

Mending Fractured Spaces: External Legibility and Seamlessness in Interface Design

by

Jamie B Zigelbaum

B.A. in Human-Computer Interaction, Tufts University, 2006

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

© Massachusetts Institute of Technology 2008. All rights reserved.

Author.....
Jamie B Zigelbaum
Program in Media Arts and Sciences
August 8, 2008

Certified by
Hiroshi Ishii
Muriel R. Cooper Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by
Deb Roy
Chair, Academic Program in Media Arts and Sciences

Mending Fractured Spaces: External Legibility and Seamlessness in Interface Design

by

Jamie B Zigelbaum

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on August 8th, 2008, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

This thesis presents *External Legibility*: a property of user interfaces that affects the ability of non-participating observers to understand the context of a user's actions. Claims of its value are supported with arguments from the social sciences and human-computer interaction literature; research in designing tangible user interfaces; and an experiment comparing the external legibility of four interaction techniques.

Thesis Supervisor: Hiroshi Ishii

Title: Muriel R. Cooper Professor of Media Arts and Sciences

Mending Fractured Spaces: External Legibility and Seamlessness in Interface Design

by

Jamie B Zigelbaum

The following people served as readers for this thesis:

Thesis Reader
Robert J. K. Jacob
Professor of Computer Science
Tufts University

Thesis Reader
Henry Jenkins
Co-Director
MIT Comparative Media Studies Program

Thesis Reader
John Maeda
Professor of Media Arts and Sciences
Program in Media Arts and Sciences

ACKNOWLEDGEMENTS

Thanks to Hiroshi Ishii and Rob Jacob, I've been very fortunate to find your guidance; your support has meant so much to me. Thanks to readers Henry Jenkins and John Maeda, I am grateful for your time and thoughtfulness.

Many people had an impact on this work, I'd like to thank Nadav Aharony, Marcelo Coelho, Michael Horn, Adam Kumpf, Anita Lillie, Aninna Rüst, Orit Shaer, Mikey Siegal, and Alyssa Wright in particular for your friendship and help.

Thanks to colleagues, teachers, collaborators, and friends Leonardo Bonanni, Kyle Buza, Michael Carvajal, Manuel Castells, Angela Chang, Chris Csikszentmihalyi, Rich Fletcher, Audrey Girouard, Jim Gouldestone, Keith Hampton, Leanne Hirshfield, Seth Hunter, Marisa Jahn, Jeevan Kalanithi, Tim Leeman, Henry Lieberman, Lisa Lieberson, Pattie Maes, David Merrill, Joshua Jen Monzen, Takashi Okamoto, Amanda Parkes, James Patten, Gena Peditto, Michael Poor, Hayes Raffle, Seth Raphael, Coach Brandon Roy, Kevin Rustagi, Erin Solovey, Thomas Taylor, Jim Trankelson, Cati Vaucelle, Alejandro Vazquez, Rebecca Waber, and Jordan Weymer. Thanks also to MIT, the Media Lab, the Tangible Media Group, Tufts University HCI group, Knocked-Up, the third floor of E15, and 1369 Coffee House where I wrote most of this thing.

To my family, I couldn't have been luckier than to end up among you. To mom and dad, I love you, thank you.

CONTENTS

Section 1: thinking	13
prelude < the sadness of hidden riches	15
INTRODUCTION.....	17
Public/Private.....	18
Communication.....	19
Collaboration.....	21
Thesis.....	22
RELATED ARGUMENTS	25
The Decline of Social Capital.....	26
Local Information Processing	27
Reality-Based Interaction	30
Technological Determinism.....	34
Summary	36
Section 2: doing	37
interlude < dorks	39
TANGIBLE USER INTERFACES.....	41
The Tangible Video Editor.....	42
Slurp.....	50
SpeakCup.....	60

EXTERNAL LEGIBILITY	65
Feedforward.....	66
Feedthrough.....	66
Feedback.....	67
Experiment.....	67
Section 3: dreaming	81
interlude < the past.....	83
THE FUTURE	85
Design Principles.....	86
Sketch 1: Trail of Important Glass Beads.....	87
Sketch 2: Cafe of Tomorrow	89
Sketch 3: Seamless Creativity	90
CONCLUSION.....	93
BIBLIOGRAPHY	97

Section 1: thinking

prelude < the sadness of hidden riches

1369 coffee shop can get pretty busy on weekday evenings. Sitting in the middle of Central Square in Cambridge, Massachusetts, it's in the heart of my fair city, right between MIT and Harvard. Patrons come in various flavors, but as you might expect, there are a lot of students.

The intellectual atmosphere is rich. As I sit here now, writing my Master's thesis, I take pause to glance around and see what my fellow Cantabridgians are up to.

There is a pretty young lady two tables over who has been trying to get the sizing right on some type of architectural model for the past two hours; there is an older gentleman writing a book about birds, he is surrounded by reference materials and ornithography; there is a man reading a blog, it's about politics; three college kids are writing papers, one appears to be about American History, one is on finance, and I can't tell what the other one is but all of the books piled around the writer seem dusty and old.

The intellectual atmosphere is rich, but I didn't find that out by glancing around as I said. In fact I had to walk through the cafe and interrupt people, asking them what they were doing, letting them know why I was asking: because it's for my research. It's generally rude to look at the screen of a stranger's laptop, so I couldn't find out that way. Also most of the screens aren't pointed

towards me, and the few that are too far away for me to tell what's happening there.

All the richness I described, the architectural model, the pictures of various birds, the rest, is invisible to the general population of the café. A casual observer doesn't know if a laptop user is writing a novel, designing toothbrushes, watching a movie, or in the middle of a business meeting. The richness is virtual, called up from the internet and hard drive platter, flashed out 83 times a second by LED illumination, filtered through thin matrices of liquid crystal, and delivered to the retina of a couple dozen individual users.

INTRODUCTION

In today's semi-public and public spaces—in our cafes and parks, our buses, our universities' common areas, our airports—space is fractured. The difference between the *space of place* and *space of flows* described by Manuel Castells [Castells 1996] can be seen all around us: when you overhear intimate details of a stranger's relationship as they are spoken into a nearby cell phone, or when you see a cafe full of laptops and blank faces. It is the dissonance between the space that we share when we are collocated in the physical world and the virtual spaces that we share when we are plugged into any of our various communications technologies.

I will take the position in this thesis that this fracturing is problematic, for various reasons that I will explore more deeply in the following pages, but I would not say it is to be eliminated. Fractured space can also characterize the difference between your mind and mine, between private and public, and between cacophony and quiet—differences that are important. For instance, most people who use their laptops in shared spaces would prefer it if their emails were not accessible to all those around them.

A problem that I will address in this thesis, generally speaking, is that computer interfaces are often perceived solely as instruments for accomplishing designed tasks. The effect of this perception is that our technologies are less suitable for re-appropriation than nature or the rest of our built environment—their capacity as media for secondary uses is impoverished. These secondary uses range from things like aesthetics (a property of computer interfaces that was nearly absent until just a few years ago

when the beige box became extinct) and external communication, and by this I mean communicating with other collocated people who may not be a part of one's network of social connections. It is this second point that I will address in the following pages.



Figure 1. 1369 Coffee House in Central Square, Cambridge, Mass. USA.

Public/Private

In both the physical and virtual worlds we have many techniques to move along the axis of public/private. On the web we have a surplus of tools for building community and communicating with each other, but these tools are accessed through single user interfaces like laptops and cell phones that segregate people from each other in physical space. These single user interfaces provide a rich toolset to allow users to construct various interactions and nuanced communications with other Internet users out in the virtual world, yet they are deficient when used for communication with other collocated people in the physical world.

Subway riders who would rather keep the title of the book they are reading private from nearby passengers on the train can cover the jacket of their book with newspaper, changing information from public to private. When reading the web, Internet users can bookmark an article using a social bookmarking service like del.icio.us, they can choose not to keep their bookmark private and thus make their connection to the article public—moving from private to public. But someone quickly surfing the web for helpful information during a lecture who wants to make it public to the other people in the lecture hall would find it difficult to do so, although there are currently at least a few research projects attempting to solve this problem [Aharony 2008, Harry 2008, Wifitti].



Figure 2. Diesel Cafe in Davis Square, Somerville, Mass. USA.

Communication

Most of the work trying to provide paths from the virtual world to the local, physical world is rooted in the use of existing interface technologies such as cell phones, laptops, and displays. They take a chunk out of the global Internet, creating pools of locally relevant information that becomes visible and accessible

through physical proximity. This is a very promising direction for research, and it opens up the possibility of creating truly seamless physical/virtual spaces. We will also need to create better paths from physical to virtual if this future is to become a reality. There are a number of ways to do things like this, sensor networks are one of them [Lifton 2007], they can automatically bring properties of the physical world into virtual spaces. Another way to do this for more direct interaction between worlds would be to create multi-user interfaces that use interaction designs based in reality and blended into physical space. For interfaces like this to be successful, they would need to offer value beyond the brute efficiency of single-user graphical interfaces like the ubiquitous laptop and cell phone. One possible value that interfaces such as this could offer is a form of holistic efficiency; a fluency that is manifest not only as efficiency of communication between user and machine, or user via machine to remote user, but of many users and many machines existing in an ecosystem of collocated and remote, historic and realtime people and things [Löwgren 2007]. A key to accomplishing a goal like this is the ability to re-appropriate the computer interface into one's existing social context.

To illustrate this point imagine an attractive middle-aged man smoking a cigarette and sitting in an outdoor cafe in the Neusdadt section of Dresden Germany. He is wearing a sport coat and sunglasses, and reading the New York Times. (Newspapers, it should be mentioned, are information and communications technologies very similar to modern electronic information delivery systems, one large difference being the timescale at which they function.) In the morning while the sun is rising and it's less busy, he has the newspaper folded loosely and held in his right hand as he reads it while reclined in his chair, the paper resting against his folded legs. In the afternoon when it's busy and he has found a particularly interesting article he spreads the newspaper wide in front of his face. Then in the evening, he is late to return home for dinner, but he has to finish reading an editorial about the intentional manipulation of the consumer price index in America, so he has the newspaper wrapped tightly in his hands leaning forward in the chair. While perhaps overly colorful, this description shows how the form of the paper is weaved into the atmosphere around the reader and how its use becomes an extension of the reader's internal state. The major function of the newspaper, to move information from the world through the newspaper staff and to the reader exists alongside secondary functions, including the moving of information from the reader to those other individuals around him.

In many ways laptops are not as versatile as the aging newspaper medium. While laptops process information at a much faster rate, they lack the affordances, or qualities, for re-appropriation that the newspaper has. This becomes especially apparent when laptop users are working collaboratively in the same physical space. We have become adept at working around the social limitations of the laptop in group settings, finding ways to grab information online quickly and show it to those around us without disrupting the flow of discussion. However, most people have not yet developed or learned an elegant style for networked collaboration while collocated in physical space. More often than not, the use of computers during this form of collaboration is disruptive [Newman 2006]. This is in large part due to the fact that we lack computational tools to support face-to-face interactions. Laptops are inherently single user interfaces, designed within a value structure that privileges the ability for a single operator to access and manipulate a wide range of information efficiently. Succeeding at this task has been the goal of interface designers and researchers in human-computer interaction for over 30 years.

Collaboration

Things have changed a bit in recent times. We have realized the value of building tools for collaboration and understand that humans and technology exist in an ecosystem that will more and more require interfaces that are designed for many-to-many interactions [Harper 2008]. In 1991 Mark Weiser, then head of the famed Xerox PARC research center which spawned the first commercial graphical user interfaces, presented his conception of Ubiquitous Computing. Weiser threw down a gauntlet for human-computer interaction (HCI) researchers, challenging the field to re-evaluate the state of interface design. He said “There is more information available at our fingertips during a walk in the woods than in any computer system, yet people find a walk among trees relaxing and computers frustrating” [Weiser 1991]. Through presenting this concept of calm computing, Weiser pointed out the problems with interfaces that demand all of a user’s attention; interfaces that are not integrated into the environments and use contexts surrounding them.

There have been many research programs initiated since the introduction of ubiquitous computing as a driving vision in HCI that have pushed farther towards seamless interaction. Tangible user interfaces (TUIs) [Ishii 1997] are a notable example. Through

the use of strong metaphorical, analogical, and literal mappings between physical objects and digital properties tangible interface designers create interactions that are based in the physical, non-digital world [Jacob 2008, Dourish 2001]. Qualities such as this enable tangible interfaces to support high levels of collaboration between people and devices. The ability of an interface to support collaboration across the divide of multiple minds, different spaces (physical, virtual, remote, collocated), and various tasks and devices may be the most important feature of interfaces over the next decades.

Collaboration can function on many levels. There is the strong collaboration of a workgroup, working together to solve a common problem, for example a team of urban planners working with a model of a city in which they are rearranging buildings and roads embodied as physical miniatures on a table surface. Collaboration can also happen in other ways. For example, imagine two people seated at the same table, each using a laptop to work on unrelated tasks. This example is not the archetype of collaboration, yet these two people do collaborate—they must work together to share their common resource: the table surface, and they share the same physical space which can also be seen as a type of collaboration.

Collaboration is defined as: *the action of working with someone to produce or create something*. In the second example from the paragraph above the two collaborators can be seen to be creating a shared work space. One could go so far as to say that two people looking at the same painting could be collaborators in as far as the presence of the other helps each to construct their own understanding of the work, that they are working together to create meaning. It is these weaker forms of collaboration that I am interested in this thesis, they become important as we move from thinking about computation as something that happens in the interface between a person and a machine to thinking about it as something that happens everywhere, between many people, things, and machines.

Thesis

Over the remainder of this document I will present the program of research that I have undertaken for my Master's degree in Media Arts and Sciences as a part of the Tangible Media Group at the MIT Media Laboratory. In the next chapter I will explore

some social aspects of interface design through a survey of related ideas and arguments within HCI and the social sciences.

In Section 2 I will present three tangible interfaces that I have developed and then describe an experiment to assess an important and undervalued feature of user interfaces, what I call *external legibility*. External legibility is a property of user interfaces that affects the ability of non-participating observers to understand the context of a users' actions. It is composed of three parts: *feed-forward*, *feedthrough*, and *feedback*. I make two claims: 1) that external legibility is a valuable property of interface design and deserving of greater attention, and 2) that some of the emerging interaction styles within HCI, such as multi-touch, gestural, and tangible, have higher external legibility than current computer interfaces like laptops.

In Section 3 I will present a discussion of future work and a conclusion.

RELATED ARGUMENTS

In this section I will review a number of arguments and ideas within the social sciences and human-computer interaction (HCI) discourses. My goal is to support my first claim, that external legibility is an important feature of interface design. External legibility fits roughly within the scope of seamless interface design in HCI and is a subcategory of general interface legibility. Seamlessness in interface design has usually been thought of in terms of fluency between user, physical interface media, and computational architecture, not necessarily including the non-participants who may coexist among technology users.

In broad strokes, I suggest that the decline in social capital evidenced through empirical evaluation and theory has a possible countermeasure in the form of local information processing. Local information processing describes the way in which collocated people share information through networks. Technologists who have preferred to focus on asynchronous technologies for remote communication have undervalued these networks and consequently we lack the interfaces to support face-to-face interaction.

I propose that recent advances in HCI characterized by the notion of Reality-Based Interaction have value for designing seamless interfaces capable of high external legibility that could support local information processing. I do not agree that technologies progress outside of social contexts, arguing against technological determinism, and suggest that we have only begun to realize the value of next-generation interaction techniques such as tangible user interfaces.

The Decline of Social Capital

In the information age, when networks are dominant and information and communications technologies are high speed and global, we can see a marked change in human society. The individual has become the networked individual [Wellman 2003], Me++ [Mitchell 2003]; an individual who no longer knows his butcher, barber, and neighbor but instead has a concentrated network of close relationships that are never farther away than a phone call, text message, or email [McPherson 2006, Caron 2007].

One can see a push-back starting against these perceived and realized changes. The drive to create multi-user, collaborative interfaces in HCI may come first from a better understanding of interpersonal psychology and observations of existing work practices, but it also reflects on the concept of the networked individual and perhaps offers a solution to some of the negative trends that have arisen concurrently, like the decline in social capital that Robert Putnam called to attention in *Bowling Alone* [Putnam 2000].

According to studies of “discussion networks” by [McPherson 2006] it appears this trend has continued at least through to 2004. Putnam paid most attention to electronic technologies rooted in the home, such as television and telephones. He criticized these technologies for their separating and isolating effect on modern society. The newer technologies of the time (Internet and cell phone) were not well addressed, but he does write:

“What will be the impact, for example, of electronic networks on social capital? My hunch is that meeting in an electronic forum is not the equivalent of meeting in a bowling alley—or even in a saloon—but hard empirical research is needed.” [Putnam 1995].

Whether an electronic forum is the equal to a saloon is still under debate, but there are some interesting data points in this question. The Internet user of today may be part of numerous online communities [Jenkins 2006]. There are gaming communities such as World of Warcraft, communities that surround popular media sites like youtube and digg, social networks like myspace and facebook, and politically like-minded communities involved in blogs like the dailykos.com, talkingpointsmemo.com, and the drudgereport.com.

The role of the Internet in the decline of social capital is still being investigated. There is evidence to suggest that the Internet acts as a countermeasure to the social isolation that appears coincident with the network society [Hampton 2008]. Although some of the early studies on the effect of the internet on social behavior were problematic, for instance a 1998 study by Kraut [Kraut 1998] that showed the internet as an isolating force was shown in 2002 by Kraut [Kraut 2002] to be flawed since the time it took for people to learn to use the web was not taken into account. Studies of internet use by Hampton and Wellman [Hampton 2003] and Hampton [Hampton 2007a] have shown that the web can increase community action and civic participation, that the web can act locally as well as globally.

As traditional media outlets such as television and newspapers are becoming redefined by the rise of the web there is pressure to understand what the differences between these new forms of media and the older ones are. As Lisa Gitelman points out, all media were at one time new [Gitelman 2006]. Each new media innovation has a disruptive effect on culture, and it may be that “media determine our situation” [Kittler 1999], though such words would attribute an agency and prerogative to media that is problematic.

Local Information Processing

Recently the Knight Foundation began the News Challenge [News Challenge] and is offering a series of grants each year to projects that work at the intersection of new media and local, civic participation. The first major grant winner was for the establishment of the Center For Future Civic Media, a collaborative venture between the MIT Media Laboratory and the MIT Comparative Media Studies program.

“The Center for Future Civic Media is working to create technical and social systems for sharing, prioritizing, organizing, and acting on information. These include developing new technologies that support and foster civic media and political action; serving as an international resource for the study and analysis of civic media; and coordinating community-based test beds both in the United States and internationally.” [C4FCM]

This new interest in using the web for civic action comes amidst an explosion of efforts to link the web to geographically defined communities. The rise of hyperlocal web portals is an indication that many web users desire services more tailored to their physical locations. Popular sites such as Gothamist.com cater to cities and act as a collection point for information ranging from public policy decisions to night life attractions and fashion. Chicagocrime.org was started in 2005 as one of the original Google Maps mashups that added in crime statistics made public by the Chicago Police Department [ChicagoCrime]. Users could go to the website to find out how dangerous a given neighborhood is or to simply get information about aspects of the social climate in their city. Everyblock.com [EveryBlock] was created as a successor to chicagocrime.org by the same developer, Adrian Holovaty, and is funded by the Knight Foundation. Everyblock expands beyond crime tracking by offering all types of information about Chicago on a block-by-block basis. In Boston, Social Capital Inc. (a company started after the founder David Crowley read *Bowling Alone*) acts along similar lines [SCI]. They have started web communities surrounding a number of districts within and near Boston with the goal of fostering local non-profits and becoming an umbrella group for web-based participatory civic involvement in the greater Boston area, their aim is to connect real people in physical space.

These sites rely on the existing infrastructure for the web and the existing culture of interfaces for accessing the web. If the web browser then is seen as the medium we are using to access the content online, then the strategy for these sites is to provide new content in order to build community. While the question of significance between medium and content is not entirely clear [Gitelman 2006, Kittler 1999], there is value in McLuhan's reading of the relationship between the two, that "the medium is the message" [McLuhan 1994]. McLuhan's oft recited and remixed phrase is the branding for a larger theory that he presents on the structure and function of media. He suggests that it is the form of the media that proves to have the largest impact on users of that media, not the content contained by that media. For example the television itself would be the change agent for the effects of television on society, rather than the television programs themselves.

Although it now seems quite plausible that websites can build local community and lead to meetings in physical space (take meetup.com for instance), they do so through the same single-user interfaces that separate people when they are collocated

within physical space. This points back to the fracturing of space that Castells references in his theory of space of place and space of flows. It seems there is a need for web services that can both facilitate community online and in the local cafe. Currently, there are many cities that have plans to roll out city wide municipal WiFi in the next few years. Part of the thought behind these plans is that municipal WiFi will serve to give access to all urban dwellers as well as allow community formation around public areas. It is not clear whether such services will accomplish either goal. From studies of existing public parks with free WiFi Hampton et al. found the opposite result, that the parks were dominated by a homogenous demographic of young, well educated males and that there was little room for serendipitous encounters [Hampton 2007b]. They attribute this second finding to both the architecture of the public spaces, the benches and seating areas, as well as the limitations of single user interfaces.

The value of face-to-face interactions in physical space has been underestimated. The thought was that information and communications technologies would cancel distance and that the 21st century would see workers telecommuting from homes outside the frenzy of cities. This has not been the case. In 2008, for the first time more than half of the world's population was measured as living in cities and the trends show that by 2050 around three quarters of the planet will be city dwellers [Cohen 2003].

Rather than a move away from the city, cities have become major information processing nodes [Hall 2006]. The value of clustering near affinity groups in business and leisure has not declined due to high speed information flow, and while the design and function of our computer interfaces and services have changed dramatically over the years to accommodate the way that information flows through virtual networks, they have changed little to accommodate the way that information flows through the slower, yet still extremely important, physical networks. Looking more closely at this second point, one could postulate that the popularity of laptops and cell phones as computational devices of choice has come to pass as an accommodation to the physical demands of users—that we like to be mobile, to work from cafes and buses, and that these qualities are artifacts of the value of physical space and its various properties. In fact the average laptop processes information more slowly than the average desktop computer, so there is evidence of a tradeoff shown there, people are willing to sacrifice some speed of digital information processing to gain some speed of physical information processing—the

Although some may see these interaction styles as disparate innovations proceeding on unrelated fronts, there is evidence that suggests they in fact constitute a third generation of human-computer interaction. Along with colleagues from Tufts University I was part of a research group under the direction of Professor Robert J. K. Jacob, we proposed that these interaction styles share salient and important commonalities, which can help us understand, connect, and analyze them [Jacob 2008].

These new interaction styles draw strength by building on users' pre-existing knowledge of the everyday, non-digital world to a much greater extent than before. They employ themes of reality such as users' understanding of naïve physics, their own bodies, the surrounding environment, and other people. They thereby attempt to make computer interaction more like interacting with the real, non-digital world. By drawing upon these themes of reality, emerging interaction styles often reduce the gulf of execution [Hutchins 1986], the gap between a user's goals for action and the means to execute those goals. We proposed that these emerging interaction styles can be understood together as a new generation of HCI through the notion of Reality-Based Interaction (RBI) and that viewing interaction through the lens of RBI might provide insights for design and uncover gaps or opportunities for future research.

We use the term "real world" to refer to aspects of the physical, non-digital world. However, the terms real world and reality are problematic and can have many additional interpretations, including cultural and social reality. For that matter, many would also consider keyboards and mice to be as much a part of today's reality as any non-digital artifact. Thus, to clarify, our framework focuses specifically on the following four themes from the real world:

- > **Naïve Physics:** people have common sense knowledge about the physical world.
- > **Body Awareness & Skills:** people have an awareness of their own physical bodies and possess skills for controlling and coordinating their bodies.
- > **Environment Awareness & Skills:** people have a sense of their surroundings and possess skills for negotiating, manipulating, and navigating within their environment.
- > **Social Awareness & Skills:** people are generally aware of others in their environment and have skills for interacting with them.

To a greater extent than in previous generations, these four themes play a prominent role in emerging interaction styles. They provide a basis for interaction with computers that is markedly closer to our interaction with the non-digital world. While we believe these themes apply to most people and most cultures, they may not be entirely universal.

Naïve Physics (NP)

Naïve physics is the informal human perception of basic physical principles, or in other words, common sense knowledge about the physical world. This includes concepts like gravity, friction, velocity, the persistence of objects, and relative scale. In the field of artificial intelligence, naive physics refers to an attempt to formally describe the world as most people (rather than physicists) think about it [Hayes 1983]. In the context of emerging interaction styles, user interfaces increasingly simulate or directly use properties of the physical world. For example, a tangible user interface (TUI) may employ physical constraints such as a rack or a slot to guide the way in which physical tokens can be manipulated (e.g. [Ullmer 2005]). Emerging graphical user interfaces, such as the Apple iPhone (see case study below), employ physical metaphors that add the illusion of gravity, mass, rigidity, springiness, and inertia to graphical widgets.

Body Awareness and Skills (BAS)

Body awareness refers to the familiarity and understanding that people have of their own bodies, independent of the environment. For example, a person is aware of the relative position of his or her limbs (proprioception), his or her range of motion, and the

senses involved in perceiving certain phenomena. Early in life, most people also develop skills to coordinate movements of their limbs, head, eyes, and so on, in order to do things like crawl, walk, or kick a ball. Emerging interfaces support an increasingly rich set of input techniques based on these skills, including two-handed interaction and whole-body interaction. For example, many emerging virtual reality applications allow users to move from one place to another within a virtual environment simply by walking on a special track or treadmill (e.g. [Mohler 2007]).

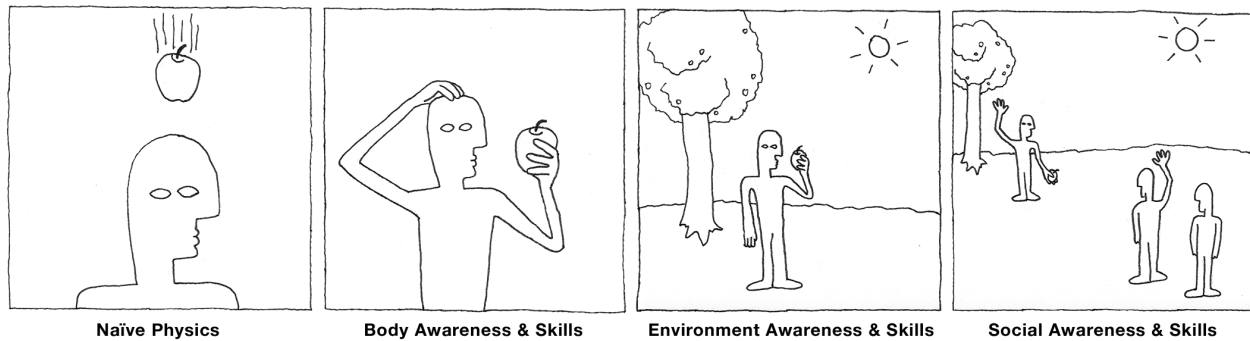


Figure 4. The four RBI themes.

Environment Awareness and Skills (EAS)

In the real world, people have a physical presence in their spatial environment, surrounded by objects and landscape. Clues that are embedded in the natural and built environment facilitate our sense of orientation and spatial understanding. For example, a horizon gives a sense of directional information while atmospheric color, fog, lighting, and shadow provide depth cues [Bowman 2004]. People develop many skills for navigating within and altering their environment.

In the context of emerging interaction styles, many virtual reality (VR), mixed reality (MR), and augmented reality (AR) interfaces along the reality-virtuality continuum [Milgram 1994] use reference objects and artificial landmarks to provide users with clues about their virtual environment and simplify size and distance estimations in that environment [Vinson 1999]. Furthermore, by representing users' bodies in the virtual world, VR interfaces allow users to perform tasks relative to the body (egocentric). Context aware or sensing systems may compute users location and orientation, and display information that corresponds to the user's position in physical space [Bellotti 2002, Benford 2005].

People also develop skills to manipulate objects in their environment, such as picking up, positioning, altering, and arranging objects. Emerging interaction styles often draw upon users' object manipulation skills. For example, in VR and in TUIs, users often select an object by grasping it, either virtually or physically. Many lightweight or tacit interfaces also track manipulation of objects. (Some of these object manipulations draw from the naïve physics and body awareness and skills themes as well.)

Social Awareness and Skills (SAS)

People are generally aware of the presence of others and develop skills for social interaction. These include verbal and non-verbal communication, the ability to exchange physical objects, and the ability to work with others to collaborate on a task. Many emerging interaction styles encourage both social awareness and remote or collocated collaboration. For example, TUIs provide both the space and an array of input devices to support collocated collaboration. Virtual environments (e.g. [Second Life]) exploit social awareness and skills by representing users' presence, by displaying their avatars, and by making the avatars' actions visible.

These four themes serve to clarify the RBI framework although they are not mutually exclusive, nor necessarily comprehensive (though it is hard to find work that does not fit into at least one of them). The fourth theme, Social Awareness and Skills is perhaps the least explored and most rich space for furthering understanding. It is by developing our understanding here that new interfaces will become proficient at local processing.

Technological Determinism

Thinking about the form of technologies and the form of social interaction as so closely linked has not always been commonplace. Before continuing on to describe the design and construction of new technologies in the next chapter it makes sense to step back momentarily and examine some higher level questions. For instance, is it even possible to shift the value structure of information and communications technologies? What is the importance of designing new technologies within a clearly defined framework or constraint system if technology just progresses along its own path, one design being replaced by another that does the job better according to simple metrics?

In thinking about questions like these I am inspired by the work of humanist theorists in the domain of Science and Technology Studies (STS), and in particular the philosopher of technology Andrew Feenberg [Feenberg 1999]. Drawing on the work of two generations of critical theorists and philosophers such as Heidegger, Marcuse, Marx, de Certeau, Weber, and Habermas (among others) Feenberg has spent the last decade developing a critical theory of technology.

Feenberg examines in detail the fetishizing of technoscience and explains how it is incorrect to believe that technological progress proceeds one-dimensionally. He shows that technology is not ahistorical, that it is in fact a part of our social matrix and that choices made in the design of new technologies are influenced by social and political factors. Calling technology ambivalent but not neutral, Feenberg elucidates a democratic rationality where individuals can shape technology and build an alternate modernity. This means that as designers we are not constrained by incontrovertible laws such as efficiency or a natural technical progression. For many of us this is not news; there have been fantastic examples of alternative technologies [Dobson 2004, Dunne 1997]. One interpretation of Feenberg's work is that as designers of new technologies it is up to us to build the world that we want to live in and not accept the often-unexamined constraints of dominant sociopolitical frameworks.

STS is primarily an analytical field, but examples of alternately structured technology development that references the knowledge and techniques from STS may be found in the realms of design, art, and the HCI community. The design team of Fiona Raby and Tony Dunne have developed technologies that question the basic tenets of design and engineering—that products should make people's lives better. By acknowledging that products (automobiles, fast food) often do not enrich lives but often complicate them, Dunne & Raby have freed themselves to design appliances, furniture, and architecture that they claim resembles film noir rather than a typically cheerful, resolved Hollywood blockbuster [Dunne 2001].

The Computing Culture Group at the MIT Media Lab actively developed techniques to defamiliarize and reveal tropes of technology development, borrowing from artistic and counter-cultural practices like the situationists' detournement. Projects have focused on personal and political reinterpretations of technology, from regendered home appliances [Dobson 2004] to technologies

that renegotiate the relationship of the public to the government [GIA] and the US Department of Defense [Afghan eXplorer].

Phoebe Sengers and the Culturally Embedded Computing Group at Cornell have also developed computational systems that integrate knowledge from STS. Building on Phil Agre's Critical Technical Practice [Agre 1997], a work that binds critical reflection with technology development within the field of AI, Sengers presents the idea of Reflective Design, writing: "Drawing on existing critical approaches in computing, we argue that reflection on unconscious values embedded in computing and the practices that it supports can and should be a core principle of technology design." [Sengers 2005]

These examples of thought and work suggest that the practice of technology production is intricately linked with social systems. It then makes sense to read and analyze the value of new technologies using metrics beyond simple measures of efficiency—they should be understood within a larger social context.

Summary

There is evidence that the value of local information processing is gaining import, that new interfaces that function more like the real world are being developed, and that the path of technological development is linked to the social-matrix and nondeterministic. There is room to create works that operate at the nexus of these developments. In the next chapter I will describe three tangible interfaces that my colleagues and I created, each of which address a different aspect of social awareness and skills.

Section 2:

doing

interlude < dorks

Why is it that people who wear bluetooth headsets cannot fail to look dorky? Even the headset made by one of the most famous product designers on the planet, Yves Behar [Behar], still looks funny. It is a gorgeous work of industrial design—the electronics, materials, and form all composed together. Yet whenever I look at someone wearing a bluetooth headset, even a beautiful one, it looks ridiculous if they aren't holding their hand up against it (as if channeling their correspondent bodily) while speaking.

Why does this physical gesture make the technology begin to feel okay? The problem, as I see it, is one of signals. There is very little state change between a bluetooth headset that is on and one that is off, sitting idle until a call comes through. Because the medium of the headset is so limited in the way that it can be used to communicate to the people around the wearer (those who one might think are not the target audience of this technology's use context at all) it fails at its goal of communicating to the person on the other end of the phone. It is not because the headset itself is not beautiful—Behar's headset is gorgeous, suitable for mounting on a pedestal and staring at—it is because our understanding of the work these devices do is limited, we don't understand how the device fits within the holistic social context of the wearer.

Imagine, instead, that the bluetooth headset has two physical states: off and on. In the off state the microphone end of the headset is flipped up and rests against the wearer's head, idle. In the on state the microphone end flips down towards the wearer's mouth, so it is closer to the sound coming from there. Imagine that

the physical difference between these two states is pronounced enough that observers can see the difference easily from a good distance. Imagine some more, that it is a nice looking headset, and well made. Now will it seem as dorky? If not, perhaps that is because it can now accomplish one of the headset's hidden purposes—communicating to the non-participating observers that we share physical space with. It's functioning has become externally legible.

TANGIBLE USER INTERFACES

In 1997 Ishii and Ullmer presented their paper *Tangible Bits: Towards Seamless Interfaces between People, Bits and Atom* in the proceedings of the SIGCHI conference on Human Factors in Computing Systems in Atlanta Georgia [Ishii 1997]. Although there were a number of physical computing interfaces made before this time [e.g. Fitzmaurice 1995], this is commonly considered the birth of Tangible User Interfaces (TUIs, tangibles). Tangibles are computer interfaces that blend the physical world and the virtual world together by using metaphor and design principles from outside the traditional scope of human-computer interaction.

Over the past ten years there have been many instances of tangible interfaces developed within the research community, as well as a number of frameworks [Ullmer 2001, Fishkin 2004, Sharlin 2004, Hornecker 2006], toolkits [Klemmer 2004, Greenberg 2001], special issues of journals, and an ACM conference [TEI]. Hornecker and Burr present four themes for tangible interaction in their 2006 framework [Hornecker 2006], these are: tangible manipulation, spatial interaction, embodied facilitation, and expressive representation. These themes illustrate the variety of ways in which tangibles have been thought of and developed over the years.

In this section I will present three of the tangible interfaces that I have created over the past few years: the Tangible Video Editor (TVE) [Zigelbaum 2007], Slurp [Zigelbaum 2008a], and SpeakCup [Zigelbaum 2008b]. I created the Tangible Video Editor for my undergraduate thesis at Tufts University along with colleagues Michael Horn, Orit Shaer, and my undergraduate advisor Professor Robert J. K. Jacob. I created Slurp and SpeakCup in the Tangible Media Group at the MIT Media Lab under the guidance

of Professor Hiroshi Ishii. My collaborators for the Slurp project were Adam Kumpf and Alejandro Vazquez; my collaborators for the SpeakCup project were Angela Chang, James Gouldstone, and Joshua Jen Monzen.



Figure 5. Working with the Tangible Video Editor.

The Tangible Video Editor

Traditional film editing systems such as Moviola and Steenbeck provide editors with a rich sensory environment that allows them to utilize enactive (muscle) memory and haptic (force) feedback. The task of cutting and splicing film into sequences for playback involves the use of physical tools such as cutting arms, and taping stations. The affordances of these tools help convey both their purpose and their means of use. In contrast, current state-of-the-art, NLE (non-linear editor) software such as Final Cut Pro, Premiere, and AVID provide filmmakers with little of the physical richness employed by their predecessors. However, they provide editors with powerful features such as the ability to reuse shots without making a new print from a negative, to undo

actions as quickly as they were done, and to initiate new projects with the click of a button.

The Tangible Video Editor (TVE) project is, in part, an attempt to combine the benefits of traditional, physical film editing with the advantages of digital, non-linear film editing. The implementation of the TVE presented here is a tool for basic video editing and storytelling, designed to engage multiple users in a collaborative process, and encourage the exploration of narrative ideas. We developed an active token approach to tabletop interaction to support this goal and created a malleable user interface that improves upon some of the limitations we found in similar projection-based interfaces.

Related Work

There are many video editing interfaces in the HCI literature, and a few that use physical interactions for related tasks. An early example of a physical interface that could be used for editing video is Video Mosaic by Mackay and Pagani [Mackay 1994]. Essentially an enhanced storyboarding tool, Video Mosaic addresses the issue of representing time spatially by combining the benefits of real-world interaction (in this case with paper) and the power of digital manipulation. A more recent example is the mediaBlocks [Ullmer 1998] system, which provides an interface for capturing, editing, and displaying multi-media content using passive wooden blocks as tokens. Although mediaBlocks is not a video editing system, the media browser and media sequencer functions are particularly relevant to our work, they allow users to view and sequence digital video by arranging the blocks within a physical device.

Tangible Video Browser [Tangible Video Browser] is another interface that uses physical tokens to represent video clips. When the tokens are placed on the browser interface, they become objects that can be manipulated to navigate through the clip itself. TellTale [Ananny 2002] and EnhancedMovie [Ishii 2003] are two other relevant interfaces. TellTale, a toy caterpillar whose body is made up of segments that can be arranged in any order, is an interface for creating sequences of audio clips. EnhancedMovie features an augmented desk, which allows users to make a movie by editing a sequence of pictures using hand gestures. Hand gestures (such as closing all fingers above an image) allow users to grab, release, select, and browse clips displayed as images on the table. Finally, Moving Pictures [Vaucelle 2005] is a tangible interface

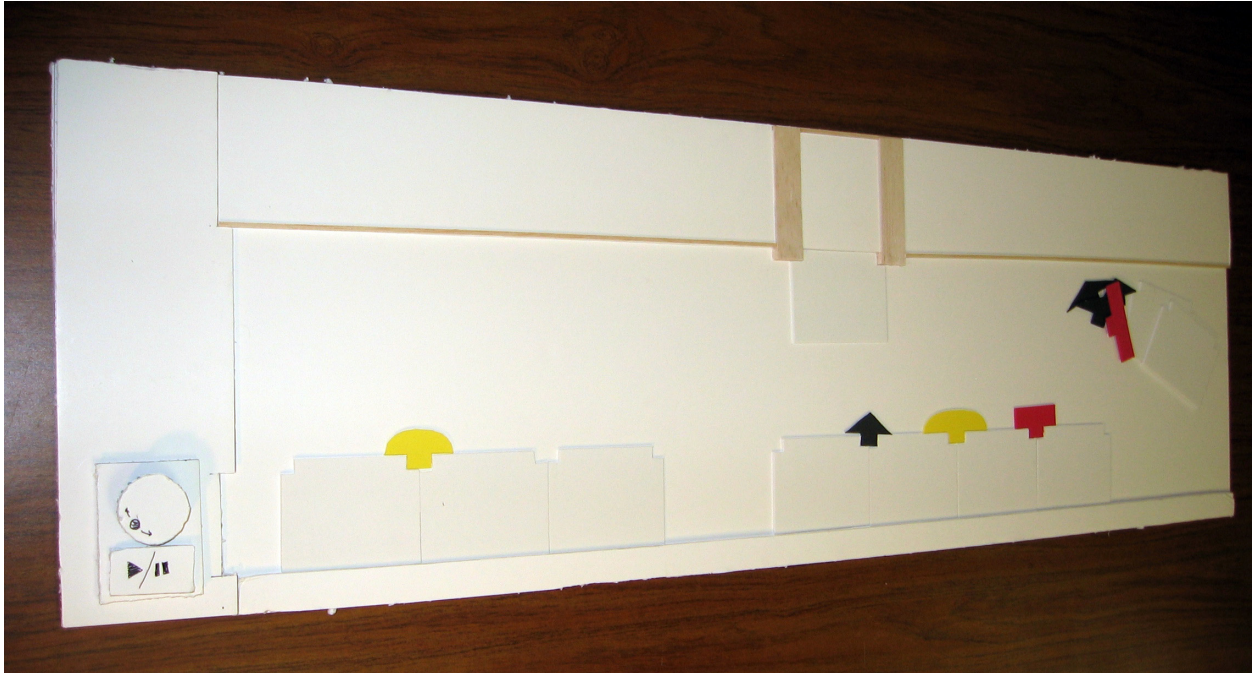


Figure 6. Foam prototype.

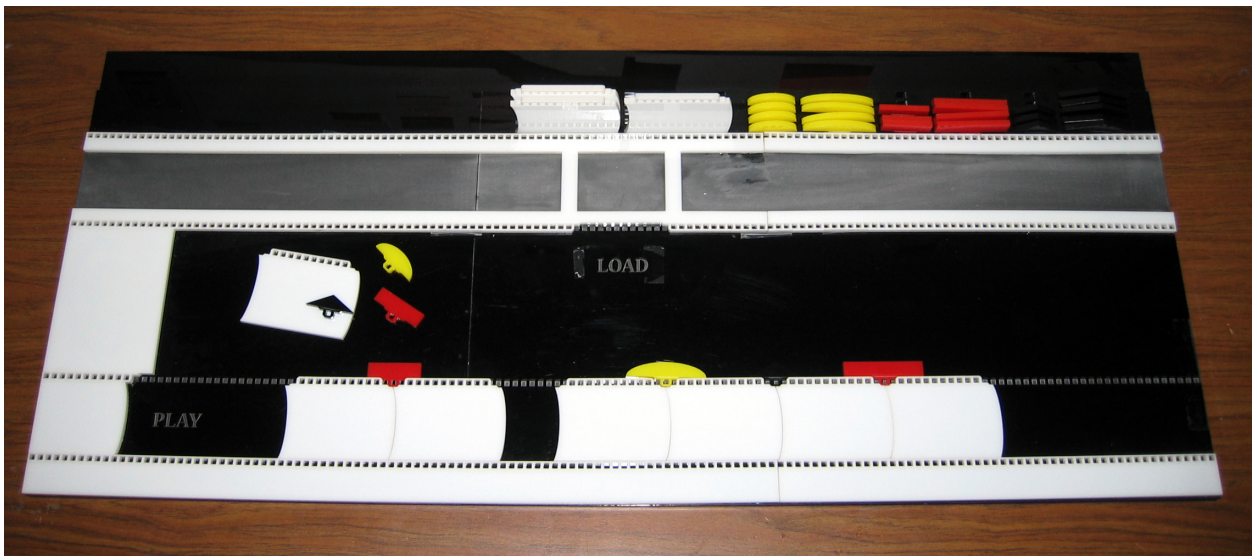


Figure 7. Plastic prototype.

aimed at allowing young users to collaboratively create, manipulate and share video content. The system consists of a video station, a screen, an interactive table, and a set of tokens. To create a movie, users can move between three modes of operation: Shooting, VideoJockey and Storyboard. Each of these modes supports collaborative physical interaction guided by a GUI.

Design Process

The current version of the TVE is the result of an iterative design process where we constructed and tested multiple prototypes with user groups. Our first studies took the form of low-fidelity prototypes made from paper, foam core, and laser-cut acrylic (Figures 6 and 7). We used principles of tangible interaction in our design as they are widely cited to aid in collaboration [Ananny 2002, Hornecker 2002, Suzuki 1993]. After testing these prototypes along with another implementation using standard projector and computer vision techniques, we hypothesized that a reduction of global constraints would foster a freedom of interaction that would aid collaboration. In short we wanted to provide an interface that users could organize and manipulate in their own ways [Hollan 2000]. Projector-based systems have limited resolution and confine users to a predefined workspace and set of interactions, so we developed a new interaction technique based on active tokens.

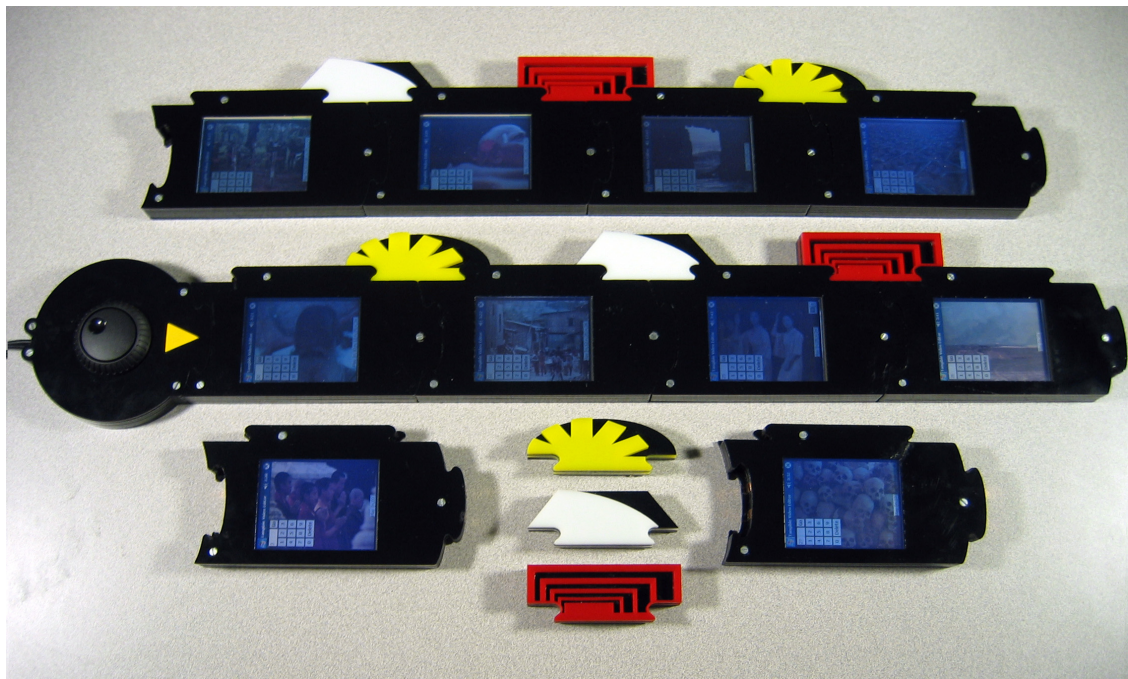


Figure 8. The components of the TVE.

Implementation

The current version of the TVE features a play-controller, 10 clip-holders, and 15 transition connectors (Figure 8). These components operate in conjunction with a desktop PC and a monitor for viewing video content. Each clip-holder consists of a Compaq iPaq Pocket PC mounted inside a plastic case. By using multiple

small screens, the TVE interfaces' resolution increases with the addition of each clip.

Physical Syntax: Designers of tangible interfaces often use form and mechanical constraints [Ullmer 2005, Shaer 2004] to express digital syntax. We designed the TVE around a jigsaw puzzle metaphor to offer both a cultural hint and a physical constraint that suggests that components can be connected in a sequence [Fishkin 2000]. Furthermore, it is not necessary for users to learn special commands to perform operations such as removing a clip from a sequence [Gutwin 1998].

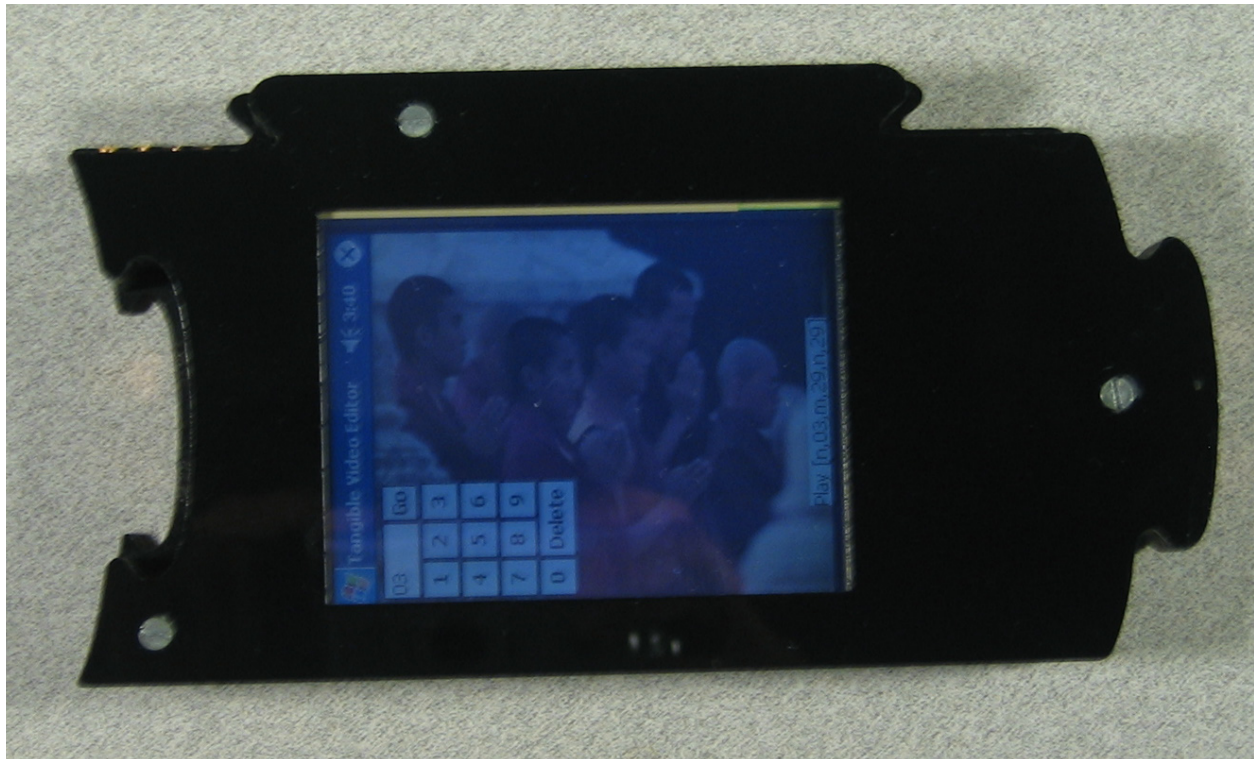


Figure 9. A clip-holder.

Case Design: The cases for the clip-holders, transitions, and play-controller (Figures 9 and 10) were constructed from layers of 1/8 inch thick extruded acrylic sheets, cut with an industrial laser cutter. The iPaks were removed from their original cases and mounted inside the clip-holders. A top layer of acrylic holds the iPaq in place and hides its internal components. Only the touch screen display is visible to the user. Copper connectors run along the outside edges of the cases where the clip-holders interlock with transitions and other clip-holders. The copper connectors are wired to the audio-out, audio-in, and application shortcut buttons of the iPaks. When two clip-holders are joined, an electrical con-

nection is formed between the audio out channel of the right clip-holder and the audio in channel of the left clip-holder. Likewise, when a transition is placed between two clip-holders, it closes a circuit to press one of the iPaq's application shortcut buttons. The software running on the iPaqs registers these events and inserts the corresponding transition data into the clip-holder data stream. The TVE supports three types of transitions: minimize, rotate, and fade.

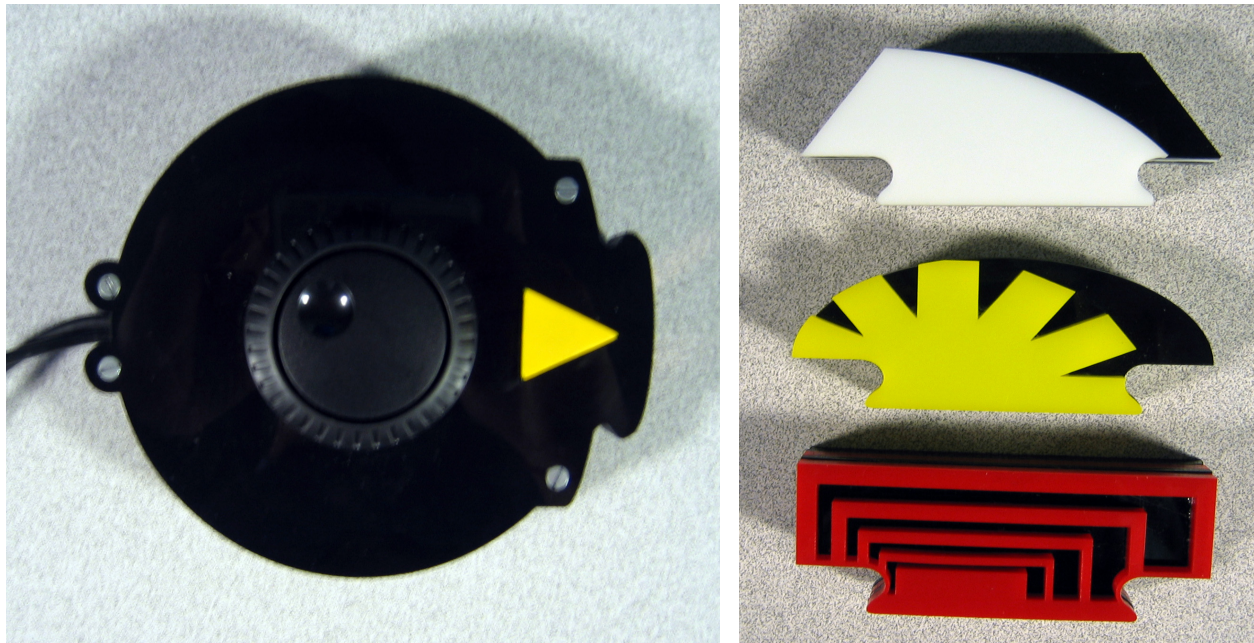


Figure 10. The play-controller (left) and three transitions (from top-right: fade, rotate, minimize).

Play-Controller and PC: The play-controller is a circular device that includes a yellow play button and a jog wheel. Users connect the play-controller to a clip-holder or the beginning of a sequence of clips holders that they want to view. Pressing the play button forwards the data stream to the desktop PC and triggers an application to play the clip sequence with transition effects on the desktop monitor. The application, written using Macromedia Flash, dynamically composes the movie from pre-loaded video clips.

Clip-to-Clip Communication: A clip-holder must be able to pass information to adjacent clip-holders about its current clip ID number and transition. This information flows along in a sequence from the rightmost clip-holder to the leftmost clip-holder, through the play-controller to the PC. Each clip-holder in the sequence receives a data stream from its neighbor to the right, appends information about its own clip and transition, and then passes the

new data stream to its left. This data stream is transmitted in a continuous loop. We use the iPaq's audio channels (microphone and speaker jacks) to encode ASCII data in a frequency modulation format.

Observations

We observed 36 subjects working in pairs to evaluate whether the TVE provides benefits in supporting collaborative editing, users' engagement, and exploration of alternative narratives. Eighteen subjects interacted with the TVE and the other 18 subjects interacted with a typical graphical video editor. The subjects were 18-74 years old and were not paid. They came from a variety of backgrounds including college students, professionals, and retired persons.

The subjects' task was to work with a partner to create a short movie from pre-existing video clips. We provided 28 short video clips for participants to work with and three types of transitions to add between clips. Participants were randomly assigned to one of two editors: the TVE and Microsoft Movie Maker (MMM). MMM is a typical non-linear editing application for home video. Participants filled out both a pre-test and a post-test questionnaire about the task and the editing process. Each session was filmed for further analysis.

Overall results were positive for the TVE. Subjects reported that they were more excited about creating their films and more enthusiastic about the editing process in the TVE group than the MMM group. They also reported that they had more fun.

During our observations we noticed very different behaviors between groups working on the two interfaces. As expected we found that teams working with the TVE shared the work more evenly. Our analysis of the session video recordings showed that subjects in these teams took more turns speaking than MMM subjects, i.e. one team member did not dominate the conversation in the TVE condition as frequently as in the MMM condition. Subjects working with MMM would most often designate one person to operate the keyboard and mouse for the duration of the task. TVE subjects, on the other hand, most often shared all of the work, including operation of the play button.

An unexpected behavioral difference between the groups became apparent within the first few minutes of each session. Subjects working with MMM often began the sessions in silence

watching the available clips one after another where subjects working with the TVE often began the sessions discussing possible story lines together and organizing subsets of clips in their work space. Subjects using the TVE organized clip-holders in their workspaces in meaningful ways to add a layer of information, e.g. clips nearer to their bodies were often more important.

MMM users reported that they became more familiar with the content of every video clip than participants working with the TVE. This was confirmed by our observations—we noticed that TVE users would often view only half to three quarters of the available clips, and these they rarely watched more than once. MMM users would often watch the clips over and over again, and the teams usually watched every available clip. This was probably due to the fact that users of MMM only had to double click on a thumbnail of a clip to watch it where TVE users had to attach a clip holder to the play controller to view it. Although the TVE users did not view all of the clips they began discussing their narrative very early and explored multiple possibilities. One of our goals for the TVE was to encourage this type of exploration, so we were happy to see this during the evaluation.

Some of the users complained that the TVE did not offer enough functionality. Their main suggestion was to add the ability to cut and merge clips. Other common suggestions were to add audio editing and more of the functions found in GUI editors such as color correction, cropping, and speed control.

Conclusion

The TVE demonstrates a new implementation for tabletop interactions that uses active tokens. By embedding handheld computers in the interface components the TVE can be used on any surface under various lighting conditions and without the need for an augmented surface, computer vision system, or video projector.

Our observations of users working with the TVE during a collaborative editing task suggest that our departure from projection-based tabletop interactions was successful and warrants further investigation. We were particularly interested in the way users arranged the TVE components in their workspace and believe that this behavior suggests that the malleability of the TVE interface was beneficial.

Slurp

As ubiquitous computing continues to spread, researchers have looked to the features of the world in which computation takes place in order to inform the creation of new interfaces [Jacob 2008]. Tangible user interface (TUI) [Ishii 1997] has emerged as a powerful concept for blending computation with the real world. Much of this power comes from the use of metaphor [Fishkin 2004], affordances [Djajadiningrat 2002, Norman 2002], physical space [Sharlin 2004], and physical syntax [Shaer 2004]. Nevertheless, we have not begun to throw out our laptops and cellphones. The very properties that make tangibles strong also limit them—solid forms embedded in persistent physical space are less mutable than pixel-based displays. Tangibles don't scale well, and although capable of manipulating abstract data [Holmquist 1999, Ullmer 1998, Ullmer 2005], the use of indirect mappings reduces the benefit of physicalization, as shown in [Fishkin 2004].



Figure 11. Slurp, held for use.

GUIs are strong where TUIs are weak. They scale well, they are great for manipulating abstract data, and they have high plasticity (capable of doing very different tasks through the same interface). How can we get the benefits of both paradigms in a seamless interaction design?

In this section we present Slurp, a tangible interface for interactions with locative media, and discuss the design issues that arise when attempting to physicalize abstract digital information. Based on the affordances of an eyedropper, Slurp provides haptic and visual feedback to extract digital media from physical objects in everyday environments. Once extracted, media can be injected into displays such as computer monitors or speakers. Our goal is to explore a novel interaction technique for the future of ubiquitous computing and reflect on the ideas and challenges encountered along the way.

Locative Media Now and in the Future

As computation spreads further into the real world one can envision a future where every physical object is created with a digital object attached to it. For example, it would be nice to have a spec sheet for the light bulb you just bought incorporated directly into the light bulb itself, or to have media files showing the history of an antique couch come embedded in the couch rather than on external media. These media files could be modified or added on to; in the couch example the new owners could add their own experiences to the couch's history. The information stored in the physical object could be a simple url, allowing for the participatory culture of the current Internet to extend into physical space and objects.

Currently RFID tagging can be used to achieve the above scenarios, but in the future other technologies may become more prevalent. Regardless of the technical details for how digital information will pervade the physical world we will have to develop new ways to interact with it. Imagine that every physical object in your living room is a container for digital information and you want to access the digital object attached to a mug on your table. One could quickly imagine a couple of ways to use a GUI for this task. A combobox or other list generating widget would work, but there could be hundreds of items in the list, or more, and if there were a few mugs on the table it might be difficult to know which list item corresponds to the correct mug. Another method would be to use a graphical map of the room and its contents

with all of the physical objects correctly identified and located by the computer, this is an interesting possibility though it has some drawbacks.

Before detailing the issues with the second case imagine a third alternative, rather than using a GUI the user just points to the mug, loading the embedded digital media onto a nearby computer. This third option makes use of the existing spatial relationships that human beings are well suited to understand, and points to some of the problems with the graphical map solution. Even if the map were implemented perfectly the user would have to resolve the translation from 3D physical space to graphical space, relying on a virtual target that is not coincident with the physical object in question—the mug. It is not too difficult to imagine using the graphical mapping interface, and in some cases it may be preferable, but why not go to the source when it's right in front of you?

Tangible Interfaces and Abstract Digital Media

A central question in this work is how to use physical affordances, metaphor, and spatiality to bridge the intermediary space between the graphical world and the physical world. This is not a new question. Ishii and Ullmer asked it when they presented their vision of Tangible Bits [Ishii 1997], as have many researchers since then. Terrenghi's work examining the affordances of gesture-based direct manipulation [Terrenghi 2005] points to relevant differences between interaction with the physical world and graphical displays. The widgets common to GUI desktop environments are not necessarily suitable for extension into physical space, nor are the metaphors that they rely on.

The use of metaphor in human-computer interaction (HCI) has been widely noted by researchers [Blackwell 2006, Djajadiningrat 2002, Fishkin 2004, Jacob 2008, Norman 1988]. Functioning as something more than a literary trope, the use of metaphor in HCI is problematic—"Novel metaphorical UIs, despite their popularity, have seldom been natural or intuitive" [Blackwell 2006]. When a designer employs metaphor to create an interface based on existing interactions, a third thing is born. The use of metaphor in HCI, though not necessarily intuitive, can serve to inform users by building on existing schemas (collections of generic properties of a concept or category) making "it easier for you to learn a new concept by tying it to a concept that you already know" [Heath 2007].

Liquid Metaphor

The digital objects that we use on a day-to-day basis must be manipulated indirectly with specialized tools and, in practice, can never be touched. Humans have many sensory channels for interpreting the world; however, due to practical constraints GUIs have remained the dominant interaction technique. In confronting the problems of how to touch digital media we must choose methods to physicalize that media, this is particularly challenging when considering abstract digital objects. Haptics has proven an exciting field for this end [Smith 2007] as has the use of tangible interaction and physical metaphor. One approach is to treat abstract digital media as water. Water, like some digital objects, is difficult to manipulate with bare hands. We can splash it around, but we need specialized tools to perform precise operations with it.

Abstract Digital Media

It has been easier to physicalize certain types of digital media in tangible interface design than others. Digital objects with spatial properties (such as building models in CAD software [Underkoffler 1999], molecules [Fjeld 2007], or geographic maps [Arias 1997]) lend themselves to physical form. Abstract digital media is difficult to embody tangibly and is therefore usually confined to screen-based interaction techniques, such as GUIs. More abstract digital objects (such as music, video, text, or data sets) can benefit from association with physical form through the use of containers and tools as defined in [Holmquist 1999]. In the musicBottles interface [Ishii 1999], glass bottles are used to contain sound, in one scenario three bottles are used, each linked to a musician in a three-piece jazz ensemble. Open one bottle and you hear the drummer, open another and the pianist joins in. In the Tangible Query Interface [Ullmer 2005] wheels, pads, and racks are used as tools for parametric viewing of a data set.

A problem with physical interface treatments of abstract digital information is that the mappings between digital and physical objects lack the tight coupling and affordances found in the use of phicons or tokens. We have tried to mitigate this issue by using haptic feedback in an active tool (Slurp) that treats abstract digital media like a fluid that can be slurped up and squirted out. Our approach is to embody abstract digital media in physical form, in the hopes of providing difficult-to-quantify benefits for users, such as enhanced feelings of ownership, improvisational support, and changes in user relationships with, and planning of, inter-

actions. Some of these benefits have been studied already [Terrenghi 2007], and although not attempted here, we feel there is much value in future studies.

Related Work

David Merrill's invisible media project [Merrill 2005] does something very similar to the example mentioned earlier, where the user is gesturing at a mug. He used IR beacons, headsets, and pointing devices to enable users to access digital media that is associated with physical objects by pointing or gazing. There are a number of related projects that use RFID, graphical symbols, or other addresses to link to digital information [Yellow Arrow, Want 1999]. These systems allow users to access digital information from tags using cell phones or custom hardware such as Merrill's headset which plays audio content related to the object targeted.

There are other systems that allow the user to choose both the input and output for their media, such as mediaBlocks [Ullmer 1998] and Pick-and-drop [Rekimoto 1997]. In mediaBlocks small wooden blocks are associated with digital media and are used for transferring images or video from one device to another, or sequence slides in an editor. Users of Pick-and-drop can move files between touchscreen displays by tapping them with a stylus; this transfers the file across the network. TOOL DEVICE [Ikeda 2003] is similar to Pick-and-drop in that they are used to move songs and other media files between touchscreens, they differ by providing local haptic feedback and using the affordances of a syringe, chopsticks, and a ladle.

Contribution

Slurp differs from existing work in a few ways. Slurp allows for the extraction of digital media from physical objects and the selection of an appropriate display device to access it from. It contains the digital information rather than working as a physicalized hyperlink. Slurp also provides local haptic and visual feedback removing the need for visible tags on accessible physical objects.

Slurp combines the properties of containers and tools for manipulating digital objects. There are two parts to the system: Slurp (digital eyedropper) and the IR nodes (Figure 12). Users hold Slurp in one hand with its bulb between the thumb and forefinger. They can extract (slurp up) media by touching Slurp to a screen, pointing it at a remote display or object and squeezing

Slurp's bulb as if the user were sucking up a volume of water. After a digital object has been acquired by Slurp via the extraction process users can inject (squirt out) the digital object by touching Slurp to a screen or pointing it at a remote display and again squeezing Slurp's bulb. A small pointer is passed between Slurp and the IR node; the related files are transferred in the background over the network.

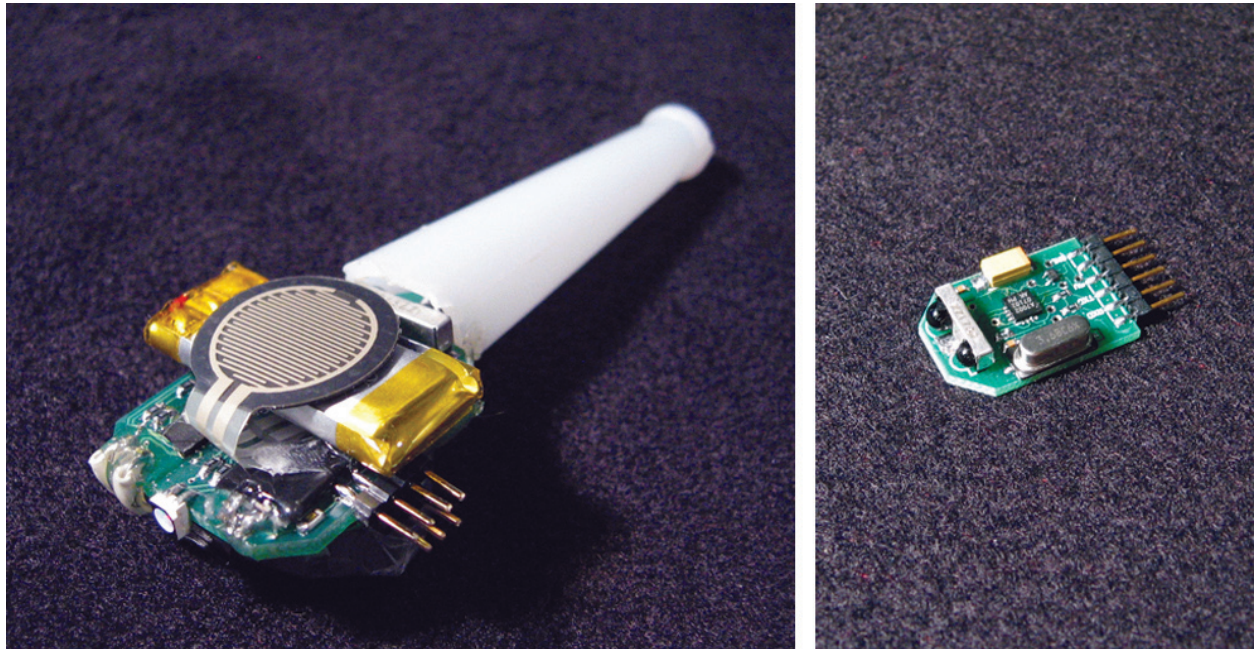


Figure 12. Left: Slurp hardware before cast in silicone. Right: Infra red communications node (IR node).

Slurp, A Digital Eyedropper

Slurp has two parts, a stem and a bulb. The stem houses a tri-color LED to represent the state of targeted displays. The bulb contains the printed circuit board and batteries to run Slurp, a force sensitive resistor (FSR) to measure the pressure of squeezes, a vibrotactile actuator for haptic feedback, and a tri-color LED to represent digital objects extracted by Slurp. The physically rigid hardware (PCB, sensor, etc.) is fully encapsulated in a soft silicone rubber to afford squeezing and to mimic the experience of using a standard eyedropper with a rubber bulb.

IR Nodes: The IR nodes use infrared data communication (IrDA) to act as gateways between Slurp and the objects or devices with which it communicates. Each IR node is attached to an object or display (visual, auditory, or other) powered by a PC. Less expensive, self-contained IR nodes running from a microcontroller are also possible and could be attached to computation-

ally passive unidirectional objects such as buildings, artwork, or trees for locative-media interactions.

Multisensory Feedback: The vibrotactile actuator is used to generate a haptic narrative that provides feedback on Slurp's state and mirrors targeted objects. Users can seek out digital signals in a given space; this interaction is similar to the beeping of a metal detector or the sounds from a Geiger counter to indicate the presence of objects invisible to the user. Once a digital object has been targeted, Slurp displays different feedback for discrete or continuous objects. Discrete objects generate a short burst of vibration and a static color in the stem. Continuous objects (such as video media) generate continuous feedback to mirror their current state. For a video playing on the screen, the color of each frame is averaged to a single pixel and displayed in Slurp's stem while the audio amplitude is converted to vibrations in the bulb. For playing audio objects (like a song on the radio) only continuous vibration feedback is generated in Slurp, the stem displays a static color.

When Slurp is empty and pointed towards an IR node Slurp's stem illuminates and mirrors the color of the target object in the same way that the stem of an eyedropper takes on the color of the liquid it is placed in. During extraction light moves from the stem to the bulb, staying in the bulb until injected. The silicone bulb acts as a diffuser for the LED; the light appears to fill the bulb. After informally testing Slurp with users we added a subtle flashing light in the stem for extra feedback, when Slurp is full and aimed at an IR node, the stem lights quiver as if the liquid inside is bubbling to get out. During injection, light moves from the bulb to the stem and then fades out. When Slurp is full, soft presses on the bulb injects the data object while retaining it in the bulb (which remains lit) for further injections. Hard presses inject and clear the data. This feedback is directly based on the use of an eyedropper; when it's full small presses release only some of the fluid.

Locative Media

As computers become more pervasive through the physical world, the spatial relationships between computational devices gain importance. Interfaces that make use of spatial relationships can reduce the ambiguity associated with navigating multiple devices through common GUI widgets.

Part of location-based or locative media is linking digital objects to locations in the physical world. This is often accom-

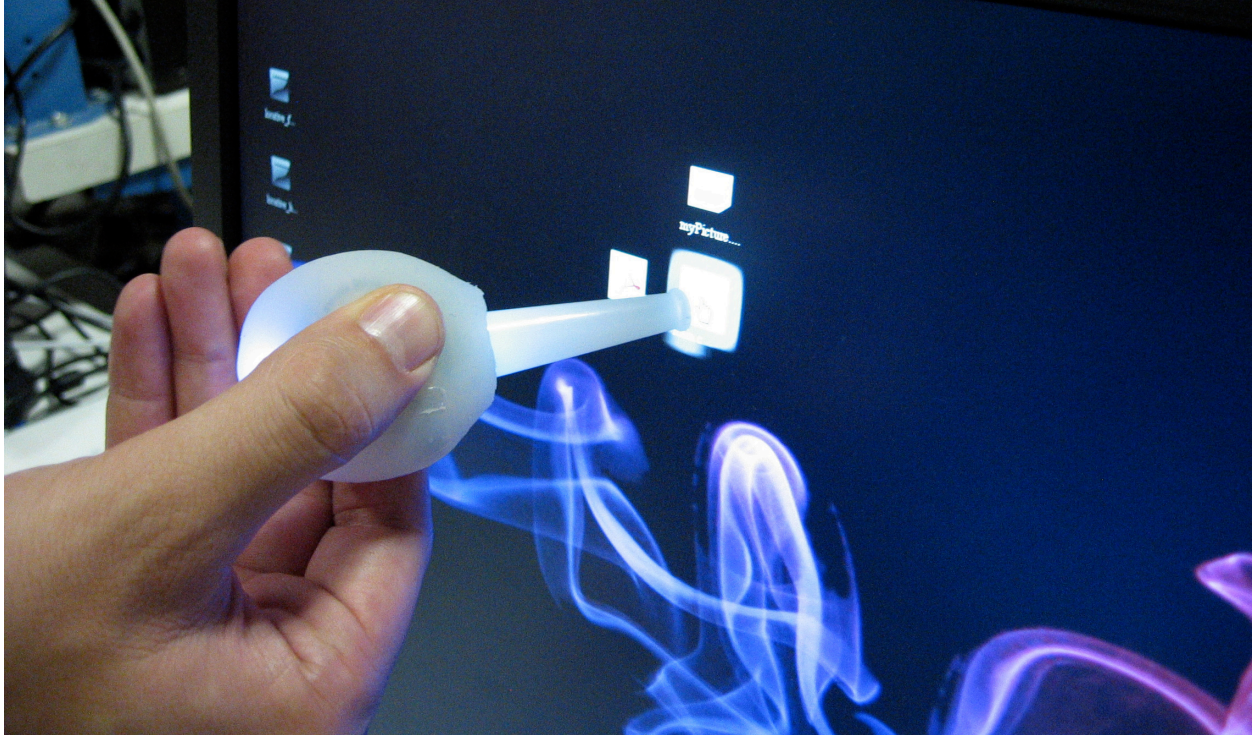


Figure 13. Slurp injecting a digital object onto a screen.

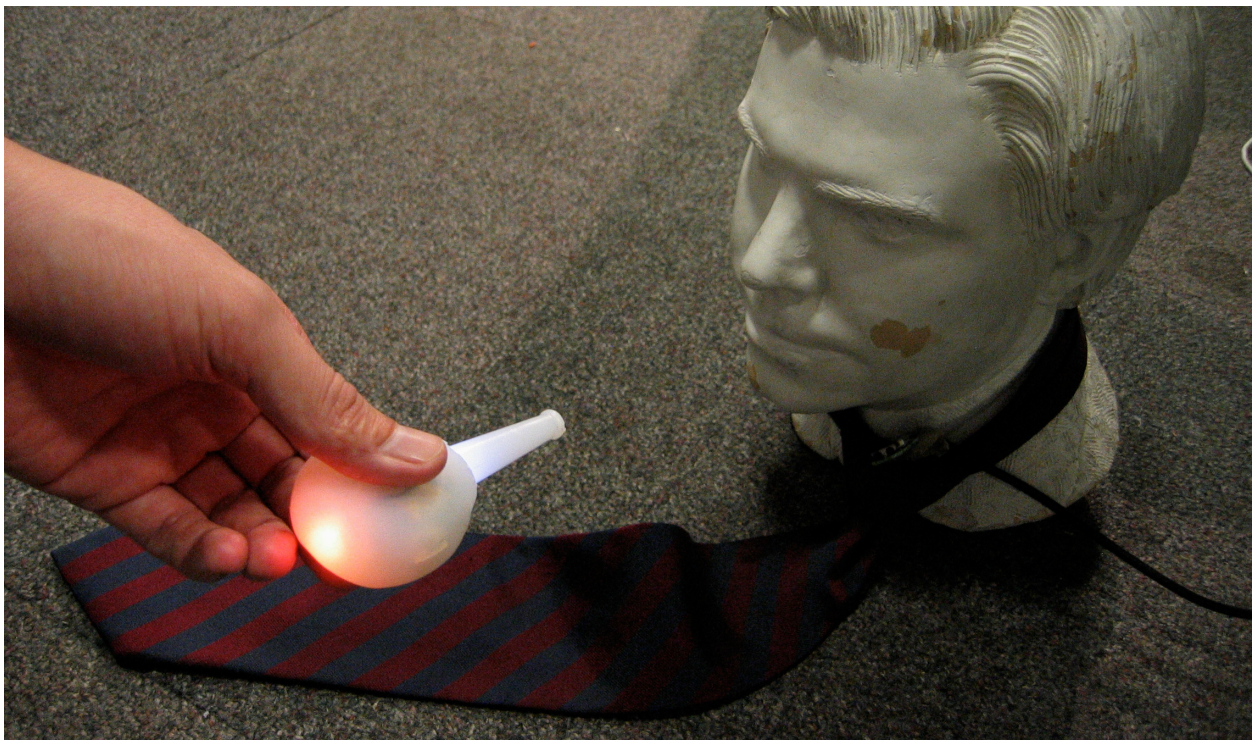


Figure 14. Slurp extracting a digital object from a sculpture.

plished using cameraphones and 2D barcodes or text messaging. The barcodes act as pointers to locations on the web or a type of physical hyperlink. In the future (and perhaps in the present) rich media will be linked to all types of physical objects, locations, and people. Slurp can be used to aggregate these digital objects.

We attached IR nodes to objects in our lab space. Since the nodes project IR out into space the user can wave Slurp around and point it at various objects to remotely identify where digital objects are present in a physical version of exploratory search [White 2007]. When Slurp is pointed at an object that is digitally active, in this case an image from a music video, Slurp reacts similarly to the previous scenario, by vibrating and lighting up. Then the user can extract the object and inject it into a container for later. This container could be a watch or cellphone with extended features for immediate viewing, but as proof-of-concept we used a PC.

Smart-Office

In developing Slurp we realized it could also be used similarly to a USB drive or Pick-and-drop for moving files directly from one screen to another. In the smart-office scenario it is common for workers to use digital whiteboards, large shared displays, PDAs, smartphones, laptops, PCs, and audio systems collaboratively and concurrently. The problem of how to move and share data objects across these displays has been well studied [Nacenta 2005, Rekimoto 2005]. In a detailed study comparing techniques for multi-display reaching by Nacenta et al. [Nacenta 2005] the authors found that systems with local feedback, 1-to-1 mapping between digital and physical space, accuracy, and remote operation were preferable to other systems.

We set up two desktop PCs and an audio system with IR nodes. We tested Slurp by moving video, audio, and files between the displays. A touchscreen display would be able to identify the position of Slurp against its screen, but since we didn't have any we simulated touchscreens by using the mouse and moving it to match Slurp's position. This allowed us to get a sense of screen-to-screen operations. By using Slurp's IR channel to tell the computer when to extract and inject the files along with the mouse position we could grab files directly off of one screen and deposit them onto the other. To provide graphical feedback we built a desktop in Adobe Flash. We created icon animations for extraction and injection of files as an additional notification of the system's state.

These animations also enhanced the feeling that Slurp was pulling something out of the screen or depositing it into the screen, rather than just triggering a file transfer in the background.

In addition Slurp can work remotely with playing video and audio (in this case these media types filled the screen) by pointing in the direction of a display. Notably, Slurp works with non-visual displays (in this case speakers), a feature not implemented on many other multi-display reaching systems.

GUI—TUI Blending

A logical next step for Slurp would be to add it to existing tangible interfaces. Siftables [Merrill 2007] is a tangible sensor network platform based on multiple, small graphical displays. By adding Slurp to the Siftables system users could navigate large libraries of video media on a GUI and extract them directly from the monitor. Slurp could be used to move video between devices, leveraging the scalability of GUIs and the spatial, tangible properties of Siftables. We could also add Slurp to musicBottles, extending its capabilities in a similar fashion. We are currently exploring these options for future work.

Discussion

We presented Slurp at our lab's open house. Around 50 people used it informally during the 2-day event. Through this qualitative demonstration we received numerous suggestions and critiques. One user wasn't sure why someone would want a separate device just for accessing digital information from physical objects; he wondered why it wasn't part of a cell phone. It seems reasonable to think of adding similar functionality to a cell phone or camera, though there would be tradeoffs in doing so. Special purpose, limited-functionality devices have compelling benefits over convergence devices, but they can be less practical.

One could use a gestural interface, cell phone, or camera for locative media, though the presence of a single purpose, tangible tool simplifies the interaction. In Zhang, Fishbach, and Kruglanski's recent paper about multi-purpose devices [Zhang 2007] they showed that a pen that also functioned as a laser pointer was less likely to be used by participants than a pen that was just a pen. By adding additional functionality to a device it adds confusion. Gestural interaction requires remembering which gesture is used for which action, and the possibility of other gestures could confuse the user. The same could be said for multi-touch displays. Simple

physical devices may be preferable to multi-featured interfaces in an age of complex interactions. Rather than add additional functionality to Slurp, such as the ability to store multiple files, we feel that creating richer and clearer feedback would be the preferred next step.

Some users questioned the use of a liquid metaphor as the basis for Slurp's interaction design. The use of a liquid metaphor cannot account for all of the functionality found in the digital world. For instance, liquids are difficult to separate once mixed. On the other hand some users found the liquid metaphor to be magical, and gasped as Slurp spit out files directly onto a monitor. We have used the metaphorical or analogical use of liquid as a point of departure for touching abstract media; in practical use design tradeoffs must be made. Basing an interaction on existing physical models will always be problematic if the interface doesn't function exactly in the same way as its model. Nevertheless, as show in the recent work on Reality-Based Interaction [Jacob 2008], when thoughtfully applied, reliance on existing skills and knowledge in an interface design can provide benefit for users.

Conclusion

Digital objects come in many shapes, sizes, formats, packages, and levels of complexity; it is this very dynamism that makes digital technology so compelling. Abstract digital media resists being captured by physical form for good reason—the constraints of static physicality could overly constrict such media's use.

Slurp is an approach towards physicalizing abstract digital media. We did not design Slurp to be a more efficient method of accessing information than existing systems (although in a future where digital media is far more pervasive it may be very efficient). Our goal was to explore a novel interaction technique through prototyping, use, and reflection in order to better understand some of the current issues in tangible interface design.

SpeakCup

As an experiment in simplicity we created SpeakCup (Figure 15), a digital voice recorder in the form of a soft silicone disk that relies on shape change as input. The design arose from our earlier investigations in reducing interface BABL (buttons and blinking lights) [Chang 2007]. To record sound with SpeakCup the user

deforms the disk into a cup, so that it can hold the sound. Pressing the cup inside out will then release the captured sound (Figure 17). Our contribution is a demonstration of using shape change as an interface strategy. We are interested in shape change as a way to privilege simplicity in interaction design and also as a means to incorporate physical metaphor or analogy into new devices.

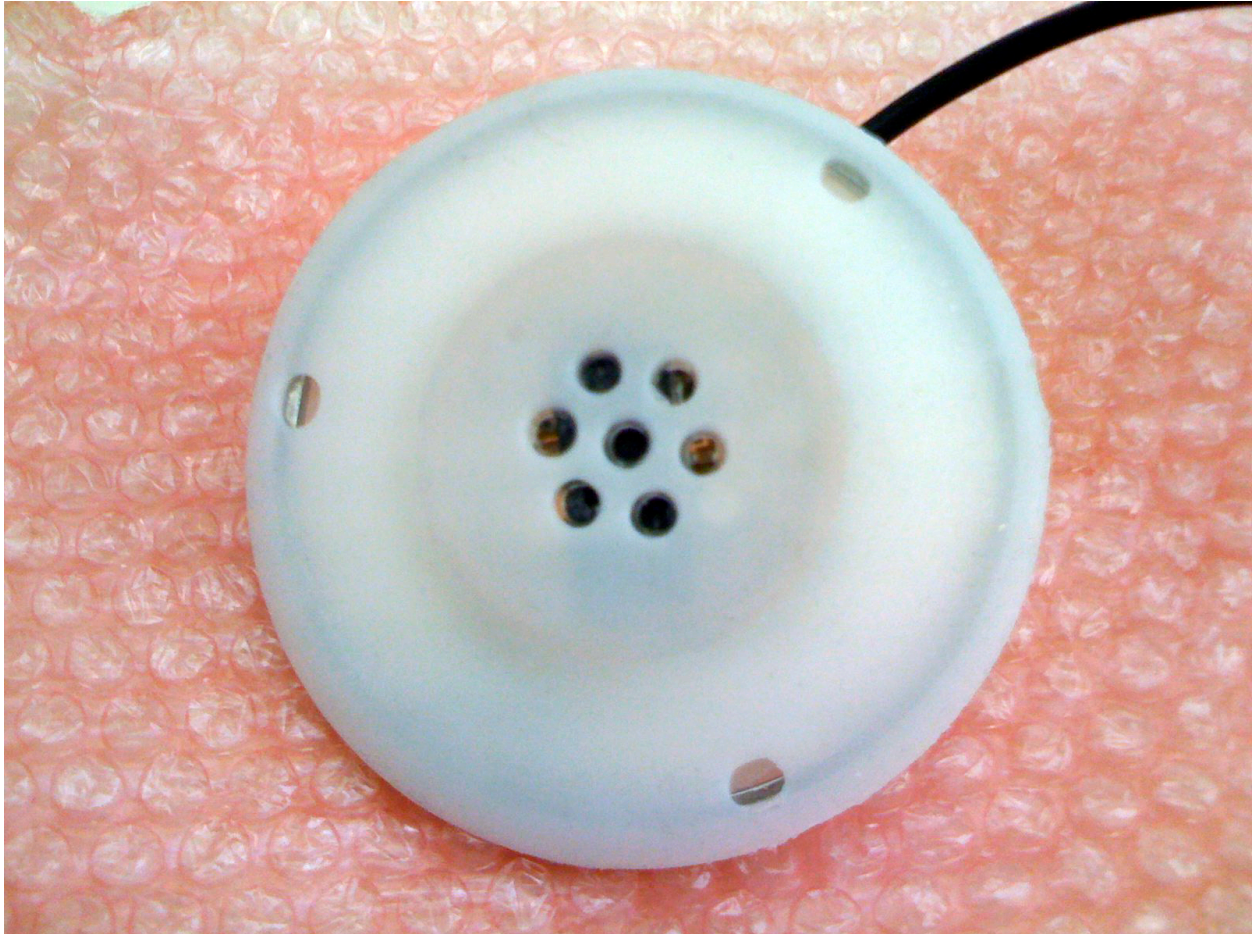


Figure 15. SpeakCup version 2.

Using shape change as an input to computational systems is nothing new, the mouse changes shape when you click it and so do keyboards. Shape change is the dominant form of human-machine interaction but in most cases the change in form and the action incurring the change are only loosely connected to the desired response. Hutchins, Hollan, and Norman described this as the gulf of execution [Hutchins 1986], in other words it's the gap between a user's goals for action and the means to execute those goals. Interfaces (by definition and observation) get in between users and their goals. Weiser's vision of calm computing [Weiser

1997] has been a serious challenge—it is not easy to make functional interfaces disappear!

One way to create invisible interfaces is to design devices that function like the real, un-digital world [Jacob 2008]. But in some cases the un-digital world is not as versatile as the digital world, so how do we go about providing valuable expressive power and reducing the gulf of execution at the same time? There are some answers for this question in Ishii and Ullmer’s tangible bits [Ishii 1997] as well as in [Jacob 2008]. In this paper we describe a small exploration in the use of shape and metaphor to address these questions.

We first imagined SpeakCup during a design exercise in which we challenged ourselves to create interfaces with minimal feature sets without relying on abstract buttons or blinking lights [Chang 2007].

Implementation

SpeakCup’s body is made from a six-inch disk of platinum cure silicone rubber. A ring of aluminum is embedded inside the outer rim of the disk so that it stays round when deformed. A combination microphone/speaker is embedded in the center of the disk. Red and green LEDs are wired around the perimeter of the microphone/speaker. Two flex sensors that change resistance when bent span the diameter of the disk embedded in the silicone and facing in opposite directions to sense deformations on either side. SpeakCup is wired to a protoboard where the electronics and batteries that power it are located. We used a cheap RadioShack voice recorder module with 20 seconds of memory to store and playback sound. We rewired this module’s buttons, running them through analog electronics to connect the flex sensors and drive the LEDs.

Interaction Design

We rely on metaphor to inform the use of SpeakCup. Sound is imagined to be a physical substance that can be contained and absorbed into the medium of a malleable disk. The disk has seven holes on one side, deforming SpeakCup so that these holes become the bottom of a cup triggers the sound recording. Once in cup form, red LEDs pulse within SpeakCup’s body. The LEDs get brighter as the user’s voice gets louder, mirroring the amplitude of recorded sound.

When the user removes pressure from SpeakCup's center it springs back to disk form. To playback the recorded sound the user then presses in the opposite direction, pushing the holes out to release the stored sound. During playback the brightness of the green LEDs portray the amplitude of the sound.

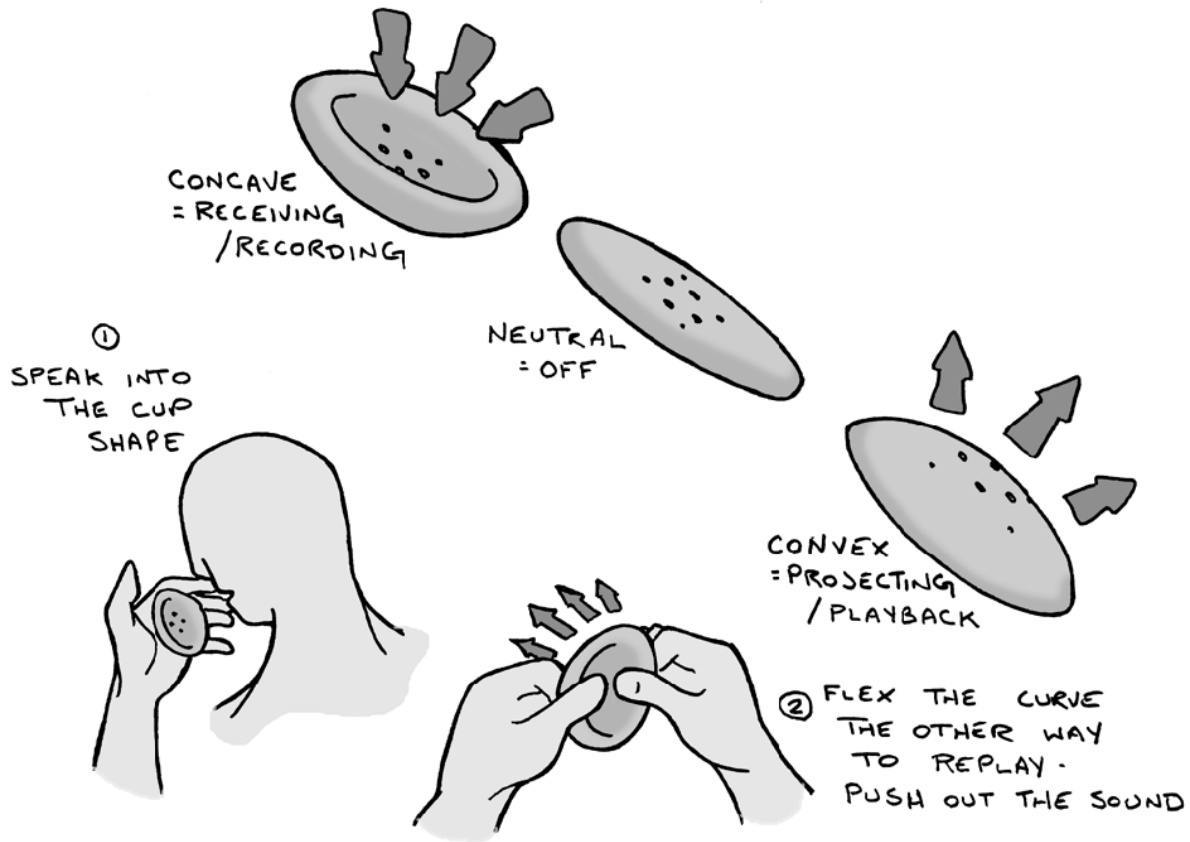


Figure 16. SpeakCup interaction design. (Illustration by James Gouldstone.)

Although our goal for this work was simplicity of function and form we experimented with adding more expressive power. We added a tri-axis accelerometer and dual-axis gyro board and built gesture recognition software so that users could store multiple voice files by linking them to parts of their body similarly to [Angesleva 2003]. This method was interesting but it added an extra level of complexity that we were unable to integrate into the existing metaphor so we did not develop it enough for testing.

The simple version of SpeakCup turned out to be best. Although limited in functionality, we have found some promising uses for it, for example as a toy for social interaction in public

spaces—users drawn to SpeakCup’s novel form could record messages and stick them to a wall for someone else to experience.

Conclusion

After reflecting on SpeakCup and trying it out with colleagues and friends we are optimistic about using soft forms and shape changes in interaction design. From our informal testing we found that users enjoyed SpeakCup. This type of testing is very subjective, but we are encouraged by the simple pleasure that this novel form is able to elicit in users. At the same time we are skeptical about using these techniques for feature rich applications. One way or another we will have to abandon the abstract buttons and blinking lights used in today’s interfaces if we hope realize the vision of calm computing. Many great strides have been made towards this goal; in this paper we have presented a small exploration into this space.

Summary

Although these three projects were not planned to explore three separate aspects of tangible interface design, in hindsight it has become apparent that they fit into the larger research program that is described in this thesis. We created the Tangible Video Editor as an interface to allow video producers to edit collaboratively in a group without sacrificing either the power of digital video editing techniques or the tactile richness of traditional film editing. Here we explored the ability for tangibles to function under directed collaborative activity. We created Slurp as a way to interact with abstract digital information as well as reflect on some issues in tangible interface design through a design study. Here we began to realize the value of seamless interface design and rich metaphor as giving rise to external legibility for non-participants. We created SpeakCup as an experiment in extreme seamlessness. Here we wondered what a computer interface would be like were it to be born fresh into the world like a tree or other organism—like in nature where form and function are indistinguishable.

EXTERNAL LEGIBILITY

One of the most interesting concepts arising from my research and development of tangible interfaces is the idea of external legibility. While the HCI literature is full of examples of studies of interface legibility or how well an individual user or a group of users can interact with or understand an interface or interaction techniques that they are directly involved with using (what could be called internal legibility), there are hardly any examples of studies to examine the impact of interface design on non-participating observers (Hornecker comes closest to this quality in her concept of visibility from [Hornecker 2006] and [Hornecker 2007] and Patten addresses similar concerns in [Patten 2002]). I define this property of interface design as external legibility.

External Legibility: a property of user interfaces that affects the ability of non-participating observers to understand the context of a user's actions.

One reason why external legibility is important in interface design has to do with its relationship to semantics. Although it may never be possible to truly understand another's mind, communication is based on shared understanding. Without a context in which to base understanding, inferring meaning or semantics becomes difficult.

Think of watching a master craftsperson working on a cabinet. You can see her hammering a nail to join two two-by-fours, you can see how she makes precise cuts along the edge of a piece of plywood. The context that the craftsperson works within is highly legible to an observer—the feeling of the wood, the knowledge of why a hammer is used, the memory of experiences of doing things like what the craftsperson are doing are available to

many of us, but unless you too are a master craftsperson you may not know why she is doing the things that she does. The specific content of her actions are private, her thoughts and strategies, but the context of her actions are public.

Without the ability to move from observation to inference accurately, it is hard to create shared understanding. External legibility is a measure of the reliability of the connection between observation and inference in interface design, but not in the traditional framing of one person and one machine—what could be called legibility. External legibility is a property of the space between one person observing another person using a machine.

I have identified three main components that make up a given interface's external legibility, they are feedforward, feedthrough, and feedback.

Feedforward

A term coined by Djajadiningrat, Overbeeke, and Wensveen in their paper *But How Donald, Tell us How?: on the creation of meaning in interaction design through feedforward and inherent feedback* the authors describe it in the following way: “With feedforward we mean communication of the purpose of an action.”[Djajadiningrat 2002]. Feedforward is very similar to Norman's concept of affordances introduced in his book *The Psychology of Everyday Things* [Norman 1988] which was derived from Gibson's use of the word in *The Theory of Affordances* [Gibson 1977].

Feedforward is differentiated from Norman's concept of affordances in that affordances are meant to communicate what the right action to do is, while feedforward instead communicates what the purpose of a given action will be. This difference results in two ways to approach an understanding of interaction design, the latter being that interface components should describe their use and hint at the consequences of their use, the former that interfaces should be designed in a way that guides the user to accomplishing preset goals.

Feedthrough

Alan Dix describes feedthrough in *Human-Computer Interaction* [Dix 1993] as “how quickly other people see the effects of

your actions (and you see theirs)”. Feedthrough has to do with an understanding of causality between a user’s input actions and the effect on the computer or technology being used.

Feedback

The output from a display can often be enough for an observer to see in order to understand the content of a given interaction. Feedback is the third quality of external legibility and in some cases can be the most important. Problems arise in understanding the connection between user input and feedback when there are multiple users engaged in an interface simultaneously and feedforward and feedthrough are not well maintained by the interface.

Experiment

To test if certain types of interaction techniques will have higher external legibility than others, I conducted an experiment to quantitatively assess non-participating observer’s understanding of computing tasks carried out on four different interfaces.

The experiment tested the following two hypotheses:

- > **Hypothesis 1:** tangible, gestural, and multitouch interfaces will have higher external legibility than a laptop
- > **Hypothesis 2:** Tangible interfaces will have the highest external legibility, then gestural, then multi-touch, then laptop.

Tangible User Interfaces

I describe some of the history and properties of tangible user interfaces in the introduction to the third chapter of this thesis. Tangible user interface describes a field of HCI that can also be referred to as tangible interaction, or tangibles, among other related names. Tangibles can take many forms (from constructive assemblies like topobo [Raffle 2004] or roblocks [Schweikardt 2008], to tabletop interfaces like PICO [Patten 2007] or reacTable [Jordà 2007]) and can be designed at many scales (from something small like Slurp [Zigelbaum 2008a] to architectural scale).

Gestural Interfaces

Although various forms of gestural interface have been around for years, and in fact one could consider Englebart's mouse [Engelbart 1962] to be a gestural interface, I am referring to a specific form of interface in this thesis. The type of gestural interaction described here is free handed manipulation of computing systems input through gestures in mid air and directed at remote screens or objects. An early example of this type of interface is Charade by Baudel and Beaudouin-Lafon [Baudel 1993] and a newer example is the g-speak gestural computing platform by Oblong industries [g-speak].

Multi-Touch Interfaces

Multi-Touch interaction has been a hot area in HCI recently, finding it's way into the consumer space and generating buzz in 2007 with the release of the Apple iPhone, the presentation of Microsoft's Surface computing program, and the popularity of some online videos of Jeff Han's multi-touch interfaces [FTIR]. Bill Gates has gone on record saying that multi-touch was the future of computing [Gates] and Windows 7 is said to be incorporating numerous new features to support multi-touch. Although the feedback systems and widgets for multi-touch have yet to be refined or standardized, the model is rather consistent across different implementations: users interact with a graphical interface similar to traditional GUI systems but instead of using a mouse they use their finger tips, two or more at a time, to enter commands. For a good survey of research on multi-touch systems see Bill Buxton's article: *Multi-Touch Systems that I have Known and Loved* [Buxton 2008].

Experimental Design

External legibility is a new concept and not something that has been tested for in the HCI literature so I designed an experiment based on a mixture of existing experimental psychology methodologies. Hypothesis 1 is a subcase of hypothesis 2, so in order to test both I needed to have participants watch a user interact with four separate interfaces and compare how well the participants were able to understand the tasks carried out.

For this initial study I chose to focus on the effect of feedthrough on external legibility rather than feedback or feedforward. Although I do not believe that these three properties can be completely disentangled, the choice to focus on feedthrough

resulted in decisions that would have been made differently were I to have focused on feedback instead.

Minimizing Error and Confounds

To minimize the chance that differing body language and user or interface behavior would confound the data I created a series of video clips for each condition of the study. Having participants watch sets of videos served to reduce error but also to narrow possibilities for interpretation of the data. To reduce differences from the latency or design of different interface technologies (typically different versions of computer vision tracking for multi-touch and gestural, and electromagnetic sensing for tangible) the graphical feedback for each interface was created ahead of time in Adobe Flash and then the user synchronized his actions along with the feedback as if he were directing it.

Method

I used a mixed design, with three independent variables. This design has 36 separate conditions, each a different combination of the 3 independent variables (3x4x3). To reduce the possibilities of a learning effect from seeing the interfaces from different viewpoints each participant was randomly assigned to one viewpoint only, they would then watch all 12 videos in a random order for that viewpoint.

Independent Variables

> Viewpoint (3 levels):

- » *Near (control)*: the video is shot directly over the user's right shoulder. Graphical feedback and physical input are as visible as possible. This viewpoint served as the control, the hypothesis being that participants should understand the majority of tasks shown at this viewpoint.
- » *Far*: the video is shot from 15 feet behind the user, over the user's right shoulder. Graphical feedback and physical input is largely obscured by the user's body and the screen is very small.
- » *Side*: the video is shot perpendicularly to the user's line of site. Graphical feedback cannot be seen, although the

surface of the screen is visible as a thin sliver and physical input is not obscured.

> **Interaction Technique (4 levels)**

- » *Laptop (control)*: the laptop used was a 15" Apple MacBook Pro with a matte screen displaying at 1280 x 768 lines of resolution. The laptop was used from a standing position to match the posture of the other interfaces. The laptop served as the control interaction technique for the study.
- » *Multitouch*: a 50" wall-mounted Samsung display with 1280 x 768 lines of resolution was used. The user input commands to the screen's surface with his finger tips.
- » *Gestural*: a 50" wall-mounted Samsung display with 1280 x 768 lines of resolution was used. The user input commands by moving his hands one foot from the screen's surface.
- » *Tangible*: a 50" wall-mounted Samsung display with 1280 x 768 lines of resolution was used. The user input commands by manipulating 1" x .5", black pucks on the screen's surface.

> **Task (3 levels)**

- » *Select*: the user selects a 50px by 50px red square located at (460, 300). The red square turns darker red for 333 ms after selected. (The origin was located in the upper-right-hand corner of the screen.)
- » *Drag*: the user drags a 50px by 50px red square from (460, 300) to (560, 320). The red square turns darker red for 333 ms after initially selected.
- » *Enlarge*: the user enlarges a 50px by 50px red square located at (460, 300) to 100px by 100px. The red square turns darker red for 333 ms after initially selected.

Dependent Variables

- > **Accuracy**: this is the measure from Q1 shown below. Correct answers are scored as "1" incorrect answers are scored as "0".
- > **Confidence**: this is the measure from Q2 shown below. Answers are given on a 7-point Likert scale.
- > **Ease**: this is the measure from Q3 shown below. Answers are given on a 7-point Likert scale.

Across all condition the user stood in exactly the same position, the user was right-handed, the camera position was consistent for each viewpoint, and the graphical feedback was the same

for each task on each interface. In the laptop condition the size of the screen was smaller in the video relative to the size of the other screen, this decision is addressed in the discussion. Rather than record the user interacting with the interfaces in a sterile environment the background used was a typical HCI laboratory environment with lots of possible static distractions, these distractions were consistent within each viewpoint.

The videos were created using Adobe Flash exported as SWF movies and embedded in HTML, they were displayed without available playback controls so the participants could not rewind, pause, or restart the videos once they began to play. Data were logged at the end of each session and stored in a comma delimited text file. All data were analyzed using SPSS statistical analysis software.

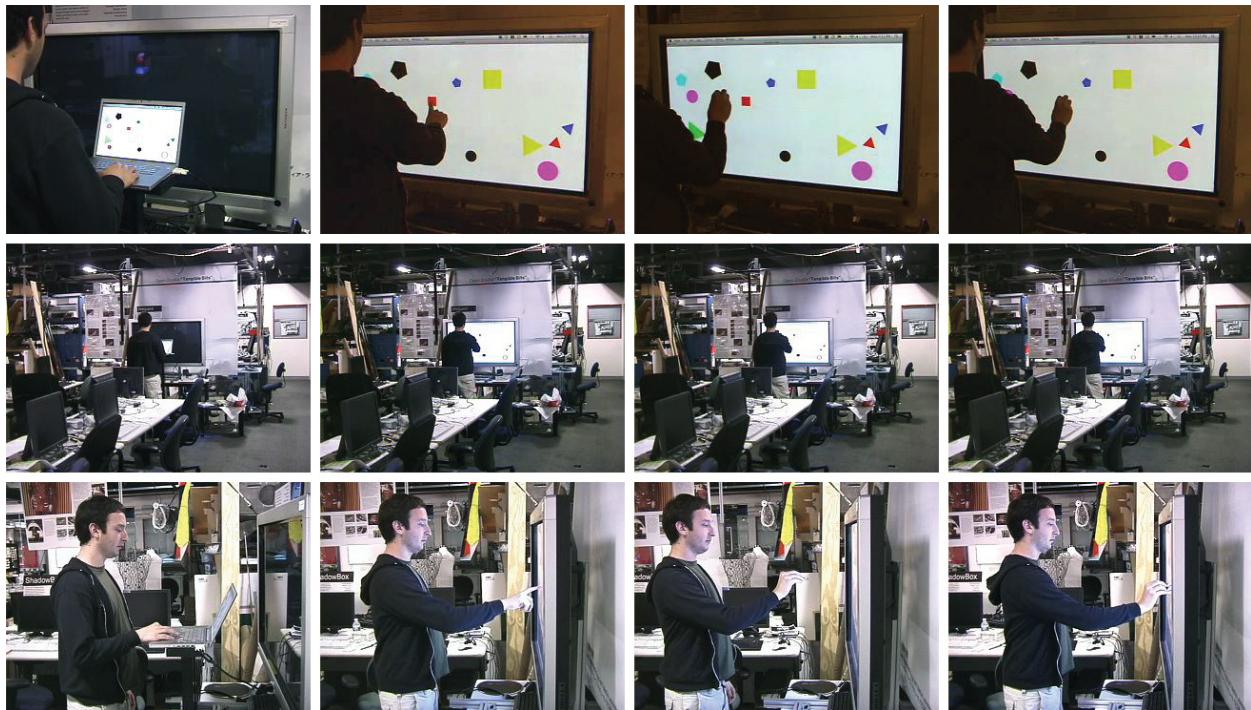


Figure 17. Stills from 00:02:00 seconds into each of the videos for the Drag task. Columns from left to right: Laptop, Multi-Touch, Gestural, Tangible. Rows from top to bottom: Near, Far, Side.

Design of Tested Interaction Techniques

Since there is as yet no standard interface design for the next-generation interfaces tested, each of the interaction techniques used for the four interfaces was based on the most archetypical instances of prior work. For the laptop interface the user interacts with the computer in the standard way, by using his finger on the trackpad, using the mouse button to click rather than tapping on the pad. This is theoretically the more legible of the two possibili-

ties for clicking with a laptop since a click to the trackpad itself could be interpreted as the beginning of a drag, this decision was made to compare the laptop in its most externally legible form.

For the multitouch interface the interaction is based on the archetypical multitouch interaction designs from Jeff Han's research at NYU [FTIR] as well as the Apple iPhone, the user selects and drags with one finger, then uses the index fingers of each hand to enlarge.

The tangible interface is based on the sensetable technology developed by James Patten [Patten 2001] and adapted for a vertically oriented screen by hybridizing the design with the form of the Designer's Outpost by Scott Klemmer [Klemmer 2001]. A small, round puck sticks onto the screen on top of the graphical feedback for the red square that serves as the target object for each condition. To select, the user simply grabs the puck. To drag, he moves the puck across the screen's surface. To enlarge, he uses both hands with two pucks attached to the corners of the target object.

Designing the interaction technique for gestural interface was the most problematic since there are fewer examples of prior art that would work for this comparison. I decided to use the simplest hand gestures for the tasks that I could imagine while keeping the user's hand at one foot from the surface of the screen. For select the user uses all the fingers and thumbs on his right hand in a brief, mid-air grasping motion. For drag the user selects the target object and then moves his hand isometrically to the feedback. For enlarge the user uses both hands with the same dragging motion, pulling the square from each corner.

Design of Graphical Feedback

The graphical feedback shown on each screen was the same for each task throughout the study. Rather than use an existing application that participants would be familiar with, I chose to design an abstract yet familiar interface where the user is selecting, dragging, and enlarging a simple red square on a white background where other shapes of various colors and sizes are also scattered around. This interface design shares features with common CAD application as well as features of common operating system interfaces such as showing a brief color change when an object is selected and drag and drop. The colors and sizes of the shapes were selected so that they would be visible on both the laptop screen and the Samsung display.

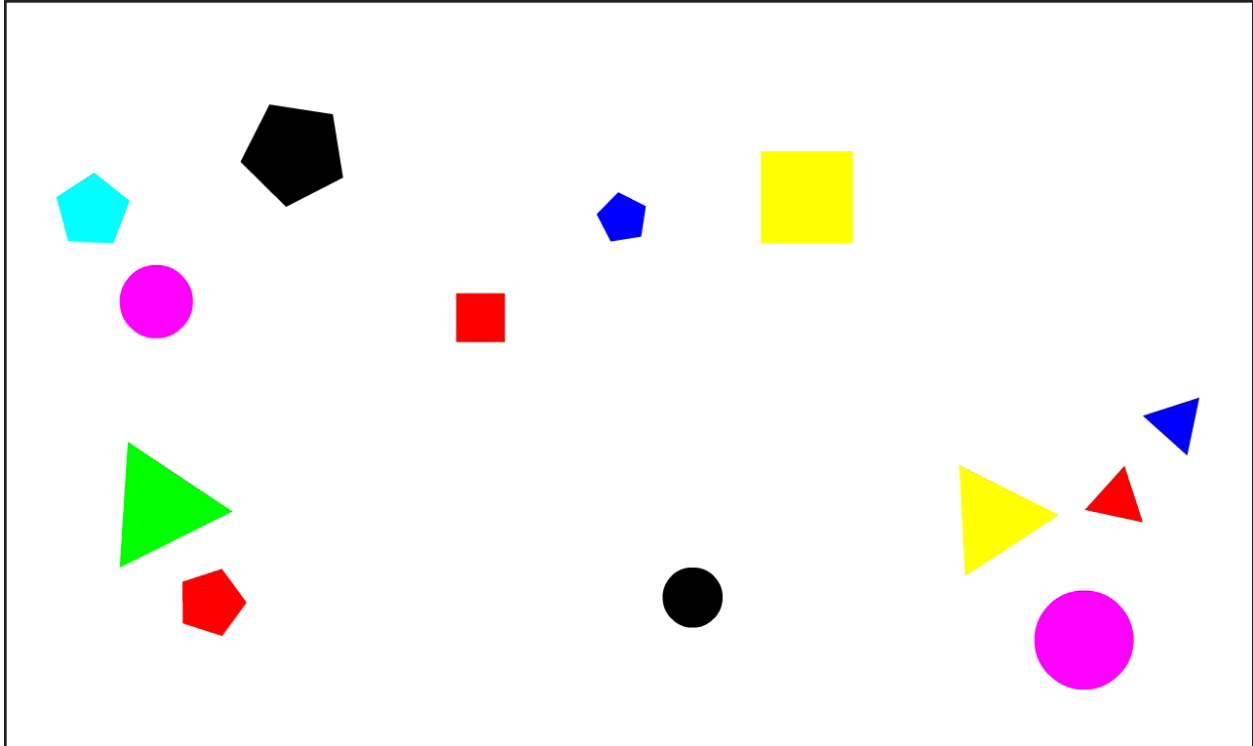


Figure 18. Screen capture from beginning of each task. The red square (the smaller of the two squares) is the target object for each task.

Subjects

136 participants were recruited via emails from a contact list of my friends and colleagues, 65 of those contacted completed the study although only 35 of those 65 were kept due to a javascript error which rendered the data from the other 30 participants useless. Of the 35 participants whose data we kept, 19 were male, 10 were female, and 6 didn't specify their sex. 29 of the 35 participants specified their age, these ranged from 19 to 53 years old. Most of the participants were graduate students familiar with next generation interfaces, this composition of participants was beneficial for the study's validity since it reduced the gap in familiarity between interface types shown. Subjects were not paid.

Procedure

The experiment was conducted entirely on the internet. After reading and signing an informed consent form participants were asked to complete a pre-test questionnaire where demographic information was collected. Then each participant was randomly assigned to one of the three viewpoints and shown 12 videos in a random order, every combination of interface (4) and task (3) for that viewpoint. After each video participants were asked the following three questions:

Q1: What task did the user perform in this video?

select

drag

enlarge

rotate

delete

copy

don't know

I didn't watch the video

Q2: I am very confident in my answer.

strongly disagree

disagree

somewhat disagree

neutral

somewhat agree

agree

strongly agree

Q3: It was very easy to understand what task the user was doing.

strongly disagree

disagree

somewhat disagree

neutral

somewhat agree

agree

strongly agree

Figure 19. Questions asked after each video.

Although there were only 3 tasks shown—select, drag, and enlarge—I included 3 other options in the multiple choice response as distractors, these were: rotate, delete, and copy. I included these distractors for two reasons: first to minimize participants ability to guess the correct answer when they weren't sure what they saw and second to more accurately mimic the state of doubt that exists when watching someone on a computer in the real world. When observing someone using a computer in the real world they could be doing any number of things, this introduces the possibility that even though you think you know what you have seen, it could be something else due to the sheer number of different tasks and applications used on standard laptops. This doubt would be difficult to recreate with only 3 options.

Participants were prompted to click on the videos to start playback, they were then shown a 3, 2, 1 countdown to make sure they were paying attention when the content began to play. Each video was exactly 6 seconds long. At first the user stands idle in front of the interface, during the first 2 seconds he reaches out to the interface, during the next 2 seconds he manipulates the interface, and in the last 2 seconds he returns to an idle position. After the video is finished the last frame is black with white text instructing the participant to answer the questions below and move on to the next video.

After watching the 12 videos some final questions are asked to ascertain the participants experience with computers and famil-

ilarity with the interfaces shown. The time it took each participant to finish each section of the study was also recorded and matched against a baseline time acquired by moving through the study as quickly as possible without thinking about the questions. Cases within 60 seconds of this baseline duration were removed from the data. Participants took between 5 and 15 minutes to complete the study. In order to remove the scores of possible repeat participants each participants IP address was recorded as well.

Results

Data were collected from 35 participants over the span of 10 days. Participants answers to Q1 were coded for accuracy, correct answers scored as a “1”, incorrect answers scored as a “0”. When a participant answered “don’t know” whatever score that they entered for Q2 was changed to a 1—“strongly disagree”. When a participant answered “I didn’t watch the video” (this only occurred 4 times for the total of $35 \times 12 = 420$ videos watched) their scores for confidence and ease were removed from the data.

A repeated measures ANOVA (General Linear Model Repeated Measures analysis using the Advanced Models add-on for SPSS 16) showed a significant difference between the means of the laptop conditions against all other interface conditions ($p < .001$). As a secondary metric for confirmation a paired samples, two-tailed t-test showed a significant difference between the laptop conditions and the non-laptop conditions as well ($p < .001$)

These finding support our first hypothesis that tangible, gestural, and multi-touch interfaces will have higher external legibility than a laptop. The second hypothesis, that the external legibility of the interfaces would be hierarchical with tangible having the highest external legibility, then gestural, then multi-touch, with laptop being the lowest was not completely supported. In a pariwise comparison, means for accuracy, confidence, and ease were shown to be significantly lower in all laptop conditions when compared against each of the other interfaces ($p < .001$) but the means for all dependent variables were shown to be highest in the multi-touch conditions ($p < .05$) not the tangible conditions, and there was no significant differences shown between the means of the dependent variables for tangible or gestural interfaces.

Discussion

As shown by the data, my first hypothesis, that the three next generation interfaces would demonstrate higher external legibility

Measure	Interface	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Accuracy	Laptop	.217	.025	.167	.267
	Multi	.782	.030	.721	.843
	Gestural	.650	.030	.590	.711
	Tangible	.619	.050	.516	.722
Confidence	Laptop	2.454	.187	2.071	2.837
	Multi	4.684	.160	4.356	5.011
	Gestural	4.206	.165	3.868	4.544
	Tangible	4.055	.176	3.696	4.413
Ease	Laptop	2.584	.226	2.122	3.046
	Multi	4.453	.161	4.124	4.782
	Gestural	3.683	.170	3.335	4.031
	Tangible	3.683	.170	3.335	4.031

Figure 20. Table showing means, error, and confidence interval for each measure (dependent variables) for each interface.

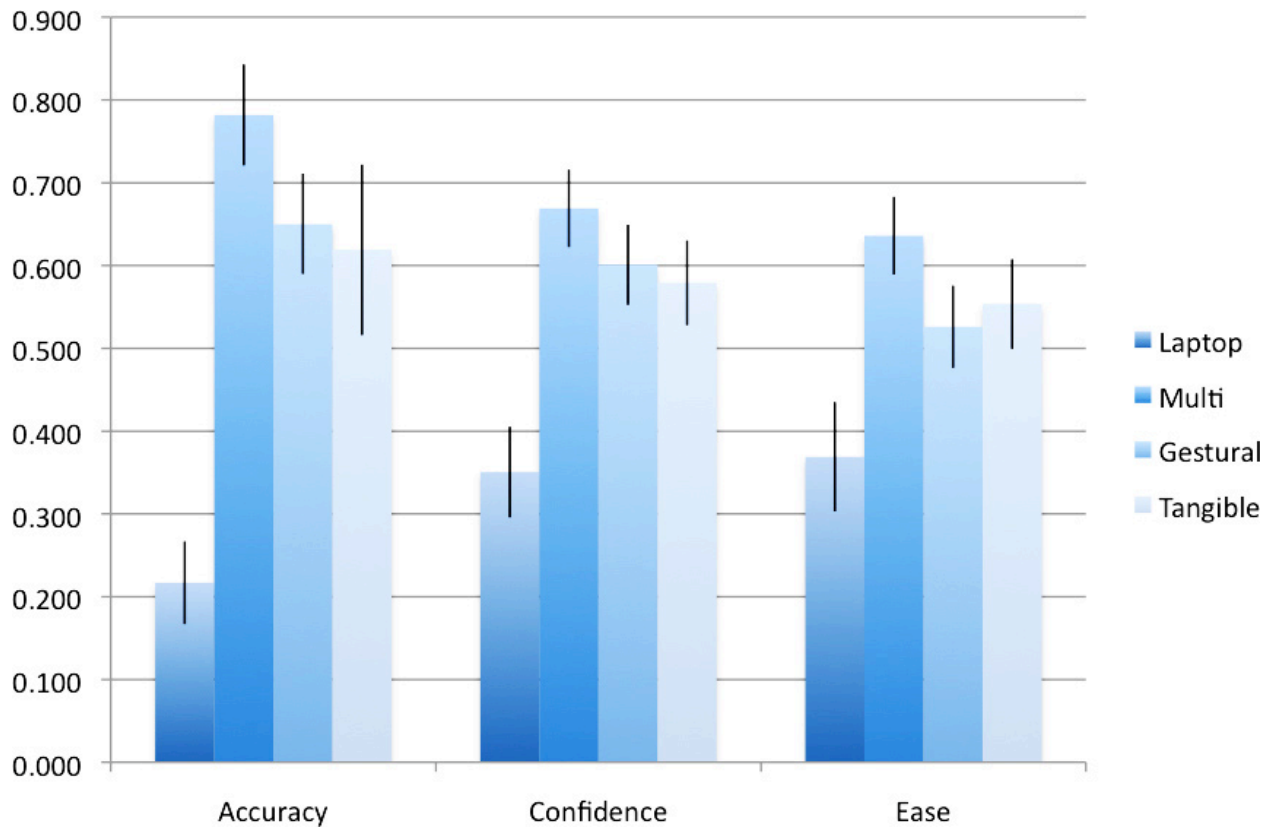


Figure 21. Chart showing means for each measure for each interface. Confidence and Ease have been normalized from a 7-point scale to a 1-point scale for consistency.

than the laptop, appears to be supported. Since I decided to standardize the camera angle relative to the user instead of the screen the argument could be made that the lower ratings of external legibility in the laptop condition are simply due to the smaller screen size. In designing the experiment I had considered standardizing the angle based on the screen for just this reason, but decided against that option since it would be more about examining the effect of graphical feedback instead of feedthrough. Instead I chose to base the angle on the position of the user and tried to minimize the effect of the screen through the experimental design.

It appears obvious that a larger screen will be easier to read than a small one—graphical feedback can sometimes be all that is needed to make a user's state transparent. But it is less obvious what the effects of different input techniques are if you remove the graphics from the equation, for instance the difference between touching a screen or gesturing at it from the side viewpoint. In the side viewpoint, none of the participants could see graphical feedback at all (refer to Figure 17). A paired samples, two-tailed t-test comparing the means of just the cases from the side viewpoint ($n = 14$) showed a significant difference in means between the laptop and non-laptop conditions ($p < .001$). This analysis would suggest that the effect of screen size of the other results is not a confounding factor and that the physical input of the user was more important for the understanding of non-participating observers.

This suggests that an important reason for the laptop's low external legibility is that observers have trouble connecting a user's input actions on a trackpad with the screen's output. All of the next generation interfaces tested utilize larger input motions than the relatively small input motions used for a trackpad, it stands to reason that this is one of the primary reasons why they demonstrate higher external legibility, the other reason may be that the input motion of the user's hand matches the graphical output more closely. This property of coincident I/O has been considered a very valuable property of tangible interfaces [Ishii 1997], it would be interesting to test how important a factor this is for external legibility. In future studies, I plan to compare large input motions that match graphical feedback to those that don't in order to begin answering this question.

While 28 out of 30 participants who answered whether they use laptops regularly responded in the affirmative, only 10 out of 35 responded that they had ever used a gestural interface, and 21 out of 35 for tangible. I compared the means of accuracy between

gestural and tangible conditions of participants who had used both gestural and tangible interfaces previously ($n = 10$) using a paired samples, two-tailed t-test but found no significant difference ($p = .678$) between the populations. For those participants who were familiar with both gestural and tangible ($n = 21$) I found the same result ($p = .540$). The fact that gestural and tangible interfaces in particular are not yet standardized or well used by the public may have contributed to their lower scores for external legibility.

In forming hypothesis 2, I imagined that tangible would be the most legible due to the use of physical objects on top of the screen that would be easier to see as linked to the graphical elements. Although this study didn't show that to be the case, it is possible that the physicality of tangible interfaces is not well represented through video (as could be the case for laptops as well). In the future I will consider running a similar study where participants observe users interacting with the interfaces live in the same physical space. Although this will be harder to control than a video study it may be important as a separate data point.

Conclusion

External legibility is only really becoming interesting in our modern era filled with so many people and machines. Back when the word "computer" was used for a person who did computations (not so long ago) and computing machines were hardly able to do math it didn't matter so much if the people watching the interface between user and machine knew what was going on. But now that every time you walk down the street you can see someone doing something on his or her cellular phone it is important. It is important if we don't want to live in a world where billions of people spend more time looking at their cell phones than at each other. That was not the vision that Weiser intended.

The study presented above is in many ways a preliminary investigation of the topic. Evaluating interfaces based on their legibility to non-participants and their capacity to transmit and support the flow of information between users in physical space will hopefully become a more commonplace activity in the future of human-computer interaction research. It may be that reality-based interfaces following from the RBI framework presented in Chapter 2 will show higher levels of external legibility, this would be a relationship worth further examination.

In the next chapter I will present some design sketches for interfaces that make use of high external legibility towards the aim of creating truly seamless and calm interactions.

Section 3:

dreaming

interlude < the past

“The Glass Bead Game is thus a mode of playing with the total contents and values of our culture; it plays with them as, say, in the great age of the arts a painter might have played with the colors on his palette.” -The Glass Bead Game [Hesse 1943]

Herman Hesse was an author of fiction, prose, and poetry; an artist of paint, pen, and mind. He first published *Das Glasperlenspiel* in 1943 in Switzerland (English title: *The Glass Bead Game*); it won the Nobel Prize for literature in 1946. In the book he describes a game, but never quite how to play it. The game is played by ascetic-scholar-artists, the most educated of their kind, it's played in the fictional city of Castalia, home to the great University. They use glass beads that represent the span of all accumulated human knowledge, they connect the beads on a physical board during festivals. These 'games' are the highest art in their world. The best games create deep connections between seemingly unrelated fields of thought, for instance matching the pattern of a butterfly wing to the progression of arguments in a work of philosophy. The point was to create a music of ideas. This music was embodied in physical form, defined by spatial relationships, and supported by formal logic.

The real content or meaning of a game was not transparent, observers needed to have some understanding of the referenced canon and indexing scheme to understand the art, but all those who watched a Magister Ludi (the name used for the highest master of the game) play could see how he held his beads and thought

about their consequences, perhaps in the way a child might watch a master carpenter observe the grain in a piece of wood.

Around the same time, in another corner of the world, an MIT scientist, Vannevar Bush, had a similar vision, but his was for this world and it was based in the technological developments of his time, such as Emanuel Goldberg's microfilm rapid selector machine, perhaps the first instance of electronic document retrieval in history [Goldberg 1932].

"First he runs through an encyclopedia, finds an interesting but sketchy article, leaves it projected. Next, in a history, he finds another pertinent item, and ties the two together. Thus he goes, building a trail of many items. Occasionally he inserts a comment of his own, either linking it into the main trail or joining it by a side trail to a particular item." -As We May Think [Bush 1945]

Vannevar Bush wrote *As We May Think* in 1939, but it wasn't until after the War, in 1945 that he published it in the *Atlantic Monthly*. While at MIT he was deeply involved in the War effort, though it seemed he was dreaming of a time when the human race could use its energy for the pursuit of knowledge rather than for the design of killing machines. He imagined the Memex, a sort of dry photography based Internet terminal, and he describes how people could use the technology to create trails through information space. Bush's Associative Trails are one of the nodal points in history that we refer back to as an origin of hypertext, yet when compared to our current global Internet Bush's trails were more personal. They told a narrative and were traded, from one person to the next.

The Memex and associative trails were a pre-vision of the Internet. There were others as well, notably the works of Paul Otlet [Rayward 1975].

These visions do more than help us understand the historical moment from which they came, but they also help us re-evaluate our current moment, to see what has changed, what could have gone differently, what we have missed, and where we can go in the future.

THE FUTURE

Hesse's glass bead game and Bush's associative trails were the inspiration for this thesis. Actually, they were the inspiration for a much earlier conception of this thesis, but their import is still relevant here. Bush and Hesse's visions were so much more physical, so much more personal than our technologies of today. While much of this thesis is about the value of some underrated properties of interface design, it is also about the way that we interact with information, in a broad sense. For some reason, the specter of the digital computer; of the personal computer; of the beige box, was depersonalizing—even dehumanizing. Katherine Hayles describes the disembodiment of technology in her work *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* [Hayles 1999] and Donna Haraway famously calls for feminists to integrate with machines in *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century* [Haraway 1991]. Both authors were responding to the way in which human culture has held digital computation at arms length, as if it were something completely outside of our natural understanding.

It has taken time, but we have begun to integrate the extensions of digital computation into our well-traveled social realm. New developments in human-computer interaction research such as tangible user interfaces and reality-based interaction hold the promise of a seamlessness between bits and atoms that is beyond anything we have seen. In this thesis I have tried to put forth arguments to support the value of such seamlessness. I have made arguments from the social sciences and from human-computer interaction, and focused in on a particularly interesting property of interaction design that I call external legibility. For the remainder

of this chapter I will present some sketches for interfaces based on the theory and perspective generated thus far.

Design Principles

In the near future, through using multi-user interaction techniques such as gestural, tangible, and multi-touch, it will be possible to develop advanced, geographically rooted interfaces. There are already some instances of next-generation interfaces being deployed into public space, such as the CityWall installation in Helsinki Finland [Peltonen 2008]. The drop in price for large display surfaces and sensing technologies will precipitate more and more of such installations. While currently the rare cases of such installations are novelties, it will be possible to structure tools and interactions that function in a similar manner to the spate of web 2.0 technologies that have made highly participatory web-culture possible. The sketches in this chapter are based on the following three principles:

Local Value

The people who inhabit the same physical space are an important affinity group. On the web you can post a question about nearly anything and get answers from people all over the world. You can find groups of people that are interested in the same things that you are interested in, you can find places that sell nearly any obscure product you could want. This long-tailed property of the web has been a noted feature of the internet economy [Anderson 2006]. It is becoming more and more evident that people want more geographical structure in their web content, this can be seen by examining the many projects to do just this (see Chapter 2). The people who we share space with are the people we vote with, the people who drink the same contaminated water, the people who have to deal with the same power structures, and the people who will require the same hospital services when we are in our waning years. Building connections between the diversity of people in an urban environment has been shown to provide health benefits [Cohen 2000], and other positive consequences [Granoveter 1974]. Local, physical space and the people in it are very important.

Local Information Processing

Following from the arguments presented in Chapter 2, small, semi-disconnected ‘pools’ of information could be a valuable commodity. “The global can overwhelm the local” writes Cas-

tells [Castells 2004]. The vast resources of information online can be too much to sort through triggering a form of risk-aversion that can freeze decision making [Iyengar 2000]. On the web we use various forms of filtering to wade through the vastness, for instance a search term on Google filters all the possible returns. The availability of more ways to filter information based on local and geographical features is a rising trend shown by services like geotagging and hyperlocal portals.

The Medium is the Message

McLuhan postulated that the power of a given information and communication technology was determined more by the properties of the medium (such as the ability for television to broadcast across large distances from small groups of producers to large groups of consumers) than the properties of the content (such as television programs). From this perspective no matter how many new websites we build to strengthen local communities, none of them will be as powerful an agent of change as completely new interfaces that eschew the single user, divisive media of laptops and cell phones.

Laptops and cell phones are valuable tools, but to connect real people in real spaces why mediate so much information flow through these narrow channels? The physical world has more bandwidth.

Sketch 1: Trail of Important Glass Beads

Nick is reading up on the history of human rights. He goes to the library and uses a pallet to search for information on the Enlightenment. He finds an article about Rousseau's Second Discourse, he places the tip of a dropper against the address bar on the browser and slurps up the URL. He grabs a bead from a rack next to the pallet and uses the dropper to squirt the URL into it. He then walks over to the canvass, a 120 inch projected display on the wall. There are some beads arranged on the right side of the canvass already, someone else's trail. Nick sticks his bead on the canvass where there is some open space, it attaches magnetically and becomes illuminated by the projection. Keywords from the article flow around the bead and related websites and books shift in and out of view surrounding it. He walks back to the pallet and examines some of the related websites which are listed there as well. He slurps up a couple web pages about the Declaration of Independence, a reference to a book about the French revo-

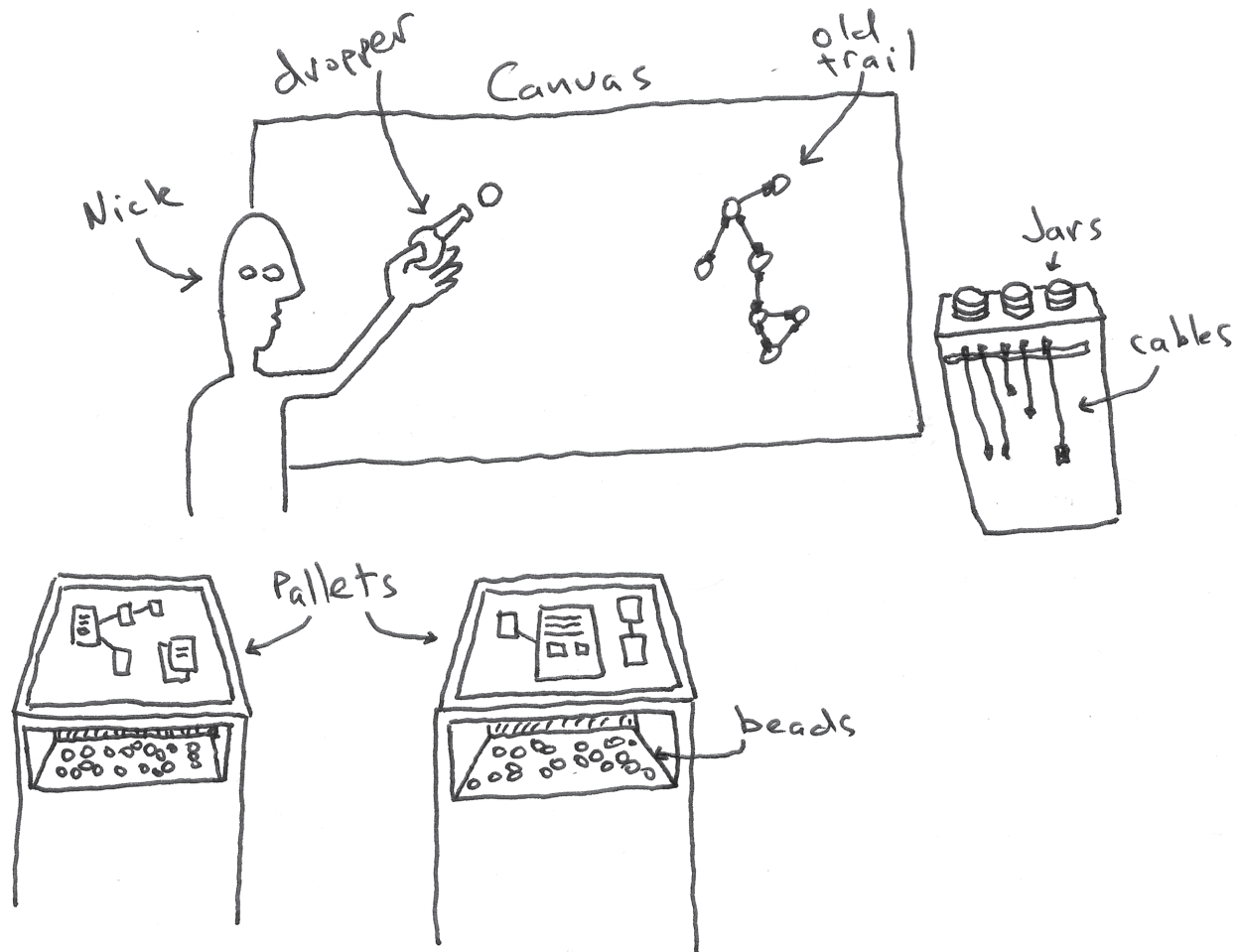


Figure 22. Trail of Important Glass Beads interface components.

lution, and some biographical information on Thomas Jefferson and adds them each to separate beads. Walking back to the canvas he arranges the beads around the Rousseau reference, using some pieces of cable to attach them and creates a subgroup for the parts about Jefferson. He then points the dropper at his face and squeezes the bulb, as he slowly releases it it records audio, he talks about the relationship between two of the beads and then squirts his new media file onto the cable that connects them, storing it there. Once he is done he holds a dropper a foot or so back from his trail to select the entire thing and slurps it up then squirts it into a jar so he can take it with him. He hops on a train home and while thumbing the empty space in his new jar he decides to add a crumpled up American flag to it when he gets home, it is a symbol that has meaning for him and will remind him of the digital trail that he has created.

Sketch 2: Cafe of Tomorrow

Lashanda is visiting San Francisco for the first time. She wants to see what the local culture is like, so she decides to spend the day hanging out in a cafe, but which one? She goes online and looks at a map of her neighborhood. She sees that there are five cafes within walking distance. One is called “Cognitive Dissonance”, the name reminds her of something so she clicks on it to find out what it’s like. She navigates through the recent public discourse from the cafe and sees that the patrons are a lively lot. There are questions about the history of the civil rights movement, papers on human rights abuse, and pictures of lolcats (cats with funny misspelled slogans in them). She loves lolcats, and has a few pictures of her own kitten (a black house-cat named Friday) that she wants to add these to the conversation. Since she has to be at Cognitive Dissonance in person to add information to its network at the level she desires, she decides to make the trip—she likes leftist crowds anyway.

When she gets into the cafe she likes it immediately. The furniture is wooden and well used, the atmosphere is dark and inviting. She notices a few urls flow by with clusters of images surrounding them. It looks like most of the people to the rear of the cafe are digging through the web, their public trails float up the wall next to them, across the ceiling and then get lost in the matrix of connected ideas and conversations displayed on the main wall. Lashanda makes her way to a seat and notices some images of her home town, Cambridge Massachusetts, float by. She is intrigued so as she sets up her laptop she gestures out to the images and pulls their shadow across the wall onto her private display. The pictures are from the street she grew up on! An older woman sitting near her grabs her attention, “are you from Cambridge?” she asks. They talk for a moment about their fair city, her new friend is a Cantabridgian as well, she was the one who had added the images.

Lashanda spends a few minutes reading emails, since she is comfortable being fairly public she allows the aura of her messaging service to illuminate the walls around her, displaying some of her contact information and public profile in case anyone nearby is interested. The content of the emails though, remains quite private. Then she finds the pictures of Friday that she wants to share and gestures at her laptop in a manner that causes the photos to spill out onto the table she is seated at. Pushing her laptop away she uses her fingers to sort through the images on the table. This cafe has some nice photo manipulation tools available to its pa-

trons so she heads to the counter to buy a double Americano and grab a set of hand tools. Back at her table she fits a loupe into her right eye and crouches over a picture of Friday with a leaf of lettuce poking out of her mouth. Through the loupe the pixels in the image float off the table surface, the least saturated pixels rising highest. She switches a graphical widget on the table surface with her finger to select brightness instead and then uses a small piece of squishy metal to knock some pixels lower, the way that she squeezes the metal determines how much spread and noise to hit the pixels with.

After fixing the image, she adds the slogan: “Can I haz pikul meow?” She flicks the image towards the main wall and walks over to stand in front of it so that she can connect her picture with related materials. There are a ton of lolcat images here! She finds some that she thinks are funny and draws lines connecting hers in with them using her finger. As she selects various images and references on the wall the layers of older conversations shimmer behind today’s content, some moving forward towards the surface of the wall as they become more relevant, others becoming clear and displaying their web of connections as other patrons in the cafe continue to generate bookmarks that splash into this local pool of information.

Sketch 3: Seamless Creativity

Alia sits with her team of ten designers around a large display table. They are brainstorming on ideas for the look and feel of a feature film project. On the walls around the table are stills and videos from other films like *The Royal Tennenbaums*, *Rumble Fish*, and *Barry Lindon*, films with thick atmosphere. Humbert, one of Alia’s most creative employees is frustrated, he stands up. “No, we are going in the wrong direction!” He presses his palms together in front of his chest firmly, but before he spreads them wide to clear the imagery off of all their screens Alia grabs the screen manager tool on the table. As Humbert hands move apart the screen manager begins to rotate but Alia stops it from clicking over—stopping the screens from clearing. The screen manager has five lenses mounted on a rotating base, each lens is a temporary storage space for a complete screen layout (more can be stored by dumping these layouts into a graphical widget on the table).

“Wait a second Humbert, I feel like we are onto something here, why do you want to change course?” says Alia.

“Because I have an idea, just bear with me,” he replies. The emotion is gone from his face and now he seems excited. Alia complies and rotates the lens on the screen manager, as the next blank lens clicks into position all of the images, video, text, notes, and other accumulated referential debris from their three-hour meeting fades away leaving the clean, blue patterned wallpaper in it’s place.

Humbert makes a grabbing motion in front of his laptop, its screen is not visible to the group. He drags something out of the machine and throws it onto the far wall, it’s an unfamiliar movie clip. Humbert motions to the screen and it starts playing, it’s a puppet animation, very dark, some dolls are dancing around another doll with a cracked and intense papier-mache face.

“This is the Brother’s Quay, Street of Crocodiles” says Humbert. Then he throws up two more clips onto the wall. One’s from Shohei Imamura’s new wave classic *The Pornographers*, the other is a scene from *The Chimes of Big Ben*, the second episode of the 1970 British television series *The Prisoner*. He then goes on to describe the depth of emotion he finds in each of these clips and how the art direction contributed to their power. But Alia is struck with another idea. As Humbert talks she picks up a clear frame of glass from the table and points the back side of her stylus at it. This projects a square editing space on the wall where the movies are playing, she adjusts it’s size, skew, and location by changing the spatial relationship between the stylus, frame, and wall. When she has selected the area she wants to draw within (large enough to encompass all three movie clips, but small enough so that her lines wont be huge) she squeezes the stylus to lock it and then moves the frame to a drawing position and begins to write.

Humbert has stopped talking and the room has begun to pay attention to Alia. She circles a section of wall in the Quay Brother’s film, pausing the playback on that frame with an outstretched palm in the air. She tosses a collection of websites from the browser on her laptop on the side wall, sorting through them quickly using swimming and grasping motions in the air, she finds one she likes. With some deft hand-finger motions she has isolated an image in the website, enlarged it and dragged it onto the wall near the Quay film. It’s a type of material sold by McMaster Carr, a large online retailer. She writes some numbers next to it, connects it to the circled wall, draws another connection to the scenery near Patrick McGoohan in the *The Prisoner* then looks around the room expectantly; her eyes half-closed.

CONCLUSION

The preceding sketches could each be implemented within a laboratory environment in under a year's time. We are on the border between science fiction and reality. For developers of new technologies the question "what can I make" is no longer as important as "what should I make". The decisions that are made when creating new products, new software, new research projects will shape the way that we relate to each other as human beings sharing air and water on a single planet for centuries to come. It is now even more important that we understand as best we can how technologies and interfaces function within their larger social context.

Through the pages of this thesis I have presented my research in understanding human-computer interaction. I have tried to look through as many perspectives as possible to reach this understanding—by conducting research in social science, philosophy, and media theory; developing new interfaces; constructing theoretical frameworks; experimenting; and designing.

Although the concept of external legibility still requires more investigation, and true seamlessness is still far from reach, the change in value systems represented by moving from thinking about computing technologies like laptops or cell phones as devices for accomplishing proscribed tasks to thinking about them as malleable information conduits existing in physical and virtual spaces within a larger ecosystem of people and things is quite substantial.

The transition into the information age has been traumatic, perhaps more so than we really understand. New theories like Reality-Based Interaction, and technologies like Tangible User Interfaces show great promise. It is my hope that continued research along these lines and across disciplines can help bring people together and that new perspectives on human-computer interaction can help create interfaces that support the highest bandwidth communication channels that are known to exist—the unmediated interface between two human beings. It is my hope that this small thesis will help us along the path towards a truly valuable goal: mending fractured spaces.

BIBLIOGRAPHY

- [**Afghan eXplorer**] Afghan eXplorer. see <http://www.usatoday.com/tech/columnist/2002/03/20/maney.htm>
- [**Agre 1997**] Agre, Philip E. 1997. "Computation and Human Experience." Cambridge: Cambridge UP.
- [**Aharony 2008**] Aharony, N., Lippman, A., and Reed, D. P. 2008. "Demonstrating SnapN'Share: Content Sharing Over a Socially Informed, Hyper-Local, Face-to-Face Networking Platform." In *IEEE CCNC'08*.
- [**Ananny 2002**] Ananny, M. 2002. "Supporting children's collaborative authoring: practicing written literacy while composing oral texts." In *Proc. Computer Support for Collaborative Learning*. Lawrence Erlbaum Assoc. (2002), 595–596.
- [**Anderson 2006**] Anderson, C. 2006. "The Long Tail: Why the Future of Business Is Selling Less of More." Hyperion.
- [**Angesleva 2003**] Angesleva, J., Oakley, I., Hughes, S. & O'Modhrain, S. "Body Mnemonics: Portable device interaction design concept." In *proc. UIST '03*, Vancouver, Canada.
- [**Arias 1997**] Arias, E., Eden, H. and Fisher, G. 1997. "Enhancing communication, facilitating shared understanding, and creating better artifacts by integrating physical and computational media for design." In *Proc. of DIS, '97*.

- [Baudel 1993]** Baudel, T. and Beaudouin-Lafon, M. 1993. Charade: remote control of objects using free-hand gestures. *Commun. ACM* 36, 7 (Jul. 1993), 28-35.
- [Behar]** The Jawbone 2 designed by Yves Behar. <http://us.jawbone.com/>. Accessed July 2008
- [Bellotti 2002]** Bellotti, V., Back, M., Edwards, W.K., Grinter, R.E., Henderson, A. and Lopes, C. Making Sense of Sensing Systems: Five Questions for Designers and Researchers *Proc. CHI 2002*, ACM Press, 2002, 415-422.
- [Benford 2005]** Benford, S., Schnadelbach, H., Koleva, B., Anastasi, R., Greenhalgh, C., Rodden, T., Green, J., Ghali, A., Pridmore, T., Gaver, B., Boucher, A., Walker, B., Pennington, S., Schmidt, A., Gellersen, H. and Steed, A. Expected, sensed, and desired: A framework for designing sensing-based interaction. *ACM Transactions Computer-Human Interaction*, 12 (1). 3-30. 2005
- [Blackwell 2006]** Blackwell, A.F. The reification of metaphor as a design tool. *ACM Trans. Comput.-Hum. Interact.* 13 (4). 2006
- [Bowman 2004]** Bowman, D.A., Kruijff, E., LaViola, J.J. and Poupyrev, I. *3D User Interfaces: Theory and Practice*. Addison-Wesley/Pearson Education, 2004.
- [Bush 1945]** Bush, V. As We May Think. *The Atlantic Monthly*, 176, 1 (July 1945), 641--649.
- [Buxton 2008]** Multi-Tocuh Systems that I Have Known and Loved. <http://www.billbuxton.com/multitouchOverview.html>. Accessed July 2008
- [C4FCM]** MIT Center for Future Civic Media. <http://civic.mit.edu/>. Accessed July 2008.
- [Caron 2007]** Caron, A. H., and Caronia, L., *Moving Cultures | Mobile Communication in Everyday Life*. McGill-Queen's University Press, 2007.
- [Castells 1996]** Castells, M. 1996. *The Rise of the Network Society*. Blackwell Publishers.

- [Castells 2004]** Castells, M., 'Informationalism, Networks, and the Network Society: A Theoretical Blueprint' in Manuel Castells (Ed.) *The Network Society: a cross-cultural perspective*, Northampton, MA: Edward Elgar, 2004.
- [Chang 2007]** Chang, A., Gouldstone, J., Zigelbaum, J., and Ishii, H., Simplicity in interaction design. In *proc. TEI '07*. ACM Press, 135-138.
- [chicagocrime.org]** Chicagocrime.org. <http://www.chicagocrime.org/>. 2005-2008.
- [Cohen 2000]** Cohen, S., Brissette, I., Doyle, W. and Skoner, D. (2000) 'Social Integration and Health: The Case of the Common Cold', *Journal of Social Structure*, vol. 1, no. 3, pp. 1-7.
- [Cohen 2003]** Cohen, J. E. Human Population: The Next Half Century. *Science*, 302, 5648 (November 14, 2003), 1172-1175.
- [Dix 1993]** Dix, A., Finlay, J., Abowd, G. and Beale, R., (1993), *Human-Computer Interaction*, Prentice Hall,
- [Djajadiningrat 2002]** Djajadiningrat, T., Overbeeke, K., and Wensveen, S. 2002. But how, Donald, tell us how?: on the creation of meaning in interaction design through feedforward and inherent feedback. In *Proc. DIS '02*. ACM, New York, NY, 285-291.
- [Dobson 2004]** Dobson, K., Blendie. *DIS '04*, ACM Press 2004, New York, NY, 309-309.
- [Dourish 2001]** Dourish, P. *Where The Action Is: The Foundations of Embodied Interaction*, MIT Press, Cambridge, Mass., 2001.
- [Dunne 1997]** Dunne A. & Raby F., *Fields and Thresholds, Doors of Perception 2*, 1997.
- [Dunne 2001]** Dunne, A. & Raby, F. *Design Noir: The Secret Life of Electronic Objects*. Berkhauser, Berlin, Germany, 2001.
- [Engelbart 1962]** Engelbart, Douglas C., *Augmenting Human Intellect: A Conceptual Framework*, Summary Report, SRI Project No. 3578, Stanford Research Institute (now SRI International), Menlo Park, CA, October 1962.

- [**EveryBlock**] EveryBlock. <http://chicago.everyblock.com/>. Accessed July 2008.
- [**Feenberg 1999**] Feenberg, A. (1999) *Questioning Technology*. Routledge, London and New York.
- [**Fishkin 2000**] Fishkin, K., Gujar, A., Harrison, B., Moran, T., and Want, R. Embodied user interfaces for really direct manipulation. In *Communications of the ACM*, 43, 9, (Sep. 2000), ACM Press (2000), 74–80.
- [**Fishkin 2004**] Fishkin, K.P. A taxonomy for and analysis of tangible interfaces. *Personal Ubiquitous Computing*, 8 (5). 2004.
- [**Fitzmaurice 1995**] Fitzmaurice, G. W., Ishii, H. and Buxton, W. A. S. Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems (Denver, Colorado, United States, 1995)*. ACM Press/Addison-Wesley Publishing Co.
- [**Fjeld 2007**] Fjeld, M., et al. Tangible user interface for chemistry education: comparative evaluation and redesign *Proc. CHI*, 2007, 805-808.
- [**FTIR**] Multi-Touch Interaction Research. <http://www.cs.nyu.edu/~jhan/ftirtouch/index.html>. Accessed July 2008.
- [**g-speak**] g-speak. <http://www.oblong.net/>. Accessed July 2008.
- [**Gates**] Gates looks into PC's future as career shift approaches. USA Today. http://www.usatoday.com/tech/techinvestor/corporatenews/2008-06-24-gates-microsoft_N.htm. Accessed July 2008
- [**GIA**] GIA, see http://www.findarticles.com/p/articles/mi_qn4176/is_20030705/ai_n14557093
- [**Gibson 1977**] Gibson, J. J. 1977. *The Theory of Affordances*. In *Perceiving, Acting, and Knowing*, Eds. Robert Shaw and John Bransford.
- [**Gitelman 2006**] Gitelman, L. *Always Already New: Media, History, and the Data of Culture*. MIT Press, Cambridge Massachusetts, 2006.

- [Goldberg 1932]** Goldberg, E. 1932. Methods of photographic registration. *British Journal of Photography*, 79: 533-534.
- [Granoveter 1974]** Granovetter, M. 1974 *Getting a Job*, Harvard University Press, Cambridge, MA.
- [Greenberg 2001]** Greenberg, S. and Fitchett, C. 2001. Phidgets: easy development of physical interfaces through physical widgets. In *Proc. UIST '01*. ACM, New York, NY, 209-218.
- [Gutwin 1998]** Gutwin, C. and Greenberg, S., Design for individuals, design for groups: tradeoffs between power and workspace awareness, In *Proc. CSCW 1998*, ACM Press (1998), 207–216.
- [Hall 2006]** Hall, P., and Pain, K. *The Polycentric Metropolis*. London: Earthscan. 2006.
- [Hampton 2003]** Hampton, K. and Wellman, B. (2003) 'Neighboring in Netville: How the Internet Supports Community and Social Capital in a Wired Suburb', *City and Community*, vol. 2, no. 3, pp.277-311.
- [Hampton 2007a]** Hampton, K. (2007) 'Neighborhoods in the Network Society: The e-Neighbors Study', *Information, Communication and Society*, vol. 10, no. 5.
- [Hampton 2007b]** Hampton, K., Trachtenberg C., Livio, O. WiFi and Public Space, a Poor Interface? An Empirical Study of Wireless Internet Use and Sociability. Unpublished manuscript.
- [Hampton 2008]** Hampton, K., and Gupta, N. Community and Social Interaction in the Wireless City: Wi-Fi use in Public and Semi-Public Spaces. To be published in *New Media & Society*.
- [Haraway 1991]** Haraway, D. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century," in *Simians, Cyborgs and Women: The Reinvention of Nature* (New York; Routledge, 1991), pp.149-181.
- [Harper 2008]** Harper, R., Rodden, T., Rogers, Y., and Sellen, A. Editors. *Being Human: Human-Computer Interaction in the Year 2020*. Microsoft Research, Cambridge, England. 2008.

- [Harry 2008]** Harry, D., Gutierrez, D., Green, J., and Donath, J. 2008. Backchan.nl: integrating backchannels with physical space. In CHI '08 Ext. Abstracts. ACM, New York, NY, 2751-2756.
- [Hayes 1983]** Hayes, P.J. The second naive physics manifesto Cognitive Science Technical Report URCS-10, University of Rochester, 1983.
- [Hayles 1999]** Hayles, K. K. 1999. How We Became Posthuman : Virtual Bodies in Cybernetics, Literature, and Informatics. University Of Chicago Press.
- [Heath 2007]** Heath, C. and Heath, D. Made to Stick: Why Some Ideas Survive and Others Die. Random House, 2007.
- [Hesse 1943]** Hesse, H., The Glass Bead Game (Magister Ludi), Henry Holt and Company, LLC, New York, 1943.
- [Hollan 2000]** Hollan, J., Hutchins, E., and Kirsh, D., Distributed cognition: towards a new foundation for Human-Computer Interaction research, In TOCHI, 7, 2, (June 2000), 174–196.
- [Holmquist 1999]** Holmquist, L., Redström, J. and Ljungstrand, P. Token-Based Access to Digital Information, 1999.
- [Hornecker 2002]** Hornecker, E., Understanding the benefits of graspable interfaces for cooperative use. In Proc. of Cooperative Systems Design 2002, 71–87.
- [Hornecker 2006]** Hornecker, E. and Buur, J., Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. in Proc. of CHI 2006, (2006), ACM, 437-446.
- [Hornecker 2007]** Hornecker, E., Marshall, P., and Rogers, Y. 2007. From entry to access: how shareability comes about. In Proc. DPPI '07. ACM, New York, NY, 328-342
- [Hutchins 1986]** Hutchins, E.L., Hollan, J.D. and Norman, D.A., Direct Manipulation Interfaces. DA Norman, SW Draper eds. User Centered System Design: New Perspectives on Human-computer Interaction, Lawrence Erlbaum, Hillsdale, N.J., 1986, 87-124.

- [Ikeda 2003]** Ikeda, Y., et al. TOOL DEVICE: Handy Haptic Feedback Devices Imitating Every Day Tools, HCI International, 2003, 661-665.
- [Ishii 1997]** Ishii, H. and Ullmer, B., Tangible bits: towards seamless interfaces between people, bits and atoms. In CHI, 1997.
- [Ishii 1999]** Ishii, H., et al. musicBottles SIGGRAPH Conference abstracts and applications, 1999, 174.
- [Ishii 2003]** Ishii, Y., Nakanishi, Y., Koike, H., Oka, K., and Sato, Y., EnhancedMovie: movie editing on an augmented desk. In Adjunctive Proc. of UbiComp 2003.
- [Iyengar 2000]** Iyengar, S. S., & Lepper, M. 2000. When Choice is Demotivating: Can One Desire Too Much of a Good Thing? *Journal of Personality and Social Psychology*, 79, 995-1006.
- [Jacob 2008]** Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., and Zigelbaum, J. 2008. Reality-based interaction: a framework for post-WIMP interfaces. In Proc. CHI '08. ACM, New York, NY, 201-210.
- [Jenkins 2006]** Jenkins, H. *Convergence Culture: Where Old and New Media Collide*. NYU Press, New York, NY, 2006
- [Jordà 2007]** Jordà, S., Geiger, G., Alonso, M., and Kaltentbrunner, M. 2007. The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In Proc. TEI '07. ACM, New York, NY, 139-146.
- [Kittler 1999]** Kittler, F. A. *Gramophone, Film, Typewriter*. Stanford University Press, 1999.
- [Klemmer 2001]** 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 Proc. UIST '01. ACM, New York, NY, 1-10.
- [Klemmer 2004]** Klemmer, S. R., Li, J., Lin, J., and Landay, J. A. 2004. Papier-Mache: toolkit support for tangible input. In Proc. CHI '04. ACM, New York, NY, 399-406.

- [**Kraut 1998**] Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukhopadhyay, T. and Scherlis, W. (1998) 'Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-Being?' *American Psychologist*, vol. 53, no. 9, pp. 1017-1031.
- [**Kraut 2002**] Kraut, R., Kiesler, S., Boneva, B., Cummings, J., Helgeson, V. and Crawford, A. (2002) 'Internet Paradox Revisited', *Journal of Social Issues*, vol. 58, no. 1, pp. 49-74.
- [**Lifton 2007**] Lifton, J. 2007. *Dual Reality: An Emerging Medium*. PhD. thesis, Massachusetts Institute of Technology.
- [**Löwgren 2007**] Lowgren, J. 2007 Dec 31. Fluency as an Experiential Quality in Augmented Spaces. *International Journal of Design*.
- [**Mackay 1994**] Mackay, W. and Pagani, D., Video Mosaic: laying out time in a physical space. In *Proc. of the Second ACM International Conference on Multimedia*, ACM Press (1994), 165–172.
- [**McLuhan 1994**] McLuhan, M. *Understanding Media: The Extensions of Man*. MIT Press, 1994.
- [**McPherson 2006**] McPherson, M., Smith-Lovin, L. and Brashears, M. E. (2006) 'Social Isolation in America: Changes in Core Discussion Networks over Two decades', *American Sociological Review*, vol. 71, no. 3, pp. 353-375.
- [**Merrill 2005**] Merrill, D., and Maes, P. Invisible media: Attention-sensitive informational augmentation for physical objects (short paper). *Ubicomp 2005*.
- [**Merrill 2007**] Merrill, D., Kalanithi, J. and Maes, P. Siftables: towards sensor network user interfaces *Proc. TEI, 2007*, 75-78.
- [**Milgram 1994**] Milgram, P. and Kishino, F., Augmented reality: A class of displays on the reality-virtuality continuum. In *SPIE*, (1994), 282-292.
- [**Mitchell 2003**] Mitchell, W. J., *Me++ The Cyborg Self and the Networked City*. MIT Press, 2003.

- [**Mohler 2007**] Mohler, B.J., Thompson, W.B., Creem-Regehr, S.H. Willemsen, P., Herbert L. Pick, Jr. and Rieser, J.J. Calibration of locomotion resulting from visual motion in a treadmill-based virtual environment. *ACM Trans. Appl. Percept.*, 4 (1). 4. 2007.
- [**Nacenta 2005**] Nacenta, M.A., Aliakseyeu, D., Subramanian, S. and Gutwin, C. A comparison of techniques for multidisplay reaching *Proc. CHI, 2005*, 371-380.
- [**Newman 2006**] Newman, W. and Smith, E. L. 2006. Disruption of meetings by laptop use: is there a 10-second solution?. In *CHI '06 Ext. Abstracts*. ACM, New York, NY, 1145-1150.
- [**News Challenge**] John S. and James L. Knight Foundation News Challenge. <http://www.newschallenge.org/>. Accessed July 2008.
- [**Norman 1988**] Norman, D. *The Psychology of Everyday Things*. Basic Books, 1998.
- [**Patten 2001**] Patten, J., Ishii, H., Hines, J., and Pangaro, G. 2001. Sensetable: a wireless object tracking platform for tangible user interfaces. In *Proc. CHI '01*. ACM, New York, NY, 253-260.
- [**Patten 2002**] Patten, J., Recht, B., Ishii, H. Audiopad: A Tag-based interface for Musical Performance, in *Proceedings of Conference on New Interface for Musical Expression (NIME '02)*, ACM Press (2002).
- [**Patten 2007**] Patten, J. and Ishii, H. 2007. Mechanical constraints as computational constraints in tabletop tangible interfaces. In *Proc. CHI '07*. ACM, New York, NY, 809-818.
- [**Peltonen 2008**] Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T. Evans, J., Oulasvirta, A. & Saarikko, P. 2008. "It's mine, don't touch": Interactions at a large multi-touch display in a city Center. In *Proc. CHI'08*. ACM Press, New York, pp. 1285-1294.
- [**Putnam 1995**] Putnam, R. D. "Bowling Alone: America's Declining Social Capital." *Journal of Democracy* 6, January 1995.

- [Putnam 2000]** Putnam, R. D. *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster, New York, NY, 2000.
- [Raffle 2004]** Raffle, H. S., Parkes, A. J., and Ishii, H. 2004. Topobo: a constructive assembly system with kinetic memory. In Proc. CHI '04. ACM, New York, NY, 647-654.
- [Rayward 1975]** Rayward, W. Boyd. *The Universe of Information: The Work of Paul Otlet for Documentation and International Organization*. (FID 520). Moscow: VINITI, 1975.
- [Rekimoto 1997]** Rekimoto, J. Pick-and-drop: a direct manipulation technique for multiple computer environments Proc. UIST, 1997.
- [Schweikardt 2008]** Schweikardt, E. and Gross, M. D. 2008. The robot is the program: interacting with roBlocks. In Proc. TEI '08. ACM, New York, NY, 167-168.
- [SCI]** SCI Social Capital Inc. <http://www.socialcapitalinc.org/>. Accessed July 2008.
- [Second Life]** Second Life. <http://secondlife.com/>. Accessed July 2008.
- [Sengers 2005]** Sengers, P., Boehner, K., David, S. and Kaye, J.J. Reflective design. in Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility, (Aarhus, Denmark, 2005), ACM Press, 49-58.
- [Shaer 2004]** Shaer, O., Leland, N., Calvillo-Gamez, E. H., and Jacob, R. J., The TAC paradigm: specifying tangible user interfaces. In *Personal and Ubiquitous Computing* 8, 5, (Sep. 2004), 359–369.
- [Sharlin 2004]** Sharlin, E., Watson, B., Kitamura, Y., Kishino, F. and Itoh, Y. On tangible user interfaces, humans and spatiality. *Pers. and Ubiqu. Computing*, 8 (5). 338-346. 2004.
- [Smith 2007]** Smith, J. and MacLean, K. Communicating emotion through a haptic link: Design space and methodology. *Int. J. Hum.-Comput. Stud.*, 65 (4). 2007 376-387.
- [Suzuki 1993]** Suzuki, H. and Kato, H., AlgoBlock: A tangible programming language, a tool for collaborative learning, In Proc. of 4th European Logo Conference 1993, 297–303.

- [**Tangible Video Browser**] Tangible Video Browser. <http://tangible.media.mit.edu/projects./tvb/>.
- [**TEI**] Conference of Tangible and Embedded Interaction. <http://www.tei-conf.org/>. Accessed July 2008.
- [**Terrenghi 2005**] Terrenghi, L. Design of Affordances for Direct Manipulation of Digital Information in Ubiquitous Computing Scenarios. *Smart Graphics 2005*, 198-205.
- [**Terrenghi 2007**] Terrenghi, L., Kirk, D., Sellen, A. and Izadi, S. Affordances for manipulation of physical versus digital media on interactive surfaces *Proc. CHI, 2007*.
- [**Ullmer 1998**] Ullmer, B., Ishii, H., and Glas, D., mediaBlocks: physical containers, transports, and controls for online media, In *Proc. SIGGRAPH 1998*, ACM Press (1998), 379–386.
- [**Ullmer 2001**] Ullmer, B. and Ishii, H. Emerging Frameworks for Tangible User Interfaces. Carroll, J.M. ed. *Human- Computer Interaction in the New Millenium*, Addison-Wesley/ACM Press, Reading, Mass., 2001.
- [**Ullmer 2005**] Ullmer, B., Ishii, H. and Jacob, R. Tangible Query Interfaces: Physically Constrained Tokens for Manipulating Database Queries, *TOCHI 2005*.
- [**Ullmer 2005**] Ullmer, B., Ishii, H., and Jacob, R. J., Token+constraint systems for tangible interaction with digital information, In *TOCHI 2005*, ACM Press (2005), 81–118.
- [**Underkoffler 1999**] Underkoffler, J. and Ishii, H. Urp: A Luminous-Tangible Workbench for Urban Planning and Design *Proc. ACM CHI, 1999*, 386-393.
- [**Van Dam 1997**] Van Dam, A. Post-WIMP user interfaces. *Commun. ACM*, 40 (2). 1997. 63-67.
- [**Vaucelle 2005**] Vaucelle, C., Africano, D., Davenport, G., Wiberg, M., and Fjellstrom, O., Moving pictures: Looking Out/Looking In, In *SIGGRAPH 2005*, ACM Press (2005).
- [**Vinson 1999**] Vinson, N.G. Design guidelines for landmarks to support navigation in virtual environments *Proc CHI, ACM Press, 1999*, 278-285.

- [Want 1999] Want, R., Fishkin, K.P., Gujar, A., Harrison, B.L.: Bridging physical and virtual worlds with electronic tags. In: Proceedings of CHI '99, New York, NY, USA, ACM Press (1999) 370–377.
- [Weiser 1991] Weiser, M., The Computer for the 21st Century. 1991
- [Weiser 1997] Weiser, M. and Brown, J. S. 1997. The coming age of calm technology. In Beyond Calculation: the Next Fifty Years, P. J. Denning and R. M. Metcalfe, Eds. Copernicus, New York, NY, 75-85.
- [Wellman 2003] Wellman, B., Quan-Haase, A., Boase, J., Chen, W., ampton, K., de Diaz, I. I. and Miyata, K. (2003) 'The Social Affordances of the Internet for Networked Individualism', Journal of Computer-Mediated Communication, 8, 3.
- [White 2007] White, R.W., et al. Exploratory search and HCI: designing and evaluating interfaces to support exploratory search interaction CHI '07 (extended abstracts), 2007, 2877-2880.
- [Wifitti] Wifitti by LocaModa. <http://www.wifitti.com/>. Accessed July 2008.
- [Yellow Arrow] Counts Media (2004). Yellow Arrow. <http://yellowarrow.net/>. Accessed July 2008
- [Zhang 2007] Zhang, Y., Fishbach, A., and Kruglanski, A. W. The Dilution model: How Additional Goals Undermine the Perceived Instrumentality of a Shared Path. Journal of Personality 2007
- [Zigelbaum 2007] Zigelbaum, J., Horn, M., Shaer, O. and Jacob, R. The tangible video editor: collaborative video editing with active tokens. In Proc. TEI '07, ACM Press (2007), 43-46.
- [Zigelbaum 2008a] Zigelbaum, J., Kumpf, A., Vazquez, A., and Ishii, H. Slurp: Tangibility, Spatiality, and an Eyedropper In Ext. Abstracts, alt.chi CHI '08, ACM Press (2008), 2565-2574.
- [Zigelbaum 2008b] Zigelbaum, J., Chang, A., Gouldstone, J., Jen Monzen, J., and Ishii, H. SpeakCup: Simplicity, BABL, and Shape Change. In Proc. TEI '08, ACM Press (2008), 145-146.