite the fact that such preliminary trials and the cognitive walkthroughs shed some light on the usability, desirable features, as well potential further developments of the appliance, it was evident that some of the goals of the project (i.e., motivating and supporting familiar learning and communication) were hard to assess in such a setting. During the cognitive walkthroughs, the participants were focused on completing the task, operating the interface, and commenting on it, but they could hardly imagine whether and how they would use the appliance in their own kitchens and with their own families. Furthermore, even though some of the testers were employees of the lab and already knew the environment of this specific kitchen, the activities they normally perform there are different from proper cooking (e.g., coffee preparation, food warm-up, and food storage). When the participants, both the employees and the external testers, were asked to use the application and to cook there, it became clear that the condition of cooking in an unfamiliar kitchen (i.e., where tools and ingredients are different, and are stored in differen places than in the participants' homes) generated stress and a sense of awkwardness. These considerations motivated the deployment of a portable version of the Living Cookbook, which allowed for testing it in the domestic kitchens of two families. The study is presented below.

In Domestic Kitchens

To better understand whether and how such an appliance can be embedded in the familiar context of a domestic kitchen, a portable Living Cookbook was designed and installed in the kitchens of two of the participants of the cognitive walkthroughs. The set-up consists of two identical tablet PCs residing in a custom crafted metal frame, so that the physical appearance of the whole appliance suggests the shape of an open book (see Fig. 5.3, a). This can be put somewhere on a table or on a counter so as to more easily fit in an ordinary kitchen, the only constraints being the power cords and the camera connection. The right tablet PC displays the user interface, while the left one displays the video during recording and playback. They can either communicate through an ad-hoc or an existing wireless network. The hardware which was used in such a set-up (see schema in Fig. 5.3, b) provided compactness, but it also implied some constraints: Only one camera could be plugged-in, due to the limitations of the hardware implementation of the USB technology on the tablet PCs; furthermore, mounting two cameras in someone's kitchen, would have implied an additional burden of cables; and finally, considering



a) b)

Figure 5.3: a) A participant is interacting with the Living Cookbook in the kitchen of his home; b) Schema of the hardware of the portable Living Cookbook.

the shortcomings of the sound hardware which were mentioned above, the component and interface for speech input were omitted in the portable set-up. Indeed, as it appeared during the cognitive walkthroughs, the speech recognition is likely to perform poorly when there is more than one person in the environment, which is often the case in a domestic kitchen.

Two families were introduced to the appliance. The device was installed in each kitchen for a week and the families were asked to report on their experience of use during the in-depth interviews which were conducted afterwards. One household consisted of a couple with a young baby (husband 30 years old, wife 31, child 15 months) and the other one of a couple with three children (husband 54 years old, wife 39, daughter 15, son 8, and another daughter 5). Five people in total (the two couples and the 15 years old daughter) used the appliance and were involved in in-depth interviews.

Every participant liked the idea of a family archive. The first family used the appliance in the teaching mode only, but actually the feature that they said they liked most was the possibility of watching family members as they cook, in the playback/learning mode. They reported that in comparison to watching professional cooks on TV, the videos with the Living Cookbook promoted fun and intimacy by the fact that they showed well-known people in ordinary activities.

In the second family the parents said that they did not enjoy using the appliance as it was perceived as an additional domestic effort, "one more task to take care of in home organization": In other words, it was too much of an engagement of time. The 15 years old daughter, on the other hand, did enjoy the experience with the Living Cookbook and even invited a girlfriend to help her cooking along her parents' video. She said that she had fun in cooking along her parents' instructions, check whether they were making mistakes in the kitchen, and stated that she would be more motivated to cook because of the entertainment added value.

The diversity of the participants' responses about the experience of using the appliance suggested to look more carefully into possible generational different attitudes and motivations towards the use of such a kind of "home video". To this end, as reported in the next paragraph, a focus group with teenagers was organized, as they were considered possible target users of the Living Cookbook.

In a Focus Group

In order to better understand how such a kind of "home video" can motivate the learning of cooking, a focus group was conducted involving 8 pupils, between 14 and 16 years old, who attend the course "Household and Nutri-

a) b)

Figure 5.4: a) The pupils discussing the learning of cooking in the focus group; b) The pupils engaged in the preparation of a sauce following the video instructions on the Living Cookbook.

tion" in a German secondary school. In such a course they learn cooking techniques and recipes: Thus, they could share with the research team their insights about learning to cook in both an institutional, as well as in a familiar educational context, i.e., in their own families. During the focus group several topics were discussed, such as: the main phases of the cooking process; TV cooking programmes; the use of cookbooks; the learning to cook at home with family members vs. in school; and how different media can support the learning activity. The pupils were then introduced to the Living Coobook appliance and were asked to cook along a simple recipe that had been authored for them: This provided them with a hands-on experience (see Fig. 5.4) which they could share and comment in real time within the group. The discussion was then focused on their impressions, ideas, and criticism about the appliance.

Their feedback was in general very positive, and they found the appliance engaging and entertaining. The pupils could envision an increase in their motivation to cook at home because of the fun factor of watching the videos of parents, siblings, as well as grandparents managing mundane activities, as a kind of "real TV". On the other hand, the pupils could not imagine using the Living Cookbook at school, where the learning environment is more rigorous and structured.

It was also mentioned that in comparison to TV cooking programmes, they would "trust" more the recipes of the Living Cookbook because they are likely to be dishes they have already tasted at home, and because they seem to be more doable than the ones shown by professional chefs on TV. Furthermore, some of the pupils said that they would feel more relaxed and less constrained to cook along a video, rather than physically cooking next to their parents: This could avoid "being bothered" by their parents' corrections and direct feedback, and it would permit them to feel more free and creative in the preparation of the dish. It was also remarked that by authoring ones' own recipe for others, learning-by-doing is indirectly supported because teaching a recipe implies practicing it. This is obviously different from writing a recipe for someone.

Finally, despite the fact that the value of video support was appreciated in general, the participants also recognized some critical aspects: If the audio output is not performing in a compliant way, this could lead to problems, because the learner might miss some important steps when following the spoken instructions (which can't be just "re-read" like in a paper-based cookbook), thus confirming the need for reversibility elicited in the previous trials.

5.1.5 Lessons Learned

The different activities and results of this project offer a chance to reflect and learn from them on different levels.

On the Interface

From observing how people interacted with the GUI, some interesting aspects emerged: It was note worthy, for instance, that the way in which the digital dial was to be manipulated was not immediately understood. Although the participants easily realized it was meant for a combined selection of a cook and a course, and they seemed to have a mental model of what they should do to trigger the desired cookbook (i.e., select both a cook and a course), they didn't understand how (i.e., dragging the pointer on the picture instead of tapping the picture and expecting the automatic movement of the graphical lever). Likewise, in order to flip pages back and forwards, a single tap was performed on the corner of the graphical page, and none of the participants expected it would be necessary to drag it from one side to the other one of the book (which was actually consistent with the way in which the interface was implemented).

These observations are somewhat consistent with the results of the trials of the Mug Metaphor Interface reported in Section 4.2.2 and suggest that, when metaphorically designing affordances for continuous physical gestures (which are proper of physical interaction in the analogue world), we need to consider that users' expectations on the behavior of digital media can be different. In other words, although the graphic design of the user interface draws upon the manipulation vocabulary and physical affordances of a physical artifact (e.g., the lever of a dial and the page of a book, in this case) in order to metaphorically suggest the physical action to be taken in the hybrid environment, people can develop different interaction behaviors at the pragmatic level of the interface (i.e., discrete vs. continuous actions). Whether this might be due to the acquaintance of interacting with digital media through discrete actions - which largely occur in the WIMP paradigm (and as hypothesized by some of the participants, cf. Section 5.1.4) - or to individual interaction preferences and attitudes, this aspect definitely raises some open issues for the transfer of elements of physical interaction in the design of hybrid ones. Furthermore, considerations of motor fatigue might sometimes drive design choices of automatic or augmented behaviors of the digital element of hybrid interactions, thus not literally transferring the manipulation vocabulary of the physical world. All in all, these considerations motivated a more careful analysis of interaction patterns through controlled studies (cf. Chapter 7).

On Methodology

Differing evaluation efforts have shown that the assessment of different aspects of a domestic appliance, which is designed for being used in family life, requires carefully selecting different methods and carefully considering the results of the tests conducted in simulating environments.

The first experience trials in the lab, for example, had suggested the use of speech command for reducing pen input, which turned out to be acoustically inappropriate in a noisy environment such as the kitchen (Terrenghi, 2006b). The issue of a good acoustic feedback (and in turn, of good acoustics for input capture) became much clearer when people really had to cook in the kitchen. The noise caused, for example, by a steak frying in a pan, or by other people chatting in the room, together with the poor audio performance of the hardware, emerged to be critical aspects of the appliance, which could be revealed only in a real cooking setting. Likewise, the intended fun and entertainment motivating factors could be assessed - and were actually confirmed - only when testing the appliance in a real familiar social context, or involving specific target users (e.g., the pupils of the focus group).

Overall, these observations indicate that the physical and social contexts in which an appliance/hybrid artifact is used have an impact on the way it is perceived, i.e., on the perception of its affordances. From a research perspective, this also implies that for an understanding of how the affordances we design are actually perceived, we need to deploy hybrid artifacts that can be experienced in the context they are envisioned for; and/or, that can be adopted as probes to be experienced and discussed by the target users we are designing for. In such a way, these prototypes provide us with an opportunity to learn more about human values and practices in those specific domains, beyond the usability and functionality of the interface per se.

On Domestic Practices

As anticipated above, the deployment of an experience prototype which could be tested in situ by different user groups permitted to learn more about home practices and members' roles within the domestic environment.

In particular, the different feedback given by different target users suggest that the affordances of hybrid artifacts for the home are different with respect to the roles of its users within the family. The younger users reported a fun factor derived from watching their parents, which as a consequence would stimulate them to use the appliance, and ultimately to cook more. Similar to that, one can speculate that elderly people, or people living alone, might enjoy watching their relatives in their domestic environments, thus providing a sense of presence and of memory as a family photo album. In terms of design, this implies that the diversity of the inhabitants of a home, and of their roles as members of a family, requires a design which offers a diversity of experiences.

Furthermore, it became clear that the space and material culture of a household affect the perception of what we consider a "familiar place". Introducing novel artifacts into such an environment needs to deal with an ecology of existing artifacts and patterns of use. The appearance, cumbersomeness, and usability of computing technology remain main issues for its acceptance in familiar places. Here, people's concern that technology could be disruptive of an aesthetic ecosystem and/or of the social intimacy of the home has fostered, to some extent, the claim for "invisible technology" of some of the Ubicomp agenda (cf. Chapter 1, Section 1.2.1). On the other hand, it is important to recognize those cases in which computing technologies can also stimulate engagement and deliberate time-spending. In these cases, other design choices can be thought of, which stage technology in the foreground as opportunity for entertainment and social binding, thus targeting users' motivation relying on values such as family relationships. In this sense, the outcomes of the project are encouraging to investigate further the potential of ubiquitous computing technologies to strengthen the social bindings and the passing on of knowledge in the home by providing novel forms of entertainment as well as capture, archival and retrieval of domestic experiences. Additionally, they provoke a reflection on the meaning of engagement in the home and on how technologies can be perceived as supportive or disruptive of domestic social practices.

Building on these insights, the meaning of portraying snippets of domestic life, as well as the potential of computing technologies for supporting engagement, are further investigated through the design and evaluation of the Time-Mill Mirror.

5.2 The Time-Mill Mirror

Similar to the Living Cookbook, the Time-Mill Mirror project draws upon the consideration of a physical artifact, a mirror in this case, and of its role and affordances within a specific domain and social context, such as the home. Its design explores novel ways in which domestic memories can be captured, archived, and displayed by combining some of the qualities of a decorative domestic mirror together with some of the possibilities provided by digital technologies (e.g., animated visualization, automatic photo shooting, and random retrieval) so as to investigate how these have an impact on the user experience of displaying and evoking family memories.

The Time-Mill is a digitally augmented multimodal mirror. Like a traditional mirror, it reflects in real time the events that occur in front of it. But unlike a traditional mirror, it also captures and retrieves snippets of those events and displays them as a dynamic collage, accompanied by a melody.

Its experience prototype served as probe to investigate the implications of braking the conventions about location and purpose of computing technologies: from office work to home family life; from an utilitarian scope to decorative and playful ones for social engagement. Such an investigation raises interesting issues about people's preconceived ideas and expectations about the functionalities and interaction possibilities of computing technologies. Furthermore, it sheds some light on how alternative paradigms at the pragmatic level of the interface can effect the experience of use with leisure home technology.

Its design rationale draws upon a field study on the way in which people capture and display family memories, as well as on the use of mirrors as home displays, which is summarized below.

5.2.1 Background and Motivation

Previous work on the different meanings of physical display-artifacts in the home (Schmidt and Terrenghi, 2007) distinguished different classes of home displays: i.e., communication, awareness, reminding, decorative, and memory displays. Considering the last category, the study showed how memories of the past are displayed in the home: These are usually embodied by framed photographs as well as physical artifacts, such as souvenirs for example, that are both evocative and decorative, as they celebrate people, places, and/or events. Their display in the home is often on the mantelpiece of the domestic walls, carefully arranged in respect of the family moral order, as disengaging witnesses of the idiosyncratic social and physical context they belong to.

This category of home displays was further investigated in a second study: This focused on the capture of family memories and on the evocative meaning that some display artifacts have in the home, as embodiment of those memories, family history, and evolving relationships. The inquiry was conducted in Cambridge, UK, in 2006, within the context of an Internship at the Microsoft Research Lab, in the department of Socio-Digital Systems. The group has been researching the design and social implications of interactive, situated displays in domestic social contexts.

This investigation was motivated by the research question: How can situated home displays support capture and retrieval of family memories? How (if possible at all) can such displays solicit reflection and evoke memories in a decorative, playful, and engaging way? The underlying idea was to investigate the domestic mirror as display artifact, and to explore the possibility of augmenting it with capturing capabilities.

In order to address those questions, 8 adults from different nationalities, professional backgrounds, and households were interviewed. Their households consisted of couples living with young children, couples with grown children, couples with no children, and single parents. The interviews focused mainly on the following topics:

- How persons and events are captured in the domestic social context: the mechanics of this, responsibility for it, storage and display of images and memories, by whom and for whom, etc.;
- How memory is displayed: how people or events are "represented" on the domestic walls (e.g., photos, prizes, cards, etc.);
- How, where, and why do people use mirrors as home displays.

Before the interview, participants were asked to take photos of their homes and in particular of the mirrors they have hung (e.g., see Fig. 5.5). During the interview, those photos were used as probes to contextualize the inquiry about mirror artifacts, their functions, why they were in that specific location, where they came from.

Concerning the observations on the use of mirrors, the inquiry suggested that decorative mirrors (i.e., not bathroom or wardrobe mirrors but rather

Figure 5.5: Domestic mirrors with different functions (e.g., augmentation of spacial perception, decoration, self-checking) conveyed by their location, frame, and format.

the ones that are placed in hallways, on the fireplace, or in the bedrooms) have mostly the following features or functions:

- They are used to augment the spatial perception of the room when placed in living or dining rooms (e.g., Fig. 5.5, a);
- They often "attract" other displays or artifacts (e.g., Fig. 5.5, b);
- They are used for "self-checking" when placed in the hallway or in the bedroom (e.g., Fig. 5.5, c);
- Their frame and form factors usually determine their location (e.g., Fig. 5.5, d and e);
- They reflect dynamism without engaging attention.

Concerning the results of the investigation on the ways in which people deal with the capture and retrieval of family memories - and the media they use - those were rather diverse among participants, but still showed some general attitudes:

Households increasingly capture family events and social contexts with digital cameras. Although several people own a device which can capture digital video, video recording happens very seldom in home life, and mostly in families with young children. Memory displays are rather embodied by still images. Only a small portion of the images that are taken are framed and hung on the walls, or archived in albums: In this case, it is usually baby photo albums to document children' evolution. The photos that are printed and framed are hardly ever changed or moved to different places in the house.

Most of the pictures that are taken are actually just stored on the PC of the camera owner. The browsing of the digital photo archives is then sporadic, and is usually done on the PC by one or two people, hardly ever in bigger groups. The random appearance of photos from personal collections is often set as screen saver for PC monitors or home multimedia systems, as kind of disengaging entertainment.

Given these facts, the question then arises as to why users make the effort of taking images in the first place. The inquiry indicated that this was motivated by several factors:

First, one could recognize a desire for safety and maintenance, which is shown by the capture and archiving of records of social family life to support memory, self-awareness, and honouring of family members. To this end, people appear to use artifacts and walls to confirm and externalize the underlying social structure, the roles and rules of the family moral order. Such a structure is reinforced by preserving a consistent spatial arrangement as shared reference system (e.g., memory displays such as framed pictures have a specific location and are seldom moved around in the house).

Furthermore, one could recognize a sense of responsibility towards the next generation to ensure that a kind of heritage and honouring of family members and their evolution is documented (e.g., baby photo albums). In this respect, a diversity of media such as digital photos, physical photo albums, as well as videos are used, thus indicating people's desire of capturing many different aspects of children's dynamic evolution for sharing them with others in different ways (e.g., via e-mail, by framing paper photos and hanging them on the walls, by publishing online photo albums, etc.).

Overall, one could motivate people's desire of "capturing the present" with the concern for the memory and honouring of the past (e.g., grandparents' pictures) in the future (i.e., for children when they will grow up). Physical artifacts (e.g., printed out photos, baby photo albums, and paper cards) become a back-up of such a record. Similar to that, people seem to put trust in the future of digital technology (especially for its capturing and storing affordances) despite a lack of immediacy (because of the limited affordances for immediate and social browsing provided by traditional PC displays). In other words, being aware that those memories are archived digitally somewhere (on a PC, online, on a CD-rom) seems to be enough to pay off for the lack of immediacy.

Reflecting on the results of this field inquiry, one can then conclude that:

- The shift from analogue to digital photography and a rather "conservative" approach to the decoration of the domestic walls (i.e., digital photos are seldom printed out, and printed photos are seldom replaced or removed once they are framed or hung) contribute to a static ecology of memory displays, despite more and more digital photos being taken, especially during family events;

- Browsing digital photos on a PC is considered as an engaging activity, in the sense that it requires a commitment of time and attention, and it is often a lonely activity. Conversely, framed physical photos and digital photos randomly appearing on a multimedia screen in the home seem to serve a decorative, ambient purpose. In other words, they appear as disengaging, in the sense that they stay in the periphery of attention, but at the same time create opportunities for casual, spontaneous social engagement;

- The mirrors in front regions of the home (e.g., on the fireplace) can often be considered as decorative displays, which are used for their capabilities of augmenting the perception of space, and reflect dynamism in real time.

Drawing upon these considerations, the design of the Time-Mill Mirror explores ways of subverting such a way of "catching" and "mirroring" the passing by of time in order to provoke and explore different ways of seeing and reflecting upon time. To this end, some of the qualities of digital technologies are exploited to:

- make the experience of capturing and rendering snippets of time a playful and engaging one;

- turn memory displays from static to dynamic and multimodal;

- provoke surprise and enhancement by delegating some of the capturing and rendering control from people to a situated device;

- make those captured snippets of time accessible in a serendipitous, esthetically arranged way, so as to stimulate new associations and perception of time.

To better understand how these goals were addressed, the design rationale and reflective analysis of the experience prototype are articulated in the remaining of this section.

5.2.2 Design

The design of the Time-Mill Mirror draws upon a consideration of the qualities of digital cameras as capturing devices, and of the qualities of mirrors as decorative displays: Thus, it explores how to merge them in a capturing/displaying situated artifact. Unlike digital cameras, which are optimized to show images after the event, mirrors are designed to reflect the "here and now". In doing so, they augment physical space through the literal reflection of the light in that space. In another sense, they reflect the activities that occur in front of these surfaces. Finally, unlike cameras, mirrors add value to where they are located in the home in different ways: Those in the dining room have a different value than those in the bathroom, for example. Embedding the capturing capabilities of a digital camera into a mirroring situated display opens interesting possibilities for super-imposing images of the past to the ones of the present. These pictures have a location in common, but can depict other people and situations, thus potentially triggering novel emotions, reflections, and associations.

In order to give people a playful means of partial control for the capturing and browsing of images, as well as to provoke intentionality in the interaction, the device was designed in such a way that a very simple physical transducer such as a wheel would work as handle for interaction. Such a design choice builds on the considerations on affordances for manipulation articulated in Section 4.3: By exploiting the physical as well as cognitive affordances of a physical wheel, the goal was to convey a feeling of control, exploration, and engagement. This choice provided an opportunity for further investigating the relationship and metaphorical mapping between the manipulation vocabulary of a physical handle at the pragmatic level, and people's construction of a system mental model at the conceptual level of the interface. In this sense, the wheel was conceived of as a mechanical engine (a kind of mill) capable to generate a flow of images, and thus capable, potentially, to generate new associations and emotions about the passing by of time. Furthermore, its manipulation vocabulary was metaphorically associated with the stereotypical conceptual model of time as a linear dimension (i.e., a "timeline") that can be unfolded back and forward like an analogue film.

Thus, the Time-Mill artifact consists of a physical wheel coupled with a mirror. A tablet PC is mounted behind a semi-transparent mirrored piece of glass, and a wide-angle digital camera is embedded in the mirror frame (see Fig. 5.6, a). As a user rotates the wheel, a melody is played, and an animation is displayed on the tablet screen. Leaves appear, flowing from left to right, visible through the mirrored glass. This animation is intended to metaphorically evoke the flowing of time, and to suggest the human capabil-

Figure 5.6: a) The components of the Time-Mill Mirror, i.e., a tablet PC, a control wheel, a digital camera, and a see-through mirror; b) Interaction with the Time-Mill Mirror.

ity of catching and remembering only impressions or snippets of the events occurring in that place. Like in a music box, the mechanical rotation of the wheel triggers the playback of the melody. When the wheel is rotated in a clockwise direction, Time-Mill takes a picture at random moments, capturing an image of the person engaged with the device (see Fig. 5.6, b). Within the leaves, these images are sporadically shown. When users rotate the wheel in the opposite direction, a similar animation and melody begin, but the pictures that are framed in the leaves are now the ones captured and stored in previous interactions, when the same (or maybe other) people played with the artifact in the past. These pictures are randomly selected from bundles of pictures created previously, and are retrieved in reverse chronological order.

Such a partially randomized photo retrieval responds to the design goal of suggesting a sort of "shuffling" experience, which is potentially somewhat different from the one that a sequence of frames/photos chronologically ordered and "scrollable" in a timeline normally provides (e.g., in several desktop applications for photo browsing). The idea, here, is rather to support the need for randomness and casualness, which can be recognized in situations such as browsing at a flea market, for example, or through old paper photos and cards in a shoe box, when we are in a state to - more or less consciously - look for random stimuli. In these situations, different cues within this randomness can trigger different thoughts, connections, and solicit communication, thus creating elastic spaces between engagement and disengagement.

The realization of the prototype was functional to explore these aspects in more detail, as discussed in the next paragraph.

5.2.3 The Time-Mill as a Probe

Though carefully designed and constructed, it is worth noting that Time-Mill is not intended as a product, but rather is an experience prototype or cultural probe (Gaver et al., 1999) designed to solicit insight onto the underlying question: How does the coupling between certain aspects of physical interaction (e.g., rotating a wheel) and other aspects of digital output (e.g., random animated visualization of images) contribute to the creation of a mental model of time flow? And beyond that, how does such a coupling affect expectations, associations, and emotions in a context which is not meant for an "utilitarian scope" but rather for the sake of dis/engaging discovery and indulgence of browsing through time?

Although it is ambitious to answer those questions, the explorative study presented below tries to underpin some of the aspects that can have an implication on the user experience, such as:

- the type of transducer;
- the dis/order of rendering;
- the aesthetics of rendering.

The explorative study was designed around those dimensions, as described in the next paragraph.

Study Design: Providing Alternatives

As suggested by Tohidi et al. (2006), presenting different design alternatives to users can engage them in a more constructive role, thus stimulating their reflection and creative elaboration of alternatives. Given the open questions that were being addressed, the study aimed at an open and articulated feedback upon the different dimensions it was focusing on (see previous paragraph). To this end, two more versions of the Time-Mill concept were implemented in order to try and unpack the factors influencing the user experience of image browsing:

- the Photo-Wheel: In this case, users can browse through the photos by rotating the wheel. A melody is played back, photos appear in a chronological order and framed in a rectangle (see Fig. 5.7, b);
- the Photo-Slider: In this case, users can browse through the photos by interacting with a pen on a soft slider on the screen of the mirror. No melody is played back and pictures appear in a chronological order, framed in a rectangle (see Fig. 5.7, c).

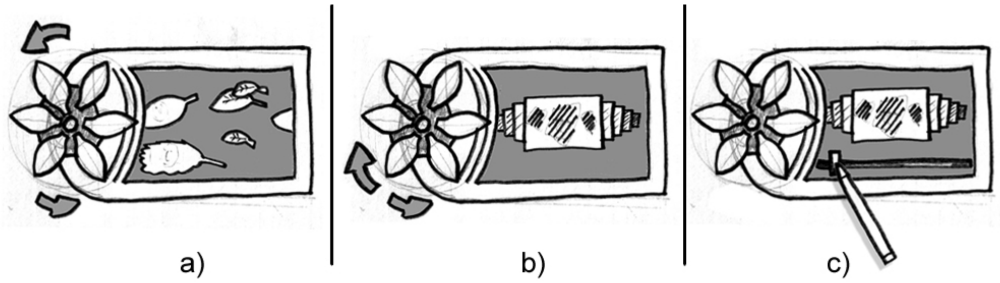


Figure 5.7: Alternative browsing techniques: a) The Time-Mill: wheel interaction, random order, leaf-shaped photos and music accompaniment; b) The Photo-Wheel: wheel interaction, chronological order, rectangular frames and music accompaniment; c) The Photo-Slider: pen interaction, chronological order, rectangular frames, no music accompaniment.

Given the different variables, the study did not aim at prescriptive answers. Rather, it aimed at suggesting how users' experiences could be affected by:

- wheel vs. pen interaction: Taking into account the considerations on haptic feedback presented in Section 4.2, the wheel was expected to convey a higher feeling of control and solicit exploration and playfulness of interaction.
- random vs. sequential rendering: Considering the way in which some devices support casualness of interaction by randomization of media rendering (e.g., the shuffling mode of the IPod), the random rendering was expected to convey a higher sense of surprise and casualness.
- animated flowing leaves vs. rectangular linear format of photo frames: The flowing leaves were expected to aesthetically better match the shape and movement of the wheel, and to convey a higher sense of flow.

The following paragraph describes how the study was conducted.

Set-up, Participants and Procedure

Although the design for the Time-Mill Mirror originated by the consideration of memory displays in the home, the questions it raised in terms of the relationship between pragmatic level of interaction, rendering of information, and outcoming evocative experience can be partially generalized to other contexts in which a defined group of people recursively shares a certain place. Responding to the goal of gaining a feedback from multiple people in a relatively limited amount of time, and considering the difficulty of testing the Time-Mill in a household with more than 3 or 4 family members that would be in a suitable age to use the device, a different context than the home was chosen as testing domain: namely, a classroom of the Academy of Arts in Munich, Germany.

The room is used as art studio and provides some desktop PCs as well. A class of 38 students in the age of 21 to 30 years old recursively visits the room, either as a whole group for attending a lecture, or in smaller teams for working on projects. Time-Mill was hung in the room for 23 days; before that, the class was briefly introduced to the functioning of the device and scope of its installation. In 23 days, 465 photos were stored by the people interacting with Time-Mill. After this phase, 12 students were involved in in-depth, semi-structured interviews (see questionnaire in Appendix A). In those sessions, the participants were invited to interact with the Photo-Wheel and the Photo-Slider browsing techniques as well. The interviews addressed the following topics:

- subjective experience of using wheel vs. pen;
- subjective experience of random vs. sequential rendering;
- connections between place, people, and time;
- expectations whether, where, and why the participants would hang the Time-Mill in their homes.

Below, the observations derived from the study are organized according to the different browsing techniques in order to recognize how those elements have an impact on each-other.

The Experience with Time-Mill (Original Concept)

When the participants were asked to name three terms they would associate with the experience of interacting with Time-Mill, some of the most recurring terms were “playfulness”, “surprise”, “simplicity”, “poetry”, and “freedom”. The interesting bits here are the values associated with those same terms, which were quite diverse among people. The simplicity of the artifact was perceived as a positive quality by some of the participants: i.e., as a different way of interacting with a computing device without an utilitarian purpose, and rather as an aesthetic and provoking one. For some other participants this simplicity was felt as limiting in terms of functionalities: For these users this fact implied a sense of passivity, as no other interactions than simply turning a wheel were possible.

Figure 5.8: Interaction with the Time-Mill browsing technique.

The wheel as physical handle was usually perceived as playful, but also in this case it elicited different values associated with playfulness: For some this was intriguing and fun, for some others it became childish and boring after a while. For most people it was obvious how to use it and such immediacy was a nice feature, which enabled them to interact with the computer in an intuitive way: For some others it was idle and the manipulation vocabulary was too limited for the functionalities they expected from a computer device. Curiously, several people also mentioned that the wheel reminded them of the wheel of fortunes, and to the same association they had different emotional attitudes: I.e., for some this was an interesting analogy to the concept of "chance" and casuality, for some others it was trivial. Additionally, some features of the wheel prototype affected the participants' reactions as well. The wheel itself was made of some leaf-shaped ashtrays commercially available, that were covered with some cardboard and radially arrayed so as to resemble a flower. The covering cardboard conveyed to some participants a feeling of temporality, which in turn motivated a sense of insecurity in manipulating the wheel, people mentioning a fear of ruining the prototype. This simple fact affected, in other words, the feeling of control for those people. This aspect is further discussed in Chapter 8, Section 8.3.

The randomness and aesthetics of the rendering was perceived differently among participants too. For some this was a form of disengagement, a way of relaxing through the browsing flow: For some others this was (negatively) distracting from the content of the picture, also due to the leaf shapes and to the animation. In these cases, there was a sense of lack of control and interaction possibilities, such as search and selection: Some people also mentioned that a technical tool such as a tablet PC "has to have a clear function".

Considering the feeling of being connected to those particular portrayed place and people, the participants' feedback was diverse too: Some perceived a stronger connection to the people portrayed in the pictures of the past because those people were intentionally framing themselves in the mirror, interacting with it, thus leaving their personal traces. In this respect, someone mentioned that people appeared "more exposed" than in traditional pictures. Someone else, considering the hidden and surprising capturing capabilities of the Time-Mill (the camera is hidden in the frame and the tablet behind a mirroring glass), felt as if this provided a kind of voyeuristic view by showing people caught in the act of discovering the device functionalities.

When asked where they could imagine to place such a display in their homes, the most diverse answers and motivations were given: in the kitchen, in the living-room, in the entry hall, as well as in the bathroom. The reasons were aesthetic (e.g. "it fits with the color of my kitchen"), social (e.g. "it would be nice in the entry hall as a kind of guest book"), as well as weird (e.g. "the Time-Mill Mirror could be hung in the bathroom to take funny clandestine pictures of the people who never used it and just play with it, thus shooting a photo of themselves in those situations").

The Experience with Photo-Wheel

When asked about their feedback on the interaction with the the Photo-Wheel (in which pictures are ordered sequentially, see Fig. 5.9), people mentioned a sense of "narrative", "sequentiality", as well as "consistency" between the move of the wheel and the visualization of photos. This fact seemed to imply a more "precise" interaction with the photos as they recognized events more easily, thus reinforcing their memory of the pictured context. Someone also mentioned this aspect can create moods: This feature for some of the participants also generated a stronger connection to the place, as the pictures all show the same background, thus sharing the same context.

It was interesting to notice how the participants appropriated the interface and were able to map the rotating speed of the wheel to the speed of photo display. Several participants enjoyed spinning the wheel quickly, exploiting its momentum, to browse fast through the photos without touching the wheel all the time. Some mentioned this triggers the effect of movies, like several frames making a storyboard or a cartoon, which is obviously not possible in the case of leaf-shaped photos.

Figure 5.9: Interaction with the Photo-Wheel browsing technique.

On the other hand, some people considered the rectangular frame as obvious, boring, and aesthetically inappropriate in relation to the shape of the wheel. Some participants suggested the appearance of leaf-shaped photographs with the same flowing motion as in the Time-Mill, but ordered in chronological order. Someone else suggested the leaf-shaped rendering for the capturing phase, and the rectangular sequential one for the browsing phase.

The Experience with Photo-Slider

The Photo-Slider, which was to be operated with a pen (see Fig. 5.10), raised several associations with terms such as "work", "control", "formality", and "responsibility". Also in this case, the values associated with those terms are diverse: A higher capability of control was felt as a too strong commitment by some people, who did not feel like "taking a decision about what is interesting" and felt disturbed by the fact that the device would make them think about work. In other words, they missed a disengaging, relaxing experience of flow. Someone also said "I would use it only if I have a goal". The very same quality of control, on the other hand, was felt as positive by other participants, who appreciated the possibility of making more fine-grained movements/decisions. In these cases, accuracy of interaction was mentioned as an added value.

An other interesting aspect is the potential for interaction associated with the pen: Several participants mentioned that while with the wheel it is ob-

Figure 5.10: Experience with the Photo-Slider browsing technique.

vious what you can do, with a pen you would expect to be able to do more, e.g., selecting, searching, as well as scribbling. Also in this case there are differences about the associations with the pen as a tool: Someone conceived of the pen as a personal tool they are able to appropriate for self-expression in the analogue world (e.g., handwriting and scribbling), others considered it as a transducer, i.e. a prosthesis for interaction with digital media (e.g., selecting, dragging), thus affecting and diversifying users' expectations in terms of interaction capabilities.

5.2.4 Lessons Learned

From this inquiry, it was possible to gain an impression of how different features of the interface at the pragmatic level seem to affect the experience of photo browsing with a similar device:

- The wheel was perceived as a more playful transducer than the pen, the latter being associated with goal-driven interactions; when interacting with the pen, on the other hand, the participants mentioned a higher sense of control and accuracy of interaction.
- The sequential rendering was considered more controllable than the random one, also when it was operated with the same transducer (i.e., the wheel), since the proximity of photos with a similar background afforded the possibility of recognizing events more easily.

- Similar to that, the framed linear rendering of the photos was perceived as more precise, and allowed for concentrating more on single photos, while the leaf-shaped rendering conveyed a sense of flow.

The relationship among these aspects is more complex than the single aspects considered as such, and the observations that were reported have shown that different combinations can convey different expectations, associations, as well as feelings. And, what probably is most interesting here, is that the values people attribute to those associations or "moods" is very diverse, and it depends on personalities, attitudes, and personal aesthetic values.

These aspects imply that when breaking the convention about the purpose (e.g., a work station or a decorative display?) and the location (e.g., in the office or in the home?) of computing technologies, we need to think thoroughly about how the affordances for the hybrid interactions we design will be perceived. In this sense, we need to consider how computing technologies have conventionally been perceived so far, in the PC instantiation, and how some people might have built a precise expectation about their purpose and location. Additionally, we need to acknowledge a diversity of values and attitudes which can't be measured or assessed in terms of performance. Like in the case of the physical display artifacts or leisure technologies (e.g., music players), their "benefits" are to be considered in terms of enjoyment, communication, coordination, as well as personal likes and dislikes. In this respect, the meaning of "feeling of control", "engagement", and "disengagement", for example, need to be understood in light of what this means in leisure, familiar domains, as well as of the users' diverse expectations of the role and functionality of computing technologies in everyday life.

5.3 Summary and Discussion

This chapter has presented two case studies, the Living Cookbook and the Time-Mill Mirror, that explore how physical mundane artifacts can be augmented by digital technologies in such a way as to partially alter their original functionality, so as to elicit and provoke new genres of communication, learning, associations, and reflections upon time.

The two case studies have shown two approaches of inquiry, namely:

- exploring the experience prototype in different contexts, as in the case of the Living Cookbook: This approach has shown how different feedback and requirements can be elicited in different environments (i.e., in a research lab, in the kitchen of a family, in a focus group), thus confirming the claim of this thesis for the design of hybrid interactions that builds on the understanding of the socio-physical context of use.

- the exploration of different alternatives of the same concept, as in the case of the Time-Mill Mirror: This approach has helped to distinguish how some dimensions of the pragmatic level can have implications for the user experience and, at the same time, it has elicited different values associated to those implications.

Both these approaches and case studies support the claim for a design acknowledging diversity, of contexts as well as of target users. The design of open experiences can leverage a deeper understanding of the factors which affect people's perception of interaction in different contexts, thus eliciting their different values and supporting conscious design choices.

The observation of people engaging with the interfaces, as well as people's subjective perception of the experience of use, suggest that as designers of hybrid interactions we need to learn about the vestiges of interaction with mouse and keyboard on the desktop PC, and take into account the purpose that computing technologies have mostly fulfilled in everyday life. In this way, we can better understand how these aspects potentially affect people's mental models for hybrid interaction (both at a pragmatic and at a conceptual level) and, in turn, people's attitudes and values. In this sense, the main focus of the evaluations that were presented in this chapter is not on the assessment of success of design solutions per se, but rather on the discovery and understanding of the factors that affect users' subjective experience.

Another consideration to make is that the hardware and the software can have an equally relevant weight in the acceptance and use of hybrid artifacts. As computing technologies can be embedded into physical mundane artifacts (e.g., into a kitchen cupboard or a mirror) to make them multimodal and context-sensitive, for example, it is important that such augmented artifacts are compatible with the existing ecology of artifacts (e.g., considering their size, whether they are stand alone or modular systems, whether they should be placed on vertical or horizontal surfaces) and with the particular aesthetic values of a household.

Finally, the different responses provided by different family members suggest that control allocation and negotiation are fundamental for the social use of home displays. A design addressing those issues needs to recognize the idiosyncratic moral order of the home: This relies on an ecology of artifacts and places which makes sense in that specific context, its purpose often going beyond functionality and efficiency, and rather embodying and supporting unspoken rules, patterns, and relationships, which together make a home. Hence, as interaction with computing moves from single-user and office work contexts to more casual and social settings, benefits such as fun, engagement, and improvement of group communication need to be further understood and taken into account. These might sometimes imply design choices which hinder efficiency and functionality for privileging other values.

The next chapter draws upon these considerations and presents an exploration of how new media and related technological tools can affect social practices, communication, creativity, and potentially support team-work. In doing so, values which are not merely functional, but rather aim at augmenting creativity and collaboration are further addressed.

6

Design of Hybrid Environments for Collaborative Creativity

In this chapter, the design focus moves from tablet-sized displays to large interactive surfaces such as tabletops and walls. The two previous case studies explored ways of augmenting situated physical artifacts - of a relatively small size format - for supporting asynchronous social communication. Here, the physical affordances of large surfaces are considered in light of their potential for being digitally augmented to support co-located, synchronous collaboration and collaborative creativity. The ways in which such large surfaces are normally integrated in social contexts and collaborative processes are considered. These issues are explored by means of two projects, the EnLighTable and the Brainstorm experience prototypes.

6.1 Augmenting Physical Spaces for Supporting Collaborative Creativity

Large physical surfaces such as walls and tabletops provide inherently social affordances, in the sense that they make the displayed information visually perceivable by multiple individuals. Furthermore, physical large displays, such as whiteboards, also afford the simultaneous visualization of a large amount of information, and thus the possibility of simultaneously seeing and comparing multiple alternatives. These two combined aspects (i.e., multiple observers of multiple displayed alternatives) normally offer, in mundane creative processes, a rich environment for the collaborative creation, manipulation, and assessment of different alternatives, from different perspectives.

The possibility of making such surfaces interactive and capable of displaying digital content opens up interesting and challenging opportunities for the design of social interactions across digital and physical media in space. Put differently, the blending of interactivity into shared artifacts and spaces, together with the distribution of control capabilities amongst co-located users, provides the potential for novel forms of shareability.

Based on the individuals' rights of parallel and co-located interaction with the information content, one can start to reflect upon and distinguish different levels of the concept of shareability, such as, in increasing order of engagement:

- perceiving;
- pointing;
- accessing;
- annotating;
- editing;
- creating.

A higher level of shareability usually relies on the lower ones (i.e., users need to be able to access and edit shared content in order to create new one). Different types of technologies (e.g., single- vs. multi-touch interaction), combined with different interface design solutions (e.g., shared areas of the real estate vs. personal territories of interaction) can affect the level of shareability of an interface. This chapter explores the augmentation of physical surfaces and spaces in order to support the shareability of multiple perspectives and ideas, and the synergistic creation of new ones. In other words, this work investigates different ways of supporting *collaborative creativity*, i.e., "a social and communicative transaction between people who in some ways share a mutual goal" (Sundholm et al., 2004).

A first step in this direction is to acknowledge that the layout of the environment affects the perception of ideas. As discussed by Fischer (1999), "Creativity occurs in the relationship between an individual and a society, and between an individual and his or her technical environment". Proper socio-technical settings can thus amplify the outcomes of a group of people by both augmenting individual creativities and "multiplying rather than simply summing up individual creativities" (Fischer, 2005). In this respect, the design challenge lies in creating socio-technical environments which are suitable for supporting and stimulating such synergistic processes.

From this perspective, the social, physical, and technological contexts acquire a main role in determining communication as well as cognitive processes. One can then approach the design space by reflecting and building on some of the existing CSCW theories on collaboration, as well as on cognitive theories on the use of the physical space for creative tasks.

Mutual visibility of action is a main design principle for groupware, as it provides awareness of what other colleagues are doing and how the actions of group members affect the shared artifacts, thus enhancing group awareness (Dourish and Bellotti, 1992). *Group awareness* (i.e., the condition where members perceive the presence of other group members and the possibility to communicate with them) provides chances for informal communication, which in turn can strengthen the ties of a group, as well as facilitate the transfer of essential information related to task-specific activities (Dourish and Bly, 1992). In this respect, the size of table and wall displays supports group awareness as it allows participants to see each other's actions, movements, expressions, and gestures in real time. Large surfaces also enhance a sense of common ground by affording the display of resources, as well as of the results of participants' interactions on a single shared area. Furthermore, looking at the cognitive benefits of the physical space, large surfaces support epistemic actions (cf. Chapter 2, Section 2.3) by enabling temporary spatial arrangements: These, in turn, enhance creative processes by allowing for the externalization and visualization of different alternatives (Kirsh, 1995).

Drawing upon these considerations, one can then start to explore some possible ways in which the embodiment of computer-mediated interaction capabilities in a physical and social setting can either:

- turn single tasks into collaborative ones: for example, take tasks that were previously performed alone on a personal workstation and have them performed collaboratively on a large interactive surface;
- use hybrid rather than physical technologies to support existing collaborative tasks. In this sense, some of the qualities of digital technologies (e.g., storage of data, process tracking, easy replication and distribution of data, for example) can provide novel features to traditional processes, thus potentially enhancing them and/or their management.

In this thesis, both approaches are explored, through the EnLighTable and the Brainstorm projects respectively. Both designs are intended to support some forms of collaborative creativity, as discussed in the next sections.

6.2 The EnLighTable

This section presents the EnLighTable project (Terrenghi et al., 2006a). This consists of a direct touch tabletop appliance designed for supporting collaborative picture selection and layout design, e.g., in advertising agencies, publishing companies, or catalogue production companies. Such an appliance enables multiple users to simultaneously manipulate digital photos of a shared collection and rapidly create and edit simple page layout. Its motivation, design rationale, and assessment of design choices are discussed below.

6.2.1 Motivation

The motivation underlying the design of the EnLighTable builds on the author's personal experience in design practice. Nowadays, the work of graphic designers is largely based on the use of digital photos and of computer applications for their manipulation. The massive transition from analogue to digital photography, indeed, has affected both the consumer market and professional graphic design. In the past, tools like light tables were usually adopted to visualize and select analogue film slides in photographic and design studios. Nowadays, the availability of high resolution images in a digital format has made designers' interactions with computer screens much more common than with that kind of light surfaces. Additionally, graphic designers are not necessarily working on paper-based communication projects only: Those projects can rather be webpages or multimedia presentations, thus relying even more on the use of digital images.

Selecting the right picture for a flyer, an online catalogue, or an advertising campaign usually involves several project stakeholders in the discussion (e.g., creative directors, art directors, graphic designers, as well as clients at times), in iterative phases of communication in the decision process. The patterns of communication in the project team may vary according to the size of the company and of the project, the hierarchical organization of the agency, or the source of the photo collection (cf. Section 6.2.3). In most cases, a strategic concept for the communication is initially defined together, in a co-located team meeting around a table, using paper, sketches, print-outs as well as whiteboards. Then, the different professionals "migrate" to their workstations (normally desktop PCs) to work with digital media on their individual sub-tasks. Most of the creative phase of image selection happens in parallel, in different locations, with limited communication among stakeholders. Afterwards, they gather again with print-outs of ideas, thus using the affordances of paper for elaborating on the concept (e.g., passing on, annotating, and sketching paper-based displays); or use digital material projected on a whiteboard; or move to each-other's PCs to visualize and discuss some concepts on the screen.

In such a domain, as one can see, the technologies that currently support the communication and work patterns are both physical and digital, in different phases of the workflow. Now a question arises: Given that the change of media (from analogue to digital photos) has affected the tools to interact with it (i.e., from light tables to computer monitors), as well as the way in which project stakeholders gather together in a creative process (e.g., from around a light table, to around the computer screen or print-out layouts) can new tools support the workflows and communication patterns around those media in such a way as to foster the synergistic co-located work of multiple designers/stakeholders? This project explores how the possibility of embedding interaction into large sharable surfaces/interfaces can be exploited for the design of tools that can better support the collaborative communication and creativity of team-work.

6.2.2 Design

In the EnLighTable project, the table is considered as a surface for stimulating the gathering and active participation of a creative team around digital media content. Working on the design of hybrid interactions for such a setting, and referring to the different types of affordances distinguished in Chapter 2, Section 2.7, this implies:

- considering existing physical, as well as social affordances provided by the table as a physical artifact;
- the design of cognitive affordances for the manipulation of digital information;
- the design of social affordances for collaboration around digital information.

Physical tables create a shared space for the manipulation of objects. Humans tend to divide this space into different functional areas. In particular, the center of a table is often used as a shared area for the exchange of objects, while perimetric areas closer to a specific user are considered increasingly private by this user (Scott et al., 2004). When dealing with rotation-dependent objects, such as text or pictures, users tend to reorient them for their own viewing direction, which can cause conflicts of interest with other users. On the other hand, when a user reorients an object towards another person, this action shows his/her intention to discuss the object with this person, thus fostering communication.

The design of the EnLighTable appliance copes with those issues by relying on the metaphor of a "set dining table" (see Fig. 6.1). The visual design of the user interface builds on some main elements: a shared pond of photos; single photos, which can be dragged and copied out of the shared collection in

a) b)

Figure 6.1: Affordances for social interaction on the table: a) Arrangement of physical artifacts in everyday life; b) Arrangement of digital information items on the EnLighTable appliance.

order to be manipulated; and an "Imagetool", a kind o virtual lens for basic editing of single photos, such as zooming, cropping, rotating, and flipping.

The perception of a personal area of interaction is initially suggested by the pre-defined placement of three Imagetools oriented towards the sides of the table as default (see Fig. 6.1, b). By analogy to plates, this is expected to suggest "guests" (i.e., designers) where to sit, thus defining their personal area of interaction (similarly to the cognitive and social affordances illustrated in Chapter 2, Fig. 2.3). In the center, a larger shared container of information is displayed. It contains the thumbnails of a shared picture collection, e.g. the pictures of a photo shooting or a shared library. The tray can be dragged with a simple gesture towards the personal area of interaction (see Fig. 6.1, b). When such a shared tray is dragged in the direction of one of the Imagetools, the pictures reorient towards the dragging user. In this way, she can have a better view of the pictures, scroll to see more of those, and drag the preferred ones out of it: i.e., she can "serve" herself with the desired information items. After the user ceases interacting with the shared tray, the latter automatically returns to the original orientation and location, thus becoming available for others. In this way, a consistent spatial arrangement of the pictures is mostly maintained, thus supporting spatial memory for every user. At the same time, the personal reorientation of the picture collection towards each user supports individuals' view and selection of the available photos.

With respect to the design of single images, these are metaphorically represented as analogue film slides. In the real world, negatives or slides are the original sources for analogue photography. In design and photographic working scenarios, slides are projected, printed, scanned, i.e., they are edited

a) b)

Figure 6.2: Affordances for manipulation of pictures: a) Physical slides on a light table: b) Digital slides for the representation of image content in the EnLighTable appliance.

and processed. Furthermore, they are often arranged on light tables, so as to visualize and compare multiple pictures simultaneously (see Fig. 6.2, a). By analogy to this, in the EnLighTable digital slides represent the original source of information and provide affordances for manipulation and editing. Their white graphical frame enhances their manipulation so as to handle them without occluding the image, as well as providing a place for the labeling of the source and date. This is similar to the function of the plastic frame of physical negatives. When users touch the original slide on the shared tray and drag it to empty areas of the table surface, a copy of the slide is created. This gets a yellow frame so as to distinguish it from the original source and to remind the user that she is manipulating a copy (e.g Fig. 6.2, b).

A metaphorical representation was also used in the design of the Image-tool, whose appearance resembles a photographic lens. Similar to the approach taken for the Mug Metaphor Interface (cf. Chapter 4, Section 4.2), this choice aimed at providing cognitive affordances that would suggest the types of physical gestures to perform for interaction, thus metaphorically drawing upon the manipulation vocabulary of a physical tool. To this end, its design seeks to define and represent the relationship between graphic appearance/function/direct touch gesture in a way that is somehow different from the one typical of other mouse-based graphic applications.

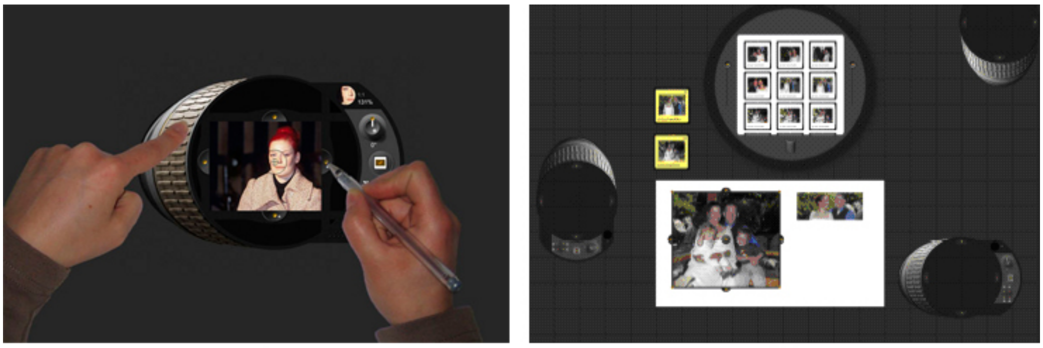
The graphical user interface of those applications (e.g., Adobe Photoshop) usually provides different icons, representing different tools, e.g. a lasso for drawing a selection, or a bucket to fill color regions. The appearance of these icons suggests their function, but not the type of physical interaction (i.e., the gesture) at the pragmatic level. When a user, for example, selects the lasso in the toolbar and moves the pointer to the picture area, the cursor turns into a lasso. In order to draw a selection, the user is supposed to click and drag, while keeping the mouse button pressed. Different from this interaction, when a user selects the bucket from the toolbar and positions the pointer on the picture, with a single click she will change the color of the pixels within the same color region. In terms of affordances, this means that no cognitive affordances are provided by the interface to suggest the kind of physical actions (e.g., single click vs. dragging) that the user is supposed to operate with the mouse.

A tighter coupling between visual representation and functionality is provided by Kai's Power Tools[1]. Based on the metaphor of a Toolglass or Magic Lens (Bier et al., 1993), these Plug-Ins for Adobe Photoshop and Corel Photopaint enable the direct application and visualization of filters on the image. As a virtual tool, the lens can be dragged over different areas of the picture with the mouse, and different filter parameters can be entered with the keyboard. Similar to Kai's Power Tools, the Imagetool relies on the conceptual model of a Magic Lens, which in this case is controlled by two hands directly on the surface of the table. The Imagetool is indeed envisioned to support the cooperative work of the non-dominant and dominant hands, the latter interacting with a pen.

The appearance of the Imagetool aims at providing cognitive affordances for direct manipulation relying on the way in which we manipulate certain physical objects. As illustrated in Figure 6.3, a, for example, the zooming gear on the left side of the tool can be scrolled with a continuous movement of one hand, as one would do with a physical photographic lens. Discrete interaction, such as tapping, is suggested by the 3D effect of the buttons for mirroring and saving changes, on the right side of the tool. A selected image can be cropped by dragging the semi-opaque ledgers within the lens, which resemble the blades of a four bladed easel used in dark-rooms. Furthermore, graphic cues are designed as yellow dots to suggest to the user where to position the pen-tip for interaction. Once edited, pictures can be saved, and remain on the shared surface of the table, available for others.

Other basic functionalities are provided by the tool, such as visualization of picture elements in 1:1 scale. Furthermore, designers can create frames for drafting a layout: By tapping and dragging the pen-tip on the table surface, so as to draw a diagonal line, they can define a rectangular layout area in which they can place and resize pictures by simply dragging edited slides into the frames (see Fig. 6.3, b).

[1] Kai's Power Tools is a plug-in published and distributed by Corel Corporation, http://www.corel.com



a) b)

Figure 6.3: a) Interaction with the Imagetool. b) Frames for drafting page layout on the EnLighTable.

6.2.3 Assessing the User Experience: Setting and Observations

In order to assess the design of the user interface, as well as to gather more insights about the requirements for the design of such an appliance, the experience prototype of the EnlighTable was tested in the lab. As already anticipated, due to the novelty of the technology as well as the size of the physical table it relies on, it was not possible to test it in the field, i.e., in the context of creative agencies. To the end of coping with such limitations, while still gathering an understanding of how such an appliance could enter the design practice, potential target users (i.e., people working in creative agencies) were invited to test the appliance.

The user interface of the prototype was realized in Flash, and run on the interactive table of the FLUIDUM instrumented room. Such a table consists of an LCD monitor with a resolution of 1366 x 768 pixels, embedded into a wooden table and equipped with a DViT[2] overlay panel for interactivity. Hence, the team members share an overall table space of 1.6 x 1.2 meters.

Whilst this simulation set-up didn't provide the full multi-handed input intended in the design, nor a contextualized evaluation in the field, it still supported a preliminary exploration and the sharing of design ideas with potential users. Thus, a qualitative approach was adopted and 7 in-depth interviews were conducted with participants who covered different roles within creative team-work: two managers of creative agencies, two creative directors, a free lancer working on conceptual design for adverting agencies, a graphic designer, and a professor of art education. The size of the agencies

[2]http://www.smarttech.com/DViT/

the participants worked for was heterogeneous (up to 20 people), as well as the focus of the agencies (print as well as web communication).

Every interview lasted about one and a half hour and followed a predefined format. Each interviewee was first asked some questions concerning the specific workflow of her/his creative team: e.g., how many people usually work on a project, their roles and professions, how activities of picture selection take place, what activities are worked out in co-location, etc. (see Appendix B). The participant was then introduced to the vision of ubiquitous displays and novel collaborative scenarios. In this context a video was shown, which illustrates how the EnLighTable works. Afterwards, the participant was led to the instrumented room of our lab (cf. Chapter 1, Section 1.2) and the EnLighTable appliance was presented. In this phase, the participant was first given a short demonstration of how to use the different functionalities, and then s/he was invited to play around with it. The interviewee was then engaged in a discussion about the interface (e.g., what s/he found easy or less easy to understand and use in the prototype); as well as about the potential use and impact of such an appliance in her/his creative team-work.

Even though the number of testers is too limited to make generalizations, those interviews provided some useful insights about the design of the user interface, and about further requirements for such a kind of appliance.

Considering the pragmatic level of the interface, the participants mentioned that they found very intuitive to use the graphic gear for continuous zooming and enjoyed the feeling of "directly touching and directly working on the photo", using their fingers and the stylus to do so. In this respect, they mentioned it was a more immediate interaction with the media than the one afforded by a transducer such as the mouse. Despite recognizing the limitations of the technology (which supports single input only), they seemed to understand right away that the design was intended for two-handed interaction, and they actually used both the non-dominant as well as the dominant hands to interact, alternatively. Furthermore, 6 of them expressed the wish of using the pen with more freedom and in a more analogue way in the interface. Indeed, the interface was designed in such a way that a tapping and dragging action with the pen-tip (i.e., a diagonal stroke) would draw a frame for drafting a layout on the blank area of the table (similar to the interaction necessary for drawing a rectangle in most drawing applications, see Fig. 6.3, b). Those participants mentioned that they would have liked to use the pen for sketching and annotating as well. This, in terms of design, would imply the distinction between two modes to recognize the way in which the pen is used (i.e., similar to an analogue pen or to a mouse); or to create a gesture vocabulary for the creation of rectangles similar to the one implemented in the Brainstorm appliance (see Section 6.3.2 and Fig. 6.5, a, in this chapter). This dichotomy between analogue and mouse-like use of the pen as transducer is further discussed in Section 6.4.

Additionally, the interviews provided a deeper understanding of how such an appliance might enter actual collaborative design practices, and how its use can be affected by existing social and professional contexts. The possibility to simultaneously view multiple pictures on a table and to easily drag them within the focus of attention and visual angle of team members - so as to discuss them together with colleagues - was highly appreciated by each interviewee. For participants representing print agencies, the resolution of the display and the possibility to visualize photos in a 1:1 scale was very important. The interaction technique, which as already said allows direct manipulation without a mouse, was positively assessed by all participants: Some of them also mentioned that such an interaction could make the process of image selection and simple editing more accessible to everyone in the creative team, and potentially to clients as well, thus having clients' feedback earlier in the design process. Such a broader collaboration, though, was not recognized as necessary or desirable by every participant. In particular, the ones representing web agencies mentioned that the selection of photos in this kind of business is often managed by a single graphic designer, who sometimes directly interacts with the client: This fact implies minor communication flows and iterative meetings within the agency. Furthermore, in the field of web communication, ad-hoc photographic shooting is quite rare, and mostly digital libraries are consulted, so that collaborative selection might result unlikely. In these cases, the need and/or benefits of such an appliance for collaborative design were less evident. In other words, the different media of design (i.e., print vs. web communication) seem to imply different workflows and hierarchical structures, at least amongst the agencies represented by the participants. This fact, in turn, seems to affect the creative decision process, and it would potentially determine different uses/users of an appliance such as the EnLighTable in different organizations.

In summary, the assessment of the experience prototype in collaboration with the target users was informative on two different aspects mainly:

- in terms of current design practices: The inquiry sheds some light on how the target media (i.e., paper vs. digital communication products) affect the functional requirements for such a kind of appliance, as well as organizational and collaborative patterns of the creative team-work;
- in terms of the user interface at the pragmatic level: The inquiry suggests that novel interaction techniques may affect aspects of collaborative work by including different stakeholders at an early stage of the

project (e.g., clients) thanks to the way in which the media can be visualized and manipulated. Furthermore, it suggests how the use of a certain tool, i.e. the pen, seems to create diverse expectations for the way in which it can be used in a hybrid setting of interaction (e.g., as an analogue vs. a mouse-like tool).

Building on these insights, the inquiry extended from the table to a multi-display environment. The goal was to further investigate how the embedding of computing capabilities into the surfaces of a physical space can support collaborative creative processes and how. The possibility of connecting multiple displays in the environment, indeed, provides additional opportunities for externalization and spatial mapping of information/ideas, which is fundamental in creative processes. Hence, in the next section the design focus moves from the table as confined interactive surface, towards a multi-display set-up as socio-technical ecosystem, and explores how the combination of the different qualities of large interactive surfaces with different orientations can support the generation, externalization, shareability, and elaboration of ideas. In this way, the physical and social affordances of different surfaces are considered in relation to each-other, and cognitive affordances are designed for enhancing communication and creativity, striving for a high level of shareability (cf. Section 6.1).

6.3 Brainstorm

Brainstorm is an appliance which relies on a multi-display environment and is intended for supporting co-located collaborative problem solving. Collaborative problem solving requires knowledge and information to be exchanged among team members; different skills have to be coordinated and the information communicated by others needs interpretation, so that new ideas can be created and new solutions can be found. This process - with its core requirements of communication, coordination, and interpretation - is called *collaborative creative problem solving* (Amabile, 1983).

The design of the Brainstorm socio-technical system is meant to explore the possibility of merging the physical and social qualities of a traditional face-to-face collaborative creative environment together with some of the benefits of digital technology, such as persistent data storage, distributed information access, and the possibility to review previous processes or to undo certain actions. Thus, this section discusses the design challenge of embedding digital technology in a collaborative creative process without causing communication breakdowns, while still taking advantage of some of the qualities of Electronic Brainstorming Systems (EBS).

6.3.1 Motivation

Brainstorming is a technique for divergent thinking. It can be individual, although the term more usually refers to a group process for generating as many ideas or options as possible in response to an open question. Thus, it is frequently used for collaborative creative problem solving and it builds on a few main principles: i.e., quantity over quality of ideas, elaboration on others' ideas, and absence of criticism (Osborn, 1957). The technique relies on the communication among group members to stimulate idea generation, and on coordination to maximize the individuals' involvement and interpretation of ideas in order to create new intellectual associations, i.e., to increase the production of ideas.

Although Osborn (1957) claimed synergy effects of brainstorming, which positively affect the productivity of ideas, other studies (e.g., (Diehl and Stroebe, 1987), (Dennis and Reinicke, 2004)) have shown that these benefits are apparently outweighed by several negative social implications of the technique, such as apprehension of social judgments in face-to-face conditions. Those studies claim that nominal brainstorming groups (aggregating ideas from separate individuals) outperform face-to-face groups. In this context, EBSs that support distributed collaborators have been successful in increasing productivity of ideas, apparently because they allow for anonymity, which reduces evaluation apprehension.

On the other hand, face-to-face collaborative creative problem solving is still very common in practice, and its value is probably not to be associated with the number of generated ideas only. The individuals' subjective perception of the outcome of the process plays indeed an important role as well, and depends on the degree to which personal interests are represented and valued in the group's output. Furthermore, the face-to-face brainstorming situation has qualities which, in the long run, might even outweigh pure productivity measurements, namely the positive social aspects of team building, group awareness, and a shared sense of achievement.

In such contexts of face-to-face brainstorming, EBSs seem to perform poorly in comparison to nominal (i.e., distributed) brainstorming settings because of their disruptive effect. Using single-user systems, such as laptops, in a co-located collaborative setting leads, in most cases, to a communication breakdown since the user's concentration has to shift away from the group towards the device in order to use it. Furthermore, the size of a personal computer screen or the keyboard, as well as the turn taking implied by devices for single usage, seem to hinder the communication process (Applegate et al., 1986). In other words, referring back to the levels of shareability discussed in Section 6.1., this kind of interfaces offer a very poor level thereof, largely due to the physical format of the device.

For this reason, in face-to-face contexts digital technology is very often absent or shut down because it results disruptive of the group communication and of the "creative flow" (Csikszentmihalyi, 1996). Therefore, instead of relying on digital technologies, co-located creative meetings commonly rely on the physical benefits (e.g., gathering around a shared space) and social benefits (e.g., having equal access to information) afforded by physical surfaces, such as tables and walls, to exchange and visualize different types of information artifacts (e.g., Post-its, paper documents, pictures, etc.) in an immediate way. The results of such processes are often turned into a digital format in the end, by taking pictures of whiteboards and posters, or typing notes in digital documents in order to distribute and archive those results. Several transitions from physical to digital media occur which require additional work, and are mostly unable to capture and represent how the creative process has unfolded in time.

Considering this trade-off between the social and physical benefits of face-to-face collaboration, and the benefits of storage in digital format provided by EBSs, the design of the Brainstorm electronic system tries to combine those benefits in order to:

- maintain the social benefits of face-to-face collaborative creativity: These depend on the individuals' subjective perception of the group process;
- exploit the benefits of EBSs which are normally recognized in distributed collaboration, such as the capability to archive and easily review the collaborative process in different locations and points in time;
- explore how such a combination can affect creative collaborative patterns in terms of generation and organization of ideas, communication, and subjective experience.

The design choices which are presented below build on these goals.

6.3.2 Design

In order to cope with the issues above, the design of the Brainstorm appliance seeks to:

- blend the interface with digital information into the physical space, operational tasks, and conceptual models which normally support paper-based collaborative creative problem solving: In other words, drawing upon the social and physical affordances of physical technologies (e.g., table, wall, and paper) is expected to enhance the flow of idea generation and communication;

- through digital technologies, augment the group possibilities for spatial mapping and externalization of ideas normally afforded by physical technologies, so as to provide cognitive affordances for the association of ideas.

In a first stage, one should consider the different physical and social affordances of the large surfaces which are normally embedded in the process of collaborative creative problem solving, such as tables and walls. The horizontal plane of a table affords writing, face-to-face communication, territoriality (Scott et al., 2004), body language, and group awareness, while a wall display allows and supports shared visualization, overview, and context awareness. In face-to-face creative meetings, these surfaces normally support different stages of collaborative creative problem solving. Tables are commonly used for the generative phase, in which different ideas are created (i.e., divergent thinking), whilst walls allow for stepping back and considering the displayed alternatives, towards the reductive phase (i.e., convergent thinking).

Building on these considerations, we (i.e., the FLUIDUM research team) developed a system that combines the different affordances of table and wall, and augments them by automatically displaying content on both of them for supporting externalization and spatial mapping. The Brainstorm system relies on the interactive table and wall of the FLUIDUM instrumented environment (see Fig. 6.4 and cf. Section 6.3.4 for more technical details), whose tracking system is suitable for implementing a two-input interface, and hence a two-users appliance.

The design of the user interface builds metaphorically on paper Post-its, on the ways in which they are socially used, as well as on their manipulation vocabulary in the physical world in order to suggest ways in which ideas can be generated and manipulated as information units in the EBS appliance. Studies of paper in work practice, in fact, show that paper continues to be widely used (Sellen and Harper, 2003), some of the reasons including its spatial flexibility (it can be quickly arranged in the physical space), sociability (it facilitates face-to-face communication by being passed on), and tailorability (it is easily annotated) (Cook and Bailey, 2005). Post-its, in particular, are commonly used in the the *idea card method* for brainwriting, which is based on Geschka's (1978) and Van Gundy's (1981) "Interactive Brainwriting Pool Technique". In this method, group members write their ideas on a piece of paper that is then placed in the center of the table for another member to read prior to writing their next comment. In this way, Post-its afford the

Figure 6.4: The Brainstorm multi-display environment for creative collaborative problem solving.

externalization and record of ideas in written rather than just verbal form in the generative phase. Furthermore, they support a certain territoriality and the creation of semantic regions (e.g., Fig. 6.6, a). When participants are given a stack of Post-its and start sticking them around their working area, they define their personal region, which remains visible to others, thus creating a mutual awareness among participants. Using Post-its on vertical surfaces supports the convergent thinking phase too, when group members stick and move Post-its on flip charts or whiteboards in order to recognize relationships and create clusters (e.g., Fig. 6.7, a).

Drawing upon these considerations, the Brainstorm interface was designed around the paper-based brainwriting technique. Users can start creating ideas by drawing a square on the table surface (see Fig. 6.5, a). This event triggers the appearance of a large yellow square, resembling a Post-it, thus defining the area to write in (as in Fig. 6.5, b). By tapping a designated area marked as a small square in the center of the Post-it, the latter shrinks to a smaller size and becomes moveable (Fig. 6.5, c). The user can then create new Post-its/ideas by drawing new squares in a blank region of the

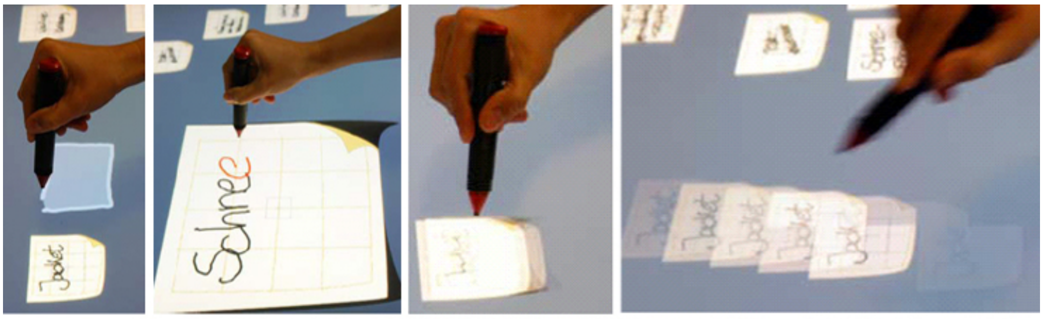


Figure 6.5: Interactions with the Brainstorm on the table: a) Creating a Post-it; b) Writing and Idea; c) Moving a Post-it; d) Skidding a Post-it to pass it over to the partner.

table and writing within the yellow region. This choice was made in order to create visual constraints for writing, so as to identify ideas as units, and to create visual cues for distinguishing territories and patterns. When the Post-it is shrunk, its content is still readable. Such graphical Post-its can be edited, moved, deleted, and copied by any participant after they have been created. To delete a Post-it, this can be dragged to the edge of the screen till it disappears. To copy a Post-it, it can be virtually flipped by tapping on its plied bottom right corner (see Fig. 6.5, b); then, it can be duplicated by tapping on the "copy" icon displayed on its back side.

Additionally, a mechanism to encourage users to build on each other's ideas was implemented: With a quick movement of the pen, each user can deliberately skid one idea to the other participant (see Fig. 6.5, d). The Post-it slides quickly across the table and smoothly reorients itself towards the other user: This was intended to support the explicit sharing of ideas so as to encourage the creation of association chains.

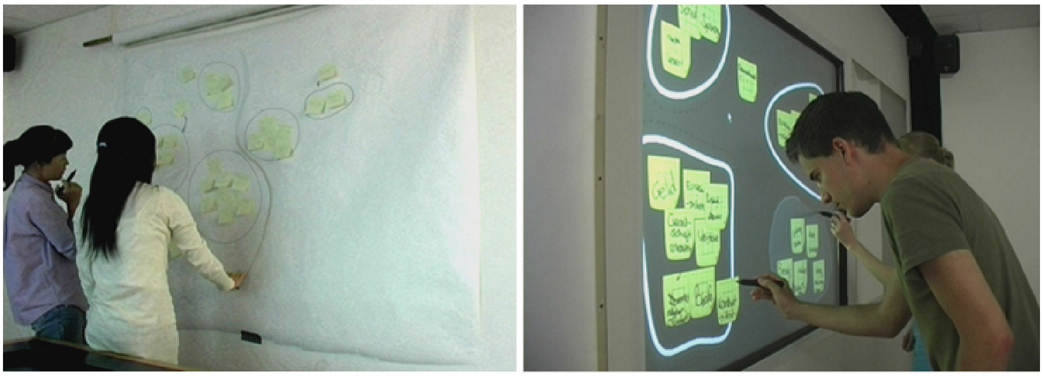
The choice of using handwriting and gestures to enhance fluid generation of ideas and interaction in a group process builds on some of the related work on collaborative creativity considered in Chapter 3 (e.g., (Klemmer et al., 2001) and (Guimbretière et al., 2001)), as well as on the work on chunking and phrasing (Buxton, 1995) considered in Chapter 2. In this respect, Buxton (1995) examines the effects of compound tasks on the users' cognitive load: These are tasks that usually can be expressed in one sentence (i.e., a phrase, such as the gesture of writing text onto a Post-it), but in standard Desktop applications normally have to be broken down into multiple steps (e.g. select target, choose text tool, type text). This may result in additional cognitive burden on top of the actual task (idea generation). To this end, the design of Brainstorm implements a limited manipulation vocabulary and relies on

Figure 6.6: Generative phase of a creative collaborative problem solving process. a) With the paper-based technique; b) With the Brainstorm electronic technique.

simple marking gestures for direct manipulation such as drawing a square, writing text, and stroking for moving, whose direct feedback is augmented by the coincident spatial mapping of input and output (i.e., there is no such device as a pointer or a remote controller). This creates a transparent causal relationship between gestures and output, and supports visibility of gestures. Furthermore, by simply dragging the Post-its, temporal spatial arrangements can be created, which are crucial to epistemic creative actions (Kirsh, 1995).

The immediate and visible change of the shared visual landscape is supported by the system in additional ways. As the participants create Post-its in their working areas, thus already creating a distinct territorial set-up, the Post-its appear simultaneously on the vertical display, which is located next to the table. On the vertical display, the Post-its are reoriented upright, i.e., readable for both readers, but they maintain a spatial mapping to the territorial set-up on the table display. In this sense, the perception of territoriality and group awareness are supported: A participant will recognize his/her own "territory" (i.e., contribution) on the wall, and at the same time will gain an overview of the ideas created by the group. This feature was intended to augment the possibilities of ideas association by replicating and reorienting ideas/units in the perceivable space (thus augmenting the shareability of the interface, cf. Section 6.1). Furthermore, it was intended to support the different phases of creative problem solving by building on the different affordances of table and wall.

Indeed, when users move from the interactive table (generative phase, divergent thinking) to the wall display (structural phase, convergent thinking), they can spatially organize the ideas which were automatically displayed by rearranging them on the wall. In addition, they can create clusters by draw-



a) b)

Figure 6.7: Convergent phase of a creative collaborative problem solving process. a) With the paper-based technique; b) With the Brainstorm electronic technique.

ing a circle around some Post-its (see Fig. 6.7, b). In this way, the same transducer (i.e., the pen) was associated with different interaction vocabularies when used on different surfaces of the same appliance. In the generative phase, as already mentioned, digital Post-its can be created on the tabletop by drawing a closed line, but this function was not enabled on the wall. Such a constraint was a design choice to cope with the impossibility of distinguishing, from a system perspective, whether a stroke of a closed line was meant for creating a new Post-it or as a marking line for creating a cluster. Thus, the gesture vocabulary that was implemented on the wall was designed for structuring ideas only, and not for generating new ones.

Clusters are merged by dragging them close together. Drawing a cross on the border of a cluster causes it to dissolve into single Post-its again. Clusters can be connected to each other or to single Post-its by drawing a line from the border of one cluster to the border of another one or to the center of a Post-it. Finally, whole clusters can be moved across the display, thus moving all the Post-its they contain. This set of clustering techniques clearly extends the functionality of a physical whiteboard or flip chart (e.g., Fig. 6.7, a) while it maintains the direct manipulation characteristics thereof. This aspect, in turn, can facilitate the creation of a structured knowledge representation (e.g., a mind map) easily editable by every participant.

6.3.3 Evaluation

In order to assess those design choices with respect to their underlying motivations (cf. Section 6.3.1), the Brainstorm EBS was compared to a paper-based collaborative problem solving process. The goal was indeed to understand the implications of blending such an EBS in the physical space, in comparison to a traditional paper-based brainstorming technique. More details about such an evaluation were published in (Hilliges et al., 2007). What is relevant, in this context, is how such a comparison elicited the different ways in which those choices affected the process and perceived experience of the collaborative problem solving process. To this end, the focus of analysis and discussion is here set on:

- the interaction with the Brainstorm interface at the pragmatic level, and whether this supports a fluid generation and spatial organization of ideas in the electronic brainstorming in the same or different ways in which physical technologies do so in paper-based brainstorming processes;
- the patterns of association and organization of ideas, and whether they are similar or different in the two modalities;
- the subjective experience of collaborating with the Brainstorm appliance in comparison to paper-based brainstorming.

Below, the set-up and procedure of the study are briefly described, and the observations are discussed in light of their implications for the design choices.

6.3.4 Technical Set-up and Procedure

The system was deployed in the FLUIDUM instrumented environment, containing an interactive meeting table (cf. Section 6.2.3) as well as displays embedded into an interactive wall. Such a wall consists of an interactive surface with a width of 5 meters and a height of 2.5 meters containing three back-projected displays. The two side displays as well as the rest of the wall are tracked by four cameras. The center display additionally provides high precision input through a DViT panel (see Fig. 6.4, and (Boring et al., 2007) for more details). Two single points can simultaneously be tracked both on the table and on the wall, thus enabling parallel interaction of two users.

Our research team conducted a within-group comparative study between the Brainstorm EBS and the original paper-based brainstorming technique with 30 participants in 15 teams, dealing with two tasks each. The participants represented a variety of professional backgrounds (e.g., computer science students, architects, designers, civil engineers, musicians, as well as journalists), ages (between 20 and 50 years old), and nationalities (5 countries). The participants had to deal with one different task/problem per technique, and the order of the techniques was counterbalanced. In the first task, the teams were asked to take care of an Inuit coming to a foreign country neither speaking the country's language, nor having any useful equipment for the new environment. In the second task, the teams had to discuss their own needs when they would leave their home country for emigration into harsh, icy arctic territories. The subjects were asked to collect all material and immaterial items they would consider necessary for survival under these conditions. Taking into account the broad variance in participants' professional education, these tasks were chosen because they could be addressed without any domain specific knowledge. Furthermore, they were rather simple tasks to which everybody could relate, and thus contribute a significant number of ideas.

Adopting a qualitative approach, we ran questionnaires before and after the task (see Appendix C) to evaluate the subjective expectations, perception and assessment of the Brainstorm system (for more details on the procedure see (Hilliges et al., 2007)). Furthermore, the video recorded sessions were analyzed to recognize association chains and communication patterns. Some main observations originated from such a video analysis, as discussed below.

6.3.5 Observations and Discussion

When looking at the the ways in which people generated ideas and structured them in the Brainstorm appliance, a number of observations could be made, concerning both the pragmatic and the semantic level of the interface.

The participants seemed to grasp very quickly the manipulation vocabulary afforded by the interface, both on the table and on the wall, and did not appear to have difficulties in performing the necessary actions for creating, organizing, and sharing digital Post-its/ideas. This is confirmed by the results of the questionnaires in which the participants were asked about the ease of use of each interaction gesture that was implemented (cf. (Hilliges et al., 2007)). Thus, the design choice of changing the functional vocabulary of the same transducer (i.e., the pen) in the two different display orientations/phases (i.e., from creating and handwriting units/ideas to connecting and clustering them) did not appear to cause any difficulty during the tests.

Some distinctions could be observed, though, between the use of the pen in the electronic vs. the paper-based organization of ideas on the wall. As already mentioned, the interaction vocabulary of the stylus on the wall was such as to allow for the creation and deletion of clusters and connecting lines, but it did not allow for handwriting because of the impossibility of recognizing, from a system point of view, whether a stroke was intended as a writing gesture or rather as a clustering or moving one. These restrictions did not seem to cause any disruption though, and none of the participants actually tried to create new Post-its when interacting on the wall display. Likewise, in the paper-based condition, people organized the Post-its they had written while they were at the table, but did not write new ones when they moved to the wall. On the other hand, they labeled their connections and clusters by writing annotations on the paper sheet underlying the Post-its, or drew lines to visually separate different semantic regions (e.g., see Fig. 6.7, a). Such use of the stylus was obviously not possible in the Brainstorm appliance because of the constraints previously mentioned.

The difference in the interaction vocabulary of the stylus between the electronic and the paper-based brainstorming appeared to affect the process of ideas organization in additional ways. Whilst in the paper-based brainstorming no ideas/Post-its were physically discarded, nor marked clusters were usually edited/moved, in the Brainstorm appliance the participants structured their ideas in a much more dynamic way, moving clusters to central areas as well as peripheral ones of the displays; creating and deleting connections; and apparently using the real estate in such a way as to visually support a semantic mapping. In the paper-based setting, on the other hand, a cluster initiated by one pair member usually grew in the same region and few changes were made, likely due to the effort of physically moving the Post-its one by one, and the impossibility of deleting marking lines. One can then speculate that the action of annotating and labeling clusters in the paper-based setting is a strategy to cope with such a limitation of easily moving and rearranging clusters.

Further observations could be made concerning how the interface affected association and communication patterns. One first consideration regards the way in which participants used the interactive wall display during the generative phase, as an additional external reference. The wall was indeed showing ideas written on the desk immediately, and thus conversely to what is possible with purely physical technologies. By analyzing the videos of the recorded sessions, it appeared that when participants wanted to visually (and, probably, mentally as well) step back and obtain an overview of the stage of the process, they looked at the wall where all ideas generated so far were available (see Fig. 6.6, b, for example). Such availability of a common reference apparently increased group awareness and overview of the shared workspace. Furthermore, it was observed that this fact often led to resuming discussion after a pause or a dead end. When a team got stuck in the generative phase, indeed, it often occurred that both members looked at the wall rather than at the table, and then started writing additional ideas: In this way, it was easier to review all the ideas generated so far and start over by elaborating on earlier ones. Likewise, the possibility of skidding digital Post-its across the table provided additional opportunities for mutually enhancing associations. This feature was used by several participants for explicitly stimulating communication over an idea, and in turn its elaboration.

When asked about their subjective experience of interacting with the electronic vs. the paper-based brainstorming system, most of the participants (24/30) stated they would favor the electronic version over the paper-based one. Concerning the reasons of their preferences, those participants mentioned the possibility of going back to previous actions, as well as automatic storage. In this sense, most of the test users seemed to recognize a strong value in the possibility of having a shared reference, a documenting record that can track the decision process: Some of the participants even suggested that it would be good to record the discussions during the sessions. All in all, these results on the subjective experience suggest that the electronic version did not just maintain the social benefits of traditional face-to-face collaborative problem solving during the creative process: Rather, it was perceived as it could provide additional benefits later on, as a shared archive.

In summary, the observations on the use of Brainstorm as a context of collaborative interaction indicate that the system positively affected the association patterns by replicating and spatially mapping ideas/units on the different displays in concomitance. These considerations foster existing theories on creative processes and collaborative work. When a person is exposed to stimuli from a variety of contexts, she is more likely to have novel associations (Santanen et al., 1999). The activation of such associations can be automatic (without intentional conscious awareness) or depend upon the context of the stimuli (conscious capacity spreading activation (Barsalou, 1982)). In this sense, the Brainstorm interactive context seems to generate additional opportunities for stimulating associations.

Furthermore, considering that the human capability of keeping chunks of information in short-term memory is very limited, the possibility of generating, externalizing, and spreading ideas in the physical space, whilst exploiting some of the advantages of digital technologies, seems to positively affect the cognitive creative process. Typical techniques for freeing cognitive resources are indeed externalization, e.g. through the use of space (Kirsh, 1995), epistemic action (Kirsh and Maglio, 1994), visual output, and spatial mapping (Norman and Bobrow, 1975), (Hutchins, 1995). Thus, more resources remain available for creative associations. In this sense, the spatial distribution of visual cues (i.e., cognitive affordances) about the generated ideas, as well as users' possibility to explicitly exchange ideas (i.e., social affordances, such as for example the possibility of skidding and reorient the Post-it/idea to the opposite user) and easily rearrange them (i.e., functional affordances, such as the possibility of moving whole clusters), seem to enable and foster the generation of new unexpected associations, as well as the creation of semantic visual structures.

6.4 Summary and Discussion

This chapter has presented two projects investigating the possibility of enhancing collaborative creativity by digitally augmenting physical large surfaces such as tables and walls, which normally support shareability and communication of information/ideas in team-work. To this end, the design goal was to embed hybrid interaction in the collaborative process without generating disruption of the social context, nor of the creative flow. In this sense, those design choices were somewhat consistent with Scott et al.'s recommendations (2003) for the design of tabletop applications: "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 the design approach that was adopted in this work can be seen as consistent with those principles, the observations of how people actually interacted with the experience prototypes only partially confirm those claims, and elicit some issues that deserve further investigation. In particular, they suggest that the way in which people use pen and paper-like interfaces in hybrid interaction is not completely consistent with the way in which they do so in the physical realm. Rather, interaction vocabularies that come from both the physical and the WIMP environments seem to concurrently shape people's expectations of functionalities of the transducer, and possibly their mental models too. In the EnLighTable appliance, for example, dragging diagonally the pen-tip on the surface of the tabletop for drawing a frame (similarly to a mouse-based interaction) was easily understood, but for some people the use of a pen also seemed to suggest the possibility (or desire, at least) of sketching and handwriting.

As designers, we then need to think thoroughly about the interaction vocabulary that we design when using a transducer which already has a manipulation vocabulary in the analogue world. One needs to consider, on the one hand, how to take advantage of the additional functionalities allowed from the digital environment: E.g., a single stroke can create a rectangle, like in the case of the EnLighTable; tapping and dragging a cluster can move multiple objects at once to another location, like the digital Post-its in the case of the Brainstorm system. On the other hand, one should keep in mind that the functionalities of such a transducer are embedded in existing mental models (i.e., a pen is used for writing and scribbling in the analogue world). Put differently, it needs to be carefully considered where to draw the line between a mouse-like transducer for single input point, and a physical tool for analogue interaction.

Based on these considerations, it becomes evident how the integration of aspects of physical interaction in the design of hybrid ones raises some open questions about the expectations, mental models, and attitudes that people build around such interactive systems, thus implying designers to ask themselves to which extent can emulating the physical world result in a behavior similar to that exhibited in the analogue world (given that this is desirable). For different kinds of interactional experiences in the digital world, what specific aspects of the physical one should be mimicked? What are the consequences of this for people's behaviors and expectations about a given system? How do different aspects affect people's mental models and behavior in interaction?

Taking a step back from some of the claims of the related literature about the benefits of simulating physical interactions, which are not always grounded in empirical analysis, the next chapter looks more systematically at the relationship between different aspects of physicality and interactional patterns. The goal is to help guiding design decisions about how and to what extent it makes sense to apply aspects of physical interaction to digital interface design.

7

Controlled Comparative Studies

This chapter presents two controlled studies focusing on some of the issues raised by the evaluations of the sketches and prototypes previously illustrated. In particular, it reports on the observations that could be made by comparing similar systems integrating different aspects of physical interaction, and discusses them in light of their implications for the design of hybrid interactions.

7.1 Reflecting on Design

When looking at the sketches of interactions and experience prototypes that were presented in this dissertation, one can make a number of considerations in relation to users' response on the design choices that were made and evaluated in these sketches and prototypes, and to the projects for different scales of devices previously reviewed (cf. Chapter 3). To this end, in Fig. 7.1 the different projects elaborated in this thesis are mapped to the physical qualities distinguished in Chapter 1, Section 1.1.2, and illustrated in Figure 1.3, consistently with the approach that was taken for the analysis of the related projects reviewed earlier in the text (cf. Chapter 3, Figg. 3.25, 3.26, 3.27, and 3.28).

At first, one can notice a consistency with some of the design choices emerging from the survey of the related work. The design of hybrid interactions for interactive surfaces such as the Mug Metaphor Interface, the Living Cookbook and the EnLighTable integrate aspects of directness and continuity, similar to most of the related work on this class of devices (see Figg. 3.25, 3.26, and 3.27). Differently from most of those related projects, though, they also metaphorically refer to physical artifacts in order to suggest cognitive

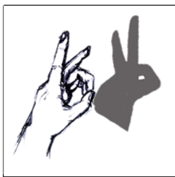
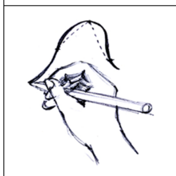
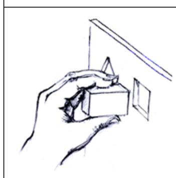
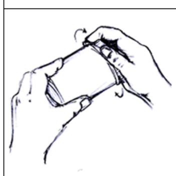
		Mug Metaphor Interface	Learning Cube	Living Cookbook	Time-Mill	EnLighTable	Brainstorm
metaphor		■	■	■	■	■	■
directness		■	■	■	■	■	■
continuity		■	■	■	■	■	■
3D space		□	■	□	■	□	□
contstraints		□	■	□	□	□	■
multimedia		□	■	■	■	□	□
two-hands		■	■	□	□	■	□

Figure 7.1: Physical qualities integrated in the hybrid interactions designed in this thesis.

affordances for the manipulation of digital media in 2D (cf. Chapter 4, Fig. 4.1, for more details on this concept).

Similar observations can be made for the Brainstorm appliance, as an example of a multi-display environment. Like other work in this field (see Fig. 3.25), the physical constraints of the real estate of different displays are integrated in the design of the interface for suggesting cognitive affordances about the functionalities and qualities of different screens (e.g., digital Post-its can be deleted by dragging them to the edge of the tabletop display; lateral screens have a lower resolution for suggesting periphery of information). But also in this case, and differently from other related projects, the interface integrates metaphorical references to the manipulation vocabulary of physical artifacts, such as Post-its in this case.

The two designs of interactive tangible objects, i.e., the Learning Cube and the wheel of the Time-Mill Mirror, have aspects of directness, continuity, 3rd physical dimension and multimodal feedback which are in common with other work previously considered (see Fig. 3.27 and Fig. 3.28). Their metaphorical meaning, though, does not simply consist of a symbolic embodiment such as in the examples of the Metadesk (Ullmer and Ishii, 1997) and Tangible Viewpoints (Mazalek et al., 2002) discussed in Section 3.2.4. Rather, the manipulation vocabulary of the 3D objects (i.e., a cube and a wheel) is metaphorically mapped to the behavior of the digital output to provide physical affordances which are consistent with the cognitive ones. By exploiting the redundancy of multimodal feedback (including the visual and the haptic one), the design goal was here to convey cognitive as well as physical affordances at the pragmatic level, which would lead to a consistent mental model of the conceptual one (cf. Fig. 4.3 in Chapter 4).

When now focusing on the different designs that were presented in this thesis (Fig. 7.1), looking at them in relation to each-other, one can notice similarities and distinguish differences among them, and reconnect those to the different lessons that were learned in their evaluation of experience of use. As one can see in Figure 7.1, all the projects share the aspects of metaphorical representation (as they reference to the manipulation vocabulary of physical objects), directness, and continuity. Two of them, though, the Learning Cube and the Time-Mill Mirror, differ in the way in which they integrate the metaphorical link: Those two projects, indeed, provide physical affordances on top of cognitive ones because of the 3rd dimension they integrate, which in turn affords graspability and active haptic feedback. Now, when referring back to the results of the evaluations of the different projects, one can notice that:

- In the interaction with interfaces on 2D surfaces, the metaphorical reference to the manipulation vocabulary of physical objects was not always understood correctly at a pragmatic level. In particular, despite the interface being designed for continuous gestures (e.g., rotating the mug in the Mug Metaphor Interface, or dragging the graphical lever on the dial of the Living Cookbook), the testers seemed to have different expectations about the behavior of the digital media (i.e., automatic changes of the graphical interface after single touch).

- In the interaction with physical 3D objects, such as the Learning Cube and the Time-Mill Mirror, the participants seemed to show an immediate understanding of how to physically manipulate the object, as well as a sense of engagement and explorative attitude in the way in which they appropriated the interface. With the Learning Cube, the children tried out different gestures to discover how these could affect the digital output; with the Time-Mill Mirror, people varied the browsing speed by spinning or carefully rotating the wheel.

Reflecting upon these aspects, one can then start to ask: What are the fundamental differences between interaction in a 3D world and a 2D one that metaphorically mimics the other, physical one? Answering this question requires a thorough understanding of the the real/phyisical and perceived/cognitive affordances for interaction (cf. Chapter 2, Section 2.7 for a distinction) in these different situations, as well as a consideration of people's expectations and mental models about physical versus digital media.

Referring to the aspects of physical manipulation that were distinguished and used as a framework for analysis in this thesis (cf. Fig. 1.3), one can then assess the implications of their integration (or not) in the design of hybrid interactions. In other words, one can meticulously analyze the effects of mimicking some of the aspects of physical interaction by comparing the interaction with hybrid interfaces to the interaction with physical artifacts. To this end, a controlled study was conducted: This compares the manipulation of physical objects with those on a digital tabletop, which emulates analogous physical tasks in a number of important dimensions.

The next section reports on this study and on the differences between the two, physical and digital contexts of interaction. Building on those, it discusses how the results can inform interface design for future interactive surfaces and stimulate further investigations along this line.

7.2 Manipulation in 3D vs. 2D

Wanting to understand at a deeper level how interactions with digital and physical objects might differ, an experimental comparison of interactions in the two modalities for the same tasks was designed (Terrenghi et al., 2007b). The key differences between the physical and digital conditions were: i) the lack of the ability to manipulate objects in the 3D space in the digital condition; ii) the corresponding lack of multimodal feedback (through proprioception, for example); iii) and the lack of physical constraints such as the ones given by the thickness and friction of 3D artifacts, resting on a horizontal surface (see Fig. 7.24). The next paragraph presents the design of the study in further detail.

7.2.1 Study Design

The study was designed so that participants were engaged in both a puzzle and a photo sorting tasks. The two tasks were chosen to explore a diversity of potential interactions with artifacts, 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 interface of digital artifacts was mapped as closely as possible (in terms of appearance and interactive nature) to the physical objects they represented. In this case, the digital tasks on an interactive tabletop were 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 were 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 tasks which were chosen facilitated two forms of analysis. First, they allowed for basic measurements of performance in order to map out broad differences in physical and digital interactions. These include 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 were categorically coded and statistically compared.

Second, a deeper and potentially more informative qualitative analysis of behaviors at the interface could also be derived from the video record, this latter analysis allowing for the interpretation of the observations in grounded instances of interaction with digital and physical artifacts.

This combination of both quantitative and qualitative analysis is consistent with the exploratory approach adopted in this study.

7.2.2 Technical Set-up

The prototype of a projection- and camera-based interactive tabletop was used as platform for the study. 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 the 2D space on screen. To translate an object, 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 center of the rotation and another relative contact point on the object to specify the angle of the rotation (i.e., similarly to the "two-point rotation and translation mechanism" described by Hancock et al. (2006)). The choice of these techniques was deliberate in order that the digital objects could be translated and rotated in a fashion similar to their paper-based physical counterparts, using one or two hands, and multiple fingers.

7.2.3 Participants and Procedure

The 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 5 centimeters 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 enjoyment and 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.

7.2.4 Quantitative Results and Qualitative Observations

The results of the study consist of a statistical analysis of both the videos and the answers to the questionnaires, as well as of a descriptive video analysis. 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), as shown in Fig. 7.2, a.

The puzzle task completion time and the total of the two tasks were 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

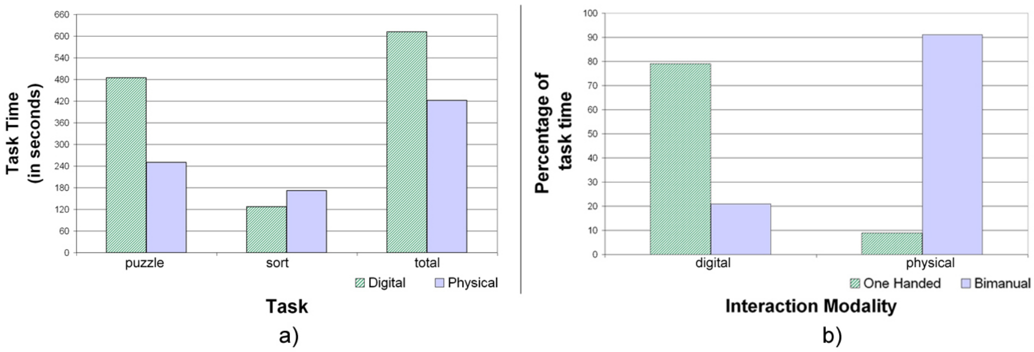


Figure 7.2: Quantitative results: a) Time to complete each task for digital and physical conditions, for both tasks and in total; b) Percentage of task time (both tasks) engaged in either bimanual or one-handed interaction.

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 and aligning the small 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 from 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 vs. 1.25 respectively), although overall frustration levels were lower for the digital sorting task than for the digital puzzle one.

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 artifacts was observed and coded.

One aspect which immediately became evident was the degree to which one versus two hands was used, and the nature of these interactions. One-handed interaction was mostly 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 artifact), 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 (Guiard, 1987). From video records of both the puzzle and sorting tasks a log was made of the time spent in one-handed and bimanual interactions (see Fig. 7.2, b).

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 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, as previously remarked, 9 out of 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 actions were more prevalent in the digital condition than asymmetric actions. Conversely, 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.

Concerning the qualitative observations, the video analysis provided rich material for recognizing salient differences in the way each participant engaged in both physical and digital interactions, and highlighted differing strategies for task completion between participants. 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 were 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.

The qualitative observations of the videos are organized on the basis of those patterns.

General Posture and Patterns of Manipulation

First of all, a remarkable difference emerged 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 as well as in 2 dimensions (e.g. sliding the puzzle tiles on the board).

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, moving items with the right hand only (see Fig. 7.3, b).

Another difference could be noticed in the relationship between the dominant (right) and non-dominant (left) hands in the two domains. According to Guiard's (1987) description of asymmetric bimanual interaction, 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

a) b)

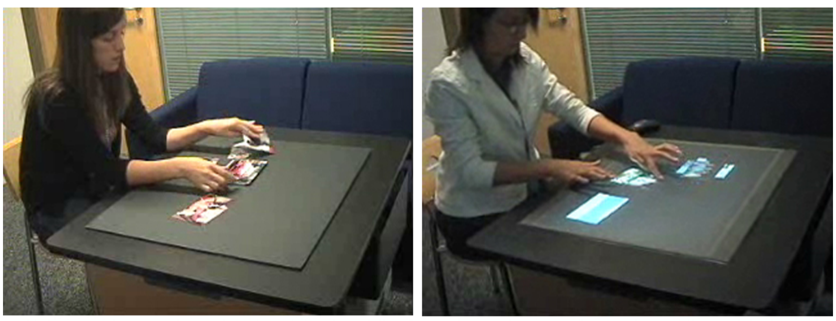
Figure 7.3: General posture adopted in the physical (a) and digital (b) tasks.

were partly confirmed by the video analysis of the physical interaction. However, some differences between the physical puzzle and the physical sorting tasks were noticed, namely:

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 artifact (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" (1987)).

In the physical sorting task, the photos were initially piled in a stack instead of being disarrayed on a 2D 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 artifact to the other one.

The remaining 4 participants left the pile on the table during the sorting task (see Fig. 7.4, a, 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,



a) b)

Figure 7.4: Two-handedness.

while 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 2D than in the 3D 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.

In the digital modality, interaction in the 2D space resulted in a different distribution of actions between hands. The range of actions observed in the physical tasks (e.g., holding a puzzle tile with one hand whilst 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 indeed not possible in this condition.

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 moving items to a pile on the left side (i.e., similar to the physical interaction shown in Fig. 7.4, a). 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, thus exhibiting an attempt for higher accuracy of action (for example see in Fig. 7.4, b). Coarser and faster symmetric movements of both hands (and multiple fingers) were sometimes used to simultaneously move multiple objects (e.g., see Fig. 7.8, b).

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, a clear spatial-temporal pattern emerged: In the physical setting, participants often brought each item closer

a) b)

Figure 7.5: Focusing on an item.

to their eyes, thus facilitating a focused view while spatially isolating the item from the rest of the array of items (see Fig. 7.5, a); then, they placed the item on the tabletop in relationship to existing structures (i.e., puzzle or piles of photos).

In the digital case, a similar pattern emerges: Here, the 2D 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 Fig. 7.5, b).

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 than in the physical sorting task, since 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 (see Fig. 7.6). 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 hands could alternatively bring one photo or another one closer to the user's eyes (see Fig. 7.6, a). More than 2 photos were sometimes kept close to each other in the air using multiple fingers (see Fig. 7.6, b); 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

a) b)

Figure 7.6: Comparing items in 3D space.

a) b)

Figure 7.7: Comparing items on the physical (a) and digital (b) tabletops.

the participant moved their forearms towards 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 (e.g., Fig. 7.7, a), despite the fact that there was no obvious need for holding them.

In the digital tasks, the 2D space of the interactive table implied 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. Also in this case, fingers were often kept on the pictures during comparison (e.g., Fig. 7.7, b).

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 2D 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 photo on the top was visible. The two modalities afforded different strategies for gaining an overview of quantity and content (see Fig. 7.8).

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 being suggested by the physical depth of the pile. In these cases (for 4 of the 12 participants, as already mentioned), 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 accomplishing the sorting task: For 3 of the participants, the physical affordances of the tabletop were more extensively exploited to spread the photos out. In these cases, 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 (Fig. 7.8, a). 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.


a) b)

Figure 7.8: Getting an overview of the pile content.

The unsorted pile in the interactive table did not afford an estimation of quantity in the same way because of its 2D 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 7.8, b, 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, it was observed 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 on in the task.

In the physical tasks, different strategies were recognizable. As shown in Figure 7.9, a, 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, engaging her visual attention rather somewhere else, but this provided a physical reminder (i.e., a haptic cue) 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

a) b)

Figure 7.9: Holding items in "stand-by" mode.

other cases, the participants placed one photo in a blank region of the board, at the periphery and away from existing piles, thus creating a visual cue.

The interactive table supported this stand-by mode in 2 dimensions only. Hence, participants often dragged an item to a blank region of the display for creating a visual cue, placing or sorting the item later on. There were also some examples (e.g. Fig. 7.9, b) in which one hand was kept as kind of placeholder on the items for some time, while looking at other items (similarly to how the left hand is used in the physical modality in Fig. 7.9, a).

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 two 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. Nine 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

a) b)

Figure 7.10: Using the same spatial layout in both conditions.

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 Fig. 7.10). Another participant created just two piles (only "keep and share", and "discard") in both settings, again laid out in approximately the same way.

7.2.5 Takeaways and Implications

This section has reported on an exploratory study which compared physical and digital manipulation on a tabletop in such a way that the effects of actually being able to "get to grips" with objects of interaction could be evaluated. Thus, it was possible to focus the analysis on how the differences of manipulation of digital artifacts (cf. Section 7.2., at the beginning, and Fig. 7.24 later in this chapter) affected the nature of interaction in comparison to the physical condition. Whilst 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.

The 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 artifacts, and so on) allowed for the elicitation and reflection on the fundamental differences between interacting with tangible objects in 3D space as compared to digital objects in 2D space. This comparison showed that, despite the different physical/real affordances (cf. Section 2.7) of the physical and digital modalities, there were fundamental elements that both conditions had in common and emerged in the patterns in which the tasks unfolded. These include, for example, the need for getting an overview of content and quantity, for comparing objects, for focusing on particular ones, for holding some objects distinct from others, and for keeping some in a "stand-by" mode. However, the means and strategies by which these sub-tasks were accomplished were intrinsically different across conditions and appeared to be affected both from the physical, real affordances of manipulation, as well as from the participants' mental models.

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, predominantly one-handed interaction was performed in this modality. Furthermore, any bimanual interaction observed 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 the 3D 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 that a lack of tangibility and 3D space - which in turn imply physical forces such as weight and friction - undermines the usual allocation of hands to these commonly asymmetric roles.

Third, the presence or absence of the physicality of individual objects affected the way in which users manipulated the artifacts. 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 and proprioception. Such assessments in the digital world required other more effortful strategies and actions, such as the ones considered in Fig. 7.8, b, 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). These were mostly visualization strategies, which exploited the 3rd spatial dimension for bringing artifacts in fore- and background, i.e., in focus and periphery of the their visual angle (and hence of their visual attention). The 3rd dimension also meant that participants could either bring objects close to the body or the body towards objects, offering a greater range of flexibility for dealing and manipulating the artifacts in the tasks.

In terms of design, these considerations imply 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 one necessarily has to mimic physical properties, but that rather it is necessary to carefully examine and recognize what those physical affordances achieve for people when working with tangible objects, and ask how one can employ perhaps different methods to attain those same ends in hybrid interactions.

An important point is the finding that hybrid interaction may not naturally engender the kind of bimanual interaction we see in the physical world (even if it supports it). This indicates that in order to confer the benefits of bimanual interaction (Leganchuk et al., 1998), one approach is to design specific tools and techniques which more explicitly require asymmetric bimanual interaction, e.g. "ToolGlass-type" interfaces (Bier et al., 1993). Furthermore, to support the need for focus plus context visualization (e.g., see Fig. 7.5), 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. Likewise, graphical visualizations that suggest depth, as is proposed in pile metaphor interfaces (Mander et al., 1992), or the emulation of physical forces, such as in the BumpTop interface (Agarawala and Balakrishnan, 2006), can support the estimation of quantity; thumbnail views of the pictures and semi-transparent overlays displaying the number of items in the pile are other examples of how design could address such issues.

Building on these considerations, another study was designed in order to explore the possibilities, design strategies, and potential benefits of supporting two-handed cooperative work in hybrid interactions. In particular, aspects of manipulation in 2D vs. 3D and their associated feedback are considered in light of their potential for suggesting two-handed interaction. Such a study is reported in the next section.

7.3 3D vs. 2D Handles at Interactive Surfaces

The previous study has shown how some elements of the physical condition, such as the ones implied by the 3D space of manipulation and multimodal feedback (proprioception in particular), seem to support two-handed, asymmetric interaction. And indeed, when looking back at the projects of interactive objects considered in the related work (cf. Chapter 3), most of them are designed for two-handed interaction (cf. Figg. 3.27 and 3.28), especially those for spatial navigation (cf. Section 3.2.1). This fact is likely due to the natural space-multiplex way in which we interact with physical artifacts, which affords the manipulation and exploration of physical objects in several contact points simultaneously.

On the other hand, 2D GUIs for interactive surfaces can as well suggest two-handed interaction, for instance by allocating different controls to dominant and non-dominant hands, or by representing the visual appearance of the interface so as to imply laterality: Two examples for this are respectively the Mug Metaphor Interface and the EnLighTable projects which were designed and presented in this thesis (see Fig. 7.1). Such experience prototypes, though, were implemented in such a way (i.e., in Flash) that it was not possible to test and carefully observe the actual two-handed interaction.

Thus, the study presented in this section investigates the implications of using a 3D tangible object as a handle for manipulation on an interactive surface vs. a purely graphical UI. In this sense, it addresses the question: If the facility for essentially manipulable 2D graphical content is concomitant, why design into a 3rd dimension? And if one does, what impact might this have on the user's interaction behavior? Considering the results of the previous study, does a 3D physical handle foster two-handed, asymmetric cooperative work?

In order to investigate the effect of tangibility and physicality more closely, a 3D and a 2D versions of PhotoLens, a system for photo browsing on an interactive tabletop, were designed. With respect to the dimensions distinguished in Chapter 1, Section1.1.2 (see Fig. 7.24), the two versions are both intended for two-handed interaction, but differ in the fact that the 3D PhotoLens integrates the possibility of manipulation in a 3D space and the haptic feedback provided by the consistency of the physical handle, while the 2D PhotoLens doesn't.

The following paragraph briefly introduces the theoretical background which motivates such an investigation. Then, the design rationale for the 3D PhotoLens is discussed in relation to what the existing research literature suggests are potential benefits of tangible devices: Those benefits create the foundations for the expected results of the comparative study. In such a comparative evaluation users explored both this interface and the 2D purely graphical alternative one (although direct touch enabled). This process allowed for the evaluation of how pushing the interaction into a tangible 3rd dimension influenced patterns of user behavior, as discussed in the remaining of this section.

7.3.1 Background and Motivation

As it was mentioned above, there are systems that support direct touch control of a graphical user interface (GUI) (e.g., (Terrenghi et al., 2006a), (Agarawala and Balakrishnan, 2006), (Wu and Balakrishnan, 2003)), and those that bring tangible physical objects (TUIs) to a computationally enhanced surface (e.g. (Fitzmaurice et al., 1995), (Jordà et al., 2007), (Patten et al., 2001), (Mazalek et al., 2006)). In each case the technology is designed such that it transfers humans' manipulation skills and mental models gained from interactions with the physical world and integrates them with the extensive possibilities of digital media.

The two approaches, though, are different in the aspects of physical interaction that are drawn upon in the design of such hybrid systems. In the case of GUIs for direct touch, designers may rely, for example, on the metaphorical 2D representation of physical artifacts and on their manipulation vocabulary in the physical world to suggest the hand gestures or marking strokes to be operated. In the case of TUIs, designers exploit the degrees of freedom, manipulation vocabulary, and haptic feedback enabled by the 3rd spatial dimension of the physical transducer.

Thus, when creating those systems, designers have mostly adopted design principles from either WIMP-based interaction (i.e., graphic design) or from physical interaction (i.e., product design). Furthermore, the benefits of physicality for interaction design are derived either from the comparative observation of physically enhanced vs. WIMP-based interaction (e.g., (Patten and Ishii, 2000)), or from the dedicated analysis of one of the two (e.g., (Guiard, 1987), (Kirsh, 1995)). This has produced valuable insights: Considerations of ergonomics, cognitive psychology, and sociology have been merged in the related literature sustaining the benefits of physicality (e.g., (Dourish, 2001), (Klemmer et al., 2006), (Hornecker and Buur, 2006)). Nonetheless, the spatial combination of physical manipulation and display of digital output in direct touch interactive surfaces creates new design challenges and opportunities which haven't been thoroughly investigated so far.

And indeed there is significant potential, given advances in technology, to construct interfaces which combine elements of both tangible interaction on computationally enhanced surfaces and the ability to perform direct touch style manipulations with their digital/graphical representations. The work on the interweaving of physical and digital aspects in interface design for interactive surfaces suggests a variety of benefits for the introduction of a physical handle: cognitive (e.g., intuitiveness and learnability), manipulative (e.g., motor memory), collaborative (e.g., group awareness), experiential, as well as in terms of efficiency. But the empirical work that supports such claims is actually limited and mostly focuses on one aspect in isolation from the others, thus taking for granted, to some extent, some of the benefits of integrating aspects of physical interaction in the design of hybrid ones.

It's a claim of this thesis that the mutual influences of the different qualities of physical interaction integrated in the design of hybrid ones cannot emerge if we do not start distinguishing *what* those very aspects of physical interaction actually are, and *how* they affect different levels of the experience of use (e.g., discoverability of the interface, ease of use, social collaboration, as well as fun). Through the comparative analysis of design solutions that integrate (or not) the 3rd spatial dimension and different types of haptic feedback (i.e., somesthesis and/or proprioception, cf. Chapter 4) we can then start eliciting the effects and implications of such integration more consciously. These aspects become especially crucial when we expect interactive surfaces to support everyday life activities, including causal and leisure interactions (e.g. Microsoft SurfaceTM[1]).

With such issues in mind, the PhotLens system was designed in order to consider more closely the implications of tangibility at interactive surfaces.

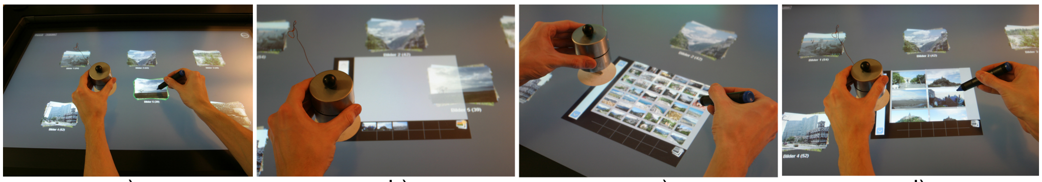
7.3.2 Design Rationale and Expectations

The PhotoLens system is a hybrid tool for browsing and organizing digital photos on an interactive tabletop. The choice of developing an interface for photo browsing is particularly linked to this notion of evolving interaction paradigms being tethered to the support of digital interactions in more social and casual areas.

The rapid shift of photography from analogue to digital, together with the reduced cost of taking pictures, has caused a substantial growth of personal photo collections, and of the technological tools that we use to capture, display and interact with them. On the other hand, the size and orientation of the displays of desktop PCs, together with the WIMP paradigm they rely on, neither provides the social affordances suitable for co-located sharing and collaborative manipulation and organization of collections (an imperative feature of users' interactions with photos (Frohlich et al., 2002)), nor the creation of temporary spatial structures, as our physical surfaces and artifacts do (Kirsh, 1995).

In the envisioned scenario, the photo collections of different users (e.g., friends, family members) can be displayed on the tabletop. Photo collections are visualized in piles, by analogy to Mander et al. (1992) and Agarawala and Balakrishnan (2006). Piles can be freely translated on the tabletop (i.e., no automatic array in a grid is present) by touching and dragging the image on the top with a finger or with a stylus (see Fig. 7.11, a). In order to save real estate and avoid clutter, the PhotoLens can be used to gain a localized, unfolded view of the pictures contained in one pile, without interfering with the information landscape of the shared display. Figure 7.11 shows in further details how the PhotoLens works: a) Piles can be moved freely on the table

[1]http://www.microsoft.com/surface



a) b) c) d)

Figure 7.11: Interaction with the 3D physical PhotoLens.

using the stylus; b) The digital lens appears when the physical tool is placed on the table; c) The pile unfolds in a thumbnail view and moving the handle up and down the scrollbar scrolls through the thumbnails; d) The view can be zoomed in and out by rotating the upper part of the tool and selected pictures can be copied onto a temporary tray (and are retained independent of the pile viewed). Additionally, a new pile containing the photos from the different collections can be created by tapping on the icon in the right bottom corner of the lens.

The design choices and the expectations about the integration of a physical 3D handle for manipulating digital media on an interactive surface build on previous and related work.

In the previous study (cf. Section 7.2.), it was observed that despite the fact that some interactive systems allow for bimanual interaction on a display (which is known to offer both physical and cognitive benefits (Leganchuk et al., 1998)), people tend to use only one hand - and preferably the dominant one - when manipulating digital media, possibly due to their acquaintance with the WIMP paradigm. Therefore, one could expect the use of a physical tool, associated with a digital frame and a stylus for interaction, to more explicitly suggest two-handed cooperative work. By providing users with both a tool and a stylus, the idea was to suggest the use of the non-dominant hand for navigation tasks (i.e., grasping and rotating the tool) and of the dominant hand for fine-grained tasks (i.e., selecting and dragging pictures), consistently with the suggestions for the design of bimanual asymmetric interactions presented by Buxton and Myers (1986). The stylus is indeed typically held with the dominant hand, hence users were expected to use the non-dominant hand for interacting with the physical tool in order to use their hands cooperatively, such as in Guiard's (1987) kinematic chain. To make this affordance even more explicit, and considering that most of the population is right-handed, the graphical lens was designed so that it would extend on the right-up side of the physical tool (see Fig. 7.11, b). Navigation functionalities, such as placing (i.e., appearing of the lens frame), scrolling, and zooming were then mapped to the manipulation of the physical tool.

Additionally, the physical affordances of the tool in 3D, like placement and rotation, were expected to support the offload of cognitive effort thanks to the haptic feedback it provides (i.e., proprioception). The tool, indeed, can be operated without looking at it, thus not hindering users' visual attention. The effect of its manipulation is mapped in real time and in the same area (e.g., zooming and scrolling of the pictures in the lens), thus providing an isomorphic visual feedback of action. In this sense, the continuity of action (rotation and translation) enabled by the physical tool, as well as the multimodal feedback it provides (haptic and visual), were expected to serve for a higher feeling of control. Buxton's (1995) work on the effect of continuity of action on chunking and phrasing, as well as Balakrishnan and Hinckley's (1999) investigation on the value of proprioception in asymmetric bimanual tasks would support those expectations in terms of the cognitive benefits associated with such a physical handle.

Since the graphical lens appears when the tool is placed on the table, and disappears when the tool is lifted, this feature was meant to support an efficient use of the real estate: Users could indeed display the lens only when required. Furthermore, the fact that the lens can be physically picked up in the 3D space and moved to another pile makes it unnecessary to drag it in 2D across the screen, stretching arms and sidling between other piles, thus providing motor benefits.

Although the related literature on tangible UIs also claims the social benefits of tangibility (e.g., (Hornecker, 2002)), those could not be assessed, because the technical set-up which was used for the study only recognizes two input points (i.e., interaction with only one PhotoLens at a time). Thus, interactions with the system are here based on individual action, which makes the social affordances of such interfaces a consideration for future work.

7.3.3 Technical Set-up

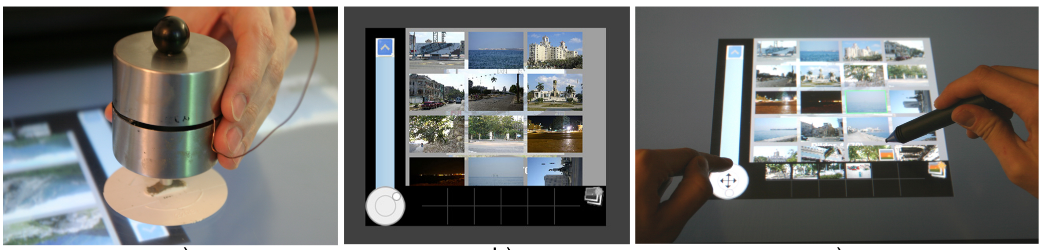
The technical set-up of PhotoLens consists of an interactive table and a modified wireless mouse for the implementation of the physical handle (see also (Butz et al., 2007) for more details). The components of the mouse were rearranged in a metal cylinder with a diameter of 7 cm and height of 9 cm, which was taken from a disassembled kitchen timer (see Fig. 7.12, a). The size of the tool is determined by the components of the mouse, as well as by considerations about its manipulability with an adult's hand.

The interactive table is the same used for the EnLighTable and the Brainstorm experience prototypes described in Section 6.2.3., which uses four cameras in the corners of the frame to track two input points simultaneously. An input transducer can either be a pen, a tool, or simply a user's finger. Such a frame shows some limitations when wide transducers are on one of its diagonals because this causes a mutual occlusion. The thinner the body of the input transducer, the lower is the risk of occlusion, and the more accurate is the tracking. For this reason, a kind of base was created for the physical tool (see Fig. 7.12, a) so that its stem creates a smaller shadow and hence provides more accurate tracking. Although such a solution enabled the use of physical artifacts "off-the-shelf" for the creation of the physical handle (i.e., a wireless mouse, a kitchen timer, a copper wire serving as antenna, a screw with a knob for disassembling the tool, etc.) its appearance was clearly "sketchy", in the sense that it made clear this was a prototype. This fact can potentially have affected some aspects of the subjective perception of experience (cf. Section 7.3.7) and is further discussed in Section 7.3.8.

7.3.4 The Comparative Graphical PhotoLens

For comparative purposes, the 2D PhotoLens had inherently the same functionality as the 3D version: It was a direct touch enabled graphical interface, but it did not extend into the 3rd dimension. Lacking a physical handle for picking it up, the 2D PhotoLens is permanently displayed on the tabletop and can metaphorically overlap and unfold photo piles when it is dragged onto them. The control for moving, scrolling, and zooming of the PhotoLens is represented by an interactive circle, as illustrated in Fig. 7.12, b.

When a user touches the small circle on the graphical control wheel and slides her finger along the circular trajectory of the graphical control, clockwise rotation zooms in, whilst counter-clockwise rotation zooms out the thumbnail view. When the user touches the center of the same graphical wheel, four perpendicular arrows appear (see Fig. 7.12, c). These resemble the symbol of movement used in the GUI of several Desktop applications (e.g., Microsoft PowerPoint, Adobe Photoshop). Sliding the finger up and



a) b) c)

Figure 7.12: Physical and graphical elements of the PhotoLens: a) The physical component of the 3D PhotoLens; b) A screenshot of the 2D PhotoLens; c) Interaction with the purely graphical PhotoLens.

down along the line of the scrollbar, the thumbnails scroll up or down, as in a Desktop GUI. When the control circle is touched and dragged away from the pile for more than 5 cm, the whole lens moves along, for example onto another photo pile or into an empty area of the tabletop.

7.3.5 Study Design

To engage users with the interface, they were asked to bring a sample of 80 personal digital photos (from a trip or vacation) to the study session. During study trials, participants completed two tasks with their photos using both interfaces (i.e. 3D and 2D, counterbalanced across trials and participants), giving a total of 4 trials. In each trial, participants were presented with 6 piles of 80 photos (80 random images from their own collection in one pile, with other piles being made up of images provided by the researchers, and used to simulate the presence of a companion's images).

In one trial, the participants were told to interact with only their pile, selecting 12 images suitable for use as desktop wall-papers. In the other trial, they interacted with all the piles, searching for 12 images to accompany a proposed calendar. These different trials were chosen to observe participants' interaction behaviors in a local region (i.e., with one pile) and across the whole real estate of the tabletop surface (i.e., with different piles arrayed on the surface). In both cases, participants were told to store selected images in the temporary tray of the PhotoLens and then create a new pile.

Before each task, the user had that current task explained and the interface demonstrated (including demonstration of the potential for using two-handed interaction). After the trials, the participants completed an evaluation questionnaire and discussed their experiences with the experimenter in charge of the session.

The participants were 12 right-handed adults (mostly university students, with different majors, in an age range from 20 to 30 years old), comprised of 6 men and 6 women, all with normal or corrected to normal vision, and all having normal range mobility in both arms.

7.3.6 Method of Analysis

To help ground a deeper analysis and to understand broader patterns of action at the interface, the research team calculated the extent of use of differing forms of interface manipulation (i.e., different forms of handed interaction) in the two conditions. To understand participants' subjective response to the two different interface styles, Likert scales from 1 to 5 (negative to positive) were also used. These focused on key characteristics such as ease of use and enjoyment, and certain specific manipulative actions such as zooming, scrolling, and placing the lens (see Fig. 7.22).

All user trials were video recorded; the majority of the evaluation is therefore based on the direct consideration of these video materials. The video recordings were studied by an interdisciplinary team and subjected to an interaction analysis (Jordan and Henderson, 1995). The main focus of the analysis was in looking for both patterns of common interaction strategies and specific moments of novel, unexpected interaction. Specific attention was also given to the initiation of moments of interaction.

This approach to the data was taken because it was more appropriate than traditional attempts to quantify behaviors at the interface. The paradigm of digital interaction that was being explored, i.e. leisure technology (photo browsing in this case), does not fit a traditional model of recording task completion times. By taking a fine-grained, micro-analytic approach to recovering patterns of activity and breakdown during interface interaction at the pragmatic level, a rich understanding could be derived of how, qualitatively, a 3rd dimension in an interface was appropriated and understood by users. Consequently, the ensuing results section seeks to articulate some specific vignettes of interaction, some moments of user activity which were considered of particular interest and were particularly illuminating in the attempt to understand the impact of tangibility on interaction behaviors.

7.3.7 Quantitative Results and Qualitative Observations

Similar to the approach adopted in the previous study, results can be distinguished into two categories: the first highlighting some broad patterns of handed interaction at the interface, and the second providing a more detailed view of some of the common elements of interaction during tasks.

In order to address the first issue, the percentage of time spent in different forms of handed interactions in the two different modalities was measured. As one can observe in Figure 7.13, participants demonstrated diverse approaches to interaction with the interface, which might suggest that they were developing different mental models of system function; or, simply, they were approaching the interface with different pre-conceived manipulation skills, habits, and preferences for physical and digital media. This observation is especially reinforced by the unexpected results of the concomitant two-handed interaction in the 2D modality (both in terms of average time results and standard deviations), which are further discussed in Section 7.3.8.

When analyzing the videos, five predominant forms of interaction with the interface could be observed, as shown in Figure 7.13, logically conforming to those actions immediately possible (none of the participants, indeed,

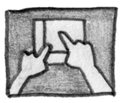
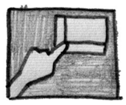
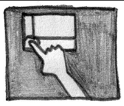
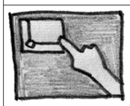
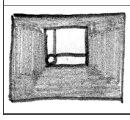
Forms of handed interactions		3D	2D
	Concomitant two-handed interaction with the PhotoLens	9.0% (11.9)	19.4% (24.7)
	The non-dominant hand interacts with the control wheel for scrolling and zooming	44.7% (15.6)	37.5% (24.8)
	The dominant hand interacts with the control wheel for scrolling and zooming	2.1% (5.6)	17.4% (24.6)
	The dominant hand interacts with the photos for selection tasks	31.3% (21.6)	17.0% (6.4)
	No hands are on the interactive area	12.9% (6.4)	8.8% (7.8)

Figure 7.13: Average percentage of time spent in differing forms of handed interactions in both Physical (3D) and purely graphical (2D) conditions (standard deviations in brackets).

selected the photos with the non-dominant hand). The interaction analysis partially draws on such a classification of conditions to identify, analyze, and describe snippets of interactions which were found relevant for what can be considered a "catalogue of interaction experiences". Such a catalogue is illustrated below, through vignettes of interaction following the common lifecycle of interface activities during elements of the photo browsing task.

Approaching the Task

At the beginning of the task, in both modalities, the participants were asked to select 12 photos from their own pile, which was displayed in the bottom right corner of the table. Piles could be moved freely across the table so as to enable epistemic actions, i.e., allow users to create spatial arrangements as they liked and found more comfortable for interaction. Despite such a feature, some interesting differences could be noticed amongst subjects in the way they approached the task and the postures they adopted.

The participant in Figure 7.14, for example, first moves away the pile in front of her using the stylus with the right hand, so as to gain space; then, she moves her pile from the right side to the center of the table. In this way she seems to create a sort of focused interaction area, where she can easily

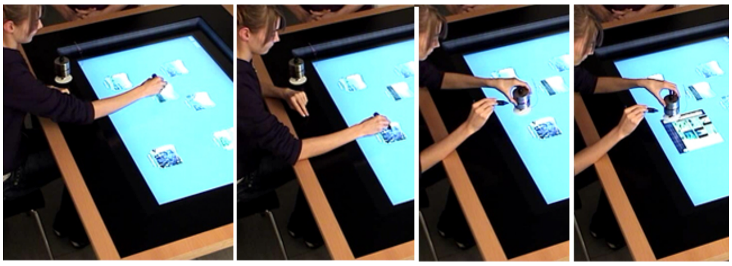


Figure 7.14: Moving the artifacts towards the body.

Figure 7.15: Moving the body towards the artifacts.

visualize and reach the photos of her collection/pile. She then grasps the physical handle from the border of the table with her left hand and starts browsing through the photos.

A very different interaction attitude can be observed in Fig. 7.15, where the participant moves her body towards the pile to be sorted, rather than the other way around. In this case, she first places the physical handle on the screen of the table with the dominant hand; she then drags it on the table towards the pile in the right bottom corner. Thus, in order to better reach the interaction area, she moves the chair to the right side of the table, in the proximity of the pile she wants to sort, and she then starts interacting with the PhotoLens.

Browsing the Photo Collection by Scrolling and Zooming

By rotating and sliding the control wheel (either the 3D or the 2D one), users could browse thought the photo collection, thus exploring the content of the pile. As anticipated in Section 7.3.2, the design choice of placing the control wheel at the left bottom corner of the lens was meant to suggest two-handed manipulation of the PhotoLens and manipulation of the control wheel with the non-dominant hand. This was not, however, always the approach taken by the participants.

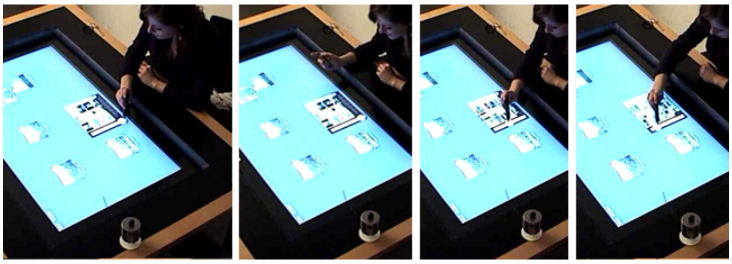


Figure 7.16: One-handed interaction with the 2D PhotoLens.

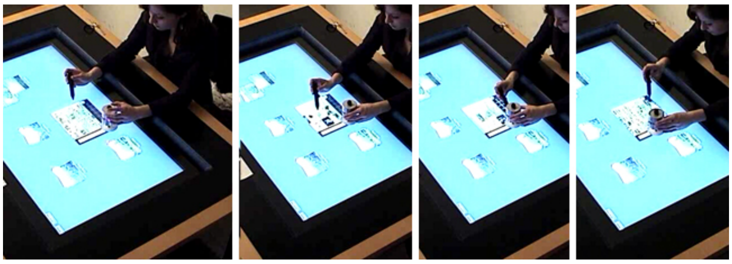


Figure 7.17: Two-handed interaction with the 3D PhotoLens.

In Figure 7.16 the participant interacts with the 2D control wheel with the pen, held in the dominant hand, while the non-dominant hand is rested on the border of the table. In this way, the participant partially occludes her own view, which brings her to alternatively lift the pen and her hand from the table to better see the pictures in the thumbnail view (e.g., second frame of Fig. 7.16). Furthermore, as she explained in the post-test questionnaire, she found it more difficult to manipulate the small sensible area of the 2D wheel for zooming, in comparison to grasping the physical handle: One can speculate that this is why, as it was observed in the video analysis, in the 2D modality she mostly used the scrolling function of the wheel to browse through the whole photo collection, hardly changing the zooming factor.

Conversely, when interacting with the 3D PhotoLens, she manipulated the physical control wheel with the non-dominant hand only, exploring the content of the collection both by scrolling and zooming (e.g., see the second and third frame in Fig. 7.17). In such an interaction pattern, both the hands were kept concomitantly on the interactive area of the table during the whole interaction with one pile.

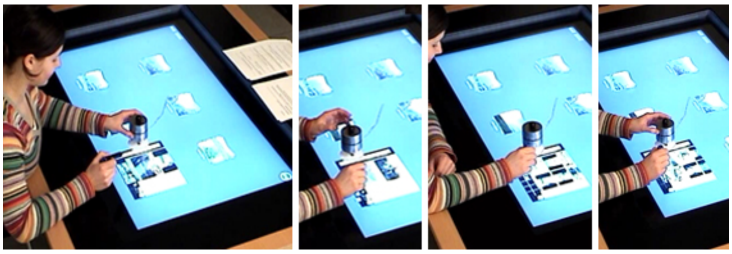


Figure 7.18: Alternate use of the dominant and non-dominant hands with the 3D PhotoLens.

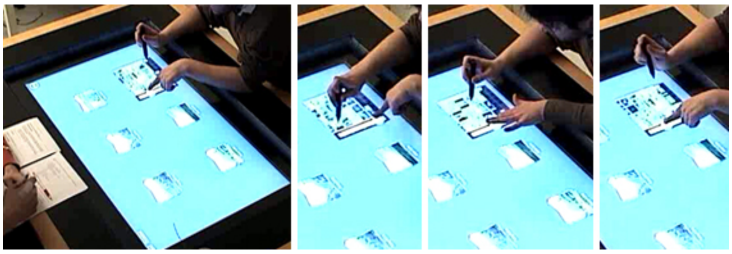


Figure 7.19: Concomitant use of the dominant and non-dominant hands with the 2D PhotoLens.

Selecting Photos in the Lens

Providing the participants with a stylus was expected to suggest interaction with the dominant hand for selection tasks: None of the participants (who were all right handed), indeed, performed selection tasks with the non-dominant hand (see Fig. 7.13).

Additionally, because of the laterality of the control wheel and of the scrolling bar with respect to the lens, one could have expected interaction patterns similar to drawing ones to emerge. In these cases, a tool (e.g., a ruler) is usually held with the non-dominant hand, while the dominant one performs micrometric tasks in the proximity of the tool (e.g., draws a line). The type of patterns that were exhibited by the participants were often rather different across modalities in the way people alternatively or simultaneously used the non-dominant and dominant hands.

As one can see in Figure 7.18, as an example of interaction with the 3D PhotoLens, the participant first positions the physical tool on a photo pile with the non-dominant hand, and then starts browsing through the photos by scrolling and zooming. In this phase, she keeps the dominant hand in the proximity of the interactive area, holding the stylus. After she has set a preferred height in the scrollbar and a desired zooming factor, she then releases the non-dominant hand (see second frame in Fig. 7.18) and rests it at the border of the table (see third frame in Fig. 7.18). She then proceeds in the task by selecting the photos with the dominant hand. Such a cycle of interactions unfolds again when the zooming and scrolling are newly set with the non-dominant hand (see fourth frame in Fig. 7.18).

Surprisingly, in the 2D modality participants kept more continuously both hands simultaneously on the interactive area (see time percentage of concomitant two-handed interaction in Fig. 7.13). As shown in Fig. 7.19, for example, the participant keeps his left forefinger on the 2D control wheel during the whole interaction with a pile: i.e., both when the dominant hand is selecting photos (e.g., second and third frame) and when it is just held in the proximity of the lens, ready for interaction (e.g., fourth frame).

Although the 2D graphic PhotoLens is permanently present on the interactive surface - and can be moved on the table only when it is dragged - several participants mentioned in the post-test questionnaire that they constantly kept their fingers on the wheel because they had the feeling that the lens would disappear otherwise.

Placing and Moving the PhotoLens

When participants were asked to create a new collection by selecting photos across several piles on the table, different strategies for moving the lens and photos could be noticed: These showed some differences amongst subjects and between modalities concerning the ways in which people took the tool to the pile or vice versa.

In Figure 7.20 one can observe how the same user interacts with the 2D (Fig. 7.20, a) and the 3D (Fig. 7.20, b) PhotoLens. To reach the piles he stands up in both conditions. In the 2D modality, he drags the lens towards different piles with a finger of the non-dominant hand. When selecting photos from one collection (e.g., third frame Fig. 7.20, a), he rests his non-dominant hand on the border of the table. He then uses it again for moving the lens towards another pile (e.g., fourth and fifth frame Fig. 7.20, a), while resting the right hand on the border this time. All in all, he never moves the piles and alternatively uses the non-dominant and dominant hands for respectively moving the lens on the table and selecting photos within the lens.

In the 3D modality, he adopts a very similar strategy. He first places the physical handle with the dominant hand on a pile; then, he swaps hands for browsing, and again for selecting. In these cases, one of the hands is always rested on the border of the table. In order to move the lens towards another pile, he slides the physical tool on the surface of the table (e.g., fourth and fifth frame in Fig. 7.20, b), rather than picking it up and placing it again.

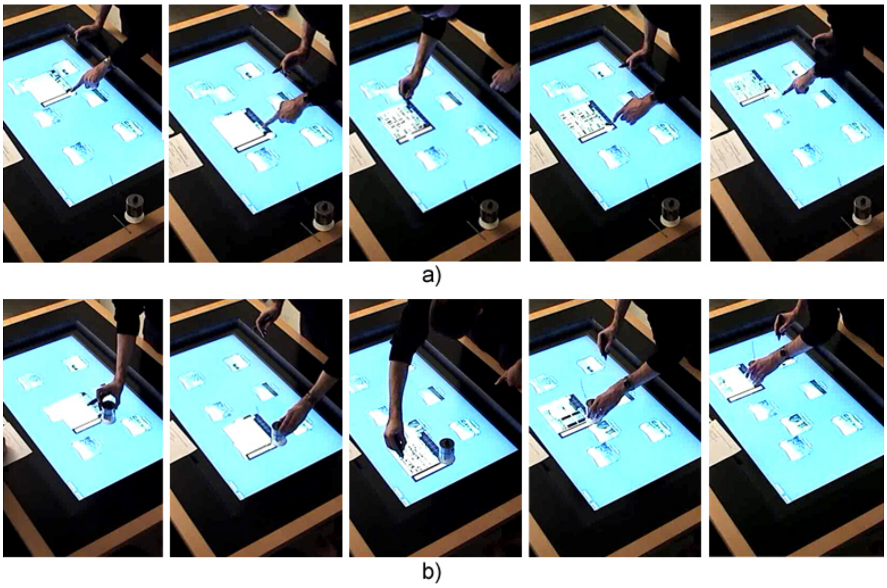


Figure 7.20: Moving the tool and the body towards the piles: a) 2D PhotoLens; b) 3D PhotoLens.

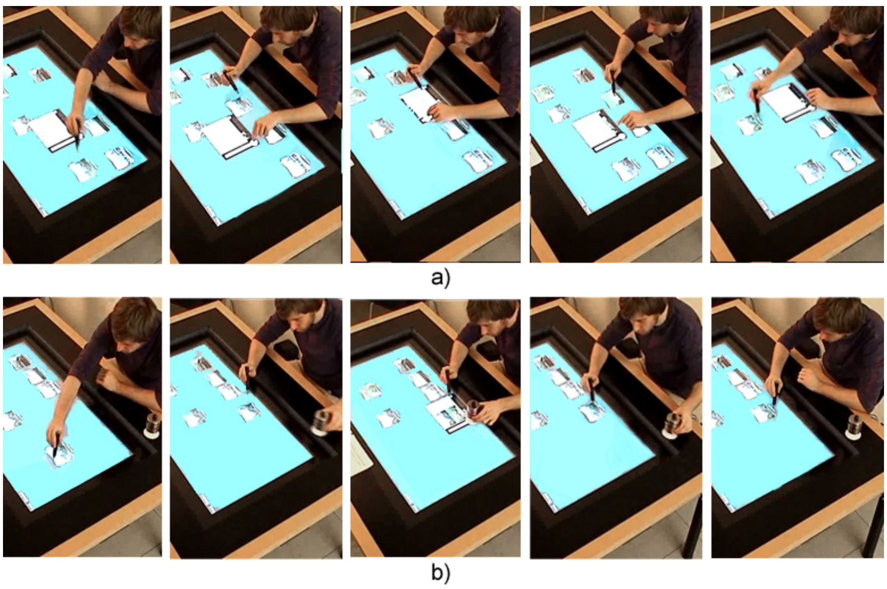


Figure 7.21: Moving the tool and the piles towards the body: a) 2D PhotoLens; b) 3D PhotoLens.

A different approach can be observed in Fig. 7.21. In this case, the participant tends to move the piles and the lens towards his body. In the first frame of Fig. 7.21, a, he drags a pile towards himself with the dominant hand; with the non-dominant one (second and third frame) he than moves the 2D Photolens towards the pile to interact with it. In the fourth and fifth frame, he moves other piles towards himself with the dominant hand, while slightly moving the PhotoLens between one interaction cycle and another one with the non-dominant hand. All in all, the interaction takes place in the proximity of his body, and the dominant and non-dominant hands are alternatively used for moving respectively the piles and the lens.

When interacting with the 3D PhotoLens (Fig. 7.21, b), he adopts a similar allocation of tasks to dominant and non-dominant hands (i.e., moving the piles and the lens accordingly). In this case, he takes advantage of the graspability and mobility of the physical handle in the 3D space to place it at the border of the table (second, fourth, and fifth frame in Fig. 7.21, b).

Perceived Experience

Figure 7.22 reports the results of the post-test questionnaires (average values on a Likert scale from 1 to 5, negative to positive). When statistically compared (see Fig. 7.23), the results of a within-subjects t-test analysis show that for most of the measures the differences between 2D and 3D usage conditions was not statistically significant, limiting the conclusions that can be drawn. However, the relative perceptions of the ease of use of the zooming function clearly showed a significant difference between conditions ($t(11) = 8.40, p < 0.01$, significant using Bonferroni correction). Interestingly, despite the physical control being easier to use on average, participants reported that overall it is more fun to interact directly with their hands on the screen than with a tool. Although this result is not statistically significant, it is worth exploring this response a little more, referring to some of the participants' comments.

Some people mentioned they found it easier to use the 3D PhotoLens, especially in the zooming function, because it does not require so much attention for accurate interaction as the graphical wheel does. With respect to this, they said: "With the physical tool you only have to rotate"; "With the physical tool you don't have to think about what you can do, you see it immediately"; "You don't need to look for the exact point where to put your finger to rotate"; "The rotation for zooming is intuitive as it reminds the use of analogue cameras"; and finally "It is easy to place it and rest it in one position: I had the feeling I needed to hold the digital lens in place".

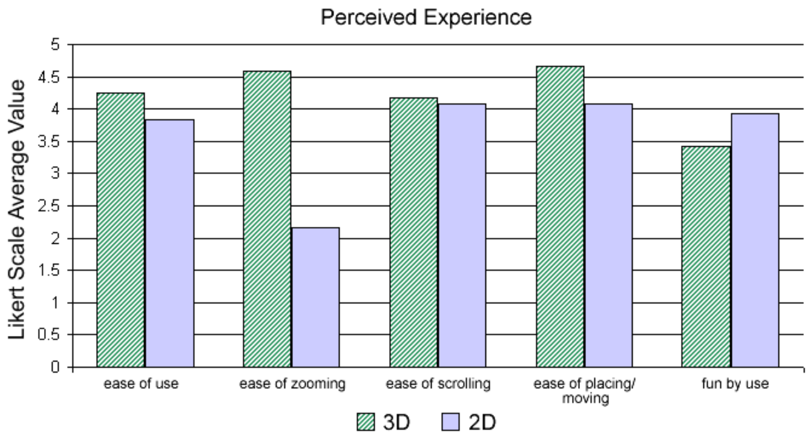


Figure 7.22: The results of perceived experience in terms of average of the Likert scale values.

Measure	3D	2D	T-Test
Ease of use	4.25	3.83	p = 0.34
Ease of zooming	4.58	2.17	p < 0.001*
Ease of scrolling	4.17	4.08	p = 0.78
Ease of placing/moving	4.67	4.08	p = 0.15
Fun by use	3.42	3.92	p = 0.17

Figure 7.23: Average responses for 5 key indicators of performance (Likert scale scores). *Significant above the $p < 0.01$ level (Bonferroni corrected level of significance).

This last kind of comment is apparently one of the reasons why the concomitant use of two hands on the interface has a higher value in the 2D than in the 3D condition (cf. Fig. 7.13). Further interesting remarks were made about the two-handedness possibilities: When asked about whether they saw any advantage in using the proposed interaction technique in comparison to a mouse, 6 participants named the possibility of using two hands as a main advantage. This indicates that despite the actual simultaneous use of two hands being limited in percentage, the freedom of alternatively and freely using both hands was perceived as an added value. At the same time, such a possibility was perceived as unusual by some participants, who commented: “People are used to interact with the mouse. You have to get used to interact with two hands on a table”; or “It would be good to have the scrollbar on the right hand, for right-handed people”. In this sense, it seems that for some people the manipulation of digital media with both hands is somewhat atypical and requires training and a familiarization process.

Finally, when considering why the graphical interface is fun to use, interesting aspects emerged, participants citing such factors as: "It is more natural to interact directly with your hand than with a device"; "With your hand you are directly on the image, the tool is too far away from it"; "You need to get used to a device, you have a better control with your hand directly"; "When you interact with the tool you don't have the feeling 'on your finger tips' of where the scrollbar ends, as with the graphical tool".

These comments raise some interesting questions about the subjective perception of directness, control, haptic feedback, discoverability, and ease as well as enjoyment of use, especially when the purposed interactions are not merely linked to models of efficiency and performance. Aspects of ease and enjoyment of interaction, for example, do not appear to be necessarily causally related. These results imply some reflections on the value and benefits of interaction at both a pragmatic and conceptual level: These considerations are articulated below.

7.3.8 Takeaways and Implications

Having presented some vignettes of action and grounded them in details of observed common practice, it is germane to discuss the implications of these observations for a discussion of tangibility in interface design. The appropriation of an experimental methodology allowed for informing a critical inquiry of tangibility by forcing users into making comparative use of two functionally similar but fundamentally altered interfaces. The choice of forcing this comparative evaluation with a direct touch enabled GUI has supported - perhaps more explicitly than in past studies (e.g., (Patten and Ishii, 2000)) - an exploration of the precise effects of pushing an interface into a 3rd dimension.

The analysis that was presented followed the common lifecycle of patterns of interaction at the interface during a photo browsing and manipulation tasks: flowing from the initiation of contact, through browsing through piles, selecting individual shots, and then moving the lens onto new piles and iterating the process.

From observing actions in each of these common stages of interaction, three key aspects of activity emerge which deserve further consideration, i.e.: i) Idiosyncracy of action; ii) Concomitant bimanualism; iii) Sequential action and laterality; iv) The "feel" of the transducer. Each of these issues has implications for an investigation of tangibility, as discussed below.

Idiosyncracy of Action

First of all, it is worth considering the standard deviations presented in Fig. 7.13. What can be seen here is that individual actions were often highly idiosyncratic regardless of the interface (3D or 2D) that participants used. Even in a first stage of analysis - considering the initiation of interaction - participants clearly approached the task (bodily) in different ways. Some users understood that piles of pictures could be dragged towards themselves, while other people moved the tool to the digital objects of interest. Those conducting this latter form of interaction were potentially demonstrating an existing mental model. This was perhaps created from years of using WIMP interfaces, where the fundamental paradigm is to manipulate an interceding tool and take that to the icon of the objects of interest to select a mode (such as using tools mediated by the mouse pointer in the Adobe Photoshop environment). This is as opposed to bringing artifacts of interest to the tool of use, such as might happen in the real world (considering uses of “examining” or “framing” tools like microscopes). Nonetheless, such patterns of interaction at the interface were not strictly consistent across all subjects, although this is perhaps to be expected with such an open interface and such a relatively open task (in terms of how it should be conducted).

This idiosyncracy of action has two implications. First, it highlights the issue of “discoverability” of the interface, that should make us reflect upon the benefits in terms of intuitiveness which are claimed in some of the literature on TUIs (Ishii and Ullmer, 1997), (Klemmer et al., 2006). On the one hand, some participants mentioned a sense of immediate understanding of how to physically manipulate the 3D handle, thus confirming those claims. With the 2D graphical handle, instead, several users thought (incorrectly) that they needed to keep a finger on the table in order to hold the lens in place. On the other hand, despite the fact that the 3D handle could be picked up and placed onto another region of the tabletop, some users dragged it on the surface in a way similar to dragging a mouse, or the 2D PhotoLens. This suggests that even if an interface is designed to incorporate a 3rd dimension, there is no guarantee that all users will appropriate it as the designer intends them to: Hence, so some of the expected performance benefits will not materialize for all users. This implies that consideration be given to possible existing mental models of interaction with a physical tool in 2D (e.g., a mouse).

Second, however, this observed idiosyncracy potentially implies that one should design for conflicting user preferences. In this open scenario, with a less constrained study task than in some previous experiments (Leganchuk et al., 1998), users showed to adapt their use of the interface to suit factors such as comfort (for example, the one-handed interactions in which the non-dominant hand was rested on the border of the tabletop). If this is how users are going to casually act with leisure technologies, perhaps designers should in the future be less concerned with the a priori shaping of the minutiae of interaction (such as appropriate handed interactions). Instead, one should more actively consider designing tangible elements that can be appropriated in various and personally defined ways by the user.

Concomitant Bimanualism

This form of interaction refers to users using both hands simultaneously to operate the interface. Relatively speaking, this did not happen that much: However, when it did happen, it was more likely to occur in interactions with the 2D interface than with the 3D interface. The reason given for this by the users appears to center on mistaken mental models of the operation of the 2D interface. Some of the users really felt that if they took their left hand away from the surface, the lens would move (similarly to a physical sheet of paper and contrarily to what they were shown) or disappear. Conversely, for these people the physical handle of the 3D interface held some form of object permanence: Once placed, the 3D handle was comfortably left alone.

Here, then, the choice of performing a comparative analysis has been particularly beneficial. Without the comparison with a 2D interface one would have been left with a poorer understanding of the effects of using a 3D handle, seeing sequential actions during its use and assuming that this was entirely user-comfort driven. From understanding the bimanual response to the 2D interface, one can see that an implication of building into the 3rd dimension - beyond apparent user comfort - is that the inherent substantiality of a 3D interface control creates assurances of consistent action. A benefit of 3D elements is possibly therefore that they suggest to the user a more consistent manipulation than a comparable 2D interface, and hence a form of control of persistence of action.

Laterality and Sequential Action

The previous study presented in Section 7.2 had indicated that users of such 2D interfaces for interactive surfaces utilize one-handed interactions. Thus, one could have assumed that the lateral interface design of the PhotoLens would promote a lateral division of handed interactions (i.e., the left hand operating the left elements of the interface and the right hand manipulating the right one). For most of the participants, this was exactly the pattern of behavior found, particularly when they were using the 3D interface rather than the 2D one (very few participants manipulated the control wheel with the dominant hand in the 3D condition). So, in this respect, such a design solution worked as expected, and one could confirm that the introduction of a tangible 3D element to the interface appeared to support the lateral division of handedness, promoting bimanualism (albeit sequential rather than concurrent). Such a sequentiality of action was possibly due to the fact that the task did not really require two-handed interaction in order to be accomplished. The potential performance benefits that could have generated by keeping the non-dominant hand in a “home position” (Buxton, 2007a) on the tool were therefore not particularly relevant in such an open task for photo organization, where issues of postural comfort were probably more important.

The “Feel” of the Transducer

Building on the last remarks, it should be considered that much previous work discussing the benefits of tangibility has taken a more engineering led approach to the evaluation. They have considered metrics of performance such as speed and task completion, and in this respect some of the questionnaire results of this study may concur with their findings. The subjective responses from the participants indicated that there were many performance benefits with the 3D interface, in terms of ease of use for example: However, this critically conflicts with their perceived preference for the 2D interface, which they found more fun to use. It is the reasoning behind this that is of particular interest here. It appears that certainly, for some users, there was a significant increase in the perception of direct engagement with the 2D interface. Contrary to the expectations that tangibility and 3Dimensionality enhance physical engagement with digital information - as it was also suggested by the informal trials of the Learning Cube and the Time-Mill Mirror - these results may suggest that a physical transducer can perhaps, in certain cases, create a perceptible barrier between user and data.

From testing the design of the PhotoLens, one can derive that if the 3D elements of an interface are not deeply considered, they can unfortunately all too easily traverse a hidden line into becoming just another tool for mediating action at the interface, another form of “mouse”. In this respect, one can speculate that the “sketchy” appearance of the tool can potentially affect such a subjective experience, suggesting a “techy” look which doesn’t suit the feeling of casualness that leisure technologies are usually meant to support. Hence, the level of direct engagement between user and digital artifact can be lower with a physical transducer than that found in direct touch enabled GUIs; consequently, it seems, this can impact users’ enjoyment of use.

7.4 Summary and Discussion

This chapter has reported on two studies investigating the implications of transferring some of the qualities of physical interaction into the design of hybrid ones, as shown in Fig. 7.24. The first study compared manipulation of artifacts in a completely physical condition to a hybrid one, where such artifacts were represented in 2D on an interactive tabletop for multi-touch interaction. Despite two-handed and multi-finger interaction capabilities being enabled in the digital condition as well, in such a modality the study showed users' tendency to one-handed interaction, and symmetric actions when bimanualism occurred. The video analysis indicated that this might be partially due to the lack of some of the physical affordances of the physical condition in the digital one, such as the physical constraints and 3D space of manipulation: These allow, for example, for shifting multiple physical puzzle tiles at once on a 2D plane; or for picking up and relocating multiple artifacts from a region to another one of a 2D plane; as well as proprioceptive feedback (see second column in Fig. 7.24). For achieving similar results in contexts of hybrid interaction one needs, hence, to design ad-hoc vocabularies which go beyond literal metaphors of physical ones: e.g., clustering and grouping gestures, such as in the case of the Brainstorm interface (cf. Section 6.3.2).

Additionally, these observations suggest that some of the functional affordances of the digital condition, e.g. two-handed interaction, were not correctly perceived or fully exploited: Put differently, the cognitive affordances for manipulation that were designed by mimicking the physical artifacts were not sufficient to suggest the same kind of manipulation behaviors. One possible explanation for this is that the acquaintance with mouse-like interaction paradigms affects such a perception and mental model: This implies design choices that more explicitly create cues and tasks so as to afford the kind of two-handed cooperative work by which we normally manipulate physical artifacts, and which has shown to provide benefits in several contexts (e.g., (Leganchuk et al., 1998)).

Drawing upon these considerations, the second study compares two hybrid interactions which are different with respect to the 3rd dimension and the type of haptic feedback (proprioception vs. somesthesis, i.e. active vs. passive) they provide (see Fig. 7.24). The design of a 3D physical interface for the PhotoLens system did suggest asymmetric use of hands, as it is normally the case when we manipulate physical artifacts, but conversely to the expectations the bimanualism that occurred was alternate rather than concomitant. On the other hand, concomitant use of two hands was more frequent with the 2D version of the interface. The comments of the participants reveal mismatching mental models of the behavior of the digital graphic lens:

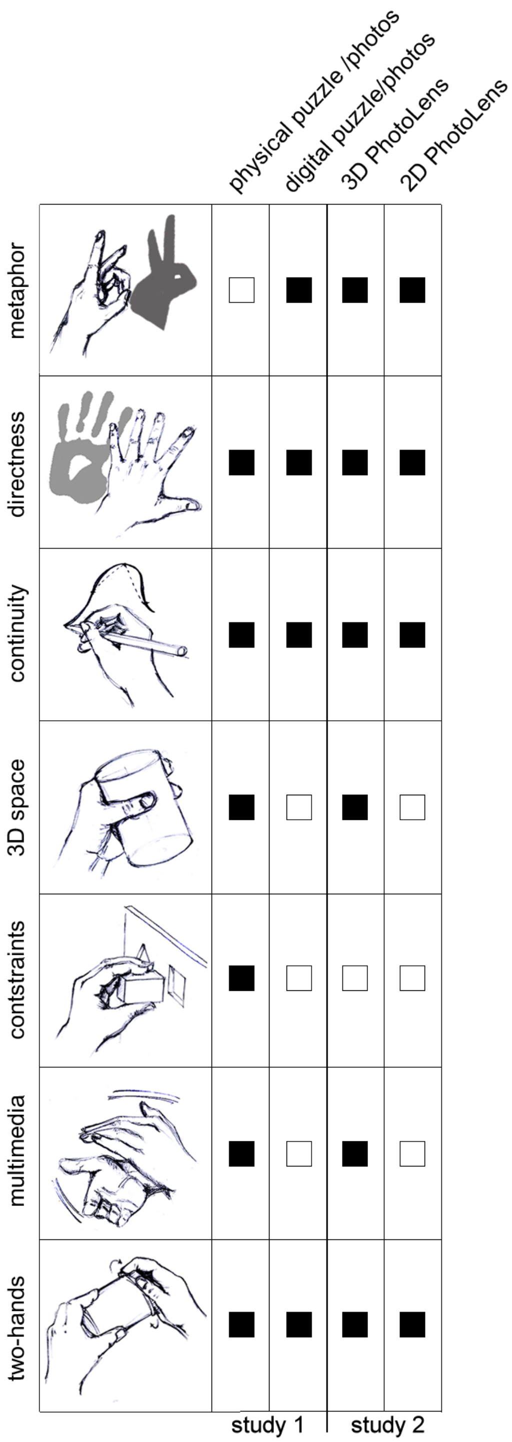


Figure 7.24: Physical qualities in the study designs presented in this chapter.

Like when holding a piece of paper in the action of writing, users kept the left hand on the digital lens for the concern it would move otherwise. Such mental models obviously affect the interaction patterns shown by the participants and should make us reflect upon the meaning of intuitiveness which is normally associated with direct touch and tangible interfaces: Although the 2D and 3D interfaces were designed to suggest a style of use (i.e., concomitant two-handed interaction), during the study (which essentially represents users' initial explorations of such interfaces) many did not use the interface as intended. Some failed to discover for themselves the prompted scheme of interaction.

These findings reveal that 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 to understand these differences, both in terms of media (as in the first study) as well as the tools we design for their manipulation (as in the second study), especially as more and more artifacts in our everyday life assume a digital instantiation (e.g., photos and documents). These changes, together with the advances in interactive display technologies, call for the design of novel ways of manipulating, sharing and integrating those artifacts with other existing ones. The Ubicomp agenda claims to research more natural interaction techniques than the one enabled by the universal desktop metaphor, but how will people really understand, learn, and experience these novel interaction possibilities? And what is a "natural" interaction in a first place, in which context and for whom?

As designers, we are called to contribute to a novel understanding of hybrid interaction and to the creation of design principles for it. To answer those questions, 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 possibilities of digital media. The studies presented in this chapter show a possible approach to the problem. Through the design of interactive systems which consciously combine physical and digital affordances - and the systematic evaluation thereof - we can then learn about people's interaction schemas. To this end, we need to investigate what the very differences, benefits and trade-offs of physical and digital qualities in the interaction actually are, and how they affect the user experience in different contexts. Which solutions provide the best mental model for bimanual cooperative work? Where shall we draw the line between graphical metaphorical representation and embodiment of the functionalities in a physical tool?

In this respect, a research agenda pursuing comparative design and evaluation can potentially contribute to a deeper understanding of human interaction behaviors: This is possible if we design comparable solutions which tackle specific aspects of the interaction (e.g., physicality and tangibility) and, at the same time, provide experiences which are open for people's expression of preferences, relating to realistic everyday life scenarios (e.g., photo browsing). In this way, the perception of engagement and control, for example, in relation to the look and feel of the interface, can be further unpacked and understood in specific contexts of leisure technologies. The work presented in this chapter has considered these issues.

Different ways of designing for diversity and comparison of alternatives have all in all been presented in this thesis, as it is further discussed in the next concluding chapter.

8

Summary and Conclusions

This chapter provides a retrospective analysis of the process and approaches that were undertaken in this thesis. Thus, it summarizes the outcomes of such a process and elaborates them to extrapolate some main remarks on methodology and design. Finally, it critically discusses the thesis contribution and its implications on future work.

8.1 Summary of the Process

This thesis has explored the possibilities and analyzed the implications of integrating aspects of physical interaction in the design of hybrid interactions for direct input. The goal was to facilitate skill transfer across physical and digital environments in specific socio-physical contexts of interaction. In this respect, a particular focus was set on the pragmatic level of the interface so as to understand how interface design can possibly solicit the transfer of humans' manipulation skills and associated conceptual models (e.g., rotation of a screw clockwise pushes the screw downwards) from physical to hybrid interaction (e.g., rotation of a graphical wheel clockwise enables a close-up of an image). In order to suggest such a transfer, the design of the interfaces presented in this work tended to maintain a semantic continuity of the representation and interaction vocabulary of the transducers across physical and digital conditions (see Fig. 8.3). Hence, physical, cognitive, functional, sensorial, and social affordances were distinguished (cf. Section 2.7), in order to recognize the different ways in which designers could borrow from the physical world. Furthermore, some of those designs were evaluated in situ, thus eliciting the implications of the social and physical contexts on the perception of those affordances and on users' expectations.

The process was largely explorative: Starting from a distinction of some of the main qualities that characterize physical interactions (cf. Section 1.1.2), these were used as framework for analyzing the related work. Building on such an analysis, which took into consideration different types of devices such as interactive surfaces, physical objects, and interactive environments, two main design approaches were conceived of: The first addressed the design of graphical interfaces for manipulation of digital media in 2D, on interactive surfaces, either with fingers or with a stylus as transducers. In this approach, the gesture vocabulary for the hybrid context of interaction borrows from physical artifacts in order to suggest (visually) the gestures to be performed on the surface, thus providing cognitive affordances (cf. Section 4.2). A second approach (cf. Section 4.3) integrated 3D physical objects in the interaction vocabulary, so as to refer to and exploit their manipulation vocabulary and associated conceptual models in existing systems and stereotypes (e.g., rotating a wheel clockwise can move forward in a timeline, like the wheel of a pulley moves back and forward an analogue film; or a knob for radio tuning moves back and forward on a frequency scale, or up and down in a volume scale). In other words, "idioms" of physical manipulation (cf. Section 2.6) were used to provide cognitive as well as physical affordances, and the sensorial affordances that derive from these (i.e., the proprioceptive haptic feedback generated by physical forces).

These two approaches were instantiated in different ways for different projects, implemented in interaction sketches as well as several experience prototypes targeting different domains. Through this process, the interfaces of those experience prototypes were designed by analogy to the artifacts of the specific domain they were meant for: e.g., a cookbook for a kitchen appliance; a mirror for a domestic environment; a photographic lens for a graphic design domain; Post-its for a brainstorming system. Furthermore, different transducers were considered, such as single input with a finger, a stylus, physical objects, as well as multi-touch; and different types of simulating environments were created. Such a diversity of domains was intended to provide an understanding of how computing technologies and the hybrid interactions they enhance can be perceived in different social and physical contexts, which do not target traditional office work and multi-purpose interfaces. Thus, those experience prototypes were evaluated in order to understand how design choices relevant to the pragmatic level of the interface were understood, and how people appropriated those interfaces. Additionally, the experience prototypes served as kinds of probes or test-beds to assess users' subjective experience of use so as to gather an understanding of the mutual influence between people's expectations and their further requirements for technologies in those specific domains.

Elaborating on the outcomes of those evaluations, two controlled comparative studies were designed and conducted: These sought to pin down which elements of physicality, and how these elements, affect different instantiations of leisure interaction such as photo sorting, both because of the real affordances as well as the perceived ones they provide. In this way, it was possible to observe more systematically the implications of integrating some specific aspects of physicality that were indicated as critical by the evaluation and analysis of the experience prototypes.

The results of this investigation on the implications of embedding aspects of physicality in the design of hybrid interfaces for specific contexts of interaction are summarized in the next section.

8.2 Summary of the Results

This thesis explored the integration of physical affordances in context-specific space-multiplex interfaces which could be either graspable or direct touch graphical UIs (cf. Section 1.1). A main difference between mouse-based GUIs in the WIMP paradigm was the characteristic of direct input, which allows for a reduction of the interaction phases (see Fig. 1.2 in Chapter 1).

The two approaches that were conceived of (i.e., manipulation of 3D vs. 2D interactive artifacts, as summarized in previous paragraph) can be mapped to the distinction between graspable and direct touch UIs, as illustrated in Fig. 8.1.

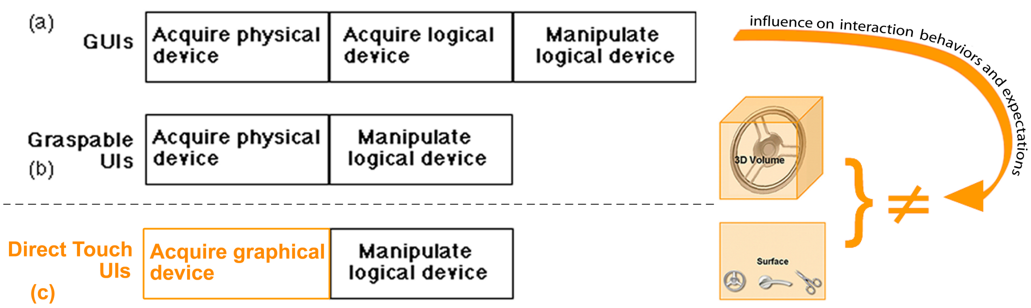


Figure 8.1: Summary of results.

Referring to those different types of interfaces, the investigation that was undertaken in this work elicited a number of observations, namely:

- Despite the integration of affordances from the physical world, the way in which people interacted with the type of hybrid interaction paradigms that were designed and explored in this thesis (i.e., Fig. 8.1, b and c) suggests the persistence of mental models and expectations coming from previous interaction with the WIMP paradigm in a desktop PC environment (i.e., Fig. 8.1, a).

- Such an influence is elicited both by the observation of users' interaction behavior at the pragmatic level of the interface, as well as by their expression of expectations and subjective experience of use. In this respect, the situated evaluation of some experience prototypes indicates that people also have a preconception about the physical location of computing technologies, as well as the type of activities they should support and the functionalities they should provide.
- Furthermore, despite the fact that both approaches to the design of direct, space-multiplex interfaces (i.e., 2D vs. 3D interactive artifacts) sought for semantic continuity of the transducer across physical and digital conditions, the effects of such approaches on users' interaction behaviors and subjective experience of use were quite different between the two.

These differences are discussed in further detail in the next paragraphs.

8.2.1 Manipulation of 2D Interactive Artifacts

The observations of how people interacted with digital media mimicking physical ones on a 2D surface showed that, in several cases, despite the user understanding the action to perform at a conceptual level, the way s/he operated the interface at a pragmatic one was not consistent with the metaphor of the manipulation vocabulary.

The way people directly manipulated 2D interactive artifacts, either with fingers or with a stylus, suggests that people's approach to digital media and their interaction with it on a screen are affected by previous interactions, not only with physical technologies, but also with digital ones in the WIMP paradigm. This is exhibited in particular in the following patterns of interaction:

- Discrete actions: Despite the fact that the interface was designed for continuous actions, this was not always understood and participants sometimes expected that discrete actions would trigger an automatic behavior of graphical elements. For example, some people tapped instead of rotating the mug in the Mug Metaphor Interface (cf. Section 4.2.2); some tapped rather than dragging the lever of the graphical dial in the GUI of the Living Cookbook (cf. Section 5.1.4).

- One-handedness: A tendency to one-handed interaction was recognized even in those cases in which spatially multiplex input was actually enabled. This occurred, for example, in the study comparing the manipulation of physical vs. digital media on a tabletop (cf. Section 7.2.4); as well as in the study comparing a physical vs. a digital handle for manipulation on an interactive surface (cf. Section 7.3.7).

- Symmetric bimanualism: In the cases in which bimanualism occurred, this showed spatial-temporal patterns which were quite different from the ones which are normally emerging in physical interaction (e.g., in Guiard's two-hands kinematic chain (1987)). In those hybrid interactions, when two hands were used in a concomitant way, their action was mostly symmetric (e.g., cf. Section 7.2.4) or anyway it revealed a mental model inconsistent with the one prompted by the interface (e.g., when users thought they should hold the 2D PhotoLens with one hand to avoid its movement, cf. Section 7.3.7).

Further observations could be made when users were provided with a pen as transducer. On the one hand, such a transducer does not allow for the design of the kind of manipulation vocabulary that two hands and multiple fingers are capable of. On the other hand, though, it appeared as if users had a clear mental model of how such a single-input transducer could be used. In this respect, it was interesting to notice how such mental models blend the use of the pen in the analogue world with its use in the digital one. For example, in the interaction with the EnLighTable, which targets a design domain, free scribbling was desired and suggested as is possible in the physical environment (cf. Section 6.2.3). In the interaction with the Brainstorm system, users seamlessly switched from a semi-analogue gesture vocabulary on the table (where handwriting was possible) to a different ones for clustering, moving multiple artifacts at once, as well as ad-hoc gestures, such as crossing-off to delete. In this cases, the transducer was used in ways which are obviously not consistent with its counterpart in a purely physical context (cf. Section 6.3.5). Furthermore, in the Photo-Slider browsing technique (cf. Section 5.2.3), the participants expressed their expectations for a richer functionality of the pen, which was both related to physical technologies (i.e., handwriting and annotating) and to computing ones (i.e., selecting and moving). In other words, different domains differently affected users' expectations and mental models of the transducer's functionality.

Considering users' responses on their experience of use, the manipulation of digital media on direct touch interactive surfaces felt in general more "natural" than the manipulation supported by a spatially detached tool (e.g., the mouse), as indicated in several stages of this inquiry. For example, during the evaluation of the EnLighTable (cf. Section 6.2.3), this interaction technique was estimated as beneficial for the involvement of multiple project stakeholders in the creative team-work; the manipulation of the 2D PhotoLens was considered by some users as more direct, and therefore more engaging, than the one afforded by the physical 3D PhotoLens (cf. Section 7.3.7); the interaction with the stylus on the Photo-Slider prototype was felt as more accurate than the one with the wheel, thus being associated with a higher level of engagement and control (cf. Section 5.2.3). Such a "naturalness" of direct touch interaction, though, needs to be understood at a deeper level to consciously take advantage, from a design perspective, of aspects of interaction in the physical wold as resources for the design of hybrid ones for direct touch.

In this respect, it becomes necessary to unpack the meaning of *engagement* and *control*, and how these concepts reveal different nuances in the way they were perceived in the two different approaches, i.e., in the manipulation of 2D vs. 3D interactive artifacts. The direct touch of media in 2D seems to provide engagement in the sense of "commitment of attention", and mostly visual attention, because the shape of the media doesn't provide active haptic feedback. Visual attention, in turn, supports accuracy and control of interaction, especially in micrometric tasks (e.g., dragging a graphical ripple along the line of a slider as in the Photo-Slider; or dragging a point along a circular path, as in the case of zooming with the 2D PhotoLens). The interesting point here is that the values associated with such a commitment can be opposite and are partially influenced by the context of use. The interaction with the Photo-Slider, for example, was considered as more accurate than with the Time-Mill by most users, but such a commitment of attention was considered as disruptive of the flow of experience for some people, and as a positive feeling of control by others. The lack of physical constraints in the digital puzzle task of the first study, where a precise goal-driven interaction was prompted, provided for frustration of most users (cf. Section 7.2.4); on the other hand, in the second study, for some users it was the physical 3D PhotoLens to be disruptive of the directness of interaction afforded by the 2D one, which some users felt "at their fingertips" (cf. Section 7.3.7).

These results suggest that similar cognitive and physical affordances can generate different subjective experiences of engagement and control according to the task and context they are designed for. Similar considerations can be made for the results of the analysis of hybrid interactions integrating a 3D component, as discussed in the next paragraph.

8.2.2 Manipulation of 3D Interactive Artifacts

The integration of physical 3D objects in the design of hybrid interfaces appeared to provide users with an immediate intuition of how such physical handles could be manipulated: This aspect could be observed in several instantiations of such an approach, as in the Learning Cube (cf. Section 4.3.2), in the Time-Mill Mirror (cf. Section 5.2.3), and in the 3D PhotoLens (cf. Section 7.3.7). These observations were also reinforced by some of the participants' comments on those hybrid systems, mentioning terms such as "immediacy" and "intuitiveness" (e.g., cf. Section 5.2.3).

Furthermore, those interactive objects seemed to support users' exploration of the interface and of its manipulation vocabulary: Children tried out several gestures with the Learning Cube in order to see how this would affect the display of the digital output; some test users tried out different speeds, as well as spinning when rotating the wheel of the Time-Mill Mirror, so as to affect the speed of photo browsing; zooming-in and -out appeared to be more frequent with the 3D PhotoLens then with the 2D one. These observations suggest that, at least in the cases that were considered in the work presented here, when a simple physical object is used as transducer and is spatially coupled to the display of the digital output, people seem to easily develop a mental model consistent with the system's conceptual one. Put differently, one could say that when the cognitive affordances providing visual feedback are reinforced by physical ones providing active haptic feedback, people seem to develop a coherent mental model. This is supported by the observation that manipulating a graspable physical artifact was perceived as easier in most of the cases, i.e.: with the Time-Mill Mirror as compared to the PhotoSlider (cf. Section 5.2.3); in the manipulation of physical puzzle tiles and photos as compared to graphical ones (cf. Section 7.2.4); with the 3D PhotoLens as compared to the 2D one (cf. Section 7.3.7).

These observations have to be carefully treated though. Providing a physical tool as transducer does not automatically mean that people will appropriate it in the same way as designers expect them to. In the comparative study of the PhotoLens, for example, it was possible to observe that some participants dragged the physical handle on the surface of the table (similarly to a mouse-based interaction) rather than picking it up and placing it somewhere else, as someone would normally do with physical artifacts. Furthermore, despite the fact that much of the related work on interactive objects claims to support two-handed interaction (cf. Fig. 3.27 and 3.28), the analysis of the interaction with the 3D PhotoLens showed an alternate bimanualism rather than a concomitant one (cf. Section 7.3.8). In those cases, people often exploited the possibility of resting the physical tool on the table and interacted with the dominant hand only, thus showing long phases of one-handed interaction. The task, indeed, was designed such that two-handed cooperative work could generate performance benefits, even though the concomitant use of two hands was actually not necessary to accomplish the task (and performance is not the primary goal of photo sorting after all). Thus, such a pattern of two-handed interaction and associated possible benefits emerged only in a limited way. It appeared that issues of comfort were driving, to some extent, the way in which people bodily approached the task.

Similar to the reflections elaborated in the previous section, it becomes clear that the task at hand and the purpose of interaction can affect the way in which affordances are perceived. Despite the fact that a 3D physical handle suggests its interaction vocabulary, some of its affordances need to be learned through exploration and experience. And again, the experience of use can be affected by subjective values associated with the purpose of technology.

A sense of playfulness seemed to be generally associated with the possibility of exploring the manipulation and interaction vocabulary of the transducer: This appeared, for example, when children shook the Learning Cube and played together (cf. Section 4.3.2); as well as when participants commented on the wheel of the Time-Mill Mirror and span it, exploiting its momentum (cf. Section 5.2.3). The values associated with playfulness were not necessarily homogeneous though: For some users, the physical interaction with the wheel of the Time-Mill Mirror, for example, was positively engaging in the sense that it provided a sense of flow; for someone else, this became trivial and the direct interaction with a stylus provided more accuracy, as it was engaging more of their visual attention (as discussed in the previous section). For some people, interacting with the 3D PhotoLens was easier because "you don't have to look", thus supporting disengagement of visual attention, and they felt more in control when zooming: But the direct touch of the 2D PhotoLens was in general perceived as more fun because it was felt to be less disruptive and more natural. One could speculate that, in these latter cases, the physical transducer was associated more with a mouse, and therefore related to office work in a Desktop environment, rather than with a tool for leisure interaction.

These results highlight once more that in order to understand the influences of design choices at the pragmatic level of the interface on the experience of use, one needs to contextualize such an experience and look into the different dimensions and nuances of it. In particular, the work presented here has highlighted the need for a deeper understanding of the relationship and mutual implications between the "look" and the "feel" of a design, and has shown how different approaches to the design of affordances for hybrid interaction can affect the subjective perception of look and feel depending on the context of use. The different implications of such approaches are schematically represented in Figure 8.2.

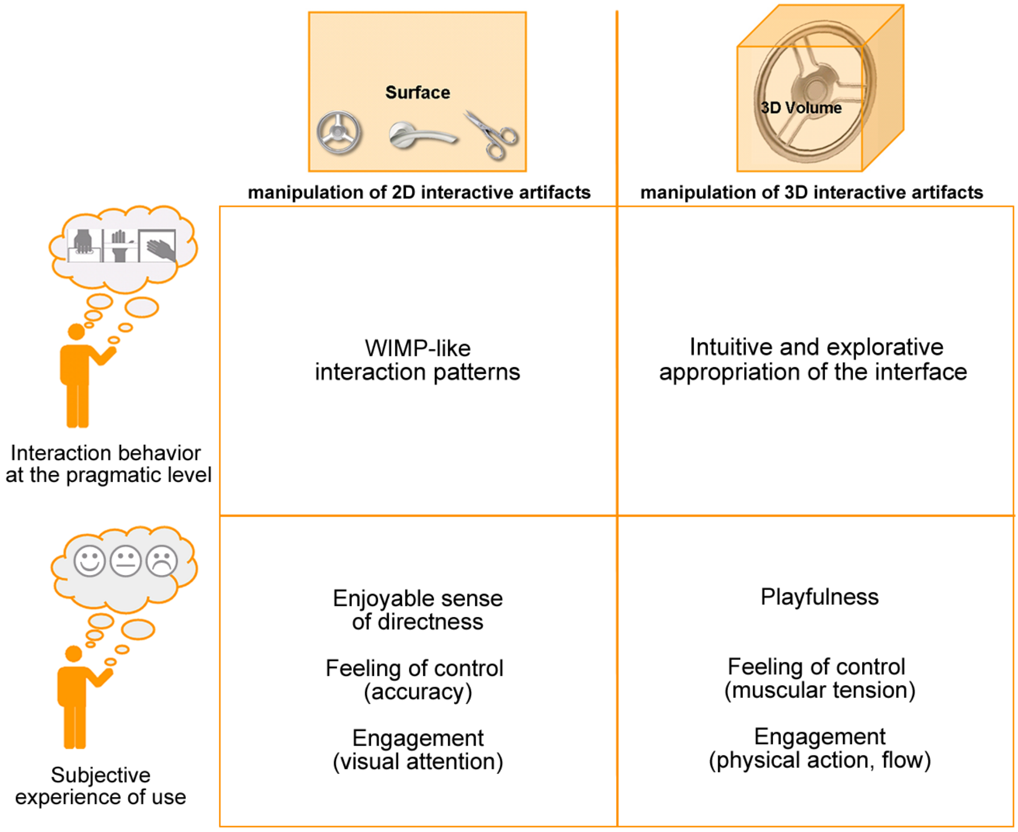


Figure 8.2: Different implications of the two approaches to the design of space-multiplex interfaces for direct input.

Drawing upon these observations one can then refine the taxonomy of hybrid interactions presented in Chapter 1, Figure 1.5, as shown in the next paragraph.

8.2.3 Refining the Taxonomy of Hybrid Interactions

By mapping the designs developed in this work to the taxonomy of hybrid interactions presented in Chapter 1 (see Fig. 8.3), one can make a number of considerations in light of the findings of the investigation that was conducted.

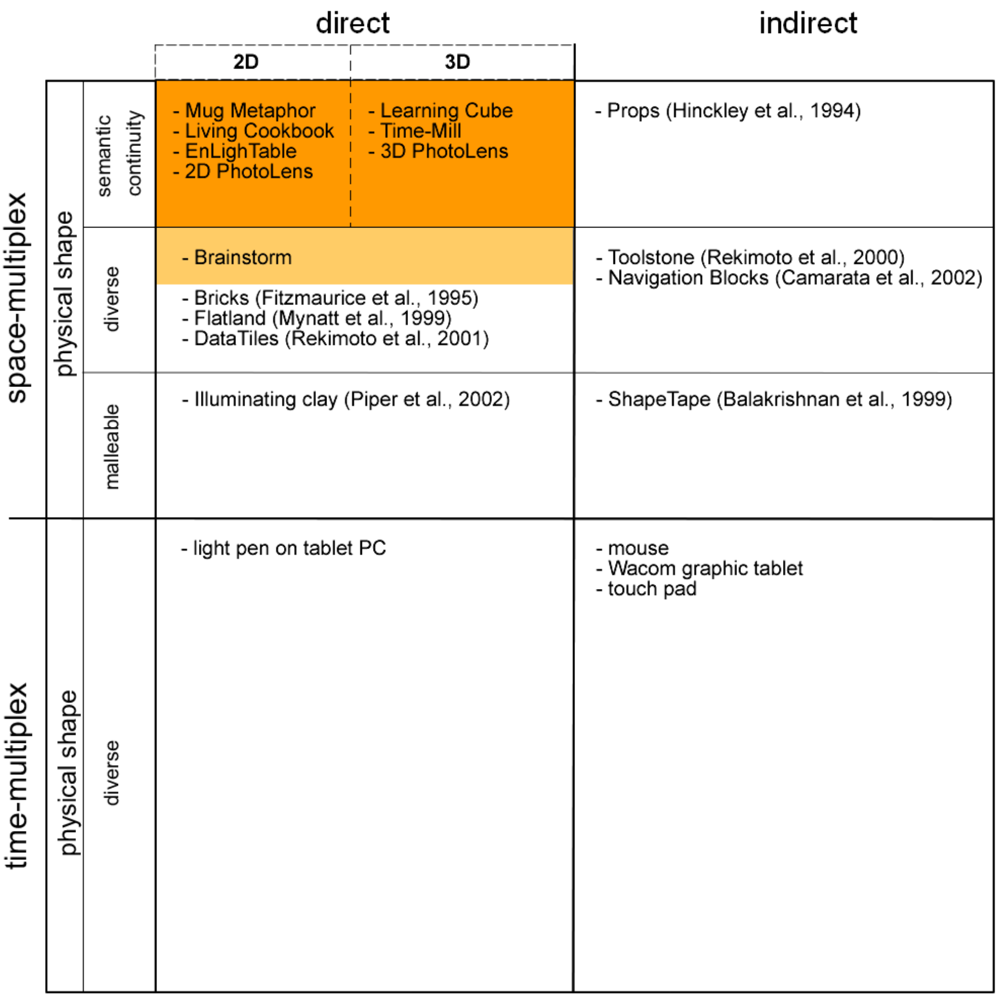


Figure 8.3: Mapping the Designs to the Taxonomy of Hybrid Interactions.

First, those findings indicate that depending whether the interfaces for hybrid interactions are 2D graphical ones for direct touch, or 3D graspable ones, different affordances can be perceived and experiences can emerge, according to the context. This suggests that this additional dimension should be considered, resulting in a more fine-grained description of the taxonomy of hybrid interactions.

Second, as one can notice in Figure 8.3, the use of a transducer such as a stylus can be considered as an example of semantic continuity in some cases (e.g., as in the case of the EnLighTable) and diverse in others (e.g., as in the case of Brainstorm system). In this latter case, indeed, the pen is used both for handwriting on the table, as well as for moving and clustering items on the wall, thus borrowing alternatively from a physical pen-like interaction vocabulary and from a light pen for direct manipulation of graphical digital media. This reinforces the claim of this thesis that for an understanding (and design) of the semantics of the transducer, one needs to carefully consider the context of use and possible ambivalent meanings of the transducer across physical and digital worlds.

In this respect, a methodological approach providing a diverse spectrum of alternatives and contexts of analysis has proved to be beneficial for the investigation of those aspects, as is discussed in the next section.

8.3 Remarks on Methodology

As anticipated, the approach adopted in this work builds on the generation of diverse contexts for elaboration and analysis of design concepts. Such a diversity is provided in a number of ways:

- **Diversity of representation.** As discussed in Section 4.4, different representation styles were shown to suit different stages of the development of design ideas. A more "sketchy" style seemed to be more appropriate for collaborative generation and elaboration of ideas, while a more refined one allowed for a more careful analysis of the way in which users appropriated the interface, and hence for the assessment of those ideas.

  An additional observation is that, when creating experience prototypes of hybrid interfaces that integrate a physical 3D object, the prototype of the physical transducer should have a level of refinement consistent with (or even superior to) the one of the graphical digital interface. Whilst graphical UIs, indeed, are not extensively manipulated with gestures such as shaking, spinning, revolving, etc., physical transducers might be. The feeling of control that haptic interfaces strive for can be achieved only if the physical prototypes convey a sufficient sense of stability and robustness. This was not always the case for the experience prototypes realized in this work. The wheel of the Time-Mill Mirror and the physical tool of the 3D PhotoLens had a more sketchy appearance than their graphical complementary parts: Some people mentioned a sense of fear of ruining the prototype in the first case, and a feeling of disruption in the second one. This aspect can potentially have affected the subjective perception of enjoyment of use and control, and should be taken into deeper account, especially in domains in which aesthetics and fun are most relevant.

- **Diversity of design domains and target users.** The design of experience prototypes for a diversity of domains provided the possibility to explore the instantiation of design concepts in a number of ways, thus gaining a more complex understanding of the implications of those design choices in different contexts. Additionally, it allowed for the elicitation of different users' needs and expectations according to different target groups (e.g., children and parents as in the case of the Living Cookbook; designers of digital and physical communication media as in the case of the EnLighTable). Furthermore, it showed how some design choices might be suitable for some domains (e.g., pen-based interaction in a design work domain, such as in the case of the EnLighTable) and less so in different ones (e.g., in a context for domestic photo browsing, such as the one supported by the Time-Mill Mirror and its alternative designs).

- **Diversity of design alternatives.** The design of comparable alternatives (e.g., of domestic photo browsing techniques, cf. Section 5.2.3) allowed for the elicitation of users' mental models of different interaction paradigms, as well as participants' self-expression of personal preferences, interaction attitudes, and suggestions too. In this respect, this approach raised an understanding of different human values and expectations in a certain context of use, and how certain design choices could suit such expectations better than others, for certain peoples' likes and dislikes.

- **Diversity of evaluation settings and methods.** Different evaluation settings, e.g. in situ (as for the Living Cookbook, for example) or in the instrumented room of a research lab (e.g., the Brainstorm system) allowed for the elicitation of different values associated with different physical and social contexts. Whilst, for example, the compatibility of a computing device in an existing ecology of artifacts was a concern in the case of a domestic appliance such as the Living Cookbook, which motivated its evaluation in domestic kitchens in addition to the trials conducted in the lab, this was not the case in the evaluation of the Brainstorm system: Here, participants mentioned that they would rather use such an electronic brainstorming system rather than the paper-based one. In other words, such an approach allowed for recognizing domains where computing technologies can be considered as more disruptive of a social and physical context rather then others, thus provoking a reflection on simulation environments for testing, and on evaluation methods. Likewise, the use of different methods of analysis, e.g. video analysis, time coding, and in depth interviews, allowed for learning about the effects of those design choices both at the pragmatic level, as well as in terms of subjective experience of use, so as to draw possible connections between those.

All in all, such an approach is coherent with a design mind-set, which builds on the generation of ideas for iterative association and elaboration of design concepts in order to recognize patterns and/or interesting points of inconsistency: that is, in order to "read" the sketches and prototypes that were designed and draw upon them to understand their implications. These are articulated in the next section.

8.4 Remarks on Design

The generative and analytical approach described above allowed for drawing some implications for the design of interfaces for hybrid experiences: These mainly address the different ways in which the look and the feel of the interface elements have an impact on the experience of use. To elaborate:

- **Look and Feel of the Media.** The way in which people bodily approach digital media which visually mimics physical ones brings with it vestiges of the WIMP paradigm, thus suggesting "hybrid" existing mental models. Nevertheless, people appeared to perceive direct input interfaces as more natural then the ones requiring a pointer, such as mouse-based interaction. Directness seems indeed to provide a higher feeling of engagement in the sense of flow, thanks to the lack of spatial disruption between the point of input and the one of output.

  Furthermore, the GUIs for mouse and keyboard interaction on a desktop PC mainly afford discrete interaction, which might have an effect on the interaction attitude people develop when approaching an interface for direct touch as well. On the other hand, continuous analogue gestures provide more opportunities for body language and visibility of action, especially in the case of shared interfaces on large interactive surfaces: This aspect, in turn, allows for people's self-expression and group awareness, as shown in the case of the Brainstorm system, for example, and should therefore be supported and enhanced by the graphical UI.

  As designers of digital media we need to consider the heritage of the WIMP paradigm and the way in which it can affect people's manipulation and, in turn, communication behaviors. Depending on the activity we want to support, this can sometimes imply the design of interfaces that explicitly suggest or force certain modes of interaction, but it can also imply the design of open interfaces, which can be appropriated by users in different ways.

The design of affordances for gesture-based direct manipulation also needs to keep in mind that with the emergence of ubiquitous computing paradigms in different domains our interaction styles are going to vary considerably, and the next users generations are likely going to be less bound to the WIMP paradigm then we are today. In this respect, metaphors that rely on the analogy with physical objects and on the way we manipulate them in the analogue world still promise to help in the design of conceptual models. The design choice of integrating physicality even more in the interface, for example with graspable 3D objects, has shown some positive effects on the "graspability" of the conceptual model, as discussed below.

- **Look and Feel of the Transducer.** When a physical object, be it a pen or a 3D tool with a different shape, conveys the physical gesture to the system, its shape has an impact on the experience of use in several interwoven ways: for example, for its visual appearance, its physical manipulation vocabulary, the haptic feedback it provides both in isolation as well as in contact with another object or a surface, and for how the coupled digital output (e.g., the thickness of a pen stroke) responds to the physical manipulation (e.g., pressure). In addition, some transducers are already embedded in a certain mental model of the way they work in the physical world, such as a pen or a wheel.

  The pen appears to be associated with both analogue interactions such as handwriting (and in this sense can be considered as a tool for self-expression, cf. Section 5.2.3), as well as hybrid interactions, such as tapping and moving, selecting, and rotating (and in this sense can be considered as a functional transducer for a direct input device, cf. Section 5.2.3). In these cases, it serves as kind of prosthesis of the human's index finger and provides the advantage of a more accurate position of the input point, thanks to its smaller tip. Additionally, as in the case of the Photo-Slider (cf. Section 5.2.3), such an accuracy and the visual attention required by the tool are associated in some cases with working or goal-driven activities, possibly because of the cognitive commitment implied by micrometric tasks.

  In these cases, we need to think thoroughly about the interaction vocabulary we design for such kinds of transducers. When carefully designed, hybrid interactions can take advantage of both the types of functionalities associated with a physical pen and with a stylus for direct input on computing devices. In this respect, the types of activities we are targeting also need to be considered (e.g., goal-driven vs. casual).

When the transducer is embodied by a physical object, its visual shape suggests the way in which it can be physically manipulated, but the mapping between physical action and digital behavior is something people find out only during the interaction itself, and in the way they explore and appropriate the physical object. An object provides physical affordances that as designers we can draw upon for digital interactions, but this does not necessarily mean that such affordances are perceived as we expect: This is the case, for example, of the 3D PhotoLens, where the possibility of picking it up for positioning it onto other photo piles was used only by some test users (cf. Section 7.3.7). Similarly, the two-handed cooperative work that such a physical handle was supposed to suggest was actually limitedly exhibited by the participants of the study trials, despite the fact that it was enabled. Also in this case, then, possible existing mental models or interaction attitudes (deriving for example from the manipulation of a mouse on a desktop) might affect the way in which people approach the tool. On the other hand one, needs to consider that the mouse itself is not something we knew how to operate from the very first time we saw it: Rather, we learned how to use it over time. When designing novel transducers, then, we need to think about the way in which they will be explored and appropriated “at first sight”, but also about the way in which a longer period of interaction might reveal different patterns of use and, in turn, different subjective experiences.

- **Look and Feel of the Device.** Finally, as the vision of Ubicomp promises to bring computing technologies within the fabrics, furniture, and on the walls of mundane environments, the look and feel of the device need to be considered in relation to the activity, social, and physical context the device is intended for. This means that, as designers of hybrid experiences, we can’t neglect the appearance of the hardware supporting the interaction we design for: In fact, such appearance must be compatible with users’ aesthetic values because this is going to influence their subjective experience of interaction.

  As seen in the case of the Time-Mill Mirror, for example, the subtle border between a mirror and a tablet PC - and the purpose and location these artifacts normally have - must be thought through in the design of the device: This aspect can affect users' perception of the affordances of the device (e.g., the transducer's interaction vocabulary) and the values associated with the experiences it enhances.

Similarly, as we expect computing technologies to be embedded into existing furniture such as tables, designers need to think thoroughly whether they will be perceived as computer screens horizontally positioned, or rather as augmented physical tables. In this respect, the type of activity we aim at supporting (e.g., casual vs. professional, individual vs. social), as well as the physical context the appliance is designed for (e.g., a living room vs. an office room), are going to affect users' feeling of engagement. In other words, we need to consider that existing ecologies of artifacts and social activities can affect the perception and acceptance of computing technologies, especially in those physical and social contexts in which computers are normally absent or have a different appearance (e.g., the desktop PC).

8.5 Outlook

This work provokes a reflection on the design issues and assessment criteria of hybrid interactions in contexts of everyday life, beyond the office. As discussed by Grint and Woolgar (1997): "Technologies, do not "by themselves' tell us what they are or what they are capable of. Instead, capabilities - what, for example, a machine will do - are attributed to the machine by humans. Our knowledge of technology is in this sense essentially social". This implies an approach to the design and study of hybrid technological artifacts that considers their social context of use, and, as it was claimed and shown in this thesis, their physical one as well.

The integration of computing technologies into mundane spaces and artifacts for supporting leisure and social activities is still in its infancy. Although existing work on interaction design provides a rich background for informing the interface design of such novel hybrid artifacts and spaces, their specific social and physical settings require a more systematic analysis of the implications of our design choices on users' interaction attitudes and subjective experience of use. In other words, the users' values and requirements we design for in the domain of PC office work can be very different in other domains: The acknowledgment of such differences is fundamental for guiding the design of hybrid systems which aim at augmenting humans' social, cognitive, as well as operational capabilities in various contexts of everyday life. This implies that criteria of assessment need to be defined for the specific context of use, so as to understand the implications of our design choices in those domains.

This, in turn, means that the benefits of integrating aspects of physicality in the design of hybrid artifacts need to be considered in relation to the specific values and activities we are targeting at. In this sense, there has been limited theoretical research that deeply looks at these issues 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? For what kind of experience?

This thesis has proposed an approach of analysis that promises to contribute to a more systematic evaluation of the effects of different aspects of physicality in different types of hybrid interactions; Hence, it supports further investigations along this line. Additionally, this thesis presented a taxonomy of hybrid interactions others can build on, and further refine in its different areas by leveraging an understanding of the implications of those and/or other dimensions.

The evaluations undertaken in this work observed interaction behaviors that occurred in a precise point in time, when people had their first contact with the interface: Thus, the evolution of people's operational skills with such novel hybrid interactions was not observed. Future work should consider the evolution of users' learning curves, as well as observe more closely computer supported collaborative interaction amongst users with different skills. These aspects can potentially affect the subjective experience of use in those domains for casual interaction, thus creating implications for the assessing criteria and focus of evaluation.

In this respect, the evaluation of user experiences implies an understanding of the different aspects which constitute an experience, and the identification of the critical factors for each of those. The HCI research community is just beginning to develop such an understanding, and design can be a powerful means for facilitating this process by generating alternative solutions: Evaluation, indeed, is normally about comparison, but instead of it being with the aim of selecting the best or optimum solution and relative experience, it could instead be used to support differentiation, a proliferation of alternatives. By acknowledging diversity, in turn, one needs to acknowledge that there are elements of aesthetic experiences that potentially cannot (or should not) be evaluated because they deal with personal values, likes and dislikes that, as designers, we can target but not judge. The contribution of this thesis is to be considered in these terms, as discussed below.

8.6 Inspirations for Design

This thesis has generated and presented a diversity of ideas, solutions, and approaches for a diversity of experiences. Diversity and complexity are the sources for different interpretations, which in turn lead to creativity, discourse, and evolution. In this sense, one can see the possibility for interpretation offered, for example, by the sketches and experience prototypes that were presented and discussed as one of the main contributions of this thesis, and as a characteristic contribution of a design perspective.

The meaning of the comparative evaluations that followed the generation of diverse alternatives, in this sense, was not the assessment of success of design solutions per se, but rather the discovery and understanding of the factors that affect the user experience. For this reason, empirical and explorative approaches were combined in the attempt to recognize patterns which can shed light on relationships between design solutions and resulting experience, so as to inform the design of hybrid systems. In this sense, the specificity of the design contribution to HCI is considered in relation to other disciplines.

The hope, here, is that by "reading" those sketches, and seeing how they were considered in this work, others can be inspired and build on them, transferring some bits in different contexts. From this point of view "inspirations for design" (Terrenghi et al., 2006b) can be then seen as an essential product of mutual fertilization amongst and between the disciplines that make up the Ubicomp research community. In this respect, one needs to acknowledge that the way in which we read sketches and experiences draws on intuitions and underlying personal attitudes, that in turn differentiate us, as researchers and designers, and finally as individuals. Thus, when moving from inspirations to more systematic observations and formal approaches of analysis, one needs to keep in mind that the same alternatives could have been interpreted in different ways, like open sketches do, thus leading to different observations. It is the diversity of alternatives, indeed, that gives people the possibility to express their interests, preferences, and in turn to express their identity. The identification of characteristic properties of diverse experiences, together with the understanding of the factors that are critical to the making of each and every experience (insofar as this is possible), promise to enable a comparison and potentially an evaluation of experiences: This is what can possibly inform and inspire design.

A first step towards conscious design choices is then to learn about aspects of interaction that, at the pragmatic level, affect the user experience: This is fundamental for combining them and conveying the system image we aim at. In this sense, the design and presentation of alternative solutions to potential users have shown the diverse ways in which different people can evaluate their experience of interaction, articulate their thoughts, as well as suggest their own solutions. Furthermore, the implementation and evaluation of these alternatives in different socio-physical contexts have deepened an understanding of people's values and expectations of computers outside the office domain and the conventional PC environment: for example, in the home and in environments for collaborative creativity. In fact, this investigation has revealed the influence of previous interactions with the desktop PC, both on people's interaction behaviors, as well as on their expectations of where a computer should be, how it should look and feel like, and what it should do within existing ecologies of artifacts, activities, and collaborative processes.

Finally, while exploring and evaluating these alternatives, it is important that we, as designers, do not commit ourselves to a defined one. Rather, when using such alternatives as probes or test-beds, it is necessary to suspend judgement of "what works better" and instead try to understand "what fits better" - where, when, and for whom. In other words, designers of hybrid interactions can be thought of as cooks of fusion cuisine: as such, designers need to know the recipes for traditional dishes and learn about the taste, scent, and dimensions of different ingredients. This allows us to consciously mix them in a chemical combination that respects our creative intentions and the enjoyment of our target consumers. Their personal taste is one of the factors we need to take into account.

Appendix A

Questionnaire on the use of alternative photo browsing techniques

INTRO		
topic	question	background
	What do you think/ feel, when you use these three Photo-Shows? associations, ideas, what comes in mind....	Thoughts on the techniques, remembering the past two weeks

PHOTO-ORGANISATION		
topic	question	background
location + content	Where are your/ personal photos in your everyday life? (at home, parents, boxes, walls, computer, mobile, etc)	different photos, different meanings
	Do the contents of the photos at the different locations differ? Why?	different emotional triggers
	Why do these photos differ? (emotions, meaning, up-to-dateness, associations)	more exact description, emotional and in terms of organization
Organization and Saving versus Usage	How are these photos organized?	Type of the user, different types of organization: sequential, casual, random
	How are the photos organized in the box / on the computer ... ?	Type of the user, differences between users
	Which organization-type is used the most?	Differences between photo-organization and organization during usage
	How do you look at the different photos? (chronologically/ swap through/ together/ alone) Why? Feelings?	better definition, background of the organization during usage
	Where are the most photos?	quantity, background

USER TESTS THE TWO OTHER INTERACTION TECHNIQUES........

	FEELINGS & ASSOCIATIONS	
topic	question	background
associations	Can you associate three terms/ words on the PhotoWheel? Explanation?	free thoughts on the photo-show-technique 1
	Can you associate three terms/ words on the linear PhotoWheel? Explanation? How does it differ from the first one?	free thoughts on the photo-show-technique 2
	Can you associate three terms/ words on the linear PhotoSlider? Explanation? How does it differ to the ones before?	free thoughts on the photo-show-technique 3
feelings	Which technique is the most fun, when you look at the photos of the past 2 weeks?	first feelings/ thoughts on the three techniques
memory of place	You can see yourself on some of these photos, do you feel more/ stronger connected to the location? Does it make the feelings stronger?	representation of the own linkage to a local history
	Does this feel different with the two other techniques?	
	You can see your classmates /friends, sometimes you were there at the same day, sometimes you weren't. What do you feel? Associations?	feelings about the group
	Do you feel a stronger connection to your classmates and friends, when you see them on the photos of the past two weeks?	representation of the own connection to a group
	Does this feel different with the two other techniques?	
	Which technique causes the strongest memories? Why? Is there a connection to the photo-organization at home?	evocation of memories
	Which technique orders the two weeks in the best way? Why? Is there a connection to the photo-organization at home?	ordering of memories
	Which technique would you use at home for showing/ organization? Why?	
	Which technique would you attach at home? Where? Why?	
Randomness	What do you think about the random playback? What does it cause emotionally? Why? Comparison to the other techniques.	feelings about random playback
	What would you improve?	
Linearity	What do you think about the linear/ chronological playback? What does it cause emotionally? Why? Comparison to the other techniques.	feelings about linear playback
	What would you improve?	

FEELINGS & ASSOCIATIONS		
Physicality/ Input/ Interaction	Which feelings does the Wheel evoke? Why? Associations?	
	Which feelings does the Pen evoke? Why? Associations?	
	Which interaction-technique suits you best?	connection interaction-technique and personal preferences
	Which interaction-technique causes the most/ strongest feelings, positive or negative?	

Tech/ Hardware		
topic	question	background
intro	Which technique did you like the most?	rational and emotional thoughts
	How did you like the different designs? Why?	rational and emotional thoughts
conclusion on interaction techniques	What do you think about the three types of photo-show?	rational and emotional thoughts
	Which interaction-technique does distract you from the photos the most? Why?	interaction-technique between user and content
	Which display-technique does distract you from the photos the most? Why?	display-style between user and content
	Which interaction/ display-technique does feel natural? Why? Associations?	
	Which interaction/ display-technique "belongs" to the photo-content?	connection content and technique
	Do you think there is a connection between these photo-show-techniques and the evocation of memories?	type of feelings
conclusion	What do you think about "leaves of memory"? Feelings?	emotional effects caused by look of device
	What do you think about the leave-shapes? Feelings? [...]	new thoughts
	Is there something we forgot?	

Appendix B

Questionnaire on the use of the EnLighTable

How many people are working in the agency?

What kind of products do you work on (advertising for press, web, external communication, corporate identity, exhibitions…?)

- Welche Bereiche eines beispielhaften Projektes bearbeiten Sie im Team?
  Which parts of a project are usually worked out in a team?

How does picture selection take place during a project? What are criteria of selection (rights, resolution, layout, size…??)? Do you agree on a conpet/parameter before (e.g. colors, mood, subjects…???)

Wie viele Mitarbeiter nehmen an der Teamarbeit Teil?
How many people are usually involved in the teamwork (max and min)?

What are their professions?

In the early stage of layout design, how many people are working simultaneously on the task?

Do they communicate with each-other? If yes, how?

- Wie ist der Workflow einer solchen Teamarbeit?
  What is a typical workflow of your teamwork?

- Was geschieht dabei sequenziell und was parallel?
  What parts happen sequentially and what parts happen in parallel?

- Sehen Sie Verbesserungspotential in Ihrem Workflow?
  Are there some points in your workflow, which you would like to change or improve?

- Können Sie sich eine Veränderung ihrer Prozesse durch den Einsatz eines solchen Systems vorstellen?
  Do you imagine a change in your daily work by using such system?

- Können Sie sich vorstellen, dass diese Veränderungen Ihre Arbeit positiv beeinflussen?
  Do you see improvements by the use of these innovations?

- Wie könnte ein veränderter Workflow Ihrer Arbeit durch den Einsatz des Systems aussehen?
  How could be a theoretical workflow by using such systems?

- Sehen Sie sogar neue Möglichkeiten in der kommunikativen Zusammenarbeit?
  Do you see new possibilities for creative teamwork?

- Was würden Sie aus Ihrer Sicht bei dem Einsatz eines solchen Systems als positiv bewerten?
  Which things do you like on the prototype?

- Was würden Sie als negativ bewerten?
  Which things you don't like or would do different?

Appendix C

Questionnaire on the use of Brainstorm

Evaluierungsbogen – post

„Elektronisches Brainstorming in instrumentierten Umgebungen“

ID [______]

Angaben zum Brainstorming Tool *Fluidum BrainStorm*:

11. Wie hat Ihnen *BrainStorm* insgesamt gefallen?
überhaupt nicht ☐ ☐ neutral ☐ ☐ sehr gut ☐

12. Denken Sie, dass Sie mit *BrainStorm* mehr Ideen gefunden haben als mit Papier?
viel weniger ☐ ☐ gleich viel ☐ ☐ viel mehr ☐

13. Denken Sie, dass Sie mit *BrainStorm* bessere Ideen gefunden haben?
viel schlechter ☐ genauso gut ☐ ☐ viel besser ☐ ☐

14. Denken Sie, dass Sie mit *BrainStorm* mehr kommuniziert haben?
trifft nicht zu ☐ ☐ ☐ ☐ trifft zu ☐

Angaben zur Bedienung von *BrainStorm*:

15. Hat sich die Bedienung von *BrainStorm* organisch angefühlt?
trifft nicht zu ☐ ☐ ☐ ☐ trifft zu ☐

16. Wie fanden Sie die Bedienung von *BrainStorm*?
sehr schwer ☐ ☐ ☐ ☐ sehr einfach ☐

Angaben zur *PostIt-Erstellen* Interaktionsgeste:

17. Wie fanden Sie diese Interaktionsgeste?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

18. Wie nützlich fanden Sie diese Interaktionsgeste?
nutzlos ☐ ☐ ☐ ☐ ☐ nützlich

Angaben zur *PostIt-Löschen* Interaktionsgeste:

19. Wie fanden Sie diese Interaktionsgeste?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

20. Wie nützlich fanden Sie diese Interaktionsgeste?
nutzlos ☐ ☐ ☐ ☐ ☐ nützlich

Angaben zur *PostIt-Öffnen* Interaktionsgeste:

21. Wie fanden Sie diese Interaktionsgeste?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

22. Wie nützlich fanden Sie diese Interaktionsgeste?
nutzlos ☐ ☐ ☐ ☐ ☐ nützlich

Angaben zur *PostIt-Schließen* Interaktionsgeste:

23. Wie fanden Sie diese Interaktionsgeste?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

24. Wie nützlich fanden Sie diese Interaktionsgeste?
nutzlos ☐ ☐ ☐ ☐ ☐ nützlich

Angaben zur *PostIt-Kopieren* Interaktionsgeste:

25. Wie fanden Sie diese Interaktionsgeste?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

26. Wie nützlich fanden Sie diese Interaktionsgeste?
nutzlos ☐ ☐ ☐ ☐ ☐ nützlich

Angaben zur *PostIt-Bewegen* Interaktionsgeste:

27. Wie fanden Sie diese Interaktionsgeste?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

28. Wie nützlich fanden Sie diese Interaktionsgeste?
nutzlos ☐ ☐ ☐ ☐ ☐ nützlich

Angaben zur *PostIt-Beschreiben* Interaktionsgeste:

29. Wie fanden Sie das Beschreiben von PostIts (nur am Tisch)?
sehr schwer ☐ ☐ ☐ ☐ ☐ sehr einfach

Weitere Angaben:

30. In welcher Brainstorming Phase fanden Sie den BrainStorm Tisch bzw. Wand besser im Vergleich zum Papier Brainstorming? (Mehrfachauswahl möglich!)
    - ☐ Der *BrainStorm* Tisch war besser beim Sammeln der Ideen im Vergleich zum Papier Brainstorming.
    - ☐ Die *BrainStorm* Wand war besser beim Ordnen und Clustern der Ideen im Vergleich zum Papier Brainstorming.
    - ☐ Der *BrainStorm* Tisch war schlechter beim Sammeln der Ideen im Vergleich zum Papier Brainstorming.
    - ☐ Die *BrainStorm* Wand war schlechter beim Ordnen und Clustern der Ideen im Vergleich zum Papier Brainstorming.

31. Wie sinnvoll finden Sie den kombinierten Einsatz von dem Tisch und der Wand in *BrainStorm*?

sinnlos ☐ ☐ ☐ ☐ ☐ sinnvoll

32. Wie sinnvoll finden Sie die räumliche Trennung der einzelnen Brainstorming-Phasen in *BrainStorm* gefallen?

sinnlos ☐ ☐ ☐ ☐ ☐ sinnvoll

33. Können Sie sich vorstellen öfter gemeinsam Ideen zu sammeln, wenn der Computer in die Arbeitsmöblierung integriert (mit den oben genannten Vorteilen gegenüber der traditionellen Art) und als solcher nicht mehr erkannt wird?

☐ Nein, und zwar ____.

☐ Ja, und zwar ____.

34. Führen Ihrer Meinung nach die erweiterten Möglichkeiten der computergestützten Ideenfindung zu mehr und bessere Ideen? (Zum Beispiel die Möglichkeit der beliebigen Vervielfältigung, dauerhafte Speicherung und Kommunikationsmöglichkeit mit anderen Standorten.)

☐ Nein, und zwar ____.

☐ Ja, und zwar ____.

35. Was halten Sie vom computergestützten Brainstorming?

Das computergestützte Brainstorming…
☐ …mindert… / ☐ …erhöht…
…die Kommunikation.

Das computergestützte Brainstorming ist…
☐ …schlechter… / ☐ ...besser…
…als die traditionelle Art.

36. Falls *BrainStorm* ausgereift wäre und Ihnen die komplette *BrainStorm* Ausrüstung schon zur Verfügung stehen würde, würden Sie das traditionelle Brainstorming mit Papier oder das computergestützte Brainstorming bevorzugen?
☐ eher Papier Brainstorming ☐ eher Computer Brainstorming

37. Falls Ihnen in naher Zukunft in Ihrer Firma ausreichend Mitteln zur Verfügung stehen würde, würden Sie gar ein System wie *BrainStorm* (jedoch ausgereifter) anschaffen?

nein	eher nicht	vielleicht	eher doch	ja
☐	☐	☐	☐	☐

38. Wenn Sie weitere Kritik und Verbesserungsvorschläge haben, können Sie diese hier unten angeben:

_ _

_ _

_ _

_ _

Wir danken Ihnen sehr für Ihre Teilnahme an unserer User Study!

Bibliography

Accot, J. and Zhai, S. (2002). More than dotting the i's — foundations for crossing-based interfaces. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 73–80, New York, NY, USA. ACM Press.

Agarawala, A. and Balakrishnan, R. (2006). Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1283–1292, New York, NY, USA. ACM Press.

Amabile, T. M. (1983). *The Social Psychology of Creativity.* Springer Verlag, New York, NY, USA.

Apitz, G. and Guimbretière, F. (2004). Crossy: a crossing-based drawing application. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 3–12, New York, NY, USA. ACM Press.

Applegate, L. M., Konsynski, B. R., and Nunamaker, J. F. (1986). A group decision support system for idea generation and issue analysis in organization planning. In *CSCW '86: Proceedings of the 1986 ACM conference on Computer-supported cooperative work*, pages 16–34, New York, NY, USA. ACM Press.

Apted, T., Kay, J., and Quigley, A. (2006). Tabletop sharing of digital photographs for the elderly. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 781–790, New York, NY, USA. ACM Press.

Arnheim, R. (1993). Sketching and the psychology of design. *Design Issues*, 9(2):15–19.

Balakrishnan, R., Fitzmaurice, G., Kurtenbach, G., and Singh, K. (1999a). Exploring interactive curve and surface manipulation using a bend and twist sensitive input strip. In *I3D '99: Proceedings of the 1999 symposium on Interactive 3D graphics*, pages 111–118, New York, NY, USA. ACM Press.

Balakrishnan, R., Fitzmaurice, G. W., Kurtenbach, G., and Buxton, W.

A. S. (1999b). Digital tape drawing. In *UIST '99: Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 161–169, New York, NY, USA. ACM Press.

Balakrishnan, R. and Hinckley, K. (1999). The role of kinesthetic reference frames in two-handed input performance. In *UIST '99: Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 171–178, New York, NY, USA. ACM Press.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, 10(1):82–93.

Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A. (2003). Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch- and pen-operated systems. In *Proceedings of IFIP INTERACT03: Human-Computer Interaction*, page 65. IFIP Technical Committee No 13 on Human-Computer Interaction.

Baudisch, P., Good, N., and Stewart, P. (2001). Focus plus context screens: combining display technology with visualization techniques. In Marks, J. and Mynatt, E. D., editors, *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 31–40. ACM Press.

Beaudouin-Lafon, M. (2001). Novel interaction techniques for overlapping windows. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 153–154, New York, NY, USA. ACM Press.

Bell, G. and Kaye, J. (2002). Designing technology for domestic spaces: A kitchen manifesto. *Gastronomica*, 2(2).

Bezerianos, A. and Balakrishnan, R. (2005). The vacuum: facilitating the manipulation of distant objects. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 361–370, New York, NY, USA. ACM Press.

Bier, E. A., Stone, M. C., Pier, K., Buxton, W. A. S., and DeRose, T. D. (1993). Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80, New York, NY, USA. ACM Press.

Black, M. (1972). *Models And Metaphors. Studies In Language And Philosophy.* Cornell University Press.

Boring, S., Hilliges, O., and Butz, A. (2007). A wall-sized focus plus context display. In *Proceedings of the Fifth Annual IEEE Conference on Pervasive Computing and Communications (PerCom).*

Butler, C. G. and St. Amant, R. (2004). Habilisdraw DT: a bimanual tool-based direct manipulation drawing environment. In *CHI '04: extended* *abstracts on Human factors in computing systems*, pages 1301–1304, New York, NY, USA. ACM Press.

Butz, A., Groß, M., and Krüger, A. (2004). Tuister: a tangible UI for hierarchical structures. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, pages 223–225, New York, NY, USA. ACM Press.

Butz, A., Hilliges, O., Terrenghi, L., and Baur, D. (2007). Hybrid widgets on an interactive tabletop. In *Ubicomp '07: Adjunct Proceedings, Demo session.*

Buxton, W. (1983). Lexical and pragmatic considerations of input structures. *SIGGRAPH Computer Graphics*, 17(1):31–37.

Buxton, W. (1994). Human skills in interface design. In MacDonald, L. and Vince, J., editors, *Interacting With Virtual Environments*, pages 1–12. Wiley, New York.

Buxton, W. (1995). Chunking and phrasing and the design of human-computer dialogues. In *Human-computer interaction: toward the year 2000*, pages 494–499. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Buxton, W. (2007a). *Human Input to Computer Systems: Theories, Techniques and Technology.* Cambridge University Press, Cambridge, UK. In preparation.

Buxton, W. (2007b). *Sketching User Experiences.* Morgan Kaufmann, San Francisco, CA, USA.

Buxton, W., Fitzmaurice, G. W., Balakrishnan, R., and Kurtenbach, G. (2000). Large displays in automotive design. *IEEE Computer Graphics Applications*, 20(4):68–75.

Buxton, W. and Myers, B. (1986). A study in two-handed input. In *CHI '86: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 321–326, New York, NY, USA. ACM Press.

Buxton, W. A. S. (1997). Living in augmented reality: Ubiquitous media and reactive environments. In Finn, K., Sellen, A., and Wilber, S., editors, *Video Mediated Communication*, pages 363–384. Erlbaum, Hillsdale, N.J.

Callahan, J., Hopkins, D., Weiser, M., and Shneiderman, B. (1988). An empirical comparison of pie vs. linear menus. In *CHI '88: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 95–100, New York, NY, USA. ACM Press.

Camarata, K., Do, E. Y.-L., Johnson, B. R., and Gross, M. D. (2002). Navigational blocks: navigating information space with tangible media. In *IUI '02: Proceedings of the 7th international conference on Intelligent user interfaces*, pages 31–38, New York, NY, USA. ACM Press.

Cao, X. and Balakrishnan, R. (2003). Visionwand: interaction techniques for large displays using a passive wand tracked in 3d. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 173–182, New York, NY, USA. ACM Press.

Carroll, J. M. and Thomas, J. C. (1982). Metaphor and the cognitive representation of computing systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 12:107–116.

Cook, D. J. and Bailey, B. P. (2005). Designers' use of paper and the implications for informal tools. In *OZCHI '05: Proceedings of the 19th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction*, pages 1–10, Narrabundah, Australia.

Cooper, A. and Reimann, R. (2003). *About Face 2.0: The Essentials of Interaction Design*. John Wiley & Sons, Inc., New York, NY, USA.

Crampton Smith, G. (1995). The hand that rocks the cradle. *ID Magazine*, May/June:60–65.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. Harper Collins, New York, NY, USA.

Dennis, A. R. and Reinicke, B. (2004). Beta versus VHS and the acceptance of electronic brainstorming technology. *MIS Quarterly*, 28(1).

Diehl, M. and Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, 53:497–509.

Dietz, P. and Leigh, D. (2001). Diamondtouch: a multi-user touch technology. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 219–226, New York, NY, USA. ACM Press.

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press, Cambridge, MA, USA.

Dourish, P. and Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *CSCW '92: Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 107–114, New York, NY, USA. ACM Press.

Dourish, P. and Bly, S. (1992). Portholes: supporting awareness in a distributed work group. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 541–547, New York, NY, USA. ACM Press.

Dragicevic, P. (2004). Combining crossing-based and paper-based interaction paradigms for dragging and dropping between overlapping windows. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 193–196, New York, NY, USA. ACM Press.

Druin, A. and Inkpen, K. (2001). When are personal technologies for children? *Personal Ubiquitous Computing*, 5(3):191–194.

Erickson, T. D. (1995). Working with interface metaphors. In *Human-computer interaction: toward the year 2000*, pages 147–151. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Fallman, D. (2003). Design-oriented human-computer interaction. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 225–232, New York, NY, USA. ACM Press.

Feiner, S. and Shamash, A. (1991). Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers. In *UIST '91: Proceedings of the 4th annual ACM symposium on User interface software and technology*, pages 9–17, New York, NY, USA. ACM Press.

Fischer, G. (1999). Symmetry of igorance, social creativity, and meta-design. In *C&C '99: Proceedings of the 3rd conference on Creativity & cognition*, pages 116–123, New York, NY, USA. ACM Press.

Fischer, G. (2005). Distances and diversity: sources for social creativity. In *C&C '05: Proceedings of the 5th conference on Creativity & cognition*, pages 128–136, New York, NY, USA. ACM Press.

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal Ubiquitous Computing*, 8(5):347–358.

Fitzmaurice, G. W. (1996). *Graspable User Interfaces*. PhD thesis.

Fitzmaurice, G. W., Ishii, H., and Buxton, W. (1995). Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'95)*, pages 442–449. ACM Press/Addison-Wesley Publishing Co.

Fitzmaurice, G. W., Zhai, S., and Chignell, M. H. (1993). Virtual reality for palmtop computers. *ACM Transactions on Information Syststems*, 11(3):197–218.

Foley, J. D. and Van Dam, A. (1982). *Fundamentals of interactive computer graphics*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Freeman, S. (1996). Gemma: an open architecture for distributed user interfaces. pages 65–85. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Frohlich, D., Kuchinsky, A., Pering, C., Don, A., and Ariss, S. (2002). Requirements for photoware. In *CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 166–175, New York, NY, USA. ACM Press.

Gaver, B., Dunne, T., and Pacenti, E. (1999). Design: Cultural probes. *interactions*, 6(1):21–29.

Gaver, W. W. (1991). Technology affordances. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages

79–84, New York, NY, USA. ACM Press.

Gaver, W. W. (1996). Affordances for interaction: The social is material for design. *Ecological Psychology*, 8.

Gaver, W. W., Beaver, J., and Benford, S. (2003). Ambiguity as a resource for design. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 233–240, New York, NY, USA. ACM Press.

Gedenryd, H. (1998). *How Designers Work*. PhD thesis.

Gellersen, H. W., Beigl, M., and Krull, H. (1999). The mediacup: Awareness technology embedded in a everyday object. In *HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 308–310, London, UK. Springer-Verlag.

Gescka, A. (1978). Introduction and use of idea-generating methods. *Research Management*, 21(3).

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, New Jersey, USA.

Goldberg, D. and Richardson, C. (1993). Touch-typing with a stylus. In *CHI '93: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 80–87, New York, NY, USA. ACM Press.

Goldschmidt, G. (1991). The dialectics of sketching. *Creativity Research Journal*, 4(2).

Grint, K. and Woolgar, S. (1997). *The Machine at Work: Technology, Work and Organization*. Polity Press, Cambridge, MA, USA.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *The Journal of Motor Behavior*, 19(4):485–517.

Guimbretière, F., Stone, M., and Winograd, T. (2001). Fluid interaction with high-resolution wall-size displays. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 21–30, New York, NY, USA. ACM Press.

Guimbretière, F. and Winograd, T. (2000). Flowmenu: combining command, text, and data entry. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 213–216, New York, NY, USA. ACM Press.

Gutierrez, A. (1996). Children's ability for using different plane representations of space figures. In Batturo, A. R., editor, *New Directions in Geometry Education*, pages 33–42. Centre for Math. and Sc. Education, Q.U.T.

Halasz, F. and Moran, T. P. (1982). Analogy considered harmful. In *Proceedings of the 1982 conference on Human factors in computing systems*, pages 383–386, New York, NY, USA. ACM Press.

Han, J. Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 115–118, New York, NY, USA. ACM Press.

Hancock, M. S., Carpendale, S., Vernier, F. D., Wigdor, D., and Shen, C. (2006). Rotation and translation mechanisms for tabletop interaction. In *TABLETOP '06: Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pages 79–88, Washington, DC, USA. IEEE Computer Society.

Harrison, B. L., Fishkin, K. P., Gujar, A., Mochon, C., and Want, R. (1998). Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour and Information Technology - BIT*, 22:315–338.

Hewett, T. T., Baecker, R. M., Card, S. K., Carey, T., Gasen, J. G., Mantei, M., Perlman, G., Strong, G. W., and Verplank, W. (1997). ACM sigchi curricula for Human-Computer Interaction. Electronic document. Date of publication: April 17, 1997. Date last modified: December 6, 2002.

Hilliges, O., Terrenghi, L., Boring, S., Kim, D., Richter, H., and Butz, A. (2007). Designing for collaborative creative problem solving. In *C&C '07: Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, pages 137–146, New York, NY, USA. ACM Press.

Hinckley, K. (1997). *Haptic Issues for Virtual Manipulation.* PhD thesis.

Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. (1994). Passive real-world interface props for neurosurgical visualization. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 452–458, New York, NY, USA. ACM Press.

Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., and Smith, M. (2004). Stitching: pen gestures that span multiple displays. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 23–31, New York, NY, USA. ACM Press.

Hinrichs, U., Carpendale, S., and Scott, S. D. (2005). Interface currents: Supporting fluent collaboration on tabletop displays. In *Proceedings of the 5th International Symposium on Smart Graphics, Fraunewoerth Cloister, Germany*, pages 185–197, Heiderlberg, Germany. Springer-Verlag.

Holmquist, L. E., Redström, J., and Ljungstrand, P. (1999). Token-based acces to digital information. In *HUC '99: Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing*, pages

234–245, London, UK. Springer-Verlag.

Hornecker, E. (2002). Understanding the benefits of graspable interfaces for cooperative use. In Blay-Fornarino, M., Pinna-Dery, A.-M., Schmidt, K., and Zarat, P., editors, *COOP'02, Fifth International Conference on the Design of Cooperative Systems*, pages 71–87. IOS.

Hornecker, E. and Buur, J. (2006). Getting a grip on tangible interaction: a framework on physical space and social interaction. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 437–446, New York, NY, USA. ACM Press.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press, Cambridge, MA, USA.

Inkpen, K., Hawkey, K., Kellar, M., Mandryk, R., Parker, K., Reilly, D., Scott, S., and Whalen, T. (2005). Exploring display factors that influence co-located collaboration: Angle, size, number, and user arrangement. In *Proceedings of HCI International 2005*, Las Vegas, USA.

Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA. ACM Press.

Johanson, B., Fox, A., and Winograd, T. (2002). The interactive workspaces project: experiences with ubiquitous computing rooms. *Pervasive Computing, IEEE*, 1(2):67–74.

Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. (2007). The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 139–146, New York, NY, USA. ACM Press.

Jordan, B. and Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1):39–103.

Ju, W., Hurwitz, R., Judd, T., and Lee, B. (2001). Counteractive: an interactive cookbook for the kitchen counter. In *CHI '01: extended abstracts on Human factors in computing systems*, pages 269–270, New York, NY, USA. ACM Press.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, 73(1-2):31–68.

Kirsh, D. and Maglio, P. P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4):513–549.

Klemmer, S. R., Hartmann, B., and Takayama, L. (2006). How bodies matter: five themes for interaction design. In *DIS '06: Proceedings of the 6th ACM conference on Designing Interactive systems*, pages 140–149, New York, NY, USA. ACM Press.

Klemmer, S. R., Newman, M. W., Farrell, R., Bilezikjian, M., and Landay, J. A. (2001). The designers' outpost: a tangible interface for collaborative web site. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 1–10, New York, NY, USA. ACM Press.

Koike, H., Xinlei, C., Nakanishi, Y., Oka, K., and Sato, Y. (2002). Two-handed drawing on augmented desk. In *CHI '02 extended abstracts on Human factors in computing systems*, pages 760–761, New York, NY, USA. ACM Press.

Konomi, S., Müller-Tomfelde, C., and Streitz, N. A. (1999). Passage: Physical transportation of digital information in cooperative buildings. In *CoBuild '99: Proceedings of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture*, pages 45–54, London, UK. Springer-Verlag.

Kranz, M., Schmidt, D., Holleis, P., and Schmidt, A. (2005). A display cube as tangible user interface. In *Adjunct Proceedings of the Seventh International Conference on Ubiquitous Computing (Demo 22)*.

Kreijns, K. and Kirschner, P. (2001). The social affordances of computer-supported collaborative learning environments. In *Proceedings of the 31th ASEE/IEEE Frontiers in Education Conference*, pages 12–17.

Kristensson, P. O. and Zhai, S. (2007). Command strokes with and without preview: using pen gestures on keyboard for command selection. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1137–1146, New York, NY, USA. ACM Press.

Kruger, R., Carpendale, S., Scott, S. D., and Greenberg, S. (2003). How people use orientation on tables: comprehension, coordination and communication. In *GROUP '03: Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, pages 369–378, New York, NY, USA. ACM Press.

Kruger, R., Carpendale, S., Scott, S. D., and Tang, A. (2005). Fluid integration of rotation and translation. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 601–610, New York, NY, USA. ACM Press.

Kurtenbach, G. and Buxton, W. (1991a). Gedit: a test bed for editing by contiguous gestures. *SIGCHI Bulletin*, 23(2):22–26.

Kurtenbach, G. and Buxton, W. (1991b). Issues in combining marking and direct manipulation techniques. In *UIST '91: Proceedings of the 4th annual ACM symposium on User interface software and technology*, pages 137–144, New York, NY, USA. ACM Press.

Lakoff, G. and Johnson, M. (1983). *Metaphors We Live by.* Basic Books, New York, NY, USA.

Laurel, B. (1986). Interface as mimesis. In Norman, D. A. and Draper, S. W., editors, *User Centered System Design: New Perspectives on Human-Computer Interaction*, pages 67–86. Lawrence Erlbaum Associates, Hillsdale, NJ.

Laurel, B. (2003). *Design Research: Methods and Perspectives*. MIT Press, Cambridge, MA, USA.

Leganchuk, A., Zhai, S., and Buxton, W. (1998). Manual and cognitive benefits of two-handed input: an experimental study. *ACM Transactions on Computer-Human Interaction*, 5(4):326–359.

Long, A. C. J., Landay, J. A., and Rowe, L. A. (1998). PDA and gesture uses in practice: Insights for designers of pen-based user interfaces. Technical Report CSD-97-976.

Madsen, K. H. (1994). A guide to metaphorical design. *Communications of the ACM*, 37(12):57–62.

Mander, R., Salomon, G., and Wong, Y. Y. (1992). A pile metaphor for supporting casual organization of information. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 627–634, New York, NY, USA. ACM Press.

Marcus, A. (1998). Metaphor design in user interfaces. *SIGDOC Asterisk Journal of Computer Documentation*, 22(2):43–57.

Matsushita, N., Ayatsuka, Y., and Rekimoto, J. (2000). Dual touch: a two-handed interface for pen-based pdas. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 211–212, New York, NY, USA. ACM Press.

Mazalek, A., Davenport, G., and Ishii, H. (2002). Tangible viewpoints: a physical approach to multimedia stories. In *MULTIMEDIA '02: Proceedings of the tenth ACM international conference on Multimedia*, pages 153–160, New York, NY, USA. ACM Press.

Mazalek, A., Reynolds, M., and Davenport, G. (2006). Tviews: An extensible architecture for multiuser digital media tables. *IEEE Computer Graphics and Applications*, 26(5):47–55.

Morris, M. R., Huang, A., Paepcke, A., and Winograd, T. (2006). Cooperative gestures: multi-user gestural interactions for co-located groupware. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1201–1210, New York, NY, USA. ACM Press.

Mynatt, E. D., Igarashi, T., Edwards, W. K., and LaMarca, A. (1999). Flatland: new dimensions in office whiteboards. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 346–353, New York, NY, USA. ACM Press.

Nacenta, M. A., Aliakseyeu, D., Subramanian, S., and Gutwin, C. (2005). A comparison of techniques for multi-display reaching. In *CHI '05: Proceed-*

ings of the SIGCHI conference on Human factors in computing systems, pages 371–380, New York, NY, USA. ACM Press.

Nakatani, L. H. and Rohrlich, J. A. (1983). Soft machines: A philosophy of user-computer interface design. In *CHI '83: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 19–23, New York, NY, USA. ACM Press.

Nelson, T. (1990). The right way to think about software design. In Laurel, B. and Mountford, S. J., editors, *The Art of Human-Computer Interface Design*, pages 235–243. Addison-Wesley, Reading, MA, USA.

Nielsen, J. (1999). *Designing Web Usability: The Practice of Simplicity*. New Riders Publishing, Thousand Oaks, CA, USA.

Norman, D. A. (1983). Some observations on mental models. In Donald R. Gentner, A. L. S., editor, *Mental Models*. Lawrence Erlbaum Associates.

Norman, D. A. (1988). *The Psychology of Everyday Things*. Basic Books, New York, NY, USA.

Norman, D. A. (1999). Affordance, conventions, and design. *Interactions*, 6(3):38–43.

Norman, D. A. and Draper, S. W. (1986). *User Centered System Design; New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA.

Norman, D. M. and Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, (7):44–47.

O'Regan, J. K. and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):955–977.

Osborn, A. F. (1957). *Applied Imagination: Principles and Procedures of Creative Thinking*. Scribeners and Sons, New York, NY, USA.

Patten, J. and Ishii, H. (2000). A comparison of spatial organization strategies in graphical and tangible user interfaces. In *DARE '00: Proceedings of DARE 2000 on Designing augmented reality environments*, pages 41–50, New York, NY, USA. ACM Press.

Patten, J., Ishii, H., Hines, J., and Pangaro, G. (2001). Sensetable: a wireless object tracking platform for tangible user interfaces. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 253–260, New York, NY, USA. ACM Press.

Patten, J., Recht, B., and Ishii, H. (2002). Audiopad: a tag-based interface for musical performance. In *NIME '02: Proceedings of the 2002 conference on New interfaces for musical expression*, pages 1–6, Singapore, Singapore. National University of Singapore.

Pedersen, E. R., McCall, K., Moran, T. P., and Halasz, F. G. (1993). Tivoli: an electronic whiteboard for informal workgroup meetings. In *CHI '93:*

Proceedings of the SIGCHI conference on Human factors in computing systems, pages 391–398, New York, NY, USA. ACM Press.

Perlin, K. (1998). Quikwriting: continuous stylus-based text entry. In *UIST '98: Proceedings of the 11th annual ACM symposium on User interface software and technology*, pages 215–216, New York, NY, USA. ACM Press.

Piaget, J. (1972). *The Psychology of the Child*. Basic Books, New York, NY, USA.

Piper, B., Ratti, C., and Ishii, H. (2002). Illuminating clay: a 3-d tangible interface for landscape analysis. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 355–362, New York, NY, USA. ACM Press.

Pirhonen, A., Brewster, S., and Holguin, C. (2002). Gestural and audio metaphors as a means of control for mobile devices. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 291–298, New York, NY, USA. ACM Press.

Pirolli, P. and Card, S. K. (1999). Information foraging. *Psychological Review*, 106:643–675.

Preece, J., Rogers, Y., and Sharp, H. (2001). *Beyond Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, Inc., New York, NY, USA.

Price, S. and Rogers, Y. (2004). Let's get physical: the learning benefits of interacting in digitally augmented physical spaces. *Computers and Education*, 43(1-2):137–151.

Rekimoto, J. (1996). Tilting operations for small screen interfaces. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 167–168, New York, NY, USA. ACM Press.

Rekimoto, J. (1997). Pick-and-drop: a direct manipulation technique for multiple computer environments. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 31–39, New York, NY, USA. ACM Press.

Rekimoto, J. (2002). Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 113–120, New York, NY, USA. ACM Press.

Rekimoto, J. and Saitoh, M. (1999). Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 378–385, New York, NY, USA. ACM Press.

Rekimoto, J. and Sciammarella, E. (2000). Toolstone: effective use of the physical manipulation vocabularies of input devices. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software* *and technology*, pages 109–117, New York, NY, USA. ACM Press.

Rekimoto, J., Ullmer, B., and Oba, H. (2001). Datatiles: a modular platform for mixed physical and graphical interactions. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 269–276, New York, NY, USA. ACM Press.

Resnick, M., Martin, F., Berg, R., Borovoy, R., Colella, V., Kramer, K., and Silverman, B. (1998). Digital manipulatives: new toys to think with. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 281–287, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

Richardson, A. (1993). The death of the designer. *Design Issues*, 9(2):34–43.

Rieser, J. J., Garing, A. E., and Young, M. F. (1994). Imagery, action and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Development*, 65:1254–1270.

Ringel, M., Ryall, K., Shen, C., Forlines, C., and Vernier, F. (2004). Release, relocate, reorient, resize: fluid techniques for document sharing on multi-user interactive tables. In *CHI '04: extended abstracts on Human factors in computing systems*, pages 1441–1444, New York, NY, USA. ACM Press.

Robles-De-La-Torre, G. (2006). The importance of the sense of touch in virtual and real environments. *IEEE MultiMedia*, 13(3):24–30.

Rogers, Y. (2006). Moving on from Weiser's vision of calm computing: Engaging ubicomp experiences. In *Proceedings of Ubicomp 2006*, pages 404–421. LNCS 4206, Springer-Verlag.

Sanneblad, J. and Holmquist, L. E. (2006). Ubiquitous graphics: combining hand-held and wall-size displays to interact with large images. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 373–377, New York, NY, USA. ACM Press.

Santanen, E. L., Briggs, R. O., and de Vreede, G.-J. (1999). A cognitive network model of creativity: a renewed focus on brainstorming methodology. In *ICIS '99: Proceeding of the 20th international conference on Information Systems*, pages 489–494, Atlanta, GA, USA. Association for Information Systems.

Schmidt, A. and Terrenghi, L. (2007). Methods and guidelines for the design and development of domestic ubiquitous computing applications. In *Proceedings of the Fifth Annual IEEE Conference on Pervasive Computing and Communications (PerCom)*.

Scott, S. D., Carpendale, M. S. T., and Habelski, S. (2005). Storage bins: Mobile storage for collaborative tabletop displays. *IEEE Computer Graphics Applications*, 25(4):58–65.

Scott, S. D., Carpendale, M. S. T., and Inkpen, K. M. (2004). Territoriality in collaborative tabletop workspaces. In *CSCW '04: Proceedings of the*

2004 ACM conference on Computer supported cooperative work, pages 294–303, New York, NY, USA. ACM Press.

Scott, S. D., Grant, K. D., and Mandryk, R. L. (2003). System guidelines for co-located, collaborative work on a tabletop display. In *ECSCW'03: Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work*, pages 159–178, Norwell, MA, USA. Kluwer Academic Publishers.

Sellen, A. J. and Harper, R. H. (2003). *The Myth of the Paperless Office.* MIT Press, Cambridge, MA, USA.

Shen, C., Lesh, N. B., Vernier, F., Forlines, C., and Frost, J. (2002). Sharing and building digital group histories. In *CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 324–333, New York, NY, USA. ACM Press.

Shen, C., Ryall, K., Forlines, C., Esenther, A., Vernier, F. D., Everitt, K., Wu, M., Wigdor, D., Morris, M. R., Hancock, M., and Tse, E. (2006). Informing the design of direct-touch tabletops. *IEEE Computer Graphics Applications*, 26(5):36–46.

Shen, C., Vernier, F. D., Forlines, C., and Ringel, M. (2004). Diamond-spin: an extensible toolkit for around-the-table interaction. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 167–174, New York, NY, USA. ACM Press.

Shneiderman, B. (1987). Direct manipulation: A step beyond programming languages. pages 461–467.

Shneiderman, B. (2003). *Leonardo's Laptop: Human Needs and the New Computing Technologies.* MIT Press, Cambridge, MA, USA.

Shneiderman, B., Card, S., Norman, D. A., Tremaine, M., and Waldrop, M. M. (2002). Chi@20: fighting our way from marginality to power. In *CHI '02: extended abstracts on Human factors in computing systems*, pages 688–691, New York, NY, USA. ACM Press.

Simon, H. (1969). *The science of the artificial.* MIT Press, Cambridge, MA, USA.

Smets, G. J. F., Stappers, P. J., and Overbeeke, K. (1994). Designing in virtual reality: implementing perception-action coupling with affordances. In *VRST '94: Proceedings of the conference on Virtual reality software and technology*, pages 97–110, River Edge, NJ, USA. World Scientific Publishing Co., Inc.

Smith, D. C., Irby, C., Kimball, R., and Harslem, E. (1982). The star user interface: an overview. In *on AFIPS Conference Proceedings; vol. 50 1982 National Computer Conference*, pages 515–528, Arlington, VA, USA. AFIPS Press.

Streitz, N. A., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. (1999). i-land: An interactive landscape for creativity and innovation. In *Proceeding of the CHI '99 conference on Human factors in computing systems (CHI'99)*, pages 120–127. ACM Press, New York, NY.

Sundholm, H., Artman, H., and Ramberg, R. (2004). Cooperative systems design: Scenario-based design of collaborative systems. In *Backdoor Creativity: Collaborative Creativity in Technology Supported Teams*, pages 99–114. IOS press, Amsterdam, The Netherlands.

Sutherland, I. E. (1988). Sketchpad a man-machine graphical communication system. In *25 years of DAC: Papers on Twenty-five years of electronic design automation*, pages 507–524, New York, NY, USA. ACM Press.

Suwa, M. and Tversky, B. (2002). External representations contribute to the dynamic construction of ideas. In *DIAGRAMS '02: Proceedings of the Second International Conference on Diagrammatic Representation and Inference*, pages 341–343, London, UK. Springer-Verlag.

Suzuki, H. and Kato, H. (1993). Algoblock: a tangible programming language, a tool for collaborative learning. In *Proceedings of the 4th European Logo Conference*, pages 297–303.

Tandler, P. (2001). Software infrastructure for ubiquitous computing environments: Supporting synchronous collaboration with heterogeneous devices. In *UbiComp '01: Proceedings of the 3rd international conference on Ubiquitous Computing*, pages 96–115, London, UK. Springer-Verlag.

Terrenghi, L. (2005). Design of affordances for direct manipulation of digital information in ubiquitous computing scenarios. In *Proceedings of the 5th International Symposium on Smart Graphics, Fraunewoerth Cloister, Germany*, pages 198–205, Heiderlberg, Germany. Springer-Verlag.

Terrenghi, L. (2006a). Shaping intelligent environments: the sign of design. In *International Symposium on Intelligent Environments*, pages 137–143, Cambridge, UK. Microsoft Research.

Terrenghi, L. (2006b). Sticky, smelly, smoky context: experience design in the kitchen. In *CAI '06: Proceedings of the international workshop in conjunction with AVI 2006 on Context in advanced interfaces*, pages 49–52, New York, NY, USA. ACM Press.

Terrenghi, L., Fritsche, T., and Butz, A. (2006a). The enlightable: Design of affordances to support collaborative creativity. In *Proceedings of the 6th International Symposium on Smart Graphics, Vancouver, Canada*, pages 206–217, Heiderlberg, Germany. Springer-Verlag.

Terrenghi, L., Harper, R., and Sellen, A. (2006b). Inspirations for design. In *Ubicomp '06: Online Proceedings, Open session.*

Terrenghi, L., Hilliges, O., and Butz, A. (2007a). Kitchen stories: sharing recipes with the living cookbook. *Personal Ubiquitous Computing*, 11(5):409–414.

Terrenghi, L., Kirk, D., Sellen, A., and Izadi, S. (2007b). Affordances for manipulation of physical versus digital media on interactive surfaces. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1157–1166, New York, NY, USA. ACM Press.

Terrenghi, L., Kranz, M., Holleis, P., and Schmidt, A. (2006c). A cube to learn: a tangible user interface for the design of a learning appliance. *Personal Ubiquitous Computing*, 10(2-3):153–158.

Tohidi, M., Buxton, W., Baecker, R., and Sellen, A. (2006). Getting the right design and the design right. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1243–1252, New York, NY, USA. ACM Press.

Tse, E., Shen, C., Greenberg, S., and Forlines, C. (2006). Enabling interaction with single user applications through speech and gestures on a multi-user tabletop. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 336–343, New York, NY, USA. ACM Press.

Ullmer, B. and Ishii, H. (1997). The metadesk: models and prototypes for tangible user interfaces. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 223–232, New York, NY, USA. ACM Press.

Ullmer, B. and Ishii, H. (1999). Mediablocks: tangible interfaces for online media. In *CHI '99: extended abstracts on Human factors in computing systems*, pages 31–32, New York, NY, USA. ACM Press.

Ullmer, B. and Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Syst. J.*, 39(3-4):915–931.

Underkoffler, J. and Ishii, H. (1999). Urp: a luminous-tangible workbench for urban planning and design. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 386–393, New York, NY, USA. ACM Press.

Van Gundy, A. B. (1981). *Techniques of Structured Problem-Solving*. Van Nostrand, Reinhold, NY, USA.

Venolia, D. and Neiberg, F. (1994). T-cube: a fast, self-disclosing pen-based alphabet. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–270, New York, NY, USA. ACM Press.

Want, R., Fishkin, K. P., Gujar, A., and Harrison, B. L. (1999). Bridging physical and virtual worlds with electronic tags. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 370–377, New York, NY, USA. ACM Press.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3):66–75.

Weiser, M. and Brown, J. S. (1995). Designing calm technology. Technical report, Xerox PARC.

Wellner, P. (1991). The digitaldesk calculator: tangible manipulation on a desktop display. In *UIST '91: Proceedings of the 4th annual ACM symposium on User interface software and technology*, pages 27–33, New York, NY, USA. ACM Press.

Wellner, P. (1993). Interacting with paper on the digitaldesk. *Communications of the ACM*, 36(7):87–96.

Wenger, E. (1998). *Communities of practice: learning, meaning, and identity.* Cambridge University Press, Cambridge, U.K. and New York, N.Y.

Williams, G. (1984). The Apple Macintosh computer. *j-BYTE*, 9(2):30–54.

Wilson, A. D. (2004). Touchlight: an imaging touch screen and display for gesture-based interaction. In *ICMI '04: Proceedings of the 6th international conference on Multimodal interfaces*, pages 69–76, New York, NY, USA. ACM Press.

Wilson, A. D. (2005). Playanywhere: a compact interactive tabletop projection-vision system. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 83–92, New York, NY, USA. ACM Press.

Winograd, T., editor (1996). *Bringing design to software.* ACM Press, New York, NY, USA.

Wu, M. and Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 193–202, New York, NY, USA. ACM Press.

Wu, M., Shen, C., Ryall, K., Forlines, C., and Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *TABLETOP '06: Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pages 185–192, Washington, DC, USA. IEEE Computer Society.

Yee, K.-P. (2004). Two-handed interaction on a tablet display. In *CHI '04: extended abstracts on Human factors in computing systems*, pages 1493–1496, New York, NY, USA. ACM Press.

Zhai, S. and Kristensson, P. O. (2003). Shorthand writing on stylus keyboard. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 97–104, New York, NY, USA. ACM Press.

Zhai, S. and Milgram, P. (1993). Human performance evaluation of isometric and elastic rate controllers in a 6doF tracking task. In *Proc. SPIE Telemanipulator Technology*, volume SPIE vol. 2057.

Zimmerman, J., Evenson, S., Baumann, K., and Purgathofer, P. (2004). Workshop on the relationship between design and hci. In *CHI '04: extended abstracts on Human factors in computing systems*, pages 1741–1742, New York, NY, USA. ACM Press.

Zimmerman, J., Forlizzi, J., and Evenson, S. (2007). Research through design as a method for interaction design research in hci. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 493–502, New York, NY, USA. ACM Press.

Biography

Lucia Terrenghi is a designer and a researcher working in the field of Human-Computer Interaction. Her interests address the relationship between humans and technologies, and between humans and humans mediated by technology. In particular, she explores ways in which technological artifacts can enhance social engagement, self-expression, and creativity.

Since December 2004 she has worked as research assistant at the Ludwig-Maximilians-Universität München, in Germany. Her research has developed in the context of the FLUIDUM project - funded by the Deutsche Forschungsgemeinschaft - whose mission is to explore and develop techniques for interacting with ubiquitous computing technologies, in settings of everyday life.

In summer 2006 she did an internship at Microsoft Research in Cambridge, United Kingdom, and worked with the Socio-Digital Systems research group at the study of sharable interfaces for the home.

Prior to that, she worked for two years as user interface designer and HCI researcher at the FIT Fraunhofer Institute for Applied Information Technologies in Sankt Augustin, near Bonn. In this context, she mainly investigated issues of interface design and user requirements for mobile and e-learning applications.

Before starting her career in academic research, Lucia worked for almost two years at brand communication for the Organizing Committee of the XX Olympic Games of Turin 2006, in Italy, combining her passion for winter sports and design. This was a follow-up of her Master thesis, which addressed the design of a mobile navigation guide for the visitors of the Olympic Games.

Lucia holds a Master of Science in Industrial Design, which she gained with honor in 2000 at the Politecnico di Milano, in Italy. During the 1998/99 academic year she attended the Technische Universiteit Delft, in the Nederlands. Throughout the period of her studies, Lucia also made several working experiences in diverse fields of design, such as interior, light, and web design.

Her design and research activities were acknowledged with different prizes: In 1998 her design for the MOMIX lamp was a winner of the YOUNG&DESIGN contest, organized by Rima Publisher, and the prototype was exhibited at the furniture fair “Salone del Mobile di Milano”. In 1999 she won second prize in a Dutch national design contest titled “De Schooltas”, sponsored by the Hema enterprise in cooperation with the Dutch Public School Ministry. More recently, in 2007, she received the Google Europe Anita Borg Memorial Scholarship, which rewards women for their contribution to research in Computer Sciences and Technologies.